# Open Data and the Philippine AI Governance and Development[1]

Peter Sy
University of the Philippines

**Abstract.** Data that can be freely used, re-used, or redistributed is essential to scientific research and public governance in general—and, in particular, to the development of artificial intelligence (AI) in the country. With AI poised to drive the Fourth Industrial Revolution, this paper briefly outlines the relationship between open data and AI governance and development. Particular attention is given to the potential impact of open data on AI-powered innovations across sectors in Philippine society. What governance principles inform the use of open data in the development of AI? What specific AI governance elements would lend weight to making "public data open, freely available, and downloadable in digestible format" (Task 6 of Philippine AI Roadmap)? What benefits, opportunities and risks might open data pose on the application of AI in sectors like social media and the justice system? Beyond the current state of the Philippine Open Data Initiative, what concepts, practices, legal and governance frameworks, initiatives, and challenges must be addressed for an effective utilization of open data at least in the public sector deployment of AI?

# Introduction

The advent of the Fourth Industrial Revolution (FIRe) brings to the fore the need for massive amounts of data to power the revolution. Nine (9) major areas (i.e., big data and analytics, simulations, system integration, internet of things [IoT], autonomous robots, cloud computing, cybersecurity, augmented reality, and additive manufacturing) support FIRe (McGinnis, 2020; Piccarozzi et al., 2018). Practically all these areas require or generate data. FIRe is going to be transformative, permeating every facet of our lives and radically disrupting every sector of society (McGinnis, 2020). Klaus Schwab (2016), founder and executive chairman of the World Economic Forum, maintains that FIRe is going to be a catalyst for raising global income levels and improving the quality of life. AI is poised to drive FIRe (Department of Trade and Industry, 2021); data is its fuel. Much of such data will have to be "open".

This paper telescopes the range of issues in AI development and governance that relate to open data. These issues include AI governance principles, governance elements, risk-based AI governance, and AI application areas (including health and social care, justice system, social media, and utilities). It also maps other related areas that need immediate attention, including AI legal frameworks as well as AI practices, initiatives and challenges, for an effective utilization of open data at least in the public sector deployment of AI. However ambitious the coverage of this paper appears, I will limit myself to the singular goal of demonstrating that AI development and

---

[1] Correspondence: psy@up.edu.ph. Special thanks to Yousef Concepcion, Therese Litonjua, and Rica Mae Ramirez (University of the Philippines) for their assistance to this paper; Rhoanna Marie Dizon for the illustrations.

governance are practically impossible without open data. This paper leverages on an earlier Philippine stakeholders-consultation work on data governance commissioned by the Department of Trade and Industry (DTI).[2]

Firstly, what is “open data”? It is data that one can freely use, re-use, share, or redistribute (World Wide Web Foundation, 2016). These days such data have to be freely available online. This is especially significant, as AI algorithms can be made to crawl online materials without human intervention, thus increasing AI’s productivity and aggregative power. Open data is essential to governance in general and, in particular, to AI governance and development. To drive democracy and development, robust data is needed (World Wide Web Foundation, 2016).

In the 2016 Open Data Barometer’s ranking of countries based on their open data readiness, implementation, and impact of government, 55% of the 99 countries surveyed (including the Philippines) have open-data initiatives (World Wide Web Foundation, 2016). However, the Barometer later found a lack of critical data in open datasets available. Of the government datasets surveyed, 90% are still closed and the 10% of data that is considered open is of “poor quality.”

The 2017 Open Data Barometer reports a decelerated and stalled state in most governments’ commitment to their open data initiatives. Findings show that due to misplaced priorities and untimely policies, most governments are not able to meet the basic International Open Data Charter Principles (World Wide Web Foundation, 2017). These principles set the “gold standard” for publishing data. Adopted by seventeen (17) countries in 2015, these six (6) principles providing the benchmark for best practices in open data include (Open Data Charter, 2015):

1. Open by Default. The full access to and the use of government data is of significant economic and social value. By default, government data ought to be made available to the public.

2. Timely and Comprehensive. Open data is only valuable if it is relevant. For such data to be relevant and valuable to stakeholders, it also has to be comprehensive, accurate, and of high quality.

3. Accessible and Usable. Publicly released data must be easily discoverable and accessible. Free of any bureaucratic or administrative barriers, it should easily be accessible to people.

4. Comparable and Interoperable. The value of data in part lies in its comparability to other data within and between sectors, across geographic locations, and over time. To facilitate comparability, interoperability, traceability, and efficient reuse, structured and standardized formats for data publication are necessary.

5. For Improved Governance and Citizen Engagement. Opening data to stakeholders would enable governments and public institutions to make better informed decisions.

---

[2] See ai-ph.org/devframe

Open data would nurture an environment of trust and good governance, in line with the obligation to transparency and accountability. Through stakeholder consultations, governments could better prioritize which data to release.

6. For Inclusive Development and Innovation. Open data helps drive inclusive economic and social development. With it, researchers could help identify overlooked social and economic issues. Open data are materials to stimulate creativity and innovation.

The country's push for open data could not be made more urgent than yesterday, especially when AI is touted as potentially contributing US$15.7 trillion to the global economy by 2030 (PricewaterhouseCoopers, 2017). The sooner the country makes open data consequential in governance and innovation, the better is its chances of being able to take advantage of FIRe and AI. Open Data in the Philippines has made little progress, relative to what it could potentially contribute to the country's knowledge economy. While there are already initiatives on making more data more open and publicly accessible, work still needs to be done to align the country's open data initiatives with International Open Data Charter Principles.

Open data "means better science" (Molloy, 2011); also, better AI, as data is the "lifeblood" of Artificial Intelligence (Data Europa, 2020). Open data unlocks the potential of AI applications, allowing for the training and creation of AI models. To accelerate the development of AI as well as the development of cures for diseases and better responses to climate and other crises, data needs to be open (Austin et al., 2021). Exactly how can open data accelerate the development of AI? Two ways: (1) by reducing data monopolies and (2) by saving time and expenses related to the process of data collection. Both promote greater efficiency, allowing more time for actual problem solving and innovation (Austin et al., 2021).

The current state of open data adoption in the country leaves much to be desired. While the Philippines in 2011 became one of the founding partners of the Open Government Partnership (OGP), an organization which pushes for "transparent, participatory, inclusive and accountable governance" (Open Government Partnership, 2022), there is a long way to go to have enough open data useful to AI researchers. Data that are "timely and comprehensive", "comparable and interoperable" remain a challenge even with the establishment of the "Open Data Philippines" (ODPH) initiative, having a single portal for collections of government data that people could search, access, and use freely for research, reports, projects, or applications (Saxena, 2018). Although the ODPH announced that portal now has datasets from at least 40 government agencies (Open Data Philippines, 2017), only five (5) of them are visible in the platform (i.e., Department of Foreign Affairs, Department of Information and Communications Technology, Department of Interior and Local Government, Department of Health, Mines and Geosciences Bureau - Department of Environment and Natural Resources), and only the last two agencies have available datasets in their respective sections. Browsing through the datasets, only data on Covid-19 and mining statistics are available. The other topics are mostly empty (Open Data Philippines, 2022).

However, there appears to be other open data initiatives in the country, good enough to earn a decent ranking (18th out of 187 participating countries) in the Open Data Inventory (ODIN), rising from the 40th in 2018 (Mapa, 2020). Among Southeast Asian countries, it ranks second to

Singapore. The Open Data Barometer (Brandusescu et al., 2017) puts the Philippines in 22nd (out of 114 countries), which is also an improvement from 47th spot in 2013 (Davies, 2013).

Still, much work needs to be done. ODPH's own assessment indicates that the open data in the country is "more supply-driven rather than demand-driven" (Pacis, 2017). AI needs lakes of data, not just buckets. There is little knowledge of the open data initiative, its purposes, and importance. Open data's usefulness is hardly appreciated by the general public. Confounding these challenges are issues of data literacy and digital divide. The government has to provide broader access to the internet, facilitate the advancement of its citizens' technical and digital skills, and increase their awareness of the value of information and data (Pacis, 2017). Incidentally, these are also the same enabling conditions for AI to take root and flourish in the country.

# AI Governance Principles and Open Data

AI governance principles guide the development and utilization of AI (OECD, 2019). Inclusive growth; sustainable development and well-being; human-centered values and fairness; transparency and explainability; and trust—these guiding principles for governance also inform how the development of AI necessitates the use of open data.

AI research and development needs inclusive growth, sustainable development, and well-being. National developmental goals should be entwined with the competitiveness of micro, small, and medium enterprises (MSMEs). These enterprises have to advance the "inclusion of underrepresented populations, reducing economic, social, gender and other inequalities, and [to protect] natural environments" (OECD, 2019). In order for AI to be inclusive of underrepresented populations and reduce inequalities, AI-Solutions Providers must have access to open data that would provide necessary information—of significant volume—regarding these populations.

An AI system also has to be fair and human-centric, where fairness entails treating people with dignity and respect. AI actors should be responsible for implementing mechanisms that do not undermine the human capacity for self-determination. It is also vital for AI systems and algorithms dependent on available data to be reviewed for latent biases. It is, therefore, crucial to evaluate the fairness of an AI system, wherein both data and algorithms are assessed for bias (Personal Data Protection Commission Singapore, 2020). Two sources of bias can be identified from consultations: (1) data fed into the AI systems and (2) the model being utilized. The first source can be traced to "flawed data sampling, in which groups are over- or underrepresented in the training data." Second, algorithms are also at risk based on their training data (Manyika et al., 2019). Understanding the sources of biases necessitates the utilization of open data that is truthful, accurate, and varied. This should allow for the development of fair and human-centric AI systems.

With transparency, Microsoft identifies five (5) components, namely: (i) Full disclosure about the objective of an AI system in decision-making, (ii) Intended purpose of the AI system, (iii) Training data, (iv) Maintenance and assessment, and (v) Ability to challenge and seek redress (Jose, 2022). With training data, in particular, an AI system or an AI-driven model utilizing open data can provide greater transparency. Open data tends to have greater respect for data privacy rights

while providing an automation auditor the opportunity to probe, review, and understand the data used and collected. The use of open data lessens the complications involved in the auditing process, which overall promotes transparency.

Lastly, trust should be central to the adoption of AI systems. In doing so, however, five (5) major challenges should be addressed, including (Lockey, et al., 2021): (i) transparency and explainability, (ii) accuracy and reliability, (iii) automation, (iv) anthropomorphism, and (v) massive data extraction. Addressing these challenges would mitigate the vulnerabilities of AI users and stakeholders, the threats to their privacy, and their potential exposure to biases and discrimination.

# Key AI Governance Elements and Open Data

Operationalizing governance in AI, at least on the level of an organization, revolves around a number of governance elements requiring open data. These elements include (but are not limited to): Governance of Data Assets, Algorithms, and AI Models; Continuous Digital Transformation; Culture of Innovation; and Risk Management and Audit. Fig 1 below shows the governance areas to which these elements belong.

*Figure 1. Lobana's AI Governance Framework (Lobana, 2021)*

| Governance Area | Governance Elements |
|---|---|
| Engaged Board Oversight | Knowledgeable Board |
| | Engaged board |
| Enterprise Leadership & Planning | Competent, Committed, & Collaborative Top Management |
| | Focused AI Strategy & Risk Capital |
| | Enterprise Architecture & Coordination |
| Core AI Technical Elements | Governance of Data Assets |
| | Governance of Algorithms and AI Models |
| | Infrastructure Scalability |
| People & Culture | Strategic People Governance |
| | Culture of Innovation |
| | Change Management & Communication |
| Operational Structures, Processes & Mechanisms | Redesigned Processes |
| | Operational Structures, Policies & Practices |
| | Performance Management |
| | Stakeholder Management |
| Enterprise Risk Oversight | Risk Management & Audit |
| | Data & AI Security |
| | Regulatory Compliance |
| AI Ethics | Embedded AI Ethics |
| | Corporate Social Responsibility |
| Ongoing Evolution | Continuous Digital Transformation |
| | Evolving Holistic System |

**Core AI technical elements** include the governance of data assets, governance of algorithms and AI models, and infrastructure scalability. As Lobana maintains, data is the key to develop AI; the quality and quantity of data “can make or break " AI projects. “One cannot talk about AI without talking about data. Data [is] at the core of AI" (Lobana, 2021). Integral to AI development and training, governance of data assets includes the sourcing of data (Lobana, 2021) wherein open data is an essential part.

AI models can be built from open-source pre-built algorithms that are themselves open data. Big Data Algorithmic Systems (BDAS), which are often based on machine learning and neural networks, particularly rely on Big, Open and Linked Data (BOLD). However, open data that is streamed real-time at high speeds, dynamic, and varied is also crucial. Managing such data and utilizing opportunities for BDAS requires increasingly advanced data governance capabilities (Janssen et al., 2020).

In the evolution of any organization, a major element of data and AI governance requiring open data is continuous digital transformation. In such transformation, data begets data, and employees have to be trained to develop digital skills to deal with new technologies and processes as well as with the onslaught of massive amounts of data. For the government especially, previously analogue government services have to evolve into digital ones. The transformation process is in lockstep with data and AI governance. A common misconception of the digital evolution is that it is just a matter of moving information to a digital platform and automating previously analog processes. Digital transformation requires organizations to take stock of their data, evaluate their structure, assets, and processes while maximizing their value and utilization and generating new insights (Lobana, 2021).

Open data expands opportunities for organizations to achieve digital transformation, including the adoption of digital technologies, acquisition of digital tools, and the alteration of processes that can improve performance and reach (Carrara et al., 2020). Carrara et al (2020) outline three key areas of digital transformation where open data is utilized, namely, in customer experience, operational processes, and business models. Firstly, open data can contribute to customer experience through increasing understanding of customer preferences and demographic information. Information derived from open data, can then optimize AI-driven marketing, sales, and service processes, eventually contributing to growth (Carrara et al., 2020). Second, open data is crucial for process digitization—open data can be used in the management of performance, the addition of further features, and the provision of insight, explainability, and transparency into AI systems and their decision making processes. Finally, open data allows for the development of business models: access to multiple and varied data sources for machine learning can stimulate innovation and can further the development of AI systems and organizations that utilize such. Open data can be used to modify organization practices and processes, particularly in AI-driven businesses, transforming an organization’s practices into more novel, globalized, and competitive business models (Carrara et al., 2020).

Open data fosters a culture of innovation. Lobana (2021) stresses the importance of cultural readiness in the success of AI, and operationalizes a culture of innovation as one that involves experimentation, risk-taking, evidence-based decision-making, team spirit, acceptance of failure, celebration of change, and drive towards innovation—all of which are essential to AI deployment. Arguably, open data is instrumental in fostering these attitudes and processes,

particularly in the aspects of experimentation and evidence-based decision-making. AI development itself is experimental in nature (Lobana, 2021). Accessible data for the purposes of experimentation is one supportive mechanism that can enhance organizational capability in data science (Lobana, 2021). Open data allows for the further reproduction and verification of experiments. It provides an empirical basis for decision making (Yong, 2015).

As decision-making processes go, they are ultimately about risk management where open data is also deemed essential.

# Risk-Based AI Governance and Open Data

AI is fraught with risks. Its governance has to be risk-based, as AI proponents, developers, researchers, and stakeholders seek proactive ways of managing AI's vulnerabilities and potential benefits. Any form of risk analysis requires data. The more open the data informing the decisions of AI researchers and developers, the better is the management of the risks involved.

Traditionally risk management is best embedded in structures and functions internal to organizations. But as designs and operations of technological, organizational, and social systems continue to be driven by artificial intelligence to unprecedented levels, an overall risk-based approach to governance and regulation has to be articulated as well.

Common risks associated with AI, as inventoried by the UK Center for Data Ethics and Innovation (2020), are concerned with data. Some of these risks include, but are not limited to, bias leading to discrimination, effects of low digital/data maturity, erosion of privacy, platform and data monopolies, and low accuracy.

Biases, which often lead to an AI-driven model being discriminatory in its decision-making, is caused by flawed data sampling, in which groups are over- or underrepresented in the training data (Manyika et al., 2019). This also results from a lack of substantial and inclusive data fed into an AI's algorithms. The accuracy of the decisions rendered by an AI system is heavily reliant on the data set used to train it.

As AI systems draw conclusions from large masses of data, significant data volume, data variety and data veracity (the truthfulness of data) are crucial for an AI system to function impartially and accurately (Data Europa, 2018). The continuous efforts to open up more data is therefore essential to develop safe, accurate and reliable AI systems that do not compound privacy issues generally associated with massive data extraction.

Data maturity, on the other hand, is a measurement of the extent to which an organization is utilizing their data. To achieve a high level of data maturity, data use must be deeply ingrained in the organization, and be fully incorporated into all decision making and practices (Palmer, 2021). However, this would not be possible without big sets of high quality (open) data that analysts, researchers, and developers can use.

Open data enables an organization to achieve high levels of data maturity and unlock new insights and innovation, especially when it comes to data-dependent technologies such as AI. It

allows for the transformation of how an organization utilizes data from merely a source of information into an influencer—or even disruptor—of an organization's decision-making (Palmer, 2021).

As recommended in the 2017 Open Data Barometer Report, government data must be open by default and must adhere to the Open Data Charter principles. The practice of proactively releasing government data to create a data ecosystem would unlock more opportunities for AI applications. It allows government agencies and AI-solutions providers to share data and crucial information amongst each other. Open data, furthermore, lessens the power imbalance in data access between larger companies and MSMEs. As Critical Data Studies scholar Rob Kitchin (2014) maintains, open data democratizes information, as opposed to "confining the power of data to its producers and those in a position to pay for access."

Despite its benefits, open data initiatives can also pose a threat to the privacy of individuals. It is, therefore, necessary to have an appropriate balance between the utility of open data and safeguarding people's privacy. A way to achieve open data objectives in a privacy-aware way is through the utilization of technical tools and policy tools.

Overall, in dealing with risks associated with the widespread deployment of AI, an equally risk-based governance and regulation must ensure that organizational and societal controls are clearly identified to mitigate potential risks. A privacy-respecting, risk-based approach is only possible with an extensive use of open data.

# Open Data and AI Application Areas (Justice System and Social Media)

The UK Center for Data Ethics and Innovation (2020) has identified "common risks" associated with AI in the following areas (Fig 2): health and social care, justice system, social media, and utilities.

*Figure 2. Common Risks of AI. Key: Higher (🔴), Medium (🟡), Lower (🟢) Risks*

| Common Risks | Criminal Justice | Financial Services | Health & Social Care | Digital & Social Media | Energy & Utility |
|---|---|---|---|---|---|
| Bias leading to discrimination | 🔴 | 🔴 | 🔴 | 🔴 | 🟡 |
| Lack of explainability | 🔴 | 🔴 | 🟡 | 🔴 | 🔴 |
| Regulator resourcing | 🟡 | 🟡 | 🔴 | 🔴 | 🔴 |
| Higher-impact cyberattacks | 🟡 | 🔴 | 🔴 | 🟡 | 🔴 |

| Common Risks | Criminal Justice | Financial Services | Health & Social Care | Digital & Social Media | Energy & Utility |
|---|---|---|---|---|---|
| Failure of consent mechanisms | 🔴 | 🟡 | 🔴 | 🔴 | 🟡 |
| Loss of trust in institutions | 🟡 | 🟡 | 🔴 | 🟡 | 🔴 |
| Lack of transparency | 🔴 | 🟡 | 🟡 | 🟡 | 🟡 |
| Unequal access to services | 🟢 | 🔴 | 🔴 | 🟡 | 🟡 |
| Effects of low digital/data maturity | 🟡 | 🟡 | 🔴 | 🟢 | 🔴 |
| Erosion of privacy | 🔴 | 🟡 | 🟡 | 🟡 | 🟢 |
| Platform and data monopolies | 🟡 | 🟢 | 🟡 | 🟡 | 🔴 |
| Excessive data retention | 🔴 | 🟡 | 🟢 | 🟡 | 🟢 |
| Low ‘human-in-the-loop’ | 🟡 | 🟡 | 🟡 | 🟡 | 🟢 |
| Mis/disinformation | 🟢 | 🟢 | 🔴 | 🔴 | 🟢 |
| Loss of trust in AI | 🟢 | 🟢 | 🟡 | 🟢 | 🔴 |
| Undervaluation of public data | 🟡 | 🟢 | 🔴 | 🟢 | 🟢 |
| Low accuracy | 🟡 | 🟢 | 🟡 | 🟡 | 🟢 |
| Undermining professional judgment | 🟢 | 🟢 | 🔴 | 🟢 | 🟢 |
| Excessive trust in AI tools | 🟡 | 🟢 | 🟡 | 🟢 | 🟢 |

Due to limited space, I can only highlight a few cases that illustrate how open data might be used to address risks. In the justice system, discrimination is particularly high a risk. It could be exacerbated by biased data sources, or mitigated by greater access to open data.

**Bias leading to discrimination in Criminal Justice.** An AI-powered tool known as the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS), is a notable example of an inaccurate AI system, compounding bias and discrimination risks. The tool analyzes data such as previous arrests, age, and employment in order to determine risks for recidivism. COMPAS is deployed in courtrooms across the U.S.A to assist judges in determining whether defendants would face jail time. A study by Larson et al. (2016) shows that the AI-powered tool was biased against black defendants. It is more likely to incorrectly tag black defendants as “high-risk” for violent recidivism compared to white defendants. White defendants were also more likely to be labeled as “low-risk” compared to their black counterparts. Black defendants often received higher recidivism scores than they should have. In order to prevent or minimize instances of bias and discrimination, huge amounts of (open) data is necessary for accuracy and unbiased criminal assessment. AI-driven risk assessment and criminal profiling systems should also have access to more data such as victimization, gang affiliation, and drug arrest records to function impartially.

Numerous countries already recognize the immense capability of AI to reduce human bias in law enforcement and shape a fairer sentencing system (West & Allen, 2018). With open data, AI systems’ automated reasoning and predictive risk analysis would improve significantly.

**Erosion of privacy in Digital and Social Media.** Digital and social media are among the most familiar and widely used platforms by the general public. “Big tech” companies (e.g. Facebook, Instagram, Twitter, Youtube, etc) are often in the spotlight for their use of our personal data. There is an added fear among the public that introduction of open data and artificial intelligence to digital and social media will lead to a greater erosion of privacy. Contrary to this claim, AI and open data are not necessarily detrimental to privacy. AI can also help protect user data in social media. An app by a startup in Germany, for instance, filters gestures, persons, and objects that need to be censored or protected (Irisnet, 2020). They have trained their AI to recognize images with children, or with more than one person, nudity, violation of copyrights, defamation content, banned symbols, and personal data (e.g. name, e-mail, address) so that it is compliant with the General Data Protection Regulation in Europe. Open data can be “washed” to be free of sensitive personal information before the dataset is released to the public. The determination of an appropriately controlled research environment can also be done so that only trusted and authorized researchers, developers, and analysts and their AI algorithms are allowed access to sensitive personal data, thus mitigating potential privacy breaches while maximizing the value of data.

# Beyond AI Governance

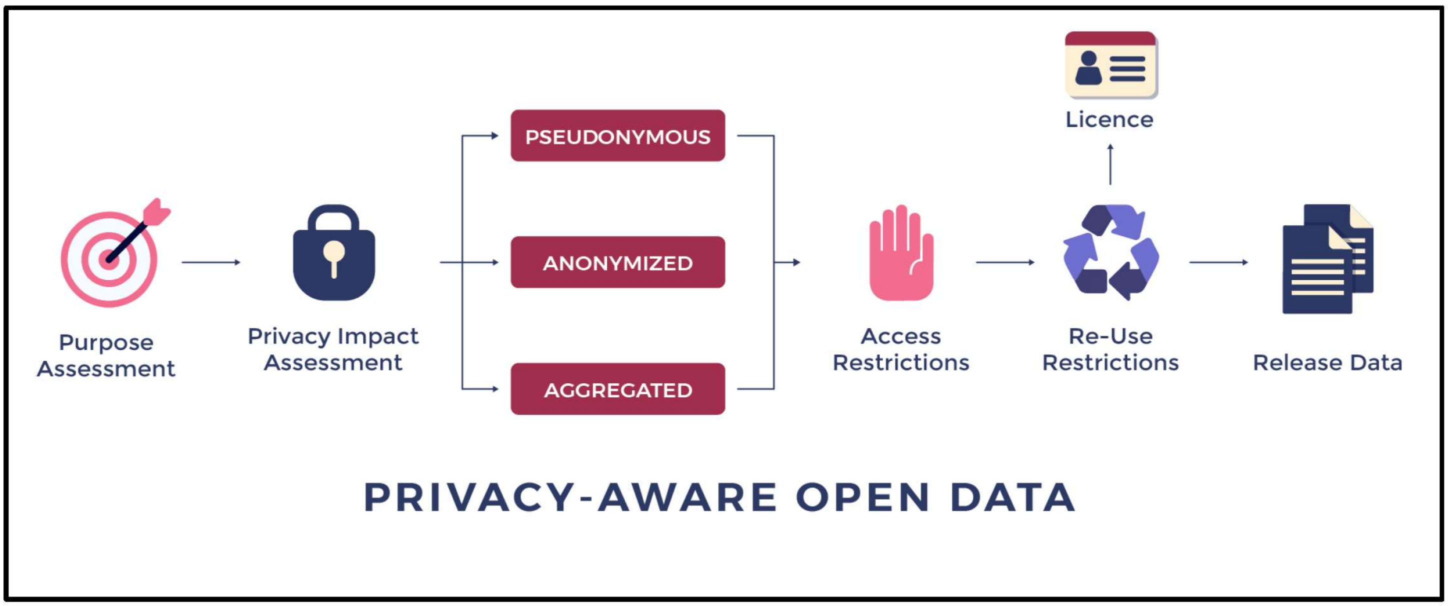
With only limited space, I can only outline here some issues and concerns beyond AI governance yet affecting the overall development and governance of AI in the country. One of such issues and concerns is government regulation. In our consultation meetings with AI

stakeholders, AI researchers and developers are particularly wary about premature government restrictions that could stifle technological innovations. CAWIL AI CEO Cherry Murillon claimed that in her experience, one of the barriers to doing big data and AI in the Philippines revolves around misconceptions about data privacy regulations that unnecessarily restrict access to important personal data (Murillon, 2021). One such misconception out there is that consent is effectively the *only* basis for processing personal data.

Even as data analysts and AI researchers and developers were to use open data involving personal information with increased risk of privacy infringement, there are technical ways to mitigate such a risk. In modeling AI, AI technology has also had the ability to re-identify anonymous data (Austin et al., 2021). It is, nonetheless, important for open data to be privacy-aware. Open data is not antithetical to data privacy, and neither are the two mutually exclusive.

There are a number of frameworks for open data that can guide governments, businesses, organizations, and stakeholders in dealing with personal data. Figure 3 below outlines the steps to consider in formulating a privacy-aware open data initiative. These steps can be undertaken to forward core privacy principles, including transparency, legitimacy of purpose, proportionality, limited use and retention, consent, accountability and security.

*Figure 3. An overview of the steps to consider to achieve a privacy-aware open data initiative (Walker et.al., 2016)*




A process for achieving a privacy-aware open data initiative begins with *purpose assessment*. It calls for the evaluation of specific objectives of the release of data and the assessment whether the data being released accomplishes such objectives or goals (Walker et al., 2016). The conduct of a privacy impact assessment (PIA) to determine the privacy implications of releasing specific data sets follows purpose assessment. That sets AI and big data analysts, researchers and developers up for the appropriate technical tools (including pseudonymous data, anonymized data, and aggregated data) to be used in order to be able to safely release

datasets into the open. Such tools help define the levels of privacy for the dataset before it is released to a safe data environment (Walker et al., 2016).

After determining the level of privacy needed to come with a particular dataset, policy tools are also used to help situate open data while protecting informational privacy. These policy tools include access restrictions, allowing the government to maintain control of the released data for a period before the data becomes fully open (Walker et al., 2016). As Borgesius et al. (2015) maintain, “To achieve a particular objective that underpins open data, it might not always be necessary to allow everyone access to government-held data.” Adopting robust processes for the determination of purpose and the assessment of privacy, coupled with the utilization of technical and policy tools, would help mitigate the privacy risks of open data while maximizing its benefits.

In conclusion, integral to the development and governance of Artificial Intelligence, especially in the Philippines, is open data. To support this claim, governance principles and elements, risk-based approach to governance, select AI application areas, and concerns beyond governance (primarily, the general regulatory framework of data privacy) have been the key areas of consideration in this paper. At best, the country would be worse off pursuing AI without the abundance of open data. AI governance would be less robust without the adherence to open data charter principles.

# Bibliography

Austin, T., Busath, K., Diehl, A., & Kamleshkumar Kishnani, P. (2021, December 10). Trustworthy open data for trustworthy AI. Deloitte Insights. https://www2.deloitte.com/xe/en/insights/industry/public-sector/open-data-ai-explainable-trustworthy.html

Borgesius, F., Gray, J., Eechoud, M. (2015). “Open Data, Privacy, and Fair Information Principles: Towards a Balancing Framework” Berkeley Technology Law Journal, 30(3), 2073-2132.

Brandusescu, A., Iglesias, C., Robinson, K., Alonso, J. M., Fagan, C., Jellema, A., & Mann, D. (2017). Open Data Barometer: Global report. 4th ed.

Carrara, W., Fischer S., Oudkerk, F., et al. (2020). *Digital transformation and open data.* Publications Office of the European Union. https://data.europa.eu/doi/10.2830/673557

Davies, T. (2013). Open Data Barometer: Global report. 1st ed.

Data Europa. (2020, February 6). The rise of AI and open data | data.europa.eu. https://data.europa.eu/en/news/rise-ai-and-open-data

Data Europa. (2018, April 7). *Unlocking the potential of Artificial Intelligence. AI and Open Data: A crucial combination*. Retrieved September 8, 2022 from https://data.europa.eu/en/datastories/ai-and-open-data-crucial-combination

Department of Trade and Industry. (2021). *Artificial Intelligence Roadmap*. Retrieved August 13, 2021 from http://innovate.dti.gov.ph/resources/roadmaps/artificial-intelligence/

Irisnet. (2020). AI Photo Screening. Retrieved September 15, 2022 from https://irisnet.de/privacy-protection/

Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T. (2020). Data governance: Organizing data for Trustworthy Artificial Intelligence. Government Information Quarterly, 37(3). https://doi.org/10.1016/j.giq.2020.101493

Jose, D. (2022). Correspondence, 15 March 2022

Kitchin, R. (2014). *The Data Revolution* (1st ed.). SAGE Publications. Retrieved from https://www.perlego.com/book/860829/the-data-revolution-pdf (Original work published 2014)

Larson, J., Angwin, J., Kirchner, L., & Mattu, S. (2016, May 23). How we analyzed the compas recidivism algorithm. ProPublica. Retrieved September 14, 2022, from https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm

Lobana, J. (2021). The governance of AI based information technologies within corporate environments. Retrieved September 9, 2022 from https://macsphere.mcmaster.ca/bitstream/11375/26685/2/Lobana_Jodie_2021July_PhD_Information_Systems.pdf

Lockey, S., Someh, I.A., & Gillespie, N. (2021, January 4-8). *A review of trust in Artificial Intelligence: Challenges, vulnerabilities and future directions* [Conference Paper]. Honolulu, HI, United States: Hawaii International Conference on System Sciences.https://doi.org/10.24251/hicss.2021.664

Manyika, J., Silberg, J., & Presten, B. (2019, October 25). What do we do about the biases in AI? *Harvard Business Review*. Retrieved from https://hbr.org/2019/10/what-do-we-do-about-the-biases-in-ai

Mapa, D. (2020, December 7). Philippines ranks 18th in Open Data Inventory 2020, up from 41st in 2018. *Philippine Statistics Authority.* https://psa.gov.ph/content/philippines-ranks-18th-open-data-inventory-2020-41st-2018

McGinnis, D. (2020, October 27). *What is the Fourth Industrial Revolution? What Is the Fourth Industrial Revolution?.* Salesforce. https://www.salesforce.com/blog/what-is-the-fourth-industrial-revolution-4ir/

Molloy, J. C. (2011). The Open Knowledge Foundation: Open Data Means Better Science. PLOS Biology, 9(12), e1001195. https://doi.org/10.1371/journal.pbio.1001195

Murillon, C. (2021). Developing a National AI Framework: Consultation with AI Startups [Zoom presentation] Metro Manila, Philippines.

OECD. (2019). *Recommendation of the Council on Artificial Intelligence*. Retrieved January 27, 2022 from https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449#mainText

Open Data Charter. (2015). *International Open Data Charter*. https://opendatacharter.net/wp-content/uploads/2015/10/opendatacharter-charter_F.pdf

Open Data Philippines [@datagovph]. (2017, October 13). *ICYMI: More than 40 government agencies are now in Open Data Philippines. Visit http://gov.ph/data now! #OpenDataPH.* [Tweet] Twitter. https://twitter.com/datagovph/status/918768104481497091

Open Data Philippines (2022). *Portal*. Retrieved September 5, 2022 from https://data.gov.ph/index/home

Open Government Partnership. (2022). *About*. Retrieved September 5, 2022 from https://www.opengovpartnership.org/about/

Pacis, J. (2017). Open data in Philippines: An issue of access and awareness. *IT for Change.* Retrieved September 5, 2022 from https://opendocs.ids.ac.uk/opendocs/bitstream/handle/20.500.12413/13008/Research-Brief-Philippines.pdf

Palmer, K. (2021, January 20). *Building a data maturity Model + the 4 stages of data maturity. Places Data Curated for Accurate Geospatial Analytics*. Retrieved from https://www.safegraph.com/blog/the-four-stages-of-data-maturity#:~:text=Data%20maturity%20is%20dependent%20on,and%20other%20data%20analytics%20capabilities.

Personal Data Protection Commission Singapore. (2020). Model AI Governance Framework (2nd ed). Retrieved December 18, 2021 from https://www.pdpc.gov.sg/Help-and-Resources/2020/01/Model-AI-Governance-Framework

Piccarozzi, M., Aquilani, B., & Gatti, C. (2018). Industry 4.0 in management studies: A systematic literature review. *Sustainability*, 10(10), 3821. https://doi.org/10.3390/su10103821

PricewaterhouseCoopers. (2017). PwC's Global Artificial Intelligence Study: Sizing the prize. PwC. Retrieved September 9, 2022, from https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html

Saxena, S. (2018), Drivers and barriers to re-use Open Government Data (OGD): a case study of open data initiative in the Philippines. *Digital Policy, Regulation and Governance*, 20(4), 358-368. https://doi.org/10.1108/DPRG-08-2017-0045

Schwab, K. (2016). *The Fourth Industrial Revolution*. https://law.unimelb.edu.au/__data/assets/pdf_file/0005/3385454/Schwab-The_Fourth_Industrial_Revolution_Klaus_S.pdf

UK Center for Data Ethics and Innovation. (2020). *AI Barometer Report*. Retrieved January 18, 2022 from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/894170/CDEI_AI_Barometer.pdf

Walker, D., McCord, M., Stradiotto, N., Zhou, M., & Singh, D. (2016). Citizen's Guide to Open Data. *GeoThink.* Retrieved September 17, 2022, from https://citizens-guide-open-data.github.io/guide/4-od-and-privacy#:~:text=The%20first%20way%20to%20balance,2122).

West, D. M., & Allen, J. R. (2022, March 9). How artificial intelligence is transforming the world. Brookings. https://www.brookings.edu/research/how-artificial-intelligence-is-transforming-the-world/#_edn1

World Wide Web Foundation (2016). *Open Data Barometer: Global Report. 3rd ed.*

World Wide Web Foundation (2017). *Open Data Barometer: Global Report. 4th ed.*

Yong, E. (2015, August 27). How reliable are Psychology studies? *The Atlantic*. Retrieved from https://www.theatlantic.com/science/archive/2015/08/psychology-studies-reliability-reproducability-nosek/402466/