GUIDELINE ON:

Data Governance
in
Data Sharing

30 November 2020

Table of Contents

Section 1: Introduction

1.1. This guideline has been issued as part of one of the strategies under the Governance pillar of the National Data Sharing Policy (NDSP)

1.2. This guideline aims to guide both public sector and private sector organisations and its members on the governance protocols and measures needed to oversee the sharing of data and to protect data shared.

1.3. It is noted that there are various data protective measures already in place in the Acts, laws and regulations of some industries to ensure proper governance around the handling of data. The Personal Data Protection Act 2010 is one such Act that has been put forth by the Data Protection Commissioner to protect the personal data of data subjects that currently resides with data users and/or data processors[1].

1.4. Please be reminded that this guideline outlines broad suggestions and recommendations for organisations to have good data sharing governance practices. Should organisations choose to follow this guideline, please ensure that the implementation of suggestions is compliant to any other obligations that organisations must meet as per regulatory requirements.

1.5. This guideline is intended to be read together with the data sharing guidelines for Data Ethics and Data Trust to get a comprehensive overview of the minimum baseline that is recommended for organisations engaging in data sharing.

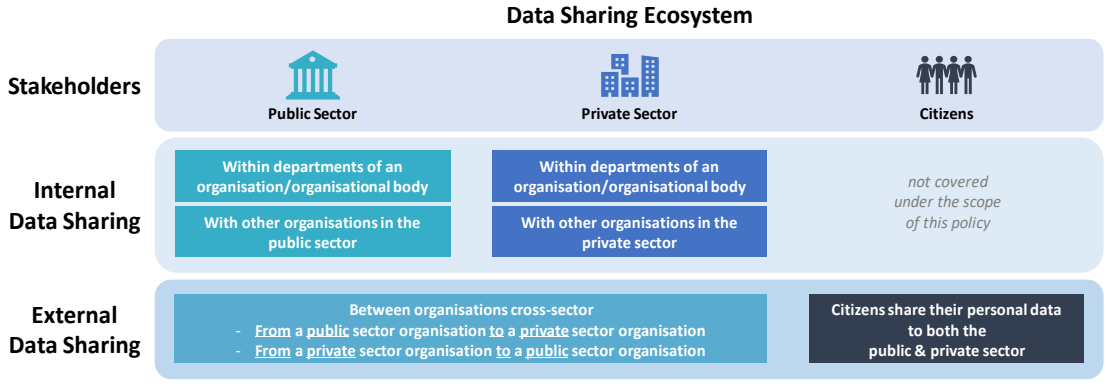
Section 2: What is Data Sharing?

2.1. As per the National Data Sharing Policy, data sharing refers to "*the disclosure of data from one or more organisations to another organisation(s), or the sharing of data between different parts of an organisation*".

2.2. The data sharing ecosystem allows organisations to share data and enables each respective organization to potentially access data they otherwise would not have been able to.

> ***Example:***
> *An organization from the healthcare industry may potentially have access to data from the telecommunications industry for research and development into telehealth opportunities*

2.3. A simplified illustration of the data sharing ecosystem is as below:

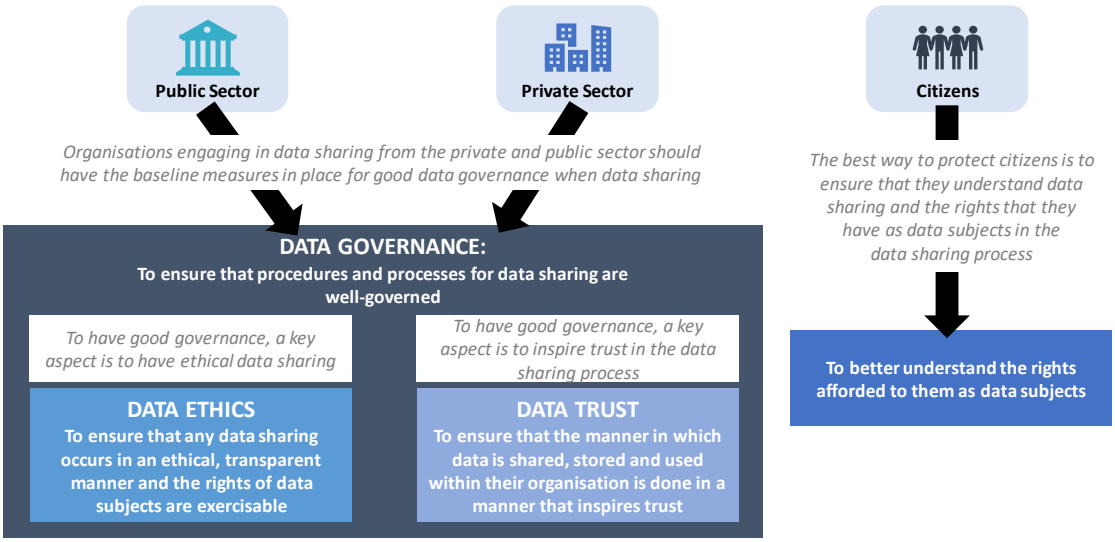


[1] Please refer to PDPA for definitions of "data user" and "data processor".

2.4. In the data sharing ecosystem, there are 3 main stakeholders:

(i) the Public Sector[2]
(ii) the Private Sector[3]
(iii) the Citizens of Malaysia

2.5. The respective roles of the stakeholders are illustrated as below:



2.6. With the ever-increasing reliance on data to feed technologies and decision-making, the benefits of data sharing to enrich and broaden insights is undeniable. With the amount of data moving between stakeholders at an unprecedented high, it is crucial for organisations to have strong data governance in place to oversee their management of data to ensure the data sharing process would occur smoothly, securely and ethically.

2.7. As such, this guideline will zoom into the expectations and recommendations on what to draft the baseline checkpoints, protocols and checklists around to ensure proper governance is in place in an organisation's data sharing processes

Section 3: Data Governance in Data Sharing

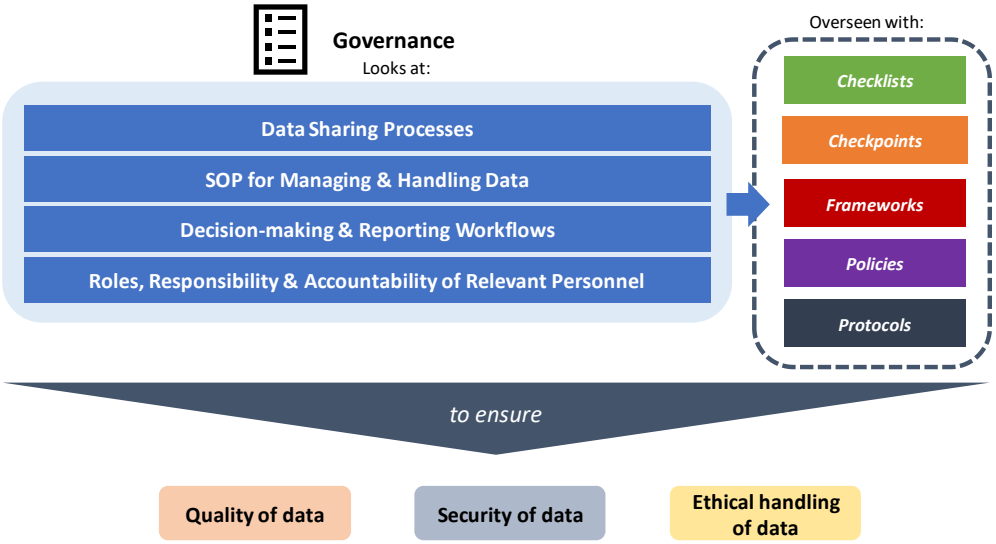
What is Data Governance in Data Sharing?

3.1. Data governance in data sharing focuses on "what are the protocols and measures needed within the organisation to ensure data is well governed".

3.2. In detail, data governance looks at the management of the data lifecycle from cradle to grave (i.e. from collection to disposal/destroy of data) to ensure data processes are running smoothly and being handled in a reliable, secure and ethical manner both by design and through the inculcation of good data governance habits in the workplace.

3.3. The illustration on the following encapsulates what data governance looks at, what it is overseen by and what are the end objectives desired.

[2] Public Sector refers to the federal and state governments, statutory bodies and local authorities.
[3] Private Sector refers to all other entities that do not fall under the Public Sector.



What does Data Governance in Data Sharing Aim to Achieve?

3.4. **For the public and private sector:** The data governance guideline will serve to guide organisations on what is recommended as base points to draft checkpoints, protocols and checklists in the data governance workflows of the data sharing process around to ensure data is being managed smoothly, securely and ethically from beginning to end.

In particular, checkpoints, protocols and checklists should specifically guide on the who, what, how, when, where and why of data. That is,

(i) Specific elaboration of the who, what, how, when, where and why of data to prevent mismanagement of data sharing process

(ii) Specific elaboration of the responsibilities, accountabilities and roles of each personnel and process

3.5. **For citizens:** Data governance in data sharing aims to assure data subjects[4] that any data of theirs that are with organisations are well governed and protected by organisations.

Why is Data Governance in Data Sharing Important?

3.6. Data governance oversees the lifecycle of data when data sharing. Without strong and clear data governance in place, the management and handling of data may be weak and could potentially raise concerns on the reliability of data as well as the security and ethical handling of data.

3.7. It is important to set clear governance checkpoints, protocols and checklists for the entire data sharing process so that organisations have standard procedures and workflows to ensure good governance that personnel can easily follow without confusion.

Additional Notes for Data Governance in Data Sharing Guideline

3.8. As noted, data governance in data sharing encompasses the entire data sharing process wherein the ultimate aim is helping guide organisations data governance practices to be smooth, secure and ethical. However, as separate guidelines have already been drafted focusing on the security (i.e. Data Trust in Data Sharing Guideline) and ethicalness (i.e. Data

[4] As per PDPA, "data subject" refers to an individual who is the subject of the personal data.

Ethics in Data Sharing Guideline) of data sharing process, this guideline will focus on suggestions in terms of the resources and checkpoints needed to ensure proper governance is carried out.

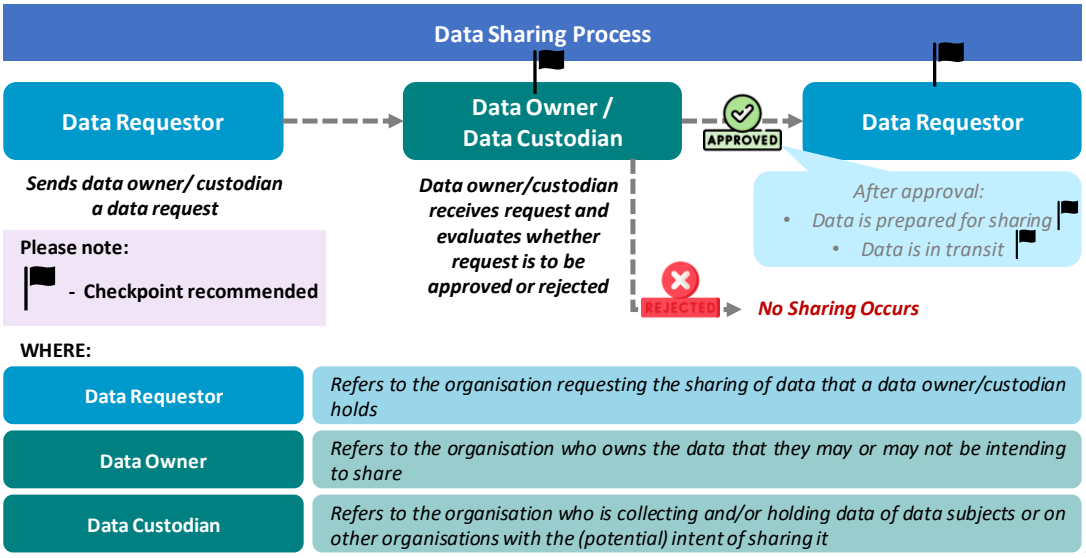
Section 4: Recommended Resources to Oversee Data Governance

4.1. With the necessity of seeing the data from end to end, it is recommended that organisations have the following resources in place as illustrated in the figure on the following page.

Recommended Resources	
Data Lead	The data lead in the organisation is the subject matter expert in all things data and is the main personnel charged with overseeing the entire data governance process. The data lead is accountable for the data decisions within the organisation, will be the lead in drafting any procedures, policies, etc. to oversee data governance and will ensure data governance in the organisation is sound, secure and ethical.
Data Steward(s)	There can be more than 1 data steward in the organisation. These data stewards are responsible for overseeing the governance of data within their domain. Data stewards are tasked with ensuring the reliability and integrity of their data domain and reports to the data lead. Data stewards work alongside the data lead to improve data governance workflows and processes.
Data Team	The data team consists of personnel who are tasked with the day-to-day responsibilities of carrying out the managing, handling, processing and sharing of data. Individuals in the data team report either to the data steward of the data domain they are in or report directly to the data lead (depending on employment contract terms).
Steering Commitee	The steering committee consists of lead executives of the organisation, the data lead, data stewards and other select personnel. The steering committee is responsible for overseeing the data strategy of the organisation and to provide ample support and resources for the data controller to execute data governance initiatives and updates. The committee also reviews and monitors the timeline and outcome of initiatives.

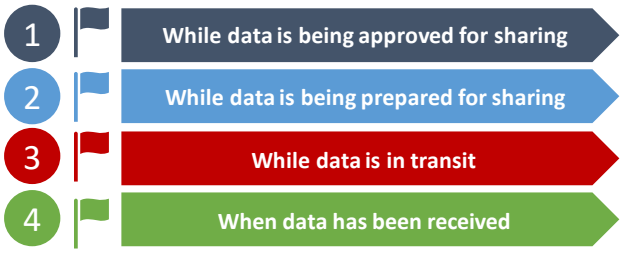
Section 5: Checkpoints for Data Governance in Data Sharing

5.1. This section will explore the recommended checkpoints at each stage of the data sharing process to ensure good data governance in the data sharing process. Please see an illustration as of the data sharing process as below:



[5,6]

5.2. As observed in the illustration in the previous page, there are 4 checkpoints that is recommended to be integrated in an organisation's data sharing process. Specifically, these checkpoints are as follows:



5.3. The following sections are intended to guide organisations on what to keep in mind when drafting the checklists/protocols/policies for each checkpoint.

Checkpoint 1: While Data is Being Approved for Sharing

5.4. When data is shared, there is potential for data to be:

(i) Shared to the wrong hands, where data may be misused

(ii) Shared to someone with insufficient governance measures in place, where data is at risk

(iii) Data that should not have shared has been shared out

5.5. As such, this checkpoint is imperative as to ensure proper vetting/screening of data request and the data requestor. In particular:

Screening of Data Request	Screening of Data Requestor
• Type of data requested is justified • Data is relatable to what the data requestor does	• Does data requestor have the appropriate governance measures in place to handle and protect data • How the data requestor will use the requested data

[5] Please note: The approved and rejected icons are sourced from Flaticon.com.

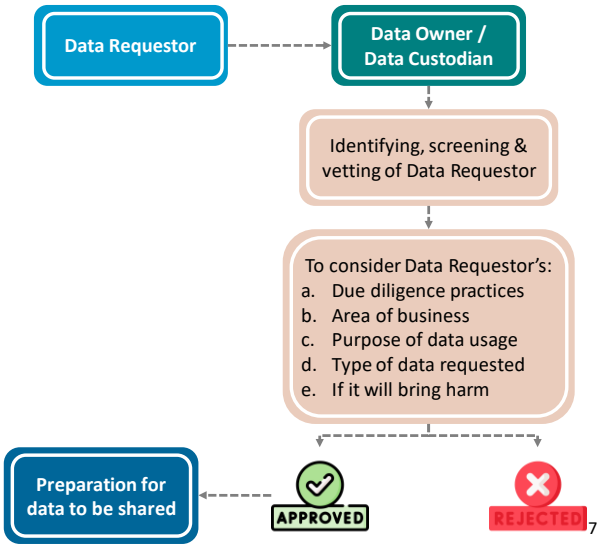
[6] Please note that PDPA defines out term "requestor" in relation to a data access/correction request by a data subject or a relevant person on behalf of the data subject. This guideline however defines the term "data requestor" in relation to an organisation requesting data from a data custodian.

• Data is safe to be shared out by data owner/custodian under the appropriate security and ethical measures	• How data requestor manages and conducts due diligence when handling data requested *(important to know as these practices are the backbone to ensure data is properly handled and not subject to misuse)*

5.6. Please find the guidelines of items to keep in mind in setting up the checkpoint for vetting of data requests and the data requestor:

- ✓ Screening of the data requestor
- ✓ Due diligence practices that will be taken by the data requestor
- ✓ Area of the business the data requestor is in
- ✓ Use purpose of data that will be attained
- ✓ Type of data requested
- ✓ Consider if sharing such data will be harmful in the hands of the data requestor

5.7. Please find a sample illustration of suggested checkpoint in data sharing process is as follows:



[7]

Checkpoint 2: While Data is Being Prepared for Sharing

5.8. Once data request from data requestor has been approved, it is essential to prepare the data for sharing, especially when dealing with sensitive information.

5.9. This checkpoint is crucial to reduce risk of data leakages (wherein data that should not have shared is shared out) and data inaccuracies. Where sensitive data is involved, this checkpoint also includes ensuring encryption of data to hedge any risks during the transference process.

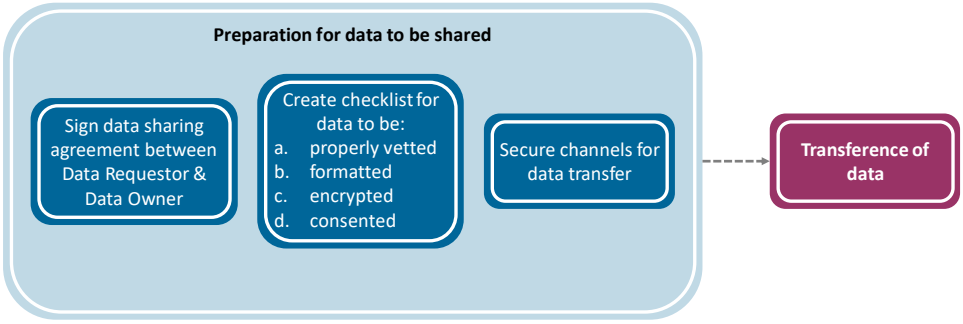
5.10. Please find the guidelines of items to keep in mind when preparing datasets for sharing

- ✓ Ensuring data being shared has been consented to be shared for specific purposes by data owners/custodians
- ✓ Ensuring data is encrypted and prepared properly for sharing purposes
- ✓ Ensuring any harmful data is omitted from the data set
- ✓ Ensuring data is tabulated accurately

[7] Please note: The approved and rejected icons are sourced from Flaticon.com.

- ✓ Creating a checklist on data fields required and which fields are to be encrypted or omitted
- ✓ Ensuring data is processed tidily and comprehensively
- ✓ Ensuring data sharing agreement is vetted by C-Suite parties and legal parties so that data shared cannot be used for anything other than what it is shared for
- ✓ Ensuring all involved parties are fully responsible and that legal action will be taken for any leakages, incoherencies, or misuse of any data within the requested datasets
- ✓ Ensuring parties involved are at an agreement and that there are safe avenues for data transference and sharing

5.11. Please find a sample illustration of suggested checkpoint as follows:



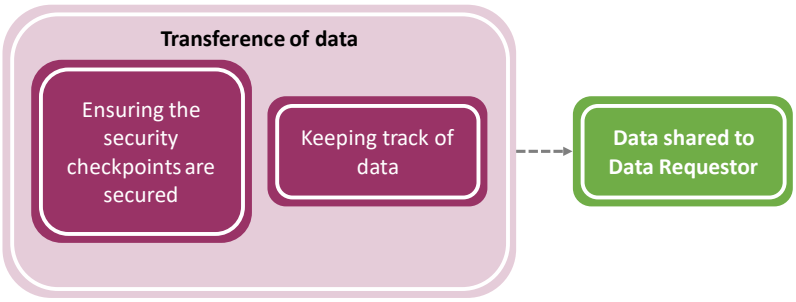
Checkpoint 3: While Data is in Transit

5.12. This checkpoint is important to monitor the transit of data wherein data is moving securely between data owner/custodian and data requestor.

5.13. Please find the guidelines of items to keep in mind in when transferring data:

- ✓ Ensuring data shared is accurate and there are no leaks of data which can be used by other external parties
- ✓ Ensuring datasets are accurate and only requested data which can be shared has been disseminated. No incoherencies or leakages are present and can be justified
- ✓ Ensuring data sharing method is secure and encrypted so that security of data is maintained
- ✓ Ensuring security checkpoint requirements are met and data transfer is tracked

5.14. Please find a sample illustration of suggested checkpoint as follows:



Checkpoint 4: When Data has been Received

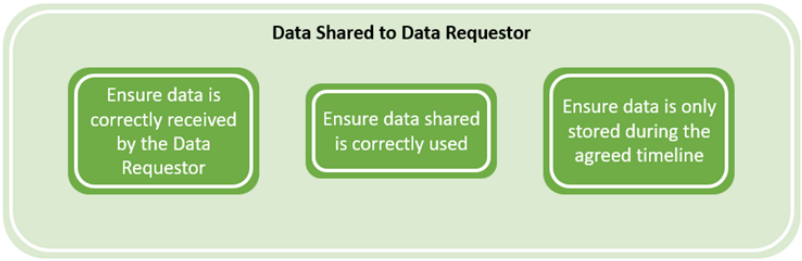
5.15. The final checkpoint in the data sharing process is in ensuring that data requestor has received the requested dataset. This entails the carrying out of additional due diligence to ensure when data is handled upon receival to ensure data is not leaked to another party.

5.16. Following on from the receival of data, it is important to ensure data misuse does not occur in the data requestor's organisation with terms as stated via the data sharing agreement.

5.17. Please find the guidelines of items to keep in mind in setting up the checkpoint for vetting of data requests and the data requestor:

- ✓ Ensuring data is collected by the correct data requestor and managed properly
- ✓ Ensuring data transfer is a success and all datasets are shared completely
- ✓ Ensuring data use is closely monitored and the Data Requestor is accountable to report use of data in any instance which is governed by the parameters of the agreement made
- ✓ Ensuring data is not misused or misconstrued
- ✓ Ensuring data is stored for agreed timeline and deleted responsibly by the Data Requestor post time agreed upon

5.18. Please find a sample illustration of suggested checkpoint as follows:



Appendix A: Recommendations for Data Categorisations

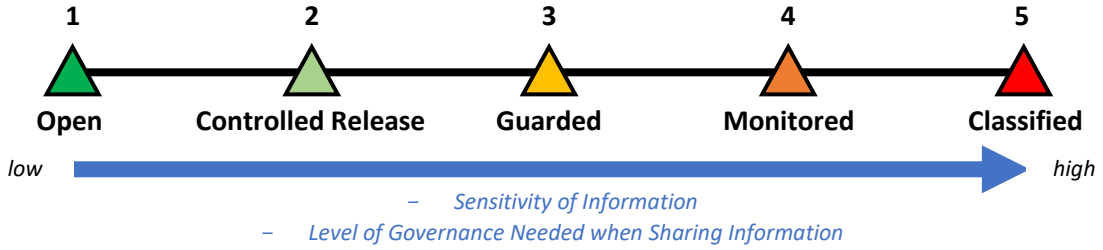
While a baseline of governance is crucial to the data sharing process, it should be noted that governance may need to vary in terms of stringency in accordance to the type of data that is to be shared.

Example:
Data that are highly personal in nature such as the medical history of data subjects or the trade secrets of an organisation will require more stringent security protocols as opposed to non-personal data such as data on climate or anonymised data

5-point Scale for Data Categorisation

In order to better help organisations differentiate the level of governance needed when data sharing, it is recommended that organisations categorise data to be shared via a 5-point scale. Please find it as follows:

Scale:



Description of Scale:

1 - Open

This categorisation should be used for all non-personal data

Non-personal data refers to data that neither pertains to any personally identifiable information (e.g. name, contact details, identity card number, etc.) nor any information pertaining to the tracking of behaviours at an individual-level. It also pertains to information that has already been made publicly available (e.g. annual reports of a public-listed company).

Sharing of such data is typically of negligible risks to data subjects if shared as it is non-personal in nature and information is on a statistical, aggregate basis.

Examples of Data Include	– Data on general conditions *(this can include conditions of climate, market, companies, etc.)* – Data that has been anonymised and aggregated – National, industry-level and organisational-level statistics
Data Shareability	Data under this category can be shared in an open manner. That is, the sharing of such data can be done on a widespread basis to all interested parties with minimal security concerns.

2 - Controlled Release

This categorisation should be used for non-sensitive personal data

Non-sensitive personal data refers to anonymised data that does not contain any personally identifiable information (e.g. name, contact details, identity card number, etc.). This data is used to

track and analyse individual behaviours of data subjects which do not pertain to sensitive information such as the health and financial conditions of data subjects.

Sharing of such data is typically of low risks to data subjects as it is non-sensitive in nature and due to the measures taken to prevent re-identification of data.

Examples of Data Include	– Data that has been anonymised but left at a row-level basis (i.e. personal data of data subjects can be seen but cannot be tied to a data subjects' identity. This type of data is used to track behaviours of individuals without identifying them) – Data that does not pertain to the personally identifiable information of a data subject – Data that does not pertain to the health and financial condition of data subjects (even on an aggregate basis) – Data that does not pertain to the geolocational data of data subjects (even on an aggregate basis)
Data Shareability	Data under this category can be shared in a controlled manner.

3 - Guarded

This categorisation should be used for personal data which contains personally identifiable information but no sensitive personal data

Personally identifiable information (PII) refers to any data that can be used to identify an individual. This includes the name, home landline, mobile number, home address, email address, identity card number, digital images and social media of individuals. PII can also refer to sensitive personal data. As per the Personal Data Protection Act of Malaysia, sensitive personal data is defined as "*any personal data consisting of information as to the physical or mental health or condition of a data subject, his political opinions, his religious beliefs or other beliefs of a similar nature, the commission or alleged commission by him of any offence or any other personal data as the Minister may determine by order published in the Gazette".* It can also pertain to the financial status/condition, geolocation and biometric information of individuals. However, the sharing of data under this category shall not include such sensitive personal data.

Sharing of such personally identifiable information that do not include sensitive personal data is of moderate risk to data subjects as it pertains to their searchable information and contact points others can reach them at.

Examples of Data Include	– Data which includes the personally identifiable information of a data subject that are deemed non-sensitive in nature – Data that does not pertain to sensitive personal data of data subjects as defined above (even on an aggregate basis) – Data that does not pertain to the geolocational data of data subjects (even on an aggregate basis)
Data Shareability	Data under this category must be shared in a guarded manner.

4 - Monitored

This categorisation should be used for sensitive personal data

As noted previously, in accordance with PDPA, sensitive personal data refers to "*any personal data consisting of information as to the physical or mental health or condition of a data subject, his political opinions, his religious beliefs or other beliefs of a similar nature, the commission or alleged commission by him of any offence or any other personal data as the Minister may determine by order published in the Gazette".* It can also pertain to the financial status/condition, geolocation and biometric information of a data subject. Other data that are classified as sensitive include any PII pertaining to children under the age of 18 years.

Sharing of such information is of high risk to individuals as it pertains to the data subject's personal wellbeing and their daily movements. The sharing of the data of children are also high risk as children are deemed as especially vulnerable to and may potentially be unaware of the potential consequences of sharing their data.

Examples of Data Include	– Data pertaining to the health and/or financial condition of a data subject – Data pertaining to the geolocation and biometric information of a data subject – Data pertaining to any personally identifiably information of children (whether sensitive or non-sensitive)
Data Shareability	Data under this category must be shared while being monitored.

5 – Classified

This categorisation should be used for data that has been classified and/or is not permitted to be shared

Classified data refers to data that have been deemed sensitive to the national security of the nation and should thus, not be disseminated. Beyond that, classified data also pertains to data that may compromise the security of an individual or the core business of an organisation as well as data wherein data subjects have not consented to its sharing and thus, should not be permitted to be shared.

Sharing of such information is of immeasurable risk due to the harm that may befall the related individuals, organisations and even the nation, and/or is in violation of the rights of data subjects. As such, sharing of such information should thus be avoided.

Examples of Data Include	– Data that data subjects have not consented for data custodian to share to other organisations – Data that may compromise the safety of the individual – Data that may compromise the core business of an organisation – Data that may jeopardise the national safety/defence of the country
Data Shareability	Data under this category should not be shared under any circumstance*. **Please note that leeway may be provided in the case of national emergencies*