

FIT5205 DATA IN SOCIETY

ASSIGNMENT 4: CRITICAL REVIEW OF OECD PRINCIPLES ON AI

SOVAN SINHA ROY | 30201152
TUTOR: *TURKI ALTHAQAFI*

EXECUTIVE SUMMARY

The **Critical Review Report of OECD principles on AI** provides a comprehensive analysis of the context, content, and stakeholders of the latest set of AI principles introduced in 2019 by OECD that has already been adopted by 42 countries. The objective of an innovative trustworthy AI is commendable and the policies for governance are applaudable too. The ethical stance of the principles is ascertained courtesy of *Beard & Longstaff's* ethical design principles. The findings suggest that most of them have been lucidly followed except for a couple like purpose. Since, OECD doesn't prescribe any rules, laws, or regulations, the report tries to investigate how these things are taken care of in countries that are adopting it. Certain limitations of OECD AI principles are also discussed which needs to be addressed for its improvement and credibility.

The report ultimately suggests a few recommendations but at the same time also acknowledges the conundrum of diversity in laws, rules, and approaches, etc across nations of the globe that may hinder it from materializing. Nevertheless, this report does succeed in an assiduous post-mortem of the OECD AI principles that can be used by anyone and everyone interested to comprehend it better!

TABLE OF CONTENTS

INTRODUCTION 3

BACKGROUND AND CONTEXT OF THE OECD AI PRINCIPLES..... 3

SUBJECT MATTER STAKEHOLDERS 4

CONTENT OF THE OECD AI PRINCIPLES & LINKAGES WITH THE ETHICAL THEORIES 4

ETHICAL STANCE OF THE OECD AI PRINCIPLES WITH RESPECT TO BEARD & LONGSTAFF
ETHICAL DESIGN PRINCIPLES 5

EXHIBITION OF COMPLIANCE 7

MEASURES & PROVISIONS FOR ACCOUNTABILITY 7

THE ESSENCE OF REGULATORY STRUCTURES & BODIES 7

BRIEF DISCUSSION ON THE OECD AI PRINCIPLES’ STRENGTHS & LIMITATIONS 8

CONCLUSION 10

REFERENCES 11

INTRODUCTION

To begin with, **Artificial Intelligence** or **AI** refers to the simulation of human intelligence in programmed machines to mimic human actions like *problem-solving*, *task completion*, etc (Nilsson, 2014). Today, AI has become an indispensable part of the modern digital technologies used to serve mankind but like any other technology, AI has its own *risks*, *perils* & *limitations* too. Now, in order to maximize the benefit to humanity and reap utmost favourable ramifications through the usage of AI, several principles, codes, guidelines, and frameworks have been concocted by many organisations over the years, one of them happens to be the recent set of 5 value-based AI principles from **The Organisation for Economic Cooperation and Development** or **OECD**. The OECD AI principles endeavour to set an international regulation standard in a wide array of areas like *privacy*, *business conduct*, etc (OECD, 2019c). So, the objective of these OECD AI principles is essentially to forge an *innovative & trustworthy* AI ecosystem that *respects human rights & democratic values* (OECD, 2019c). The governments of various nations have invariably formulated *legislations & regulations* around ethical AI *design, development & implementation* based on policies for those AI principles which they initially adopted and then amended. The aim of this report is mainly to critically analyse this set of OECD AI principles in its organic form so as to ascertain the strengths & shortcomings of it.

So, this report commences with the description & explanation of the OECD AI principles' *content, context & stakeholders*. This is followed by the mapping of the AI principles against the *Beard & Longstaff* principles to establish the *ethical stance* (of the principles). Thereafter, we will try to *canvass & discuss its governance* aspects involving accountability, regulation, etc, subsequently, a brief discussion will be undertaken elucidating on the strengths and limitations of the OECD AI principles. Finally, an insightful inference will be presented summarising all the key points with a few suggestions for the improvement of these principles!

BACKGROUND AND CONTEXT OF THE OECD AI PRINCIPLES

The 5 complementary *value-based principles* (can be taken as a framework also) that are primarily *beneficence, fairness, transparency, safety, and accountability* which AI ought to promote in the contemporary world is an integral part of the final proposal articulated by a *multi-stakeholder* expert group called **AI Group Of Experts (AIGO)** to foster trust in the adoption of AI in the society (OECD, 2019c). This proposal was submitted to the *Committee on Digital Economic Policy* (CDEP) by **AIGO** comprising 50 experts from diverse disciplines and realms like representatives from 20 governments as well as leaders from *civil society, academia, technical community*, etc. Those 5 principles were then duly adopted in May 2019 by all *The Organisation for Economic Cooperation and Development* (**OECD**) member countries along with a few non-member countries like **Argentina, Ukraine, Brazil, Romania**, etc thus making it the first set of such AI principles adopted concurrently by various nations or governments (Floridi & Cowls, 2019). Moreover, a practical *framework* to contextualize and consider ways to implement the principles in a plethora of AI systems can be easily undertaken which apparently has been done by all those OECD member countries that have adopted it. OECD is basically an *intergovernmental economic organisation* with 37 member countries that *build policies to stimulate socio-economic progress, environmental refinement, and world trade* (OECD, 2019a). Furthermore, apart from the principles, there are *four recommendations* for *national AI policies* and an *overarching principle on International Cooperation* for

trustworthy AI (OECD, 2019a) to ensure responsible stewardship towards the *betterment & uplift* of AI usage in various sectors for various purposes by various organisations (Floridi & Cowls, 2019).

The OECD AI principles apparently laid an emphatic set of *practical & flexible* standards capable of standing the test of time in a rapidly transforming world of technology (OECD, 2019c). These principles complement existing OECD standards in areas such as *privacy, digital security risk management, and responsible business conduct* (OECD, 2019c). So, in a nutshell, the objective of this set of 5 OECD AI principles was to formulate an AI-specific model code of conduct facilitating *innovation & trust* whilst being *implementable, scalable, and conducive* to sustained cooperation to *respect human rights* and uphold *democratic values* (OECD, 2019b). Furthermore, it serves as a *reference instrument* for all the adhering countries that aspire to shape a *human-centric approach to trustworthy or reliable AI*.

SUBJECT MATTER STAKEHOLDERS

The *stakeholders* in this matter of discourse concerning OECD AI principles encompass all *public & private sector organisations* and *individuals* involved in, or inherently affected in any form or manner by the AI Systems i.e., a machine-based system that is capable of influencing the environment by making *recommendations, predictions or decisions* for a given set of *objectives*, explicitly or implicitly (OECD, 2019a). They (the stakeholders) consist of *civil society, the academic & technical communities, industries, governments* (of various nations), *labour representatives*, and *several trade-unions* as well as *individuals as workers* or perhaps *mere data subjects* (OECD, 2019c). Moreover, **AI Actors** are a *subset* of stakeholders, these are elements that invariably play an *active role* in the *AI system lifecycle* (iterative cyclic AI phases). All the *public and private sector organisations or individuals* that acquire AI systems to *deploy, operate, and/or utilize* them can be deemed as AI actors.

CONTENT OF THE OECD AI PRINCIPLES & LINKAGES WITH THE ETHICAL THEORIES

There are 5 *explicit and complementary value-based principles* for responsible stewardship of *innovative & trustworthy AI* which are delineated as follows:

- 1) *Inclusive & Sustainable Growth & Well-Being*: According to this principle, AI systems should be *beneficial by empowering people* whilst including everyone *equally & unconditionally* enhancing their *capabilities & creativity* (OECD, 2019c). So, *underrepresented populations should be brought to the forefront, social & economic inequalities should be reduced across the globe, and sustainable growth and well-being should be reinvigorated through judicious practices & righteous actions* (OECD, 2019c).
- 2) *Human-Centred Values & Fairness*: Effective mechanisms should be established to *respect the human rights & democratic values* (of people) including but not limited to ***freedom, dignity, autonomy, privacy, non-discrimination, fairness, social justice, and diversity***, etc (OECD, 2019a). This can be achieved through **Virtuous Actions (Virtue)** from the AI actors wherein people should be allowed to *determine & safeguard* their *digital identity and personal data* (OECD, 2019c). Moreover, various prevalent biases

like *perception, technical*, et al should be *decreased* as much as possible to entail fair results.

- 3) *Transparency & Explainability*: As per this AI principle, people should be allowed to *understand* how an AI system is *developed, trained, and deployed*; which variables are used, and which variables impact a specific *prediction, recommendation, or decision* (OECD, 2019c). The concerned stakeholders must *disclose* these *pertinent information* regarding AI systems as it will not only help in *raising awareness* but also *facilitate informed decision making* by those adversely affected by AI systems (Hagendorff, 2020). *Transparency* in AI usage is essential to retain **trust** and **legitimacy** in the eyes of the public, this echoes with the ethical theory of **Contractualism**.
- 4) *Robustness & Safety*: AI systems should be *robust*, in the sense that they should be able to *withstand* or *overcome* any atrocious circumstances. They (the AI systems) should also be *safe*, in the sense that they should not pose any *unreasonable safety risk* during its normal or foreseeable deployment. Likewise, even when they are *misused*, they should not *wreak incessant havoc* and should have some sort of *circuit-breaker* to bring the entire system to a halt when needed (Guihot, 2019).
- 5) *Accountability*: AI actors should be *accountable* for the *proper functioning* of AI systems. The onus should be on them to uphold the other aforementioned principles according to their *individual roles, circumstances, and state of the art of the AI system* (OECD, 2019c). So, the entire principle is based on **Taking Responsibility** or in other words exercising **Existentialism** form of ethical theory.

There are some *National Policies* or *Recommendations* for trustworthy AI manifested along with the OECD AI principles too which governments of various concerned countries must develop in *cooperation* and *collaboration* with all the other stakeholders. The objective simply put is to accomplish *inequitable & propitious* outcomes for people and the planet, consistent with those AI principles (delineated before). This can be achieved through *sincere investment in AI research & development, fostering a digital ecosystem for AI, Providing an agile & controlled policy environment for AI, and building human capacity & preparing for job transformation* (OECD, 2019c). Furthermore, there is one principle on *International Cooperation for trustworthy AI* also (this can be taken as another AI Policy too) which elaborates on some potential steps to be taken by all countries for *reliable* AI.

ETHICAL STANCE OF THE OECD AI PRINCIPLES WITH RESPECT TO BEARD & LONGSTAFF ETHICAL DESIGN PRINCIPLES

- 1) In the principle of *Inclusive & Sustainable Growth & Well-Being* by emphasising on **human well-being**, they have tried to achieve the *Beard & Longstaff's* principle of **maximum benefit (Net Benefit)** (Beard & Longstaff, 2018) by **minimizing the harms** in the form of *social & economical injustices*, and *fostering inclusive growth of all alike* (OECD, 2019c). Moreover, it empowers the *governments* to *promote, guide & incentivize* responsible AI to other stakeholders as a **shared responsibility (Responsibility)** towards *innovative & trustworthy AI for posterity* (Jobin, Ienca, & Vayena, 2019).

- 2) The *Human-Centred Values & Fairness* principle encourages to ***promote dignity*** by treating people with the utmost respect like they deserve (OECD, 2019c). It also ensures that intentionally, operational AI systems ***don't manipulate*** people and ***exploit*** human rights (***Non-Instrumentalism***). Moreover, by permitting people means to ***safeguard*** their ***personal data*** and ***digital identity***, this AI principle tends to uphold ***user sovereignty*** for better ***Self-Determination*** of those who are being impacted by AI technology and seeking hindsight about the same. Furthermore, ***fairness & social justice (Fairness)*** is the ***cornerstone*** of this AI principle that strives towards ***natural justice*** from ***biases & wrongdoings*** to those who are affected and also those who are anxious that they might also be affected adversely by AI sooner rather than later (Hagendorff, 2020).
- 3) The *Transparency & Explainability* AI principle resonates with ***show your work*** rule of thumb of ***Accessibility*** whereby it suggests that the AI actors should ***disclose the relevant information*** and allow the people to ***understand*** how an AI system is ***developed, trained, and deployed*** by providing them with ***simple & proper explanation*** (OECD, 2019c). Furthermore, by exhibiting ***the sources, variables, and metrics*** used for ***problem-solving assessment*** and thereafter ***problem-solving execution*** by the AI system openly to all, ***transparent provenance*** can be seen to be conformed to a great extent by the AI systems in line with the (***Designer***) ***Responsibility*** principle (Beard & Longstaff, 2018).
- 4) The *risk management approaches* suggested in the *Robustness & Safety* AI principle to ***mitigate various AI risks*** ensures that several ***side-effects are anticipated*** and ***addressed appropriately*** to ***maximize the good (Net Benefit)*** when it comes to ethical AI usage (Hagendorff, 2020). Furthermore, by devising mechanisms of shielding susceptible audience from potentially ***dire consequences*** of ***AI misuse or malfunctioning*** by working on the safety aspect, this principle advises ***harm quarantines*** (OECD, 2019c), the principle also acknowledges the significance of mankind and ***doesn't reduce them to the status of mere 'things' (Non-instrumentalism)!***
- 5) The *Accountability* AI principle implores to enforce ***accountability*** on the concerned AI actors like ***governments, etc;*** for the proper functioning of the AI systems and thus entrusts a colossal ***Responsibility*** on them to ensure that ***ethical values*** are abided by and the ***intended purpose*** of AI usage doesn't ***compromise*** or ***jeopardize*** the ***fundamental human rights, rule of law, and democratic values*** enshrined in the constitutions of the various nations adopting it.

From the aforementioned points, it is evident that the OECD AI principles adhere massively to the *non-instrumentalism, self-determination, fairness, net-benefit, accessibility, responsibility* ethical design principles of *Beard & Longstaff* and several aspects of them faithfully but nothing much has been elucidated or insinuated on the ***ought before can*** principle nor about the ***purpose*** that preaches ***designing, developing, and deploying digital technologies*** with ***honesty, clarity, and fitness of purpose***. Additionally, no ***legal guidelines*** or ***statutory advisory*** has been furnished as a deterrent towards unethical usage of AI (Guihot, 2019).

EXHIBITION OF COMPLIANCE

Albeit OECD doesn't administer any matter of compliance regarding its AI principles, the *governments, individuals and organisations* can demonstrate that they are *acting in accordance* with the OECD AI principles or framework by adhering to the *guidance or recommendations for consideration to policy makers* suggested by OECD so as to *maximize & share the benefits from AI while minimizing the risks & concerns* (OECD, 2019a). The members of the **G20 International Forum** comprising a *group of 19 countries & European Union* follow these policies by giving *special attention to international cooperation & inclusion of developing & underrepresented participants* (Twomey & Martin, 2019). Moreover, **G20** members also *recognize the need to promote AI capacity building & skills development* which is a fundamental aspect of OECD AI recommendations (DigitalEurope, 2019). Furthermore, countries like **Singapore** have created AI frameworks like **Model Framework** way earlier than the OECD AI principles already covering areas like *Accountability, Fairness & Explainability*. Through the **Model Framework**, they have ensured that they can capitalize on all the aforesaid AI principles and other principles as well like *Accuracy & Auditability* by paying heed to the recommended AI policies like *agile & controlled policy environment for AI which is accommodated in their flexible governance framework* (PDPC, 2019). Moreover, they have also undertaken *significant investment on AI research and development and are committed to improving unceasingly in this regard for the better lives of their people* (PDPC, 2019).

MEASURES & PROVISIONS FOR ACCOUNTABILITY

OECD doesn't have any *incumbent provision* to hold *participating countries & organisations accountable for AI governance*. They *merely advise* governments to conform to certain policies that are very likely to assist them in enhancing their *accountability*. However, OECD is working on this and have launched an online **AI Policy Observatory** '*to nurture & monitor the responsible development of trustworthy AI systems for the benefit of the society*' (Commission, 2019). At a national level, the **UK's Centre for Data Ethics and Innovation (CDEI)** is an *advisory body* led by an *independent board of experts* (UK Government appointed) to *investigate & advise the accountable government on data-enabled technology matters including AI* (ArtificialLawyer, 2018). It makes *recommendations* to the government & *produces effective codes of practice & guidance for industries besides the government*. **CDEI** also provides *expert advice to the UK government* on this subject to *maximize the benefit by addressing the loopholes* (ArtificialLawyer, 2018).

Likewise, **Singapore** has *internal governance structures & measures* in place to have a *robust oversight* over an organisation's usage of AI within its *territory* to prevent misuse of AI (PDPC, 2019). An organisation's *existing internal governance structures* can be *modified*, and/or *new structures can be implemented* if necessary after a thorough evaluation by the **Personal Data Protection Commission (PDPC) Singapore** that *collaborates* with other parties on AI governance mechanisms (PDPC, 2019).

THE ESSENCE OF REGULATORY STRUCTURES & BODIES

To *regulate* the usage of AI, *autonomous regulatory bodies and structures* are quintessential otherwise it can prove *curtains to human rights* (Commission, 2019). Even though *guidance & recommendations* are there as regulatory measures to trustworthy AI adherence yet OECD AI principles are *not legally binding* or in other words, there is no *single consolidated governing*

body to *monitor & control the global proceedings, make laws, rules & regulations to be strictly followed worldwide, and take punitive actions towards those who don't obey it sincerely transcending demographic boundaries*. So, it is the *discretion* of an individual country to implement it and thereafter regulate it via its *own rules, codes, and standards for legislation, co-regulation or self-regulation* (Jobin et al., 2019).

Well, competent *regulatory or advisory* bodies like **CDEI (UK)** are *paramount* to act as an *autonomous statutory body for advising & monitoring* the government and other *public* as well as *private sector organisations* in this realm to ensure that they don't misuse AI (Dafoe, 2018). Interestingly, it seems that today there is a *conspicuous dearth* of exemplary AI regulatory bodies hence more such credible bodies should be established expeditiously to regulate AI technology all across the globe. Furthermore, formidable *regulatory frameworks* are being used comprising all the *three types of regulations along with international human rights instruments* by various countries, for instance, **Singapore & USA** to ensure the *proper functioning* of AI amidst *protection of privacy and other human rights* (Fjeld, Achten, Hilligoss, Nagy, & Srikumar, 2020). Concisely speaking, the AI regulatory frameworks ensure *that innovation & utilization* of AI are not stifled whilst all the associated *risks like biases, etc when encountered* are *immaculately mitigated*.

Therefore, for *better standardized AI governance*, it would be great if, soon, the OECD can come up with its own *legal & advisory committee* to lay some *common undisputable rules, laws & punishments* in this regard to the maximum extent feasible since a *comprehensively uniform setup is not viable* considering the *diversity in the inter alia, rules, laws, cultures, and approaches of different countries* (Arrieta et al., 2020).

BRIEF DISCUSSION ON THE OECD AI PRINCIPLES' STRENGTHS & LIMITATIONS

We have already discussed in detail how meticulously these OECD AI Principles have been formulated for *innovative & trustworthy AI* by a *multi-stakeholder expert group*. So, it is a *practical & flexible* set of principles that *respects human rights, rule of law, democratic values, etc*. It is compliant with almost all the *ethical design principles* and takes into account most of the *key elements* required for the *betterment & uplift* of AI, probably that's why perhaps more than 42 countries have adopted it and still counting!

However, there are a few limitations too, like:

- a) The OECD AI principles are *not a legal binding* or a *stringent measure*, they are just a part of an *influential recommendation* that can potentially enhance the *trustworthiness* of AI to what extent is however something that is *not quantified* yet (Guihot, 2019).
- b) The principles are *broad & don't carry the force of laws, or perhaps even rules*. They just seek to *encourage* member countries to *incorporate those rendered values or ethics* in the *development & deployment* of AI *without any overseeing* by OECD (Guihot, 2019).
- c) The OECD AI principles were *published and adopted* in 2019, before that, many AI Frameworks and principles were already in existence, still it failed to subsume core parameters like *Accuracy & Privacy*, etc as part of its main principles like other prominent frameworks like those from *Microsoft, Google, European Commission* et al (Burle & Cortiz, 2019).

- d) Very little emphasis has been given on the *purpose* and *subsequent ramifications of AI misuse*, it can be dangerous to leave it entirely at the *discretion* of all the countries without being *Watchdogs* in such a *pivotal matter* as governments have been notorious for *overlooking ethics & misusing AI systems* many a times in the past for example., *the UK Police* for *surveillance*, etc (Hill, 2019).

CONCLUSION

To sum up, initially, we saw the background & context of the OECD AI principles along with the stakeholders involved. The recently introduced OECD AI Principles strive towards *innovative & trustworthy* AI which is commendable and for that it has 5 value-based complementary principles i.e., *Sustainable Growth & Inclusiveness, Fairness, Transparency & Explainability, Robustness & Security, Accountability*. However, it fails to incorporate other key parameters strongly like *Privacy & Accuracy (of AI)*, etc. Most of the *Beard & Longstaff* ethical design principles are applicable to the principles in discourse but a couple like *Purpose*, etc; are *vaguely* and *cryptically* touched upon, this leaves a room for refinement. Interestingly, the OECD AI principles are *non-binding* in nature, so they are voluntarily adopted by 42 countries without any *enforcement* or *administration* by OECD. These countries have established their own *mechanisms & provisions* for asserting accountability and ensuring better AI governance. *Regulatory & Advisory* bodies are also there in a few countries to *monitor, advise, and assist* the *central & state authorities* in this regard but OECD is again not involved in any of them. Ergo, all in all, OECD AI principles inherently have quite a few *shortcomings* and need significant work to be done for improvement. For starters, it should develop a *holistic* view and give equal emphasis to *social* aspects vis-à-vis *economic* ones it usually does. Additionally, it should try to unveil its own *laws, rules, and regulations* to do its bit in ensuring that AI is used ethically globally for the sheer *betterment & uplift* of the people!

Therefore, when it comes to an *ideal ethical framework*, there is no *silver bullet*. The aesthetic definition & boundary of *Ethical & Unethical AI* is a matter of *grey area* because *governments & societies* differ on what they regard as *ethical behaviour* (Arrieta et al., 2020). So, the dilemma is, under these *uncertainties & inequities* that are prevalent today, how can there be a *global consensus on the ethical development & unitary guideline/framework* on AI usage (which perhaps is the need of the hour today)? (Arrieta et al., 2020).

REFERENCES

- Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., . . . Benjamins, R. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82-115.
- ArtificialLawyer. (2018). UK Launches AI-Focused 'Centre for Data Ethics + Innovation'. Retrieved from <https://www.artificiallawyer.com/2018/11/26/uk-launches-ai-focused-centre-for-data-ethics-innovation/>
- Beard, M., & Longstaff, S. (2018). Ethical By Design: Principles For Good Technology. *The Ethics Centre*.
- Burle, C., & Cortiz, D. (2019). MAPPING PRINCIPLES OF ARTIFICIAL INTELLIGENCE. Retrieved from <https://ceweb.br/publicacoes/mapping-principles-of-artificial-intelligence/>
- Commission, A. H. R. (2019). Human Rights and Technology Discussion Paper [Press release]. Retrieved from <https://humanrights.gov.au/our-work/rights-and-freedoms/publications/human-rights-and-technology-discussion-paper-2019>
- Dafoe, A. (2018). AI governance: A research agenda. *Governance of AI Program, Future of Humanity Institute, University of Oxford: Oxford, UK*.
- DigitalEurope. (2019). 2019 G20 Ministerial Meeting on Trade and Digital Economy Recommendations for Promoting Innovation, Digital Technologies, and Trade. Retrieved from <https://www.digitaleurope.org/resources/2019-g20-ministerial-meeting-on-trade-and-digital-economy-recommendations-for-promoting-innovation-digital-technologies-and-trade/>
- Fjeld, J., Achten, N., Hilligoss, H., Nagy, A., & Srikumar, M. (2020). Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI. *Berkman Klein Center Research Publication*(2020-1).
- Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1).
- Guihot, M. (2019). Will we ever agree to just one set of rules on the ethical development of artificial intelligence? [Press release]. Retrieved from <https://theconversation.com/will-we-ever-agree-to-just-one-set-of-rules-on-the-ethical-development-of-artificial-intelligence-117744>
- Hagendorff, T. (2020). The ethics of Ai ethics: An evaluation of guidelines. *Minds and Machines*, 1-22.
- Hill, R. (2019). British cops told to scrap 'discriminatory' algorithms in policing. Retrieved from https://www.theregister.co.uk/2019/02/05/cops_told_to_scrap_algorithms_for_policing_work/
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389-399. doi:10.1038/s42256-019-0088-2
- Nilsson, N. J. (2014). *Principles of artificial intelligence*: Morgan Kaufmann.
- OECD. (2019a). *Artificial Intelligence in Society*. OECD Publishing, Paris. Retrieved from <https://doi.org/10.1787/eedfee77-en>.
- OECD. (2019b). OECD Principles on AI. Retrieved from <http://www.oecd.org/going-digital/ai/principles/>
- OECD. (2019c). Scoping the OECD AI principles. doi:<https://doi.org/10.1787/d62f618a-en>
- PDPC. (2019). *MODEL ARTIFICIAL INTELLIGENCE GOVERNANCE FRAMEWORK SECOND EDITION*. Retrieved from <https://www.pdpc.gov.sg/Help-and-Resources/2020/01/Model-AI-Governance-Framework>
- Twomey, P., & Martin, K. (2019). A Step to Implementing the G20 Principles on Artificial Intelligence: Ensuring Data Aggregators and AI Firms Operate in the Interests of Data Subjects.