GUIDELINE ON:

Data Ethics
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 standard baseline needed to have good data ethics practices when data sharing.

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 ethical practices when handling data. The Personal Data Protection Act 2010 (PDPA) 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 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 Trust and Data Governance to get a comprehensive overview of the minimum baseline that is recommended for organisations engaging in data sharing.

Section 2: What is Data Sharing?

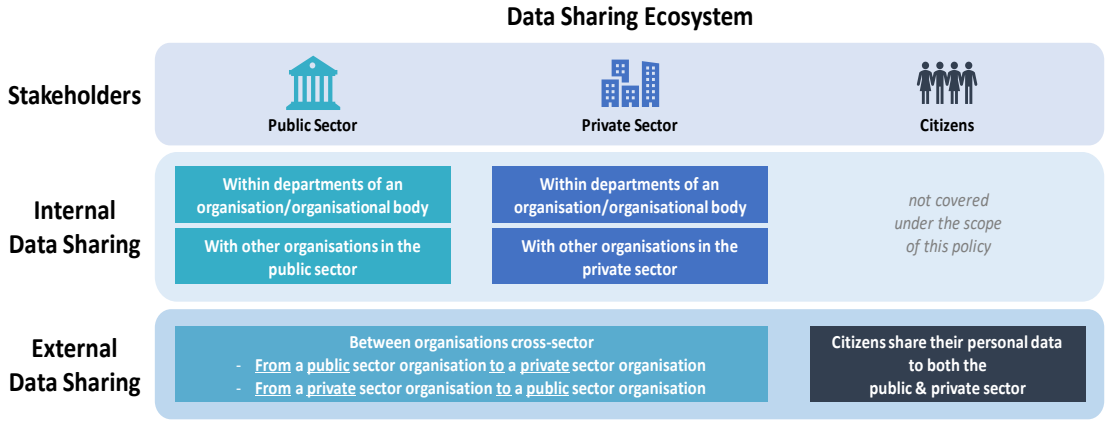
2.1. As per the National Data Sharing Policy, data sharing refers to "the disclosure of data from one or more organisations to another organisation(s), or the sharing of data between different parts of an organisation".

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:

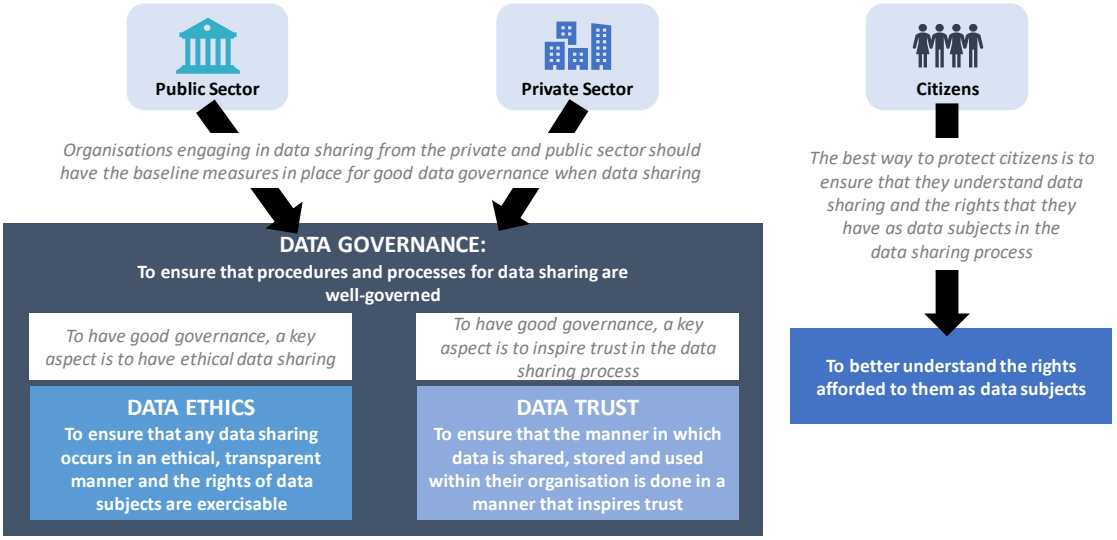
[1] Please refer to PDPA for definitions of "data user" and "data processor".



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. However, data sharing must be done in an ethical, transparent manner and should not infringe upon the privacy of an organisation whose data is being shared[4] nor the privacy and/or rights of a data subject[5].

2.7. As such, this guideline will zoom into the expectations and recommendations surrounding data ethics in data sharing for the stakeholders in the data sharing ecosystem

[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.
[4] This pertains to data that has not been voluntarily disclosed, whether via written or verbal disclosure, by the organisation.
[5] As per PDPA, "data subject" refers to an individual who is the subject of the personal data.

Section 3: Data Ethics in Data Sharing

What is Data Ethics in Data Sharing?

3.1. Data ethics in data sharing focuses on "what data can be shared within the ethical limits".

What does Data Ethics in Data Sharing Aim to Achieve?

3.2. **For the public and private sector:** Data ethics will serve to guide organisations in the data sharing ecosystem on their ethical responsibilities when data sharing. This is done through the setting of a baseline of principles and values that stakeholders can refer to when considering the ethical boundaries of sharing data.

3.3. **For citizens:** Data ethics in data sharing aims to assure data subjects that any data of theirs that are with organisations will not be shared in a manner that violates their rights as data subjects and/or their privacy.

Why is Data Ethics in Data Sharing Important?

3.4. With the increase of data sharing, it is important to set clear boundaries to ensure that any sharing of data does not infringe upon privacy and rights of the data subjects.

> ***Example:***
> - *Case 1 – Where data subject is a citizen and data to be shared pertains to their personal information, data should not be shared without their consent*
> - *Case 2 - Data to be shared pertains to a corporate company's intellectual property (IP), data should not be shared without their consent*

3.5. These boundaries will help guide the values and principles of organisations in the data ecosystem around responsible sharing of data. These guiding values, if well-instilled in an organisation, will better protect data subjects and data on companies, relative to having security watchdogs in place, as data ethics is built into the foundation of all data sharing processes of the organization.

Section 4: How to Practice Ethical Data Sharing?

4.1. This section will look at what are the baseline principles recommended to practice ethical data sharing.

4.2. This is applicable to any organisation, from the public sector or private sector, collecting and/or holding data of data subjects with the intent of sharing (i.e. data custodians).

Base Principles for Data Ethics in Data Sharing

Consent Explicitly Obtained

4.1. Where personal data of a data subject is involved, sharing is ethical if the data subject has given consent to the sharing of their respective data.

4.2. All purposes for sharing a data subject's information should be clearly communicated to the data subject and consent should be obtained for each and every purpose noted.

4.3. Consent should have the following characteristics:

- ✓ freely given by data subject
  should data subject not consent to their data being shared, it should not necessarily data subject's ability to access the product/service of the data custodian.

  Example where characteristic is not practiced:
  Adam refused to tick the box indicating his consent to allow Company A who is collecting his data to share data with other organisations. Because of that, Adam is unable to access Company A's website.

  In this situation where 'consent' of data subject is forced in order to access an organisation's product/service, this is not constituted as true consent.

- ✓ clear, informed decision of a data subject
  in requesting a data subject's consent, request should be worded in a clear, concise manner using plain language so that data subject can make a clear, informed decision.

  Example where characteristic is not practiced:
  Adam could not make sense of the terms and conditions of Company A who is requesting his consent for the sharing of his data. The terms used are overly complex and does not clearly state what his data will be used/shared for but rather addresses it in a roundabout manner.

  In this situation where data subject is unable to clearly understand what they may be consenting to, this is not constituted as true consent.

- ✓ unambiguous indication from data subject to allow the sharing of their information
  when receiving a data subject's consent, consent should be obtained in a format that is a clear indication of a data subject's agreement to the sharing of their data. Consent should not be assumed or implied, it should be explicitly received.

Example where characteristic is not practiced:

When Adam was filling in an application form for Company A's service, he noted that there was a small disclaimer at the end of the page which stated that upon his submission of the form, he also grants permission to Company A to collect and share his data if Company A chooses to.

In this situation where 'consent' of data subject is implied and does not require the explicit action of data subject to approve/disapprove request for consent, this is not constituted as true consent.

4.4. Once consent has been collected, it is important for data custodian to keep clear records of consent collected so that this may be produced as evidence of consent if needed.

4.5. Data subject should be allowed the right to withdraw consent at any time. Data custodians are to ensure that this right is clearly communicated to data subjects and can be easily exercised by data subjects.

- *Please note that withdrawal of consent shall not affect ethicalness of sharing data that has been consented to be shared prior to its withdrawal.*

Example:

Company B has shared personal data of data subjects to Company C on the 5th January 2019 and 25th August 2020. Brenda initially consented to Company B sharing her information on 1st January 2019, but subsequently withdrew her consent on 30th January 2019.

As such, when Company B shared her data on 5th January 2019, it was ethical to do so. However, should Company B share her data on 25th August 2020, it would be considered unethical and in violation of Brenda's rights as a data subject.

Transparency in Data Sharing Process

4.6. Sharing is ethical if data subjects are provided the option to understand how their data is being disseminated and circulated, and for what purpose(s) their data is being shared for.

4.7. Data custodians are to ensure that data subjects have been made aware of how and why their data is being shared to other organisations:

(a) If data shared is on aggregate level –

Organisations are to assure data subjects that their identities will be protected as data is only shared at an aggregate level. Organisations are to ensure sufficient measures have been taken so identities of individuals are properly masked.

(b) If data shared is on a non-identifiable, individual level –

Organisations are to assure data subjects that no personally identifiable information has been shared. Organisations are to ensure sufficient measures have been taken to prevent re-identification of data.

(c) If data shared is on an identifiable, individual level –

Organisations are to ensure that the explicit consent of data subjects has been received to share their personally identifiable information. Organisations are to disclose to whom the personally identifiable information of data subjects are being shared to.

4.8. Data custodians are to have appropriate measures in place to communicate details relating to the sharing of the data subject's data in a clear, concise, transparent and easily accessible manner using plain language.

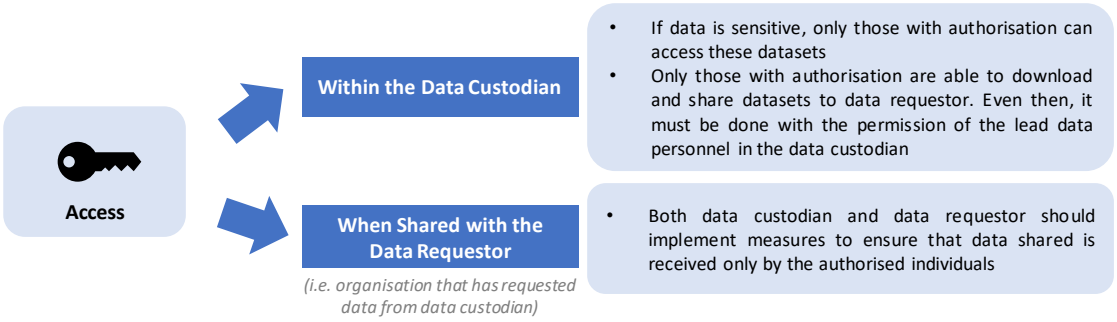
- *Such details shall be provided upon a data subject's request and can be provided in any format so long as there is proof that details communicated are in relation to the data subject.*

> ***Example:***
> *Company D allows Caleb the option to call in to inquire as to understand what data of his is being shared out and for what purpose it was shared for. Company D has ensured that they have clearly communicated to Caleb his ability to call in and ask on this.*

Accessible only by Authorised Individuals

4.9. Sharing is ethical if it is ensured that only authorised individuals are able to access and share data.

4.10. Data custodians must have measures in place to prevent data that they hold from being accessed by personnel who do not have the authorisation to access said data.



> ***Example:***
> *Delia of Company E wants to access a dataset with some sensitive information. However, since the role she is in does not have the authorisation to do so, she is unable to access the database with said information.*

Rights of Data Subjects Exercisable

4.11. Sharing is ethical if data subjects have the ability to exercise the following rights:

(a) Rights to Access

Data custodian should allow data subjects the right to inquire whether data is being shared and if yes, to request more information on:

- *the purposes of data being shared*
- *how data shared will be utilised*
- *the categories of data concerned (as per Section 5)*
- *where personal data is concerned:*
  - *the fields of data being shared*
  - *the recipient of data shared*
- *the retention period of data shared with the recipient*

(b) Rights to Edit

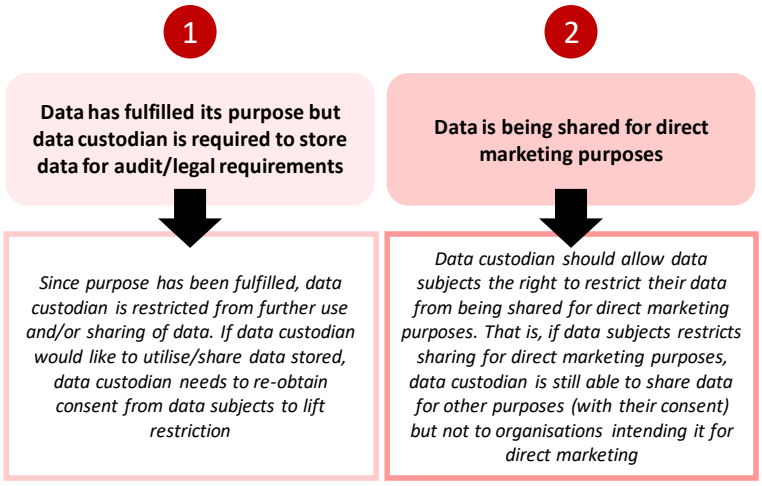
Data custodian should allow data subjects the right to rectify and/or edit any inaccurate data relating to the data subject. An acceptable timeframe for acceptance of rectification and for changes to be reflected is 2-4 working days. If more time is needed, this should be clearly communicated to the data subject.

(c) Rights to Withdraw Consent

Data custodian should allow data subjects the right to withdraw consent that they have previously given. An acceptable timeframe for acceptance of withdrawal is 2-4 working days. If more time is needed, this should be clearly communicated to the data subject.

(d) Rights to Restrict Data Sharing

Data custodian should allow data subjects the right to restrict data custodian from sharing their data in the following circumstances:



Section 5: Ethical Data Sharing for Data Pertaining to Intellectual Property of an Organisation

5.1. Data sharing of data pertaining to an organisation's intellectual property should not infringe upon the privacy of said organisation.

5.2. That is, data can only be shared if the data has been voluntarily disclosed by the organisation itself to the data custodian, whether in a written or verbal format.

Section 6: Guide on Data Ethics Needed for Each Data Category

6.1. While the general implementation of **Sections 4 and 5** are needed to have ethical data sharing, the stringency of its practice is dependent on the type of data that an organisation wants to share and should thus, vary accordingly.

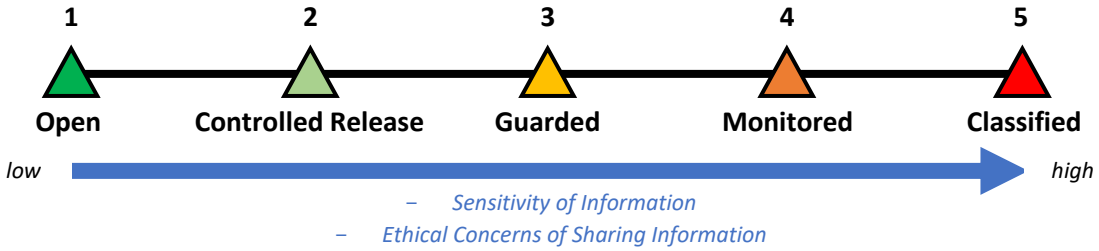
Example:

Data that are highly personal in nature such as the medical history of data subjects or the trade secrets of an organisation will face much higher ethical concerns as opposed to non-personal data such as data on climate or anonymised data

5-point Scale for Data Categorisation

6.2. In order to better help organisations differentiate the level of data ethics needed when data sharing, it is recommended that organisations categorise data to be shared via a 5-point scale. This 5-point scale will determine the level of stringency required for application of data ethics. 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 Ethics Needed for Data Sharing of Data Under This Category	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 ethical concerns. The data ethics needed for data under this category pertains to ensuring: – data shared is non-personal in nature and/or data shared has been anonymised and aggregated

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 Ethics Needed for Data Sharing of Data Under This Category	Data under this category can be shared in a controlled manner. The sharing of such information should be approved by the leading data personnel of the organisation before it can be disseminated. The data ethics needed for data under this category pertains to ensuring: – no identifiable and/or sensitive personal data is included in data shared – precautionary measures have been taken to prevent re-identification of anonymised data

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 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 Ethics Needed for Data Sharing of Data Under This Category	Data under this category must be shared in a guarded manner. The sharing of such information should be approved by the leading data personnel of the organisation before it can be disseminated. The data ethics needed for data under this category pertains to ensuring: – no sensitive personal data is included in data shared – no unauthorised individuals can access such information – no individuals can pull data out from the organisation

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 Ethics Needed for Data Sharing of Data Under This Category	Data under this category must be shared while being monitored. The sharing of such information should go through at least 2 layers of approval – via an ethics committee established within the organisation and finally, the leading data personnel of said organisation before it can be disseminated. The data ethics needed for data under this category pertains to ensuring: – no unauthorised individuals can access such information – no individuals can pull data out from the organisation

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 Ethics Needed for Data Sharing of Data Under This Category	Data under this category should not be shared under any circumstance*. As a result, the highest level of data ethics should be applied wherein organisation should ensure that data is ethically treated and have taken stringent, precautionary measures to prevent the unethical dissemination of such data. **Please note that leeway may be provided in the case of national emergencies*