Our Preliminary Points

Please clarify that this Proposed Advisory is intended for the traditional AI systems and is not intended for Generative AI systems.

Content	Comments
Part I: Introduction and Scope	
Focus of the 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. The Guidelines are advisory in nature, are not legally binding, and do not constitute legal advice. *(paragraph 1.2 & 1.3)*	No comments
The Guidelines are organised according to the stages of AI System implementation as follows: • Part II (Development, testing and monitoring): Using personal data for training and testing the AI System, as well as monitoring the performance of AI Systems post deployment. • Part III (Deployment): Collecting and using personal data in deployed AI Systems ("business to consumer" or B2C). • Part IV (Procurement): AI System or solution provider providing support to organisations implementing the AI System ("business to business" or B2B). *(paragraph 3.1)*	We recommend defining the parties who are subject to the various obligations, in particular to clarify that: • Part II of the guidelines apply to AI system developers that use personal data to develop and test AI systems ("AI System Developers"). • Part III applies to companies that deploy the AI systems developed by AI System Developers to consumers ("AI System Deployers"). • AI System Developers and AI System Deployers may or may not be the same company. • Part IV of the guidelines apply only to a subset of AI System Developers who are engaged to develop bespoke or customizable AI systems for companies. It does not apply to companies that develop AI systems in-house or who purchase off-the-self AI solutions.

Content	Comments
Part II: Using Personal Data in AI System Development, Testing and Monitoring	
Apart from seeking consent for the use of personal data to train an AI model, organisations may wish to consider if it is appropriate to rely on the (1) **Business Improvement** or (2) **Research** exceptions to consent under the PDPA, when using personal data to develop an AI System (see details below). *(paragraph 4.1)*	No comments
Under the **Business Improvement Exception** of the PDPA, organizations may use personal data without consent, for: • Improving, enhancing existing goods and services or developing new goods or services; • Improving, enhancing existing methods or processes or developing new methods or processes for business operations in relation to the organisations' goods and services; • Learning or understanding behaviour and preferences of individuals (including groups of individuals segmented by profile); or • Identifying goods and services that may be suitable for individuals (including groups of individuals segmented by profile) or personalising or customising any such goods or services for individuals. This exception is relevant when the organisation has developed a product or has an existing product that it is enhancing. It is also relevant when an AI System is intended to improve operational efficiency by supporting decision-making, or to offer more or new personalised products and/or services such as through offering recommendations to users. This also applies to data sharing for such	No comments

Content	Comments
purposes within a group of companies, as well as interdepartmental sharing of data within a company. *(paragraph 5.1, 5.2)*	
Organisations relying on the **Business Improvement Exception** will need to ensure the following under the PDPA: • The business improvement purposes cannot reasonably be achieved without using the personal data in an individually identifiable form; and • The organisation's use of personal data for business improvement purpose is that which a reasonable person would consider appropriate in the circumstances *(paragraph 5.3)*	While the practise of not using personal data is a good practice, we recommend that organizations be given the flexibility on whether to use personal data to improve AI System/ ML model. Organisations should not need be required to consider the requirements set out here.
To illustrate, the following are examples of purposes where the **Business Improvement Exception** could be relevant to AI System development: • Recommendation engines in social media services that offer users content more aligned to their browsing history; • Job assignment systems that automatically assign jobs to platform workers; • Internal HR systems used to recommend potential job candidates by providing a first cut in matching candidates to job vacancies; or • Use of AI Systems or ML models to provide new product features and functionalities to improve competitiveness of products and services. *(paragraph 5.4)*	No comments

Content	Comments
Relevant considerations for organisations on whether to rely on the **Business Improvement Exception**: • Whether using personal data contributes towards improving the effectiveness or quality of the AI Systems or ML models and their output; • Whether it is technically possible and/or cost-effective to use other means to develop, test or monitor the AI Systems or ML models without using personal data (e.g., it may not be cost-effective where the personal data constitutes a small part of a wider data set containing non-personal data and it would require disproportionate effort to anonymise it); • Common industry practices or standards on how to develop, test and monitor such AI Systems or ML models; and/or • Whether such use will contribute to the effectiveness or improved quality of new product features and functionalities that help organisations innovate, improve competitiveness, become more efficient/effective, and enhance consumer choice, experience, and usability. *(paragraph 5.5)*	Please clarify that the list here is not an exhaustive list of considerations which organisations may consider when deciding whether to rely on the Business Improvement Exception.
Organisations could rely on the **Business Improvement Exception** if personal data is used as part of the process of testing an AI System to improve or assess ML model performance e.g., to assess accuracy of the ML model in a live environment with personal data; ensure that de-biasing of the ML model is effective; or to check if privacy enhancing measures have compromised the accuracy of the AI System, taking into considerations the requirements as set out in paragraphs 5.2 to 5.5 above.	No comments

Content	Comments
Different standards for securing and protecting the datasets apply (see paragraphs 7.1 to 7.6 of the Guidelines), depending on the type of data used. *(paragraphs 5.6 & 5.7)*	
Organizations should consider the following in assessing whether the **Business Improvement Exception** applies, where personal data is used to check if protected characteristics, such as race or religion, are well represented in datasets or to assess the bias of the training dataset, so that adjustments may be made for de-biasing during AI system development: • Whether using personal data for this purpose is relevant for the effectiveness or improved quality of the AI Systems or ML models and its output; • Whether it is technically possible and cost-effective to use other means to debias AI Systems or ML models without using personal data; and/or • c) Common industry practices or standards on how to debias datasets used for AI Systems or ML models. PDPC understands that generally, industry best practice is to use personal data to debias datasets used for ML model training. *(paragraph 5.8 and 5.9)*	No comments

Content	Comments
The **Research exception** is intended to allow organisations to conduct broader research and development that may not have any immediate application on their products, services, business operations or market. This would include sharing data between unrelated companies for the purposes of jointly conducted commercial research to develop new AI Systems. The Research exception is subject to the following conditions: • The research purposes cannot reasonably be accomplished unless the personal data is provided in an individually identifiable form; • There is a clear public benefit to using the personal data for the research purpose; • The results of the research will not be used to make any decision that affects the individual; and • If results of the research are published, the organisation must publish the results in a form that does not identify the individual. • Companies will also need to assess whether it would be impracticable for the organisation to seek the consent of the individual for such disclosure. *(paragraph 6.1, 6.2 and 6.3)*	The conditions of the Research exception appear to be rather narrow and may not be applicable to most commercial purposes. We recommend that PDPC expand the conditions to allow the Research exception to be used in commercial business context.
Where using personal data, organisations are encouraged to undertake a data protection impact assessment to consider the effectiveness of risk mitigation and remediation measures applied to personal data. *(paragraph 4.2)*	Please clarify that the recommendation here is to conduct DPIAs only if required under PDPA.
When designing, training, testing, or monitoring AI Systems using personal data, appropriate technical, process and/or legal controls for data protection should be included. Where possible,	No comments

Content	Comments
organisations are encouraged to pseudonymise or de-identify the personal data used as a basic data protection control. In deciding what kind of controls for data protection should be implemented, companies should consider the following: • The types of disclosure/theft risks that the personal data would be subject to • The sensitivity and volume of the personal data used. *(paragraph 7.1)*	
Generally, privacy controls for internal use need not be as extensive as intra-group or cross-company sharing. However, companies are still encouraged to assess the risks and implement appropriate legal, technical and process controls for such use of personal data. *(paragraph 7.2)*	We recommend that the privacy controls for intra-group should not be as extensive as cross-company sharing.
In the context of developing AI Systems, organisations should practise data minimisation as good practice. Using only personal data containing attributes required to train and improve the AI System or ML model will also reduce unnecessary data protection and cyber threat risk to the AI System. To similarly reduce such risks, organisations should use the volume of personal data necessary to train the AI System or ML model and base this on relevant time periods and any other relevant filter e.g., market/customer segment, attributes, etc. Organisations may wish to refer to the PDPC's Guide to Data Protection Practices for ICT systems for further guidance in this area. *(paragraph 7.3)*	While the practice of data minimization is a good practice, we recommend that there should be no limit to the types of personal data used to improve the AI System or ML model as this would enable organisations to fully leverage the potential benefits of the AI System / ML model. We recommend that PDPC allow for organisations to adopt internal sandbox where data minimization requirement can be lifted for testing purposes.
If pseudonymisation is not possible and raw personal data has to be used e.g., facial images, organisations are reminded of their protection obligation under the PDPA. They should pay particular	No comments

Content	Comments
attention to the data security and protection measures around the development environment and are encouraged to conduct a Data Protection Impact Assessment. Standards for data protection in the development environment should meet similar standards needed for systems handling personal data. Organisations may wish to refer to the PDPC's Guide to Data Protection Practices for ICT systems, as well as the Guide on Responsible Use of Biometric Data in Security Applications where biometric data is used. *(paragraph 7.4)*	
Whether AI Systems are built in-house, externally, or using a combination of both, they will have security risks/points of weakness that can be exploited for privacy attacks to obtain information on the training data used e.g., model inversion attacks. In this regard, organisations should take a privacy-by-design approach and take steps to assess the risk of such privacy attacks as well as seek to mitigate such risks where possible within the AI System. *(paragraph 7.5)*	No comments
As per the Accountability obligation under the PDPA, organisations must ensure that their policies regarding the use of personal data in their organisations to develop AI Systems are updated and practices are established. For example, establishing policies relating to when model training should be conducted using anonymised or pseudonymised data, and when it is permissible to use identifiable personal data e.g., when model performance is degraded or for bias testing. *(paragraph 7.6)*	No comments

Content	Comments
Organisations are encouraged to anonymise their datasets as far as possible instead of using personal data. While anonymised data is not subject to the PDPA, organisations are reminded that such data still bears risk of re-identification and disclosure, and appropriate legal, technical and process controls should be instituted when using or disclosing such data. Organisations can refer to the PDPC's Guide to Basic Anonymisation for further guidance. PDPC recognises that there are trade-offs with using anonymised data when developing or training AI Systems or ML Models, such as model accuracy, repeatability, or reproducibility of results. In this regard, organisations should carefully weigh the pros and cons of using both types of data, and clearly document internally the reasons for choosing to use personal data instead of anonymised data. Organisations should employ appropriate corporate governance methods to make such decisions, including consulting relevant stakeholders and having such decisions made at an appropriately senior management level. *(paragraph 7.7, 7.8)*	No comments
In terms of what would be considered effectively anonymised data in the context of developing or training an AI System or ML model, organisations can balance their need for accuracy or retention of key target attributes by seeking to anonymise the dataset only to the extent that there is no serious possibility of reidentification. Organisations should refer to Chapter 3 of the PDPC's Advisory Guidelines for Selected Topics for PDPC's criteria on what constitutes anonymised data that falls outside the scope of the PDPA. *(paragraph 7.9)*	No comments

Content	Comments
As to whether anonymisation is sufficiently robust as to reduce the risk of reidentification, this would include considerations such as: • Whether the process of chosen anonymisation method is reversible; • The extent of disclosure of the dataset and its intended recipients (e.g., internal closed-group sharing vs. cross-company sharing); • Whether a motivated individual can likely find means to re-identify the anonymised dataset using either publicly available information or information the organisation already has in its possession; and • The extent of controls the organisation has put in place, including within the AI System, to prevent re-identification of the anonymised data It is likely that where the risk of re-identification is lower e.g., due to limited circulation or extensive controls over the anonymised data to prevent re-identification, the extent of anonymisation could be lesser to preserve the utility of the dataset. *(paragraph 7.10, 7.11)*	No comments
Part III: Deployment – collection and use of Personal Data in AI Systems	
When organisations deploy AI Systems to provide new functionalities or enhance product features by collecting and/or processing personal data, they should be mindful of the following PDPA obligations: • Consent • Notification • Accountability	No comments

Content	Comments
(paragraph 9.2)	
The raison d'etre for **Consent** and **Notification** obligations is to enable individuals to provide meaningful consent. To do so, organisations should place themselves in the shoes of consumers and craft notifications that will enable individuals to understand how their personal data will be processed to achieve the intended purpose. The notifications need not be overly technical or detailed and should be proportionate to the risks present in each use-case (e.g., taking into account potential harm to the individual and the level of autonomy of the AI System). Sufficient information ought to be provided so that individuals are able to appreciate how their personal data will be processed to achieve the intended purpose. Organisations are encouraged to provide information on the following in crafting their notifications: • The function of their product that requires collection and processing of personal data (e.g., recommendation of movies); • A general description of types of personal data that will be collected and processed (e.g., movie viewing history); • Explain how the processing of personal data collected is relevant to the product feature (e.g., analysis of users' viewing history to make movie recommendations); and • Identify specific features of personal data that are more likely to influence the product feature (e.g., whether movie was viewed completely, viewed multiple times, etc). *(paragraph 9.4 & 9.5)*	Privacy notices should be easily accessible by data subjects by being succinct. The information set out in the paragraph in red would add complexities which may be overwhelming to the data subject and would be counter to the objective of being easily accessible.
The provision of such information could be through notification pop-ups or can be included in more detailed written policies that are publicly accessible or made available to end users on request	Privacy notices should be easily accessible by data subjects by being succinct. The information set out in the paragraph in red would add

Content	Comments
Example 1: A bank uses AI to assist in credit scoring when assessing whether to approve applications for credit cards. It prepared a policy document entitled "Bank's Credit Assessment Policy Statement" which provides information about what personal data it collects from applicants and how they are processed by AI when the bank assesses applications. The policy document is provided to applicants who request for the information. **Example 2:** An organisation provides personalised recommendations for content to an individual on its online social media platform. To provide information to individuals as to why specific content is shown to them, the organisation has provided a pop up containing a link to a page to explain why this content is shown and ranked highly on the content feed for the user. The page includes information on why that content is shown, what information has the largest influence over the order of posts in the user's content feed, such as past interactions or membership in specific groups on the platform etc. *(paragraph 9.6)*	complexities which may be overwhelming to the data subject and would be counter to the objective of being easily accessible.
It may also be useful to consider "layering" of information. This would mean displaying the most relevant information more prominently, and then providing more detailed information elsewhere. For example, notification pop-ups could provide a link to publicly accessible privacy policies; additionally, privacy policies may be structured to have more detailed information organised in expanding sections or in separate tabs. The industry is also developing disclosure best practices, such as model cards and system cards. Information necessary to meet the Consent and	Allowing organisations the flexibility as set out in this paragraph is a good approach. We recommend this flexible approach be adopted going forward as it is a potential method to simplify relevant information, and it is up to organisations to determine the best approach based on factors including and not be limited to complexity of AI systems, readability of the relevant information and potential risk to data subjects.

Content	Comments
Notification Obligations may also be provided through such model and/or system cards, if the organisation adopts this practice or assesses it to be useful. **Example 1:** An organisation provides a video streaming service. It informs users that its service uses AI to provide recommendations. Through its notification pop-up, it informs users that it collects and analyses users' declaration of topics of interest, browsing activities and media consumption data to recommend videos that users may be interested in. Users are provided the option to consent or decline the use of this feature. The notification pop-up contains a link to its privacy policy, which contains a section that provides information about what declared topics of interest, browsing activity and media consumption data are collected and analysed. This includes the topic classification of videos that users watch, duration and proportion of the video that is played, how many times the video is played, whether the video is watched in a preview window or in actual size, etc. The organisation also explains that the topics of videos that users watch in full are most likely to influence future recommendations. **Example 2:** A social media platform provides an AI system card to its users to explain how its AI uses user activity data to generate recommendations for its content feed. The system card contains a step-by-step walk through on how the AI system gathers user activity data and broadly processes it in its AI system with other parameters to generate personalised output for a content feed. *(paragraph 9.7)*	
PDPC recognises that organisations may need to protect commercially sensitive and/or proprietary information, as well as the security of the AI System. To this end, where organisations	No comments

Content	Comments
assess that it is necessary for any of the aforementioned reasons to limit or omit detail and, if appropriate, provide a more general explanation instead, it is good practice for these decisions to be justified and documented clearly internally. *(paragraph 9.8)*	
Organisations that make use of AI Systems should be transparent and include relevant practices and safeguards to achieve fairness and reasonableness in their written policies. The level of detail to be provided under the **Accountability** obligation should also be proportionate to the risks present in each use-case (e.g., taking into account potential harm to the individual and the level of autonomy of the AI System). *(paragraph 10.3)*	No comments
Organisations must make information about such policies and practices available to individuals upon request. The raison d'etre for such external communications with consumers is to help organisations build trust with data subjects by demonstrating accountability in their compliance with the PDPA. Hence, organisations should consider pre-emptively making such written policies available through their website, and not only upon request of individuals. Organisations should also consider making such written policies available in the form of short policy statements that provide information to consumers simply, clearly, and concisely. Additionally, developing industry best practices, such as model cards and system cards, can also form part of an organisation's written policies. Where organisations have relied on exceptions to the requirement of consent, such as Business Improvement and Research Exceptions,	We recommend clarifying that 1) The exemption afforded under the PDPA to refuse frivolous / vexatious requests (in line with access requests) stands. 2) The use of additional policies / model cards / system cards is not a requirement but a potential method to simplify relevant information, and it would be dependent on organisations to determine the best approach based on factors including but not limited to complexity of AI systems, readability of the relevant information and potential risk to data subjects. 3) When relying on exceptions to the requirement of consent, such as Business Improvement and Research Exceptions, organisations would have conducted a DPIA that includes the practices and safeguards to protect the interests of the individuals. To that point, we suggest PDPC to consider if the requirement to include these information in written policies would be appreciably meaningful as a control.

Content	Comments
written policies can also provide information about the practices and safeguards that were adopted to protect the interests of individuals. *(paragraph 10.4 & 10.5)*	
Organisations are generally encouraged to consider providing more information on data quality and governance measures taken during AI System development. This is only if such information is deemed relevant and doing so does not compromise security, safety, or commercial confidentiality. Information that organisations can consider including are: • Steps taken to ensure the quality of personal data in the training dataset (e.g., how representative it is of the market and how recently it was compiled) to improve model accuracy and performance; • Whether model development was conducted using pseudonymised data, and if not, what organisation, process or technical safeguards were adopted to restrict access to personal data to developers and/or testers who had access; • Whether it was necessary to use personal data when conducting bias assessment to check if protected characteristics, such as race or religion, are well represented in the training dataset or to assess the bias of the training dataset; • If personal data was used, what process or technical safeguards were adopted to secure the testing environment and to limit access to testers; and • Whether data minimisation was practised at all stages of model and/or AI system development and testing. *(paragraph 10.8)*	No comments

Content	Comments
Organisations may wish to refer to the Model AI Governance Framework for further suggestions on managing stakeholder interaction (see in particular Section 3, pages 53 – 55). Organisations may also find the guiding questions and examples on stakeholder interaction provided in Section 5 of the Implementation and SelfAssessment Guide for Organisations (ISAGO) helpful in this regard. Organisations can consider using technical tools such as AI Verify to validate the performance of the AI System or ML model. Information from the testing report can be used to support information that they wish to include into their notifications or written policies. It is good practice for organisations to develop processes to regularly review the quality of the information provided, as well as the effectiveness of its notifications, policies, and practices for their intended audience. *(paragraph 10.9 to 10.11)*	Allowing organisations the flexibility as set out in this paragraph is a good approach. We recommend this flexible approach be adopted going forward as a potential method to assess AI Systems or ML models. It allows organisations to determine the best approach based on factors including but not limited to complexity of AI systems, and potential risk to data subjects.
Part IV: Procurement of AI Systems	
This section is relevant for organisations that engage service providers (e.g., systems integrators) who provide professional services for the development and deployment of bespoke or fully customisable AI Systems. *[Note: It is not relevant to organisations that develop AI Systems in-house or who purchase commercial offthe-shelf solutions that make use of AI for their product features and functions.]* Where service providers, as part of developing and deploying bespoke or fully customisable AI Systems, process personal data on behalf of their customers, they may occupy the position of data	We suggest including a section that relates to the scenario whereby the service provider is a data controller.

Content	Comments
intermediaries and may thus have to comply with applicable obligations under the PDPA. Service providers who are data intermediaries should adopt the following practices: • At pre-processing stage, use techniques such as data mapping and labelling to keep track of data that was used to form the training dataset; • Maintain a provenance record to document the lineage of the training data that identifies the source of training data and tracks how it has been transformed during data preparation. *(paragraph 11.1 & 11.2)*	
The following are best practices on how service providers may support organisations to develop policies and practices that can meet their **Consent**, **Notification** and **Accountability** obligations. **Step 1: Understand the information that customers are likely to require based on their needs and impact on users.** Service providers are encouraged to be familiar with the different types of information described in paragraphs 9.5 and 10.6 – 10.8 of the Guidelines that contribute towards meeting their customer's Consent, Notification and Accountability obligations. To do so, supply-side businesses will have to pay attention to the context and impact the AI System will have on individuals. Information that is likely to be relevant should be identified, and supply-side businesses are encouraged to engage their customers on what will be helpful for them.	We agree with Steps 1 and 2, and believe that service providers are in the best position to carry out these steps.

Content	Comments
Step 2: Design AI system to ensure that organizations can obtain relevant information As part of implementing privacy-by-design, service providers are encouraged to try to build in processes when designing their bespoke or customisable AI System that facilitate the extraction of information relevant to meeting their customers' PDPA obligations. This will enable service providers to better support their customers who may require their assistance in developing policy documents or notifications. These include the following:	

Process	Description of best practice
Providing information for user organisation's internal use	• Translate the operation of AI System and/or ML model into easily understandable language for operators so that they can understand how the outcome is arrived at. • This can be done through textual explanation, visual aids such as video, graphs or table, or a combination of all or some of these. • Use technical tools such as AI Verify to aid in the development of various types of explanations. Please refer to paragraph 10.10 above.
Supporting implementors of AI Systems to understand the acquired AI System	• Where needed e.g., where the AI System is more complex, training for human decision-makers involved in the AI-assisted decision-making process to ensure that they have adequate and appropriate knowledge of how to use the AI System.

Content	Comments
<table><tr><td></td><td>• Where the AI System is intended for autonomous decisionmaking, providing training or a clear explanation to ensure that they properly understand how the AI System operates.</td></tr></table> *(paragraphs 11.4 to 11.6)*	
While service providers can support organisations in achieving their Consent and Notification Obligations as well as Accountability Obligation, PDPC reiterates that organisations bear the primary responsibility for ensuring that the AI System they have chosen to use would be capable of meeting their obligations under the PDPA. *(paragraphs 11.7)*	No comments