Singapore Management University
Centre for AI & Data Governance
per email: jgallen@smu.edu.sg

Mr Lew Chuen Hong
Personal Data Protection Commissioner
per email: corporate@pdpc.gov.sg

30 August 2023

Dear Mr Lew

PDPC's Public Consultation on Proposed Advisory Guidelines on use of Personal Data in AI Recommendation and Decision Systems

We would like to thank you for inviting responses to the 'Proposed Advisory Guidelines on use of Personal Data in AI Recommendation and Decision Systems' (the 'Advisory Guidelines'), and we would like to take this opportunity to provide our feedback.

The SMU Centre for AI & Data Governance (CAIDG) is an interdisciplinary research centre housed at the Singapore Management University Yong Pung How School of Law (YPHSL). Working broadly on issues of data governance, emerging technology regulation, and digital economy law, CAIDG was established in 2018 with an inaugural grant from the Singapore National Research Foundation/Infocomm & Media Development Authority. Supported by the YPHSL Faculty and a global community of affiliated researchers, our work explores the governance of artificial intelligence (AI) and data use and to support Singapore's position as a global thought leader and first-mover jurisdiction in this space.

These points of response to the Public Consultation on Proposed Advisory Guidelines on use of Personal Data in AI Recommendation and Decision Systems have been collated by CAIDG Director Associate Professor Jason Grant Allen and Deputy Director Associate Professor Warren Chik. CAIDG Research Associate Mr Florian Gamper led on the drafting, and CAIDG Research Affiliate Mr Josh Lee Kok Thong provided substantive comment.

We are very supportive of the overall aim of the Advisory Guidelines, particularly the distinction drawn between the data protection issues in the **development** of AI systems and the **deployment/procurement** of AI systems. Data protection issues arising in the these two

lifecycle phases are fundamentally different, and it is important that the PDPA is interpreted along these lines. In broad terms, we do not see data protection issues arising in AI development provided the data is kept safe from misuse; the issues that arise here are more related to things like data monopolisation, for example, than to privacy. Privacy issues are far more important in the deployment of AI systems.

There are three major points we would like to highlight in our response to this Consultation:

1) In addition to the providing guidance on the Business Improvement Exception and the Research Exception, it would be useful to receive guidance on how section 11 (Compliance with the Act) and Part 3 and the First Schedule (Legitimate Interest) of the PDPA apply to the development of AI systems. As detailed below, we believe that section 11 and the Legitimate Interest Exception are relevant for the development phase of AI systems, and potentially provide a better way of dealing with the relevant issues than the Business Improvement Exception and the Research Exception.

2) We have some concerns about the guidance on the Accountability Obligation (section 10 Advisory Guidelines). Some users of the Advisory Guidelines may end up believing that the PDPA contains a general obligation for AI systems to be fair, transparent, and if appropriate, to have human oversight. Although we strongly support the existence of such an obligation, we note that it may be difficult (as a matter of statutory construction) to read these obligations into PDPA. Therefore, it may be advisable to change some of the wording in section 10 of the Advisory Guidelines to provide additional clarity.

3) We are aware that the PDPC seeks a response to the Advisory Guidelines only, however, we would take this opportunity to suggest that the data protection issues in the development of AI system may be best addressed through a statutory amendment of the PDPA. The Advisory Guidelines are an important step in the right direction but, as detailed below, an amendment to the PDPA would be the optimal solution.

There are also a number of smaller points we would like to bring to your attention, which we further detail below. We hope you will find this feedback helpful. Thank you for giving us the opportunity to provide feedback to the Advisory Guidelines. If you have any questions or would like to discuss any parts of our response further, please do not hesitate to reach out.

Yours faithfully

Jason Grant Allen
Warren Chik
Florian Gamper
Josh Lee Kok Thong
SMU Centre for Artificial Intelligence and Data Governance

Detailed Response

1) PDPC may wish to consider making the exact scope of Advisory Guidelines clearer and to provide guidance on what is to be considered an 'AI system'

The title of the Advisory Guidelines indicates that the Guidelines apply to personal data usage in 'AI recommendation and decision systems' only. In a similar manner, para. 1.2 of the Advisory Guidelines indicates that the Advisory Guidelines are focused on the case when AI is used for recommendation systems or decision-making.[1] However, para 3.1 states that Advisory Guidelines are aimed at the use of personal data in the development of AI *generally*. [2] We appreciate that when reading the different sections of the Advisory Guidelines it seems clear to what situation they apply. However, to provide additional clarity, it may be helpful for the PDPC to clarify whether Advisory Guidelines apply to AI systems that are not 'recommendation or decision systems'—such as, for example, generative AI systems.

It would also be helpful for PDPC to provide some clarity on when a system would be considered an 'AI system', given that no such definition is in the PDPA (or for that matter, the Model AI Governance Framework). Guidance on factors the PDPC would consider in analysing whether a system is considered an AI system may therefore be helpful and something PDPC might wish to consider.

2) Issues with the application of Business Improvement Exemption to the development of AI

We welcome the discussion of the Business Improvement Exception under Part 5 of the First Schedule and Division 2 under Part 2 of the Second Schedule to the PDPA. We particularly welcome the guidance in para. 5.1 that '[w]hen organisations require the use of personal data in the development of the AI System, they may consider relying on the Business Improvement Exception in lieu of consent'. This is particularly so because we believe that organisations developing AI systems will find these legal bases useful when processing data, without having to over-use and/or re-seek consent from individuals about their personal data.

---

[1] Para 1.2. states: 'The focus of the Advisory Guidelines on the Use of Personal Data in AI Recommendation and Decision Systems under the Personal Data Protection Act ("Guidelines") is to clarify how the PDPA applies to the collection and use of personal data by organisations to develop and deploy systems that embed machine learning (ML) models ("AI Systems") which are used to make decisions autonomously or to assist a human decision-maker through recommendations and predictions.'

[2] Para 3.1 states: These Guidelines are provided for situations where the design and/or deployment of AI Systems involve the use of personal data in scenarios governed by the PDPA. The aim of these Guidelines is to (i) clarify how the PDPA applies when organisations use personal data to develop and train AI Systems; and (ii) provide baseline guidance and best practices for organisations on how to be transparent about whether and how their AI Systems use personal data to make recommendations, predictions, or decisions.

However, there are some issues we would like to note. The first point is a technical one. The examples of 'AI development' found in para. 5.4 are rather examples of AI deployment rather than AI development. While there is not always a clear line between development and deployment, the PDPC may want to consider including in Advisory Guidelines a clear, paradigmatic case of AI development.[3]

To provide such an example would be useful because *inter alia* there is potentially some uncertainty in this area. For an organisation to make use of the Business Improvement Exemption there must be an improvement in or the development of new: goods or services; methods or processes; learning about and understanding the behaviour and preferences of a person of goods provided by that organisation; identifying any goods or services provided by the organisation. We believe that the Business Improvement Exemption is wide enough such the development of an AI system itself and that, for instance, the development of an Large Language Model (LLM) would fall under it. Nevertheless, there is some potential uncertainty because one could argue that an LLM is simply an algorithm and therefore either a good, service, method or process etc. We do not believe that this is a strong argument, but it would be helpful if the PDPC would add a line to the Advisory Guidelines or give an example to put the matter beyond doubt.

We further would like to point out that the Business Improvement Exception may not be conducive to the development of AI systems across organisations. Division 2, Part 2, Second Schedule PDPA provides that organizations may use personal data to improve the business (e.g. goods, services, methods etc) of said organization. Part 5, First Schedule provides that a corporation may use data collected from a related corporation i.e. a corporation within the same group. Therefore, the Business Improvement Exception would not be available to organisations who want to train AI systems collectively, unless the organisations belong to the same corporate group. For instance, this would mean that various hospitals, if they are different organizations and do not belong to the same corporate group, could not train an AI system collectively. The Business Improvement Exception not being available for cross-organizational training of AI systems could be a potential limitation to fostering an innovative environment for AI development in Singapore.

Furthermore, the development of AI systems often requires access to large amounts of data. In the recent past, significant breakthroughs in the field of AI were achieved by small independent actors. However, a limited ambit of the Business Improvement Exemption within organizations or corporate groups could lead to the situation that only large, incumbent organizations have sufficient data to develop AI systems. In this environment, it could be more difficult for smaller, start-up organizations – the very kind of organizations that have achieved many of the breakthroughs in AI in the recent past – to develop AI systems.

---

[3] For instance, it would be interesting to consider the example of messaging app using text messages generated by its users to train a large language model.

We further note that organizations that have access to large amount of data or often not primarily in the business of developing AI systems. For instance, the provider of a so-called 'super app' (where user can get transport services, food deliveries, payment services etc) may have large amounts of data. However, the operators of 'super apps' are in the business of operating an app and not in the business of developing AI systems. The providers of super apps may use AI systems, and may even try to develop such systems, however, their primary focus and area of expertise is not AI. Relying on the Business Improvement Exception does not seem to enable organizations which specialize in creating AI to access the necessary data to develop said systems. We are also concerned that relying on the Business Improvement Exemption could further strengthen large, incumbent organizations as data monopolies, rather than strengthening competition.

To be clear, the Business Improvement Exemption should be interpreted to allow organizations to develop AI system in-house. Our point is that the Business Improvement Exemption alone is unlikely to be sufficient to create the optimal environment for AI innovation in Singapore.

3) Issues with the application of the Research Exemption to the development of AI

We welcome the discussion of the Research Exemption (Division 3 under Part 2 of the Second Schedule to the PDPA) in the Advisory Guidelines. In para 6.1 the Advisory Guidelines state that the Research Exemption 'would include sharing data between unrelated companies for the purposes of jointly conducted commercial research to develop new AI Systems'. One of the challenges we foresee in applying the Research Exemption is the condition for there to be 'a clear public benefit to using the personal data for the research purpose' (The 'Public Benefit Requirement').

In practice, it may be difficult to determine what constitutes a clear public benefit. In particular, it is not clear from the Advisory Guidelines how the Public Benefit Requirement applies to the commercial development of AI, and it would be helpful to get some more guidance on this point. *Inter alia*, the Advisory Guidelines states at para 6.4(b) that one of the relevant considerations for organisations to consider whether the Research Exemption applies is: 'Potential of application of the AI System to increase innovation in products or services that benefit society by improving the quality of life'.

However, one could make the argument that innovation in relation to almost any product or service benefits society somehow (barring extreme examples, like making cigarettes more addictive). According to certain schools of thought, one could even plausibly argue that if innovation increases the profitability of a company this alone constitutes a benefit to society (for instance, because the company is an employer and provides a livelihood for its owners and workers). However, an overbroad interpretation could be contrary to the legislative intend behind the PDPA, as the exception would be triggered too easily.

Providing guidance on this issue is extremely difficult. However, it is important because we see an increasing trend in AI (but also in technology more generally) that research is undertaken not by classical research institutions (e.g. universities) but by privately-funded organizations. Further, some of the business models used in AI development 'blur' the line between private benefit (i.e. benefit which will accrue to the organization, which developed a particular AI system) and public benefit. For instance, OpenAI Inc. is a non-profit organisation; however, it does have a for-profit subsidiary. OpenAI's mission is to '… is to ensure that artificial general intelligence (AGI) […] benefits all of humanity.'[4] If a similar scenario were to arise in Singapore, the question is whether the development of an AI system under such circumstances would meet the threshold of the Public Benefit Requirement.

4) The PDPC may wish to consider providing guidance on section 11 PDPA (Compliance with the Act) and the Legitimate Interest Exception

We welcome the distinction drawn in the Advisory Guidelines between the data protection issues raised in the development phase of an AI system and in its deployment phase. We believe that this is a very important distinction because the data protection issues in the development phase of an AI system are fundamentally different from the data protection issues raised the deployment phase. For instance, a research report by the European Parliament on the compatibility of training models to develop AI systems and the EU's General Data Protection Regulation, GDPR notes as follows:

> In general the inclusion of a person's data in a training set [to train an AI system] is not going to affect to a large extent that particular person, since the record concerning a single individual is unlikely to a make a difference in a model that is based in a vast set of such records. However, the inclusion of a single record exposes the data subject to risks concerning the possible misuse of his or her data, unless the information concerning that person is anonymised or deleted once the model is constructed.[5]

Although there are basic differences between the GDPR and the PDPA, we believe that this insight also applies to the PDPA. The point is that training an AI system on personal data does not *per se* impinge on the legitimate data protection interests of individuals. The objective of training AI systems is to find correlations between the data points in the training set and to apply these correlations outside the training set.

We believe that the Advisory Guidelines takes important steps towards balancing the interest of creating conducive environment for AI development in Singapore and the legitimate interests of individuals by pointing out that the Business Improvement Exemption and the Research Exemptions may be used to train AI systems. However, we believe that relying on

---

[4] OpenAI Charter, https://openai.com/charter , accessed 25 August 2023
[5] STUDY Panel for the Future of Science and Technology, EPRS | European Parliamentary Research Service.
The impact of the General Data Protection Regulation (GDPR) on artificial intelligence
https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641530/EPRS_STU(2020)641530_EN.pdf, page 46, accessed 25 August 2023

the Business Improvement Exemption and the Research Exemptions has some drawbacks. Therefore, it may be helpful for the PDPC to consider providing additional guidance on section 11 PDPA (Compliance with the Act) and the Legitimate Interest Exception (Part 3, First Schedule). Section 11 and the Legitimate Interest Exception may provide a better way to balance the legitimate interests of AI developers and data protection needs of individuals.

Regarding section 11 PDPA (Compliance with the Act), it would be helpful if the Advisory Guidelines would specifically address the issue how section 11 applies to the training of AI systems on personal data.[6] Section 11(1) PDPA states that '[i]n meeting its responsibilities under this Act, an organisation must consider what a reasonable person would consider appropriate in the circumstances.' The central question is whether a reasonable person would consider it as appropriate to train an AI system on personal data or not. The answer to this question will always depend on the specifics of the situation. The purpose of PDPA is to strike a balance between the 'the right of individuals to protect their personal data and the need of organisations to collect, use or disclose personal data for purposes that a reasonable person would consider appropriate in the circumstances'.[7] Thus, PDPA uses a balancing approach to data protection.

In our view, the development of AI system *per se* does not impede a person's right to protect their personal data. This being so, assuming a robust data protection context, the balance should generally be in favour of the organisation who wants to train AI system on personal data. The 'risk to privacy' in the training phase of an AI system is that personal data is being misused. Therefore, the relevant test to determine whether the training of an AI system on personal data complies with section 11 is whether the risk of data misuse is so little that a reasonable person would consider it appropriate to use the data for the purpose of training AI systems.

If this is the correct interpretation of section 11(1) PDPA, then the issues mentioned above in relation to the Business Improvement Exception and Research Exception (i.e. data usage not being available across organisations and the difficulties in defining 'public/societal benefit') do not arise. We are aware that if this is the correct interpretation of section 11(1), then the training of AI systems would largely be outside the scope of the PDPA.

The PDPC may further want to consider providing guidance on Part 3 of the First Schedule (Legitimate Interest). The Legitimate Interest Exception provides that consent is not required if:
a) the collection, use or disclosure (as the case may be) of personal data about an individual is in the legitimate interests of the organisation or another person; and
b) the legitimate interests of the organisation or other person outweigh any adverse effect on the individual.[8]

---

[6] We also note that no specific guidance on this point is provided in advisory guidelines on Key Concepts in the Personal Data Protection Act (issued 23 September 2013 and revised 16 May 2022)
[7] Article 3 PDPA
[8] Para 1(1) Part 3 of the First Schedule

Further guidance on the extent to which the Legitimate Interest Exemption can be applied to the development of AI systems would be beneficial. As stated above, we believe that training of AI systems on personal data *per se* does not have 'any adverse effect on the individual'. If this is correct, then in the case of AI systems' development 'the legitimate interests of the organisation or other person' should normally 'outweigh any adverse effect on the individual'.

Part 3 of the First Schedule further provides that certain requirements must be met for the Legitimate Interest Exemption to apply, e.g. the organisation must 'identify and implement reasonable measures: (i) to eliminate the adverse effect; (ii) to reduce the likelihood that the adverse effect will occur; (iii) or to mitigate the adverse effect; and (c) comply with any other prescribed requirements.'[9] Although training an AI system *per se* does not adversely affect an individual, there is the risk of data misuse.

However, because the Legitimate Interest Exemption requires the organisation to 'reduce the likelihood that the adverse effect will occur', there is an obligation on the organisation to reduce the risk of data misuse. Therefore, the Legitimate Interest Exemption strikes a good balance between the legitimate interest of AI developers and the data protection needs of individuals.

We further note that the PDPC's Advisory Guidelines on Key Concepts in the Personal Data Protection Act[10] notes at para 12.56 that: "Legitimate interests generally refer to any lawful interests of an organisation or other person (including other organisations)."[11] Thus, the Legitimate Interest Exemption may be a suitable exception to enable the training AI systems across organisations and thus does not have the drawback of the Business Improvement Exemption, of restricting data use within organisations or corporate groups.

5) Part III Deployment of AI systems

Para. 9.2 and following provides guidance on the Notification Obligation. We are supportive of this guidance but would point out that if the Notification Obligation is applied too strictly, this may lead to the proliferation of notifications to such an extent that users stop paying attention to notifications at all.

6) Advisory Guidelines on the Accountability Obligation may cause confusion

Section 10 of the Advisory Guidelines deals with the Accountability Obligation (sections 11 and 12 PDPA). In particular para 10.6 of the Advisory Guidelines states:

---

[9] 3(b) and 3(c) Part 3, Fist Schedule
[10] Issued 23 September 2013, revised 16 May 2022
[11] Internal speech marks omitted

Written policies also play an important function in education and confidence-building, which are necessary ingredients for building consumer trust and confidence. This could include behind-the-scenes measures taken to ensure that the personal data is used in a safe and trusted manner within the AI System, such as:

a) Measures taken to achieve fairness and reasonableness for recommendations, predictions, and decisions for the benefit of consumers during model development and testing stages. These can include measures relating to bias assessment, ensuring quality of training data or other data governance measures, or the repeatability/reproducibility of results using personal data.

[…]

b) For outcomes that have a higher impact on the individual, organisations may wish to consider whether it is useful to provide information on how proper accountability mechanisms and human agency and oversight have been implemented. It may also be useful to provide information on safety and/or robustness of the AI System or ML model (i.e., how the AI System or ML model will operate when encountering adversarial or unexpected input.

We are concerned that this wording may lead to confusion among users of the Advisory Guidelines. In particular, we are concerned that this may be interpreted as the PDPA containing a general obligation for the transparency and fairness of AI systems. We are not in principle opposed to these requirements—indeed, we are strongly supportive. A reader of the Advisory Guidelines could interpret para 10.6 to mean that the PDPA contains a general requirement for AI system's fairness, accountability or human oversight. However the PDPA in its present form does not contain such obligations.

7) Procurement and the Role of Service Providers

As noted by the PDPC, service providers may fall under the meaning and definition of data intermediaries (DIs). It was also noted the DIs are subject only to the protection and retention obligations. As such, the recommended best practices do not fall under these obligations (and instead should be made contractual obligations to help customer-organisations meet *their* obligations (as required by section 4(3)), unless they may relate to the protection obligation in some way). Clearly, these are not independent obligations under the PDPA. It could perhaps be made clearer that organisations (to which these guidelines apply) should ensure that their DI service providers help them meet their obligations by making it explicit in their contractual obligations.

8) From Guidelines to an algorithm training exemption in the PDPA?

The Advisory Guidelines set out a very sensible interpretation of how to apply the PDPA to AI systems. Although it goes beyond the scope of the present consultation, we would suggest that, in the longer term, an amendment to the PDPA in the form of a new 'AI Exemption' dealing with the training of AI systems could be an attractive option to consider. This would broadly mirror the approach taken by the Singapore legislature to computational data analysis in the

context of intellectual property law, namely the text and data mining exception (TDM exception) in the Copyright Act.[12]

Before this exception was created, already a 'fair use' doctrine existed in the Copyright Act pursuant to which it was be possible to train AI system on copyright protected materials. Nevertheless, the Singapore Parliament decided that it was better to amend the CA and to create a specific TDM exception. Along the same lines, we recommend the creation of an 'AI Exception' in the PDPA, which clearly defines what it means to train an AI system and what conditions need to be met for the AI exception to apply. We believe that such an exception would reflect Singapore's broader legislative objective to foster a conducive environment for technological innovation, while protecting the legitimate interests of individuals.

An express exception would have the benefit of clarity and certainty. Further, it would deal with a potentially contentious matter—in which competing interests are balanced—in a considered, bespoke manner. For instance, the amendment could clearly define what an AI system is, what it means to train an AI system, and what requirements need to apply for the AI exception to apply—none of which are actually as clear as they may intuitively seem. The specific requirements of the AI exemption would, of course, need to be analysed further. However, the advantage of a legislative amendment is that specific issues can be addressed in a specific manner. For instance, we believe that consent should not be required to train AI systems on personal data. However, in order to foster trust for the development of AI systems among the public it may be worthwhile to require notification if personal data is used to train AI systems.

Providing a specific amendment to the PDPA Singapore would provide clarity in a complex area, which would benefit AI developers based in Singapore and it could help to attract AI developers to Singapore. There is considerable uncertainty in many jurisdictions how privacy legislation applies to the development of AI systems. An amendment to the PDPA would send signal to international technology companies that Singapore's data protection regime provides clarity.

-END-

---

[12] Division 8 Copyright Act 2021.