



DATA SHARING HANDBOOK

For Banks and Non-Bank
Data Ecosystem Partners

Foreword by ABS

Digitalisation is fast changing how new and enhanced financial services are delivered to customers. Access to data has become a key enabler of the digital economy, and banks are no exception. Yet the power of data for decision-making depends on the depth, breadth and types of data available, and there is only so much data that each organisation has access to within its boundaries. Consequently, data sharing between organisations has become increasingly important as it facilitates industry wide innovation and increases business agility. This data-oriented approach has now generated much attention among banks and their partners in this data ecosystem as they adapt to rapid changes in consumer and business expectations.

The Association of Banks in Singapore is pleased to present the "Data Sharing Handbook for Banks and Non-Bank Ecosystem Partners" to promote a standardised approach for the sharing of data. The Handbook aims to define a common language and approach to data sharing for banks, and guide their ecosystem partners in engaging with banks on data sharing. It outlines core principles to enable safe data sharing and provides an illustrative data sharing journey which outlines key considerations in the exchange of data among organisations. As data sharing programmes are initiated in other markets, we hope that this Handbook will also be a useful reference for them.

I would like to take this opportunity to acknowledge the contributions of the ABS Standing Committee on Data Management, PwC Singapore, the Monetary Authority of Singapore and Infocomm Media Development Authority that have culminated in the publishing of this Handbook.

We hope you find this Handbook a valuable resource in planning and implementing the sharing of data in your business.



Ong-Ang Ai Boon, Mrs

Director

The Association of Banks in Singapore (ABS)

Foreword by MAS

There are two key drivers that will have a major impact on the financial services industry in Singapore as the global world digitalises. They are the increasing role of data to the success of financial institutions and also the blurring of lines between financial services industry and other industries.

In the financial services industry, data will drive the use of technologies such as artificial intelligence and data analytics to offer new insights, opportunities and also risks. For example, business models and processes will transform with the use of robo-advisory services, AI driven investment management and digital banking.

To prevent information silos, which will limit the ability to maximise the use of data for innovation, greater sharing of data will avail more data for deeper analysis to generate insights that was not possible in each silo previously. In addition, as the line between financial services industry and other industries gets blur, there will be a need for data sharing between industries and this will result in better customer experience, more business opportunities, greater profitability and ultimately higher potential for the general economy. Nevertheless, the sharing of data needs to be underpinned by trust and security and in line with the existing regulations.

The Monetary Authority of Singapore (MAS) has been involved in a number of initiatives on data sharing such as the Singapore Financial Data Exchange (SGFinDex) which will enable Singaporeans to consolidate their financial information for more effective financial planning. MAS is also involved in the pilot to build a platform to securely share data with patient consent for more efficient claims processing.

This Handbook, jointly developed by industry practitioners and expert group, is a timely document which will shed light into the data sharing journey of financial institutions and lays out the data sharing principles and keys considerations such as types of data, data sharing patterns, law and regulations. It is hoped that this document will lead to more discussion and collaborative effort on data sharing going forward.

I hope this Handbook will be a useful and living guide in facilitating data sharing conversations and promotes data sharing in a safe and secure manner among the relevant stakeholders such as financial institution and ecosystem partners. In conclusion, we would like to thank and congratulate ABS, members of the financial service industry and ecosystem partners for this collaborative journey.



Vincent Loy

Assistant Managing Director (Technology)
Monetary Authority of Singapore (MAS)

Acknowledgement

The Data Sharing Handbook (“Handbook”) was commissioned by the Standing Committee on Data Management of The Association of Banks in Singapore (ABS-SCDM). It facilitates data sharing between banks and their ecosystem partners to better meet customers’ financial needs in today’s digital economy, while upholding customers’ trust in banking institutions to protect the privacy of their information as enshrined in the Banking Act.

The Handbook was crafted with the inputs of the ABS-SCDM working group members from:

- Australia and New Zealand Banking Group Limited
- Citibank NA / Citibank Singapore Limited
- Bank of China Limited, Singapore Branch
- DBS Bank Limited
- Deutsche Bank AG, Singapore Branch
- Maybank Singapore Limited
- MUFG Bank, Ltd.
- Oversea-Chinese Banking Corporation Limited
- Standard Chartered Bank (Singapore) Limited
- United Overseas Bank Limited

In addition, the Handbook was completed with the strong support of:

- PwC Singapore

We would also like to acknowledge the guidance from our regulatory stakeholders, specifically the Monetary Authority of Singapore (MAS) and the Infocomm Media Development Authority (IMDA).

Finally, we would like to thank all organisations and individuals not otherwise mentioned, for their active participation in the creation of the Handbook.

Noel D’Cruz

Head Group Data Management
OCBC Bank

Sameer Gupta

Chief Analytics Officer
DBS Bank

Richard Lowe

Group Chief Data Officer
UOB Limited



“ *Data sharing capitalises on the non-rivalrous nature of data, and allows multiple partners to enjoy different beneficial uses over copies of data. It is also important to remember that data sharing should be managed as an ongoing relationship. The Data Sharing Handbook provides a helpful framework for banks and ecosystem partners to share data in a responsible manner. I am pleased that ABS has adapted and expanded IMDA’s Trusted Data Sharing Framework for banking data.*

YEONG Zee Kin

Assistant Chief Executive (Data Innovation and Protection Group), Infocomm Media Development Authority of Singapore and Deputy Commissioner, Personal Data Protection Commission



Contents

About this Handbook	7
What is Data Sharing?	11
PLAN	
I. Principles of Data Sharing	18
II. Types of Data	26
III. What is Sensitive Data?	33
DESIGN	
IV. Data Sharing Models	44
V. Legal and Regulatory Considerations	48
IMPLEMENT	
VI. Putting It All Together	61



About this Handbook

What is this Data Sharing Handbook?

The Association of Banks in Singapore (ABS) has developed this Data Sharing Handbook (the “Handbook”) in an effort to demystify banking data and foster purpose-driven, secure, and lawful data sharing between banks and ecosystem partners. This builds upon the IMDA’s Trusted Data Sharing Framework¹ and has been written with a sharper focus on risk management considerations relating to banking data. It is primarily intended for:

- **Non-bank data ecosystem partners**, to encourage more understanding and initiatives in data sharing with banks, with a specific focus on banking data; and
- **Banks**, to address common obstacles and encourage safe and responsible data sharing.

This Handbook is structured according to a typical data sharing journey and the key considerations involved throughout the process. The following three case studies are also provided to illustrate the application of the Handbook:

<p>Page 21</p>  <p>Case Study 1: Bringing the principles to life</p>	<p>Page 41</p>  <p>Case Study 2: Managing sensitive data</p>	<p>Page 58</p>  <p>Case Study 3: Contracting</p>
--	--	--

¹IMDA has published a Trusted Data Sharing Framework, incorporating guidance from PDPC, that sets out the baseline considerations and acts as a starting point for organisations who are exploring data sharing partnerships. For more information, visit: www.go.gov.sg/data-innovation



About this Handbook

What is the Data Sharing Journey?

The Data Sharing Journey, illustrated on the following page, is an example of the steps and key considerations involved in data sharing.

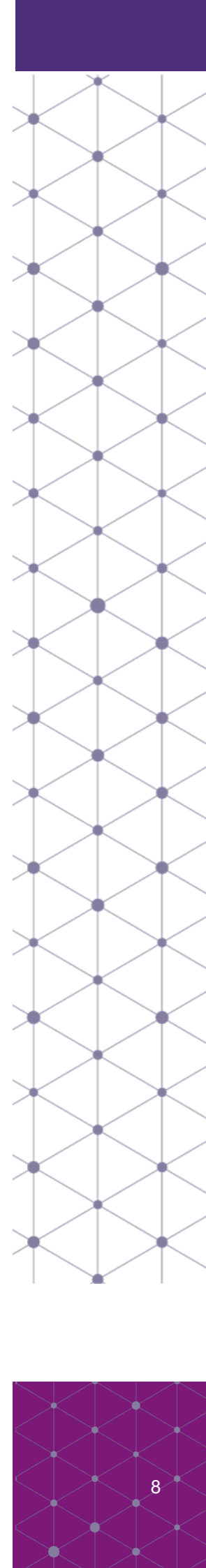
Data sharing participants generally go through three stages when engaging in data sharing:

- Planning;
- Design; and
- Implementation.

The Journey starts with identifying the purpose of data sharing and lays out five principles that should be the core tenets of any data sharing between banks and ecosystem partners. These principles arc through the Journey, for example in identifying the types of data shared, data sharing models used, and ensuring compliance with laws and regulations.

The depth and level of detail required at each stage depends on the maturity of each party involved as well as the nature of the data sharing engagement and type of data being shared. However, detailed consideration of laws and regulations and alignment with data sharing principles is important at all stages of the Journey.

The information in this paper does not constitute, and should not be construed as, legal advice. It is intended for reference purposes only.



About this Handbook

What is the Data Sharing Journey?

This Journey represents the typical steps that prospective participants take as they engage in data sharing. These steps are not exhaustive, nor are they necessarily performed in the order presented below. In fact, the Journey is often iterative, with steps being revisited and refined throughout the process. These steps can be a useful framework for discussions on data sharing and form the chapter structure of this Handbook.

PLAN – Define the opportunity and assess feasibility

I

Establish principles

- Establish data sharing vision, principles, strategy and business case
- Analyse potential use cases of data sharing
- Identify potential data sources or data consumers (due diligence)
- Identify purpose of transaction and authority required

II

Identify & classify types of data

- Understand and categorise data types and domains– this helps to identify potential regulations or consent requirements that may impact sharing
- Define clear metadata and quality standards upfront to reduce time and effort in data sharing

III

Assess considerations for sensitive data

- Determine sensitivity (e.g. regulatory requirements) for the data types
- Identify if explicit consent required from customer, obtain if necessary
- Determine if data can be desensitised, e.g. anonymised, encrypted

DESIGN – Build out the model, put contracts in place

IV

Define data sharing model

- Identify potential data sharing models (e.g. direct, marketplace)
- Evaluate potential commercial models (e.g. exchange, fee)
- Explore appropriate data sharing mechanisms and architecture (e.g. raw transfer, anonymisation, federated learning)

V

Identify legal & regulatory considerations

- Evaluate relevant regulation and legislation, internal policies, such as the Personal Data Protection Act (PDPA), the Banking Act and other MAS Guidelines and Notices with respect to data, outsourcing, technology risk and cybersecurity

- Establish a data sharing agreement
- Map agreement where required to different stakeholders
- Reference data sharing principles as a guide
- Establish data lifecycle management requirements for on-boarding & off-boarding data, e.g. archival, retention, disposal requirements

VI

IMPLEMENT – Put it all into action

Assess readiness and implement

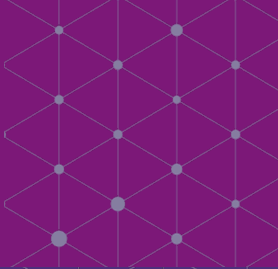
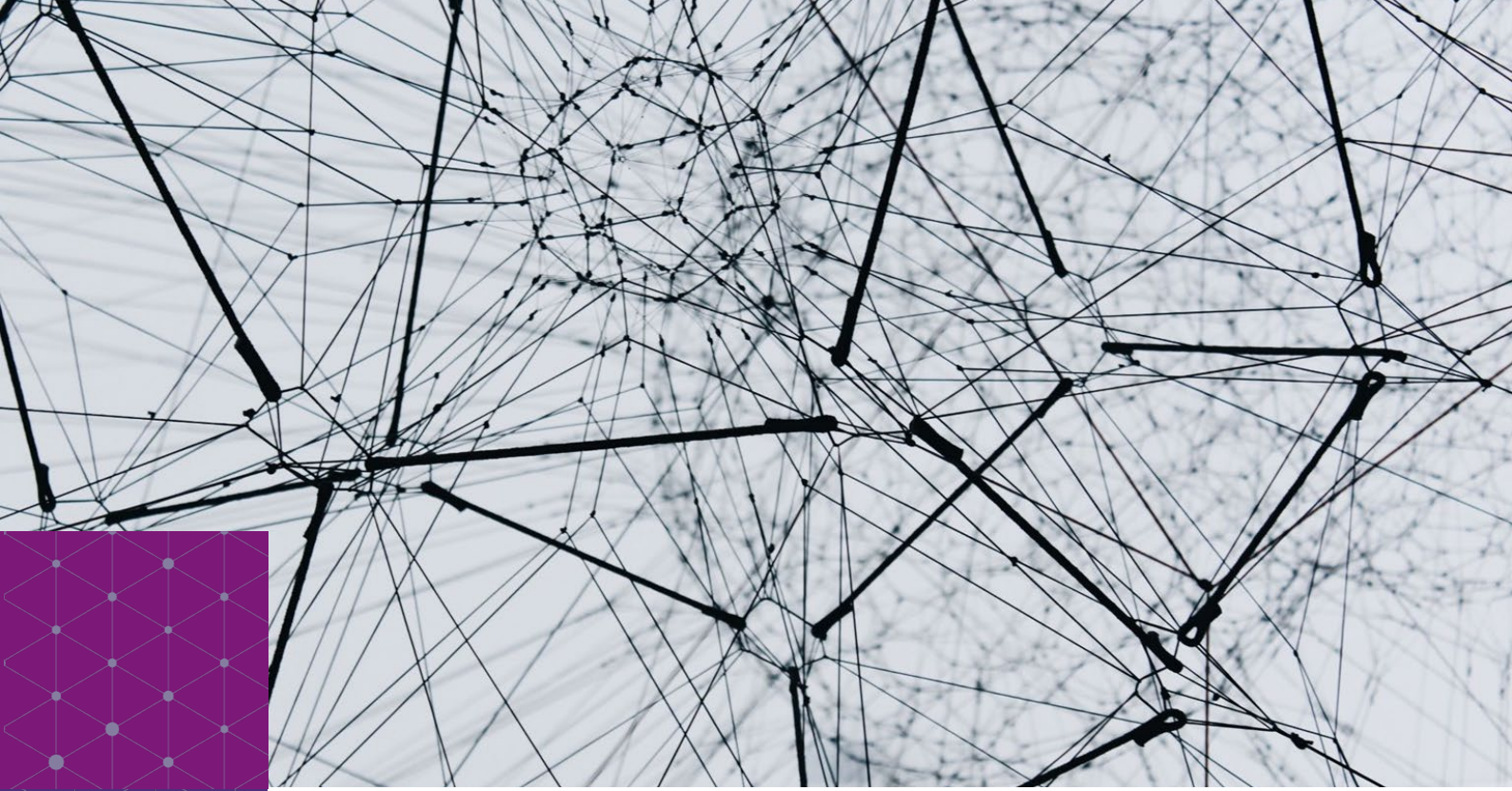
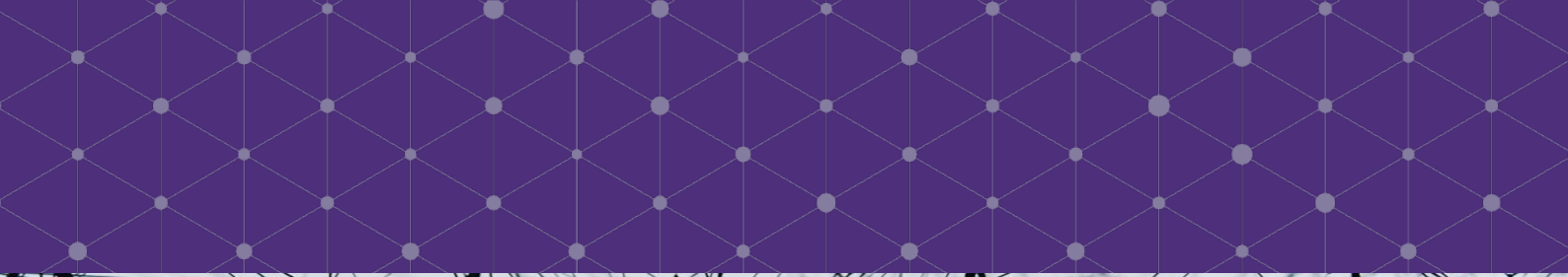
- Plan and align on activities, roles and responsibilities
- Establish data governance, protection, security and quality protocols / frameworks and controls
- Prepare, build, and test technical environments, data, algorithms
- Share and integrate data, models

Monitor data sharing

- Monitor risks and manage defects on an ongoing basis.
- Report and resolve incidents and breaches as required
- Execute roll-back or contingency plans where necessary
- Monitor and manage consent, where required, to ensure data usage is consistent with authorised purposes

Build for the future

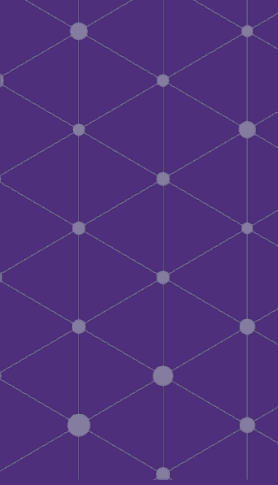
- Analyse ways to improve and enhance data sharing models.
- Explore additional data sharing opportunities
- Work with regulators, industry bodies and broader ecosystem to identify high value sharing opportunities
- Showcase success stories for new value creations



“ *Data is the key resource which continues to underpin and feed the growth of today’s digital economy. In order for organisations to better serve the needs of their customers, employees, shareholders and the communities in which they operate, there is tremendous value in increasing the flow of data and information between both private and public entities. This Handbook looks to support this endeavour in banking by providing practical considerations to ensure this can be achieved responsibly and purposefully.*



Sameer Gupta
Chief Analytics Officer | DBS Bank





WHAT IS DATA SHARING?

What is data sharing?

Data sharing is defined as the transfer, or granting, of governed access to data between entities for the achievement of mutual benefits. This purpose-driven transfer of data can deliver both tangible financial benefits to the parties sharing data, and can also indirectly drive societal benefits. Examples of broader data sharing that drives societal benefits and enables ecosystems are provided at the end of this chapter.¹

Two trends are having a major impact on how organisations need to approach data sharing

1. The volume of data generated by an organisation today is already vast and growing, and is increasingly critical to business operations.

The thirst for more data to create actionable insights, to develop new revenue streams or to gain a competitive advantage has led to a growing demand for data sharing. Even industries that traditionally may not have collaborated, and data sharing across the public-private sector boundaries for the benefits of citizens, are becoming more commonplace.

1. Data sharing is becoming more regulated, complex and challenging.

On one hand, the tools for processing data continue to evolve, improve and become more user-friendly. On the other, privacy regulation, data sovereignty requirements, cybersecurity risks and other constraints are changing how data sharing can be responsibly implemented.

Given the intrinsic value that data sharing can drive, organisations need to explore how they can effectively balance these two trends.

According to the OECD, data access and sharing can help generate social and economic benefits worth between 1% and 2.5% of GDP, equating to up to approximately S\$12.3B for Singapore.²

¹For the purposes of this Handbook, data sharing between banks and external parties for operational and/or regulatory requirements are not covered, e.g. the calculation of the Singapore Interbank Offered Rate (SIBOR) or Swap Offer Rate (SOR).

²"Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies", OECD report, 26 November 2019. For more information, visit <https://www.oecd.org/sti/enhancing-access-to-and-sharing-of-data-276aaca8-en.htm>

Benefits of data sharing

1

Broadens data pool available to the organisation

Existing data can be supplemented with partner or other 3rd party data



Enhances products and services

Enriching customer insights to develop new or enhanced products e.g. via personalisation or improved targeting



Improves decision-making

The accuracy of forecasting and predictive models can be optimised to improve operating efficiencies and margins

2

Accelerate the creation of value from data

Establishing data sharing processes and practices will unlock value from data



Enables inter-firm sharing

Processes and procedures can enable more effective data sharing between organisations



Unlock value within conglomerates

Holdings companies and conglomerates can improve data access, and the creation of value from data amongst their portfolio companies



Develop new revenue streams

New data products or data driven services for customers that build new revenue streams and uplift company valuation

3

De-risk business operations

Implement risk controls, checks and balances that were previously not available by leveraging shared data

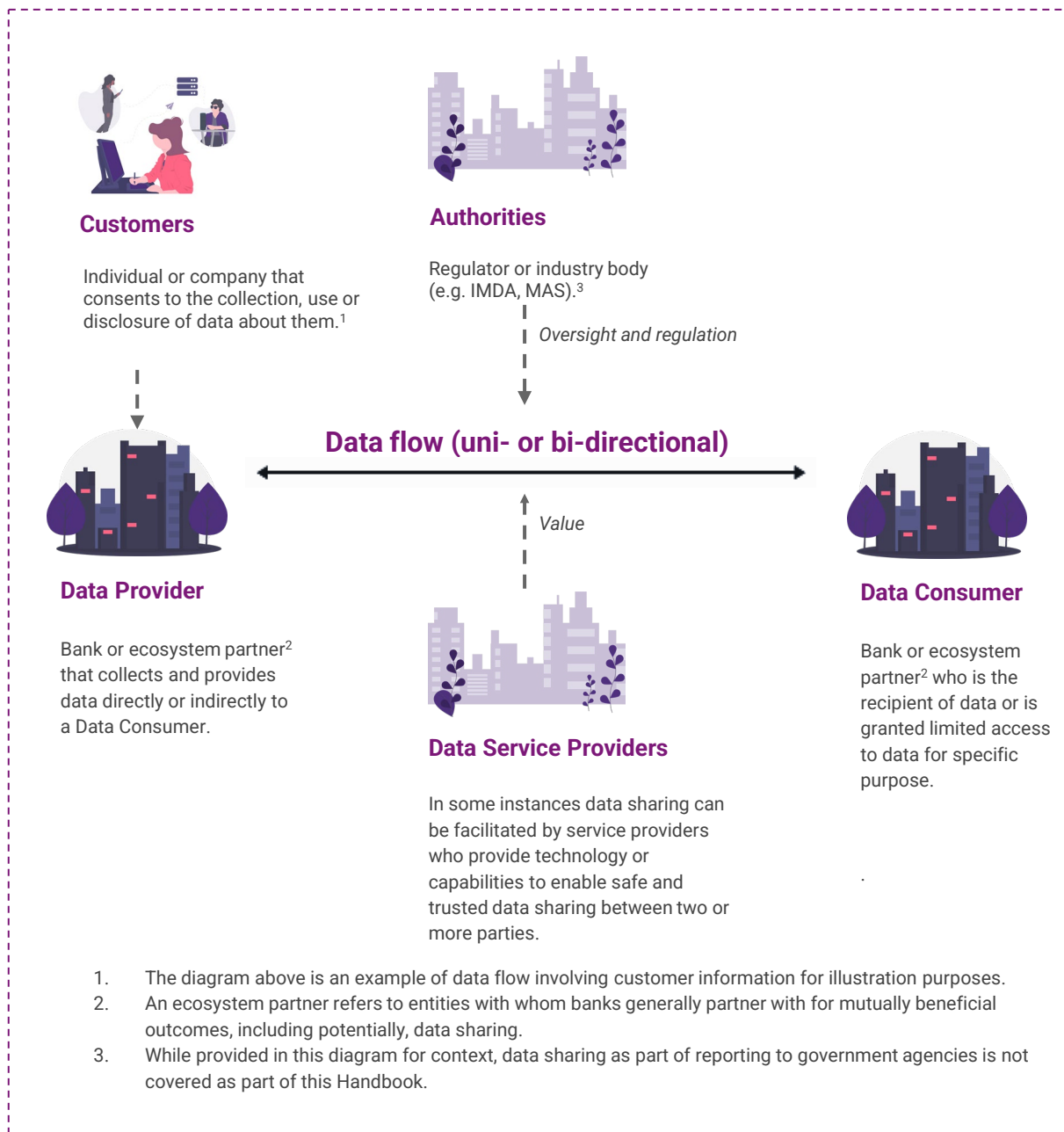


Reduce costs and manage risk

Generating efficiencies by leveraging new sources of data and associated insights to streamline operations and processes, including risk management, financial crime, fraud detection, and so on

Data sharing ecosystem

The data sharing ecosystem consists of multiple stakeholders with primary roles of data providers, consumers, service providers and authorities as illustrated below. Data sharing models and methods are covered in “Data Sharing Models” chapter of this Handbook. A key difference between the IMDA Trusted Data Sharing Framework and this Handbook is that “Authorities” now specifically includes the Monetary Authority of Singapore whose role is to oversee activities of financial institutions in Singapore.



There are increasing examples of data sharing ecosystems developing in Singapore. Two such ecosystems studied by the workstream are included on the following pages.

Data sharing driving social value

SGFinDex

Singapore Financial Data Exchange (SGFinDex) is a world's first public digital infrastructure developed by the Monetary Authority of Singapore (MAS) and the Smart Nation and Digital Government Group (SNDGG) in collaboration with The Association of Banks in Singapore (ABS) and seven participating banks, to use a national digital identity (SingPass) and a centrally managed online consent system. This enables individuals to access their financial information held across different government agencies and financial institutions.

Previously, individuals had to manually collate financial data across different data sources to understand their overall financial health and plan their finances, which is cumbersome and time consuming. With SGFinDex, individuals can use their SingPass to retrieve their personal financial information (such as deposits, credit cards, loans, and investments) from the participating banks and government agencies (HDB, CPF, and IRAS). This will help individuals better understand their overall financial health and plan their finances holistically. SGFinDex is designed with data protection, security, privacy, and consent-centricity in mind, and uses common data, API and encryption standards.

The next phase of SGFinDex will include the participation of insurers and Central Depository, allowing individuals to access information on their insurance policies held with insurers and their holdings of stocks at the Central Depository.

Key takeaways for this Handbook

- With appropriate considerations for data security and consent, large-scale, multi-party innovation and data sharing can be leveraged to enhance product offerings and benefit consumers.
- Confirm legal basis to share the data in accordance with local laws, regulation and, in the case of banks, licensing.
- Carefully consider the type of data being shared and, if the data is related to individuals, whether personal details are necessary or if anonymised or aggregated information would suffice.
- If sensitive data is being shared, take commensurate measures to protect and secure the data.

SGFinDex

<https://abs.org.sg/consumer-banking/sgfindex>



Data sharing enabling ecosystems

SGTraDex

SGTraDex, created as a result of the Singapore Together Alliance for Action (AfA) on Supply Chain Digitalisation is an initiative by the public and private sectors of Singapore to build a common data infrastructure for trusted data sharing in the supply chain ecosystem. Common pain points in the industry include significant inefficiencies in physical events, documentation and financial information flows across the value chain. For example, banks currently have limited visibility over the physical movement of goods in the supply chain, which reduces the ability of the trade finance industry to address demand from shippers. Similarly, logistics players face frequent congestion at container flow nodes, such as depots and warehouses, due to limited end-to-end visibility of container flows.

SGTraDex was launched to enable trusted sharing of trade data via a neutral and open digital infrastructure. Initial use cases developed enable participants to strengthen the financing integrity of trade flows, enhance operational efficiency by optimising logistics functions across partners, and provide visibility on supply chain transactions. The use cases have the potential to unlock more than S\$200 million of value annually when fully developed.

Key takeaways for this Handbook

- Data sharing can enable broad cross-industry outcomes and specifically address inefficiencies and pain points across the value chain.
- With appropriate considerations for data security and consent, large-scale, multi-party innovation and data sharing can be leveraged to enhance product offerings and benefit consumers.
- For broad industry sharing, data standards and legal agreements are critical enablers to allow for common infrastructure.

SGTraDex

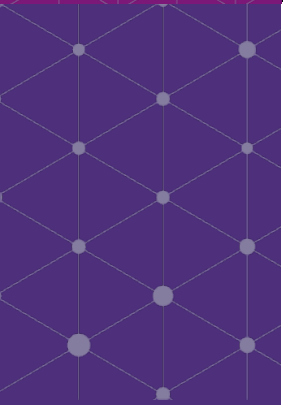
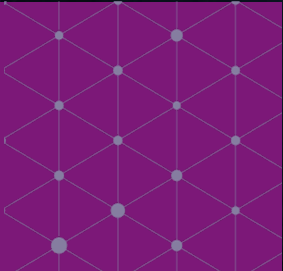
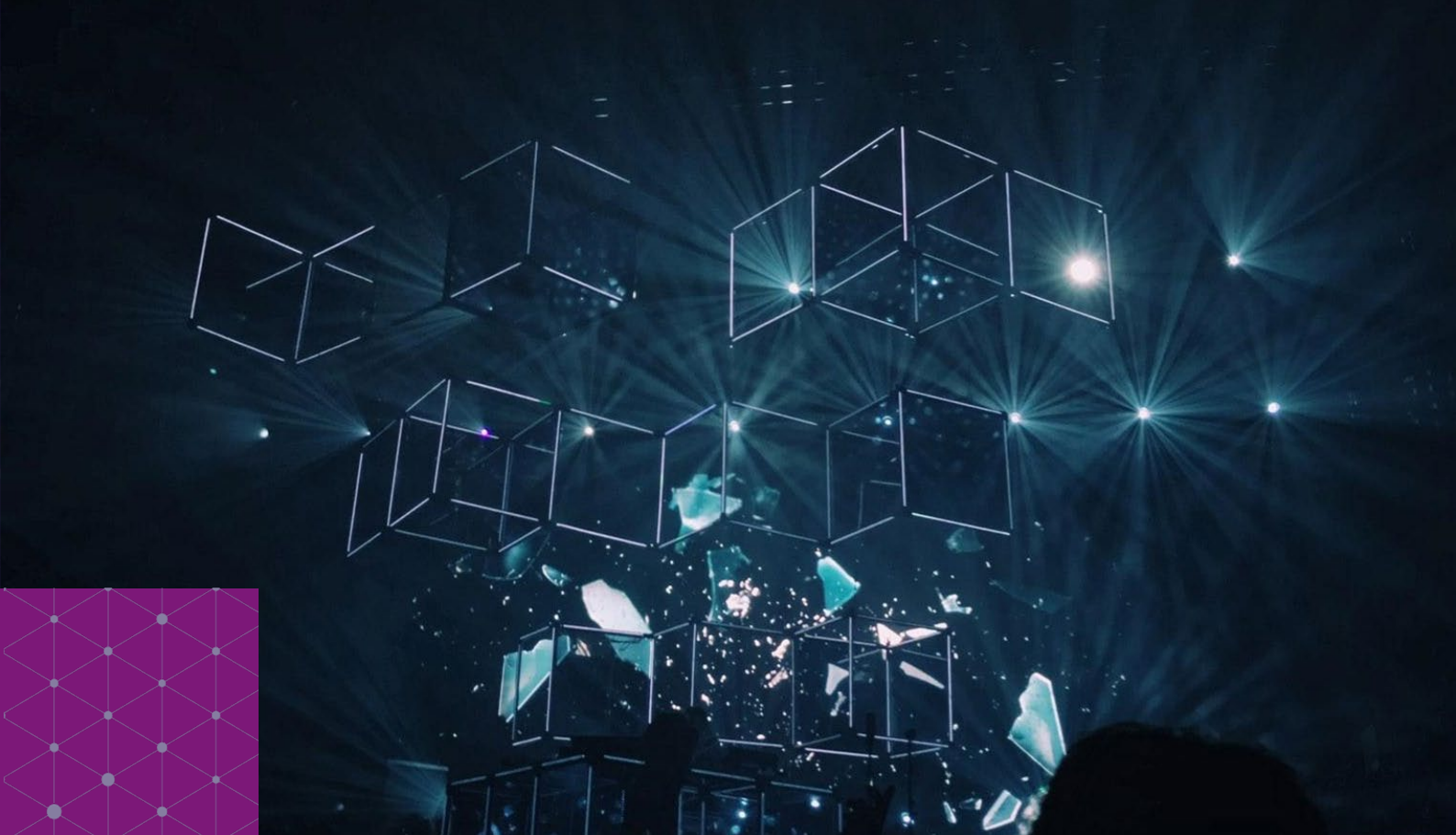
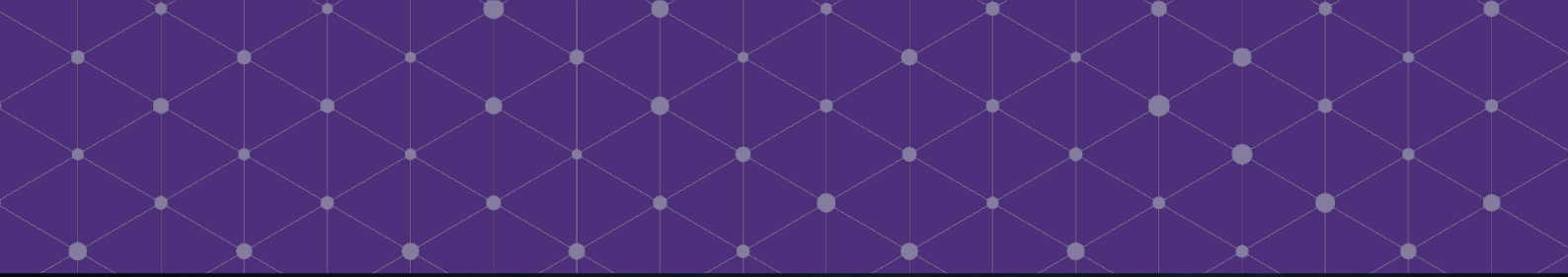
<https://sgtradex.com/>



Press release

https://www.sgpc.gov.sg/medi_a_releases/imda/press_release/P-20210713-1



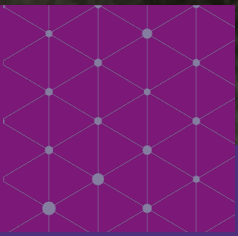


“

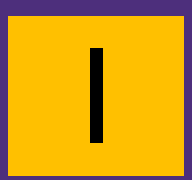
Data has become increasingly important in the way we make decisions and will continue to be at the forefront in driving new innovations. Many of us are already working with various partners to provide better service to our customers. This Handbook comes at an opportune time for us to have a common landing in data sharing, that covers the end-to-end journey including trust and security.

Dr John Lee

Chief Executive Officer | Maybank Singapore



PRINCIPLES OF DATA SHARING



Establish principles



Safe data sharing

Data sharing is increasingly topical with many global initiatives aimed at unlocking societal benefit. When banks are considering a data sharing partnership, the following factors should be reviewed:

- What is the data going to be used for?
- What is the expected benefit to the bank, ecosystem partner or customer?
- Is the use case ethically sound?
- How much data and to what level of detail is required for this purpose?
- Who is going to access the data?
- Do employees have sufficient training and authority to access and securely manage the data?
- How long will the the data retained?
- Will the data be securely protected?
- Can the use case be achieved without sharing of personal data? If not, can the data be anonymised and are there sufficient safeguards to reduce re-identification risk?
- Is there a legal basis for the data sharing exercise?
- Does the data partner have accountable and responsible data protection practices (for example, are they certified with the Data Protection Trustmark¹)?

To address these questions, the ABS has developed data sharing guiding principles as a framework for banks to ensure that data is shared appropriately and with robust safeguards. The ABS Data Sharing Principles are outlined below. These principles help guide organisations, both banks and non-banks in sharing their data. These and the guidance contained in this document are not intended to be prescriptive or exhaustive, but aim to enable a better understanding of key concepts to facilitate safe data sharing.

At a glance

For data sharing to be safe the following key principles should be observed:

- Data sharing must be purpose-driven
- Data shared must be proportionate
- Participants must have necessary skills & authority
- Data protection must be aligned with data sensitivity
- Data sharing must have a legal basis

¹Developed based on the PDPA and international benchmarks, IMDA's Data Protection Trustmark (DPTM) is a voluntary, enterprise-wide certification for organisations to demonstrate accountable & responsible data protection practices, thus strengthening regulator and consumer trust. For more information, visit: www.imda.gov.sg/dptm

Guiding principles



Data sharing is purpose-driven

Data is shared for an appropriate purpose that delivers clear value to the bank, data provider, partner or to society and can be demonstrated to be ethically sound.

Banks are only able to share data for specifically defined use cases where there is a clear benefit from providing access to the data.

Banks should assess use cases prior to any sharing and clearly justify the purpose of the sharing and the benefits to the bank/data provider, partner or to society.



Data shared is proportionate

Data shared between partners is proportionate to and sufficient for the specific agreed use cases.

If a use case can be achieved without sharing of personal identifiable data then this should always be the default position.

The type and amount of data shared should correspond to and not exceed the requirements of the use case, and should only be retained as long as necessary or in accordance with archival policies, applicable laws and regulations.



Data sharing participants have necessary skills & authority

Approved data sharing participants have the appropriate authority and required skill sets to access and use the data in the agreed and appropriate way.

Data sharing participants must understand and adhere to their responsibilities under key

regulations such as the PDPA and the Banking Act as they can be penalised for inappropriate collection, use and disclosure of data. Appropriate protocols should be maintained to ensure data is only accessed by approved users. Participants can consider certification schemes (e.g. IMDA's Data Protection Trustmark) to demonstrate accountable and responsible data practices.



Data protection is aligned with data sensitivity

Data is appropriately protected and managed to minimise risk of unauthorised use during transfer, analysis and subsequent downstream usage.

Sensitive data needs greater protection. For customer information, this can include techniques like anonymisation, aggregation, minimisation and unique record suppression to minimise potential risks of customer re-identification, risks to the privacy of customer information, and support compliance with PDPA and MAS requirements.



Data sharing has a legal basis

All sharing parties have the necessary legal basis in accordance with laws and regulations to share the data and have assessed and clearly documented the potential Data Protection impacts.

Legal contracts should be in place between sharing parties. The contracts must include the data protection required to be applied to any shared information.

A Data Protection Impact Assessment (DPIA) is recommended prior to entering legal contracts. The DPIA should clearly state the purpose of the use case along with how the sharing can legally comply with regulations such as PDPA and Privacy of Customer Information obligations under the Banking Act – either through customer consent or through existing terms & conditions.

Case Study 1:

Bringing the principles to life

Providing new mothers with personalised recommendations in an online marketplace

Data sharing is purpose-driven

A bank evaluated the data sharing purpose in the following manner:

What was the problem?

New mothers have their hands full looking after their new child. A single portal where they can access relevant information, products and services related to each stage of a new baby's development would make their lives much easier.

What was the solution?

The bank partnered two leading retailers to develop an online marketplace specifically designed to help new mothers with their newborn related needs. The platform provided mothers with personalised recommendations of relevant baby products based on the development stage of their child.

What was the outcome?

New mothers were able to enjoy a one stop platform providing them with personalised information and services aligned to their child's developing needs. The initiative enabled the three partners to be more relevant to this valuable segment at a critical moment in life – enabling business growth and deeper digital engagement for all partners.



Customer-driven use case

The platform was designed to help address pain points related to new mothers – providing them with help, information and services related to each development stage of their baby. This freed up time & removed anxiety for new mums by providing a one-stop shop which could support many of the daily needs of their developing baby.

Case Study 1:

Bringing the principles to life

Providing new mothers with personalised recommendations in an online marketplace

Data shared is proportionate

The scope of data sharing was to develop a set of initial product personalisation rules. As this could be performed with a small number of data fields, the data sharing was restricted to the minimal set of attributes necessary. Developing an understanding of which products are relevant for each development stage of a baby, does not require the model builder to know the identity of the individual and so data shared was pseudonymised to prevent individual identification. Pseudonymisation replaces direct identifiers (e.g. name, NRIC number) with a token. The data provider withholds the mapping between individuals and tokens when sharing, reducing the risk of subsequent re-identification.

What types of data were exchanged?

- **Pseudonymised Customer ID** e.g. token to enable matching of individuals across companies – without revealing the customer identity
- **Demographic data** e.g. Age
- **Transaction data** e.g. Types of products purchased in supermarket & department stores by the new mothers



Proportionate

Review the use case to understand what data is really required to support it. Only share the minimum amount of data required to achieve the use case outcome.



Pseudonymisation

Many use cases can be completed without the need for personal data. Pseudonymisation is one approach to enable joining of datasets without revealing who the individuals are.

Case Study 1:

Bringing the principles to life

Providing new mothers with personalised recommendations in an online marketplace

Data sharing has a legal basis

The three partners had existing legal agreements in place for broader financial partnership collaborations – which included data sharing elements. The Mothers Marketplace analysis scope was limited to customers that had relationships covered by these existing agreements.

Data protection is aligned with data sensitivity

Whilst personal data was not being shared, a checklist on security requirements for the transfer, management and deletion of data was defined upfront. Data was also managed in a secured, standalone environment limited to only a small number of authorised users.

Upon completion of the analysis to develop the Marketplace personalisation rules, the data was no longer required and was deleted. Evidence of deletion was produced by all parties to ensure auditability of data removal.



Know-your-partner

Working with existing partners where you have existing commercial and legal agreements in place might be a way to fast-track your Data Sharing initiatives. Existing contracts are a foundation which can then be built upon to add the additional agreements relating to data sharing.



Data retention

Data should only be retained if being actively used. Once the use case has completed, data should be removed immediately from partners environments. Proof of data deletion should be captured.

Case Study 1:

Bringing the principles to life

Providing new mothers with personalised recommendations in an online marketplace

Data sharing participants have the necessary skills and authority

Retail and Supermarket data are not traditional banking data assets. Workshops were conducted with the data domain experts within respective organisations to understand the key data definitions and determine what would be most relevant to the model build process.

A secured sandbox environment was provided for whitelisted developers to discover, understand and experiment with test data.



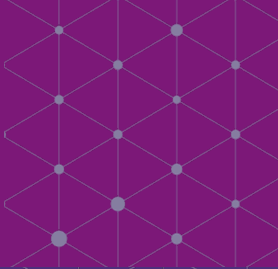
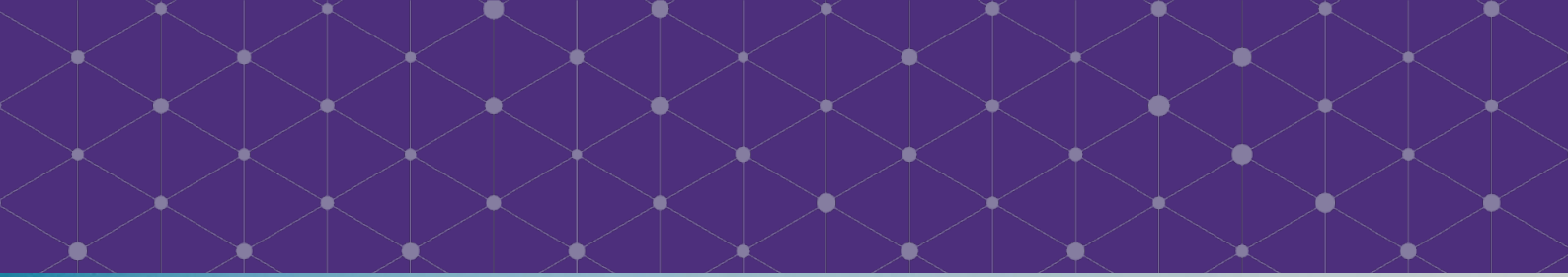
Importance of metadata

Data requirements need to be clearly defined from the start to ensure that the correct data is acquired first time – without the need for rework.

Metadata can help to:

- Accelerate data understanding and discovery process
- Minimise knowledge and technical gaps due to vague or ambiguous data parameter naming convention and its actual content.

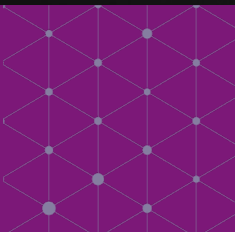
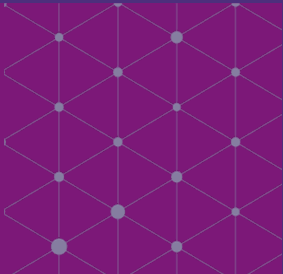
Refer to the [Types of Data](#) chapter for more information on metadata.



“ *Industry research predicts that organisations will use data sharing mechanisms as one of the tools to grow their businesses and the wider economy. For banks and ecosystem partners, often there are too many considerations when embarking on data sharing. The issuance of this Data Sharing Handbook is timely, as it gives best practice structured guidance to the banks in ensuring safe and trusted sharing of data in order to improve their offerings across the banking services.*



Richard Lowe
Group Chief Data Officer | UOB Limited



TYPES OF DATA



Identify & classify types of data



Why understanding data is important

Data is more than just data

Data sharing projects often fail due to a lack of understanding of the actual underlying data attributes. Understanding these attributes can enable organisations to identify potential data sharing use cases and determine if the data would meet their requirements. At a detailed level, this information is captured by metadata. Metadata refers to data about data. For example, the number of records, the types of data and the definitions of data fields.

Data quality is another example of metadata that is important for data sharing. Data quality standards and expectations are critical to agree on in the early stages of the data sharing journey, in addition to metadata requirements. Banks, like many other types of organisations, have vast amounts of data that might not widely be understood. Strong metadata and quality standards help enable data sharing as it allows participants to be clear and specific with regards to the key attributes of data during the sharing process.

At a glance

- Banks hold many categories of data that can provide valuable information in data sharing.
- Defining clear metadata and quality standards upfront reduces time and effort in data sharing.
- By focusing on categories and standards of data it helps define exactly what data is to be shared, and the underlying attributes including quality.

What categories of data do banks have?

Like most industries, there is no standardised taxonomy adopted or enforced for how banks categorise their data. The taxonomy below provides an overview of common categories of data assets that banks generally have. It is common that data from across multiple categories be combined to form more meaningful insight e.g. combining date of birth, demographics and transaction data points to derive spending habits of customers across different age groups.¹

Category	Sub-categories	Examples (non-exhaustive)
 <p>Party Individual or legal entity that is a customer or has relationship with the bank.</p>	<ul style="list-style-type: none"> • Legal entity • Individual • Related party 	<ul style="list-style-type: none"> • Personal data: Name, address, date of birth • Demographics: Income, sector, nationality, incorporation country, occupation • Contact information: Phone number, email, address
 <p>Account Registry of transactions; statement of business dealings or debits and credits; sum of money deposited at a bank.</p>	<ul style="list-style-type: none"> • Customer account • Internal company account • General ledger (GL) account 	<ul style="list-style-type: none"> • Accounts: Current/savings account, mortgage, corporate lending, global markets • Statements: Account balances, deposits, withdrawals
 <p>Product Product transaction and services provided to customer.</p>	<ul style="list-style-type: none"> • Corporate/principal finance • Retail/consumer • Transaction banking • Financial markets • Wealth management/private banking 	<ul style="list-style-type: none"> • Transactions: Amount, Currency, Instrument • Payments: Domestic and cross border: location, time, amount, currency • Settlements: Confirmations, margins, terms, value amount, value date
 <p>Limit & Collateral Bank's appetite to extend a loan or facility, and asset or property pledged as security for a loan.</p>	<ul style="list-style-type: none"> • Customer lending limit • Group lending limit • Account lending limit • Collateral • Guarantee 	<ul style="list-style-type: none"> • Scoring: Internal and external ratings • Limits: Credit card, unsecured and secured lending (mortgage) • Collateral: Cash, securities
 <p>Risk and compliance data Risk measurement and control activities.</p>	<ul style="list-style-type: none"> • Risk metrics • Finance metrics • Compliance • Technology 	<ul style="list-style-type: none"> • Potential future exposure • Risk weighted assets • Value at risk • Gross operational losses
 <p>Reference data Commonly used across different systems for the purpose of categorising data.</p>	<ul style="list-style-type: none"> • Market • Vessel • Country • Segment • Location 	<ul style="list-style-type: none"> • Currencies: Currency code • Industry standards: ISO country code, SWIFT code • Jurisdictions: Area code, city code

¹Banks may not share all categories of data shown above, and sharing is subject to sensitivity of the data as well. Refer to the "What is sensitive data?" chapter for more information.

Why is metadata important?

Metadata refers to data that provides information about data. It includes information about business meaning, associated technical, operational and business processes, data transformation rules as well as logical and physical data structures.

Examples of technical metadata in data sharing include number of records, tables, fields, field names, data types, file format, last updated dates, and tags. These are usually stored within a data dictionary. Examples of business metadata include field definitions and abbreviations, business logic, field sensitivity classification, unit of measure, and data quality information. These are typically stored within a business glossary.

Metadata makes it easier to **understand and interpret datasets** in data sharing. A shared understanding of data through consistent and descriptive metadata can improve data discoverability and help to improve the value data sharing is intended to bring.

1 Ensure clear meaning & consistency

By clearly defining metadata, an organisation has a structured and consistent view of what data it has, what the data represents, where it originates, how it moves through systems, who has access to it, or what it means for the data to be of high quality.

Consistency of metadata to access and use by multiple users is important to find relevant information easier. Metadata plays a crucial role in providing insights into the history of the data and how this has been transformed in this journey based on business rules.

Keywords can be used easily to search for the pool of shared data (including other formats like audio, images, and video).

Applying certified standards ensures that **quality and consistency** are maintained, as well as help to facilitate interoperability between systems and organisations. International Organisation for Standardisation (ISO) has developed metadata standards¹ that can be applied across organisations and data types.

Why is metadata important?

2 Reduce time and cost

The importance of having rich metadata will grow as data analytics becomes increasingly common in every organisation.

When it comes to discovery, in some cases only metadata itself needs to be shared across departments and organisations in an open data environment. This allows them to know what information is held, and where, but the data itself does not have to be open to everyone. This **reduces the time, security overhead and cost** during the data sharing process.

Effective metadata enables **data to be discovered** by users, systems, or AI/ML applications. Without appropriate metadata, a manual and time-consuming process may be required to physically interpret whatever data is available or provided, thereby increasing the chances of human errors and wasted time.



The role of data quality in data sharing

Data quality refers to the degree to which data is “fit”, or meets the requirements, of the data consumer’s intended purpose and use. Data sharing participants should establish and align on data quality expectations early in their data sharing journey, for example, by identifying international or industry standards to use. Doing so can help to reduce the likelihood of disputes and remediation further along in the data sharing journey.


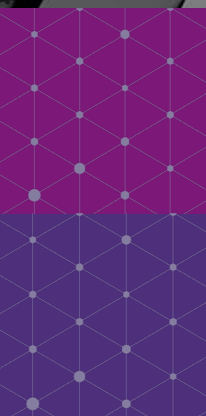
Depending on the use case, specific data quality requirements may be included in data sharing contracts, although generally these will constitute high-level requirements. In addition to these requirements, contracts may also include terms pertaining to the extent that the data can be relied upon for a specific use or purpose, as well as any associated liabilities and accountabilities.

Defining, measuring and improving data quality can involve significant effort, as a result of which participants may agree to prioritise certain data fields for detailed data quality analysis, while establishing a minimum baseline of data quality expectations for the remainder of the data being shared. Typically, data quality is assessed using dimensions, such as those listed below:

- **Completeness:** whether required data attributes and data records are available
- **Accuracy:** the degree to which data is correct and reflects actuality
- **Timeliness:** a measurement of how well the data is reflective of agreed timeframes and is functionally available when needed
- **Validity:** a measure of the conformity of data with required formats or technical requirements
- **Consistency:** the extent that the same data has the same value across data storage locations
- **Uniqueness:** a measure by which data records are distinct and with no unwanted duplication
- **Interoperability:** the degree to which the data is “joinable” with other data sources

Where personal data is being shared, participants should be aware of the accuracy obligation under the Personal Data Protection Act (PDPA) to make reasonable effort to ensure that the personal data is accurate and complete.

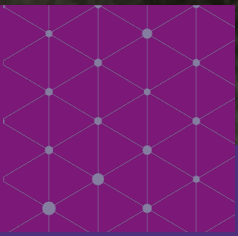
As participants progress in their data sharing journey, details on data quality can be captured and shared using metadata. This information can be used as a benchmark during the testing and eventual sharing of the data. In the event that data quality does not meet requirements, participants may discuss options including remediation of the data, complete or partial acceptance of defects, or ultimately a termination of agreements if quality requirements cannot be effectively met.



“ Data sharing will be the engine that powers the data driven experiences of the future.

Lim Kiang Tong

Group Chief Operating Officer | OCBC Bank



WHAT IS SENSITIVE DATA?



**Assess considerations
for sensitive data**



Why understanding data sensitivity is important

The focus of data sharing is often on customer information, which relates to any data that can be attributed to an individual or a corporate entity. Typically such data is considered sensitive. However, there are other forms of sensitive data which, like customer information, are subject to regulatory requirements. Sensitive data is broadly defined as data whose unauthorised disclosure will have a negative / material impact on the customer, bank, or external stakeholders.

From a regulatory perspective, when undertaking data sharing with sensitive data, there must be adequate risk management controls to reduce the impact and likelihood of unauthorised disclosure. Organisations may consider applying privacy preservation techniques to protect the data and mitigate such risks.

Where a defined data sharing use case requires a bank to receive sensitive data from a data sharing partner, the same safeguards will apply to that data as if the data was sourced internally, that is if it was the bank's own sensitive data. For an example of how a bank has handled the sharing of sensitive data with an ecosystem partner, refer to Case Study 2: Managing Sensitive Data.

This chapter covers in more detail what is sensitive data from a bank's perspective and the methods to reduce the impact and likelihood of unauthorised disclosure which aid in meeting regulatory requirements. For non-bank sensitivity classifications, refer to IMDA's Trusted Data Sharing Framework.


At a glance

- Sharing of sensitive data is subject to regulatory requirements for risk management controls to reduce the impact and likelihood of unauthorised disclosures.
- There are multiple methods to reduce sensitivity of data (including privacy preservation methods).
- Data sharing does not necessarily need the sharing of sensitive data. Data sharing projects are often less complex when sensitive data is not involved.

How to determine data sensitivity

Data sensitivity relates to the potential impact that unauthorised disclosure of the data could cause to either an end customer or data provider. It may also be driven by rules and restrictions imposed by relevant laws and regulations, such as data protection obligations for personal data under PDPA, permitted disclosure of customer information under the Banking Act, and other MAS guidelines and notices.

The table below provides an example of the classification of banking data and the materiality impact of unauthorised disclosure, which can guide the types of risk management controls required in a data sharing arrangement (terminology may differ between banks).

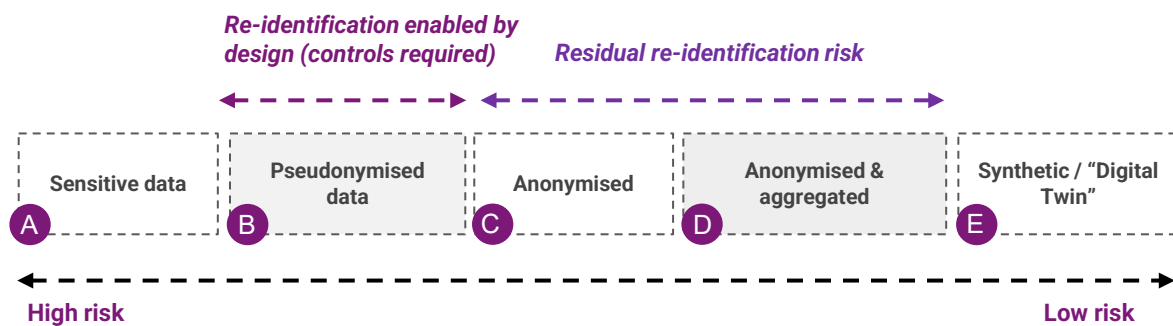
Data Classification	Examples (non-exhaustive)	Materiality of impact ¹	
Sensitive			
<p>Secret (Category 4) <i>Unauthorised disclosure will threaten the bank's existence, cause huge financial losses, seriously damage the bank's reputation, strategic position and compliance with regulatory requirements.</i></p>	<ul style="list-style-type: none"> • Bidding price • M&A details prior to public release • Passwords to critical systems and customer accounts • Encryption keys for secret and confidential data 		
<p>Confidential (Category 3) <i>Unauthorised disclosure will have adverse impact on the bank's smooth operation, business opportunities, profit, reputation, relationships with customers and partners, and compliance with regulatory requirements</i></p>	<ul style="list-style-type: none"> • Customer information, including banking details such as account number and credit card number • Personal data of individuals, including employees • Investigations or disciplinary proceedings • Patents or intellectual property rights 		
Non-Sensitive			
<p>Internal (Category 2) <i>Unauthorised disclosure will have moderate impact on the bank's operations, profitability and reputation.</i></p>	<ul style="list-style-type: none"> • Internal policies, procedures and guidelines • Internal training materials 		
<p>Public (Category 1) <i>Available publicly or can be released to the public without any detrimental effect on the bank or its customers and business partners.</i></p>	<ul style="list-style-type: none"> • Online public information • Job postings • Press releases • Business cards 		

¹Materiality in this context refers to the potential impact of unauthorised disclosure or leakage of the data on organisations, customers or individuals. For more information on materiality and its implications on controls required, refer to the MAS Outsourcing Guidelines. These Guidelines can be referenced for sound practices on risk management in handling of data by third parties.

Reducing risk of unauthorised disclosure through privacy preservation techniques

Sharing of sensitive data is often more complex and time-consuming than sharing of non-sensitive data. If the use cases are specifically defined and sensitive data is deemed appropriate to be disclosed, there must be adequate risk management controls to reduce the risk of unauthorised disclosure. Organisations may consider applying privacy preservation techniques such as anonymisation to protect the data and lower the risk of re-identification. These techniques can help to reduce the complexity of data sharing and often accelerate the process.

The diagram below illustrates that the risk of disclosing sensitive data can be reduced as direct identifiers are removed from the data through the use of privacy preservation techniques. Organisations should note that re-identification may still be possible as a result of combining separate datasets. Assessment of the risk of re-identification and the robustness of such techniques should be carried out to ensure that the identities of the customers cannot be readily inferred from data available internally or publicly. For details, refer to IMDA’s “Guide to Basic Data Anonymisation Techniques” and PDPC’s “Chapter 3 (Anonymisation) of the Advisory Guidelines on the PDPA for Selected Topics”.



A Data that may have a material impact on the bank, its customers or the broader ecosystem in the event of unauthorised disclosure. For banking data, it can refer to data that is classified as “Secret” or “Confidential”

B Data that relates to a customer or individual but has had identifiers removed and replaced by an anonymised identifier, commonly a ‘token’. This data can be re-identified with additional mapping information. Banks retain this mapping when sharing pseudonymised data to lower re-identification risk. If the mapping is shared, the data would still be classified as sensitive.

C Data that either does not relate to a customer or individual, or has had identifiers removed in such a way that the data is no longer directly (from unique personal characteristics) or indirectly (when information is linked together) referable to a customer or individual.

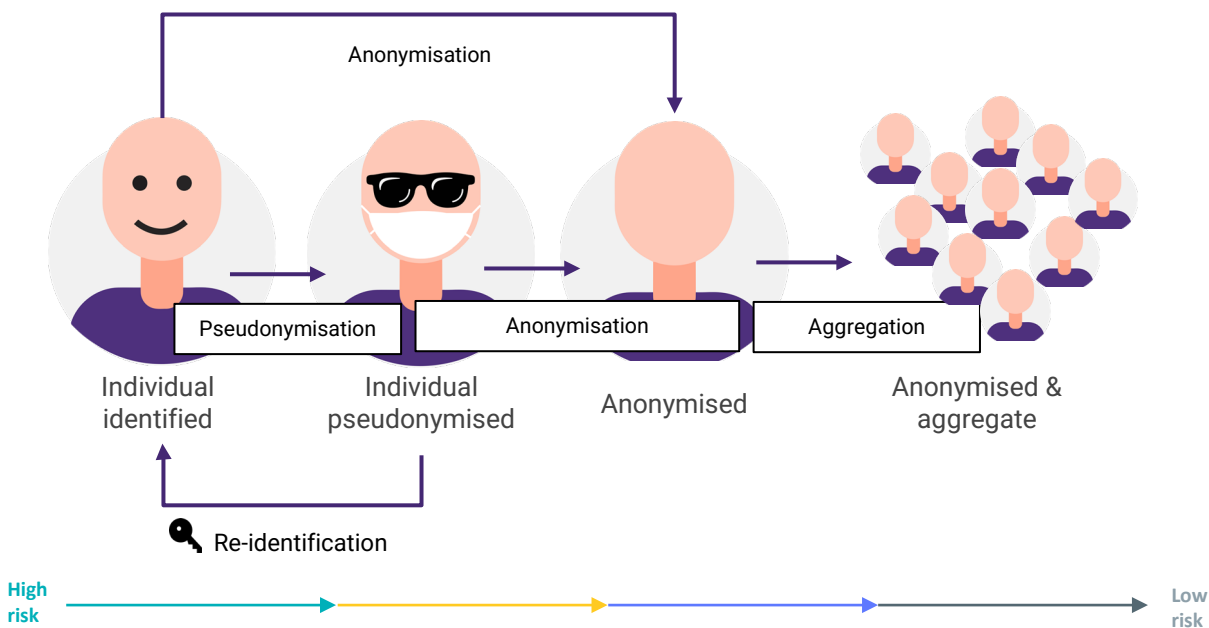
D Anonymised data that is aggregated such that the level of granularity is reduced. Aggregation is performed by expressing information in data summaries, so that the data could no longer referable to a customer or individual.

E Data created algorithmically to be highly correlated to the original data, but contains no real data or customer information.

Examples of privacy preserving mechanisms

There are multiple different methods to enable the sharing of data while minimise risks, enabling privacy, and importantly safeguarding it as it goes from one organisation to another. The overarching principle is to ensure data protection is aligned with sensitivity.

Anonymisation, Pseudonymisation and Re-identification



Pseudonymisation is a privacy preservation process that refers to the replacement of direct identifiers (e.g. name, email) with pseudonyms or tokens.

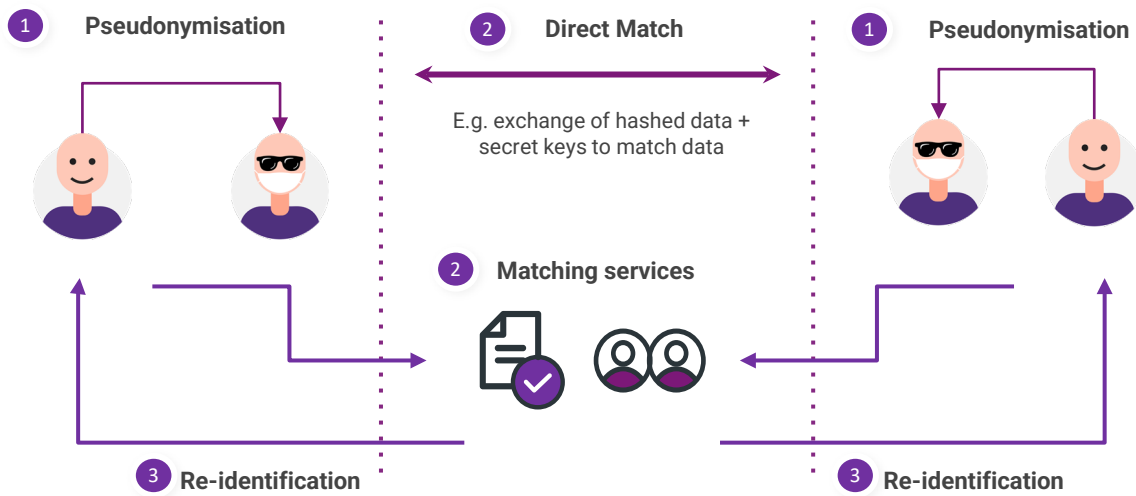
If the data has been **anonymised**, there should be no direct link back to the individual from the data alone.

Re-identification is the process in which data is re-linked to the individual to which the data belongs.

Examples of privacy preserving mechanisms

A common challenge faced by organisations sharing customer information is the lack of a trusted method to match individual customers across organisations in a privacy preserving way, that does not require personal information to be shared and facilitates adherence to privacy laws. Organisations may opt to use a service provider to accelerate the data matching process and also limit the volume of data shared directly with the consumer. Individual pseudonymised data can be used to match customers in a privacy preserving manner, where personal data is replaced with a pseudonym prior to the match being executed.

Privacy preserving via service provider



1 Pseudonymisation
Pseudonymisation at source. Segregation of personal data from non-personal data. Personal data replaced with a pseudonym.

2 Matching
Pseudonymised individuals matched directly or via a third party matching service that is regulatory compliant.

3 Re-identification
Re-identification only possible by the data provider.

Common privacy preserving techniques

Pseudonymisation

Pseudonymisation replaces direct identifiers (e.g. name, NRIC number) with a placeholder called a pseudonym or token. This is synonymous with tokenisation.

Tokenised or pseudonymised data cannot be attributed to a specific customer or individual without the use of additional information.

Hashing

Hashing is an algorithm performed on data to produce a value called a hash. Every hash value is unique.

All hash functions such as SHA256 are one-way functions, that is, a function which is practically impossible to reverse.

Hashing can be used as a way to pseudonymise data.

Anonymisation

Anonymisation refers to the process of modifying data such that the data does not refer to an individual or customer. This includes removal of all direct identifiers (e.g. name, email) as well as removal or alteration of indirect identifiers (e.g. job title, company name).

Anonymisation can be reversible or irreversible. Organisations should take measures to reduce the risk of re-identification when sharing data, such as through the use of safeguards over any data that could be used to re-identify the individual, or the deletion of such information.

Aggregation

Aggregation is performed by expressing information in data summaries, so that the data is no longer individual-level data.



Emerging privacy preserving techniques

Homomorphic Encryption

Homomorphic Encryption (HME) is an encryption technique which enables a query to be performed on encrypted data without revealing the underlying records. Only the outputs of the query are decrypted and shared.

Differential Privacy

Differential Privacy (DP) is a technique which adds randomness (i.e. noise) into your data. A certain amount of data utility is preserved as the probability distributions are similar to the original data set.

After adding noise, the data can be shared without providing the original data.

Synthetic Data

Synthetic data is generated programmatically by applying a sampling or simulation technique to create new data that does not contain any of the original data.

Synthetic data can be shared and used to train, test and develop models and solutions, so long as the data is versatile and robust enough to be useful.

Federated Learning

Federated Learning (FL) is the technology which allows for users to access and send queries to data that remains within decentralised devices or servers.

Security is preserved as the data stays in place, while still allowing the data to be utilised through a 'federated' technical method.



Case Study 2:

Managing sensitive data

One-stop portal for efficient travel arrangements

Data sharing is purpose-driven

What was the problem?

Many travellers spend a significant amount of their time on pre-vacation planning by reviewing multiple sites for flight, accommodation and travel insurance options. Personal and payment details have to be entered repeatedly for purchases across various sites.

What was the solution?

A bank worked with leading partners from the airline, hospitality and insurance spaces to provide travellers a seamless payment-enabled platform for air travel, accommodation and travel insurance needs.

What was the outcome?

Travelers were able to enjoy a fuss-free, one-stop platform that caters to their travelling needs. Such strategic partnerships allowed the bank to grow its customer base and drive digital engagement of customers across its banking products, whilst the partners realised increased revenue via scaled distribution.

Data shared is proportionate

Only data fields necessary to facilitate purchase of flight tickets, hotels or travel insurance were shared.

Types of data exchanged:

- **Individual** data e.g. customer name, passport number, date of birth
- **Account** data e.g. payment card number
- **Transaction** data e.g. origin country, destination country, departure time



Customer-driven use case

The platform was able to address pain points along the travel customer journey by offering a one-stop-shop experience for customers to browse and purchase their flight tickets, hotels and travel insurance seamlessly. Customers who logged in via their banking credentials and consented to sharing their data could also enjoy faster checkout by having their personal and payment details pre-filled automatically in their travel bookings.



Proportionate test

Partners should not request for more data than what is typically required if the customer had purchased via their own website or portal.

Case Study 2:

Managing sensitive data

One-stop portal for efficient travel arrangements

Data sharing has a legal basis

How did the bank and partners ensure that they have the right to share data?

The bank and partners had to consider relevant provisions in the Banking Act and PDPA on the disclosure of customer and/or personal data, such as obtaining customer consent for disclosure where necessary and how consent should be sought.

How did the bank and partners establish trust in data sharing?

The bank and respective partners signed a legally enforceable data sharing agreement that laid out the obligations and rights of both parties, including (but not limited to) clauses that cover data privacy rights, confidentiality, intellectual property, termination rights and exit provisions.

What were the challenges encountered during contract negotiation?

There was considerable time and effort spent to translate regulatory requirements into legal clauses. These negotiations include:

- Which legal agreement template to use
- Alignment on mandatory clauses
- Interpretation of bank-specific requirements
- Rationale of including bank-specific clauses

What could have helped accelerate the contracting process?

A high-level understanding by all parties on common regulatory considerations and sector-specific regulations could help accelerate the contracting process.



Know-your-partner

As banks operate in a regulated sector, potential partners can expect operational risk assessment to be conducted in areas such as materiality assessment, business continuity management and information security controls.



Legal agreement template

Organisations that are more familiar with data sharing may already have an existing legal template in place. The template can then be customised to include use case specific terms.



Sector-specific regulations

When sharing data with banks, potential partners can expect bank-specific regulations to translate into additional and/or more stringent clauses. Early understanding of the regulations that apply could accelerate the negotiation process and bring the risk forward for use cases that may not be viable.

Case Study 2:

Managing sensitive data

One-stop portal for hyper-personalised travel offerings

Data protection is aligned with data sensitivity

To ensure that personal data and other confidential information were appropriately protected and managed by the partners, the bank had to consider whether these partners had robust processes in place to minimise the risk of unauthorised disclosure. Some of the key considerations include application design and configurations, access controls, network security, controls to protect data at rest and in transit, and security monitoring.

Data sharing participants have the necessary skills and authority

Travel and tourism data are not traditional banking data assets. Rather than introducing new processes or manual interventions, the bank aligned the travel customer journey with partners' business-as-usual (BAU) process.

Workshops were conducted with the data domain experts within respective organisations to understand the data requirements and interoperability of the APIs.

A sandbox environment was also provided for whitelisted developers to discover, understand and experiment with test data.

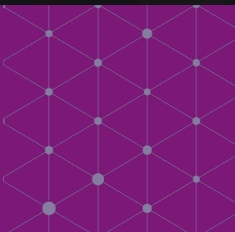
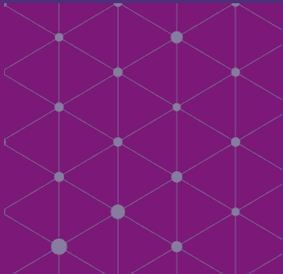


Importance of metadata

Data requirements need to be clearly defined from the start to ensure that systems on both sides are able to cater for the sharing and collection of all data fields that have been agreed upon.

Metadata can help to:

- Accelerate data understanding and discovery process
- Minimise knowledge and technical gaps due to vague or ambiguous data parameter naming convention and its actual content.



DATA SHARING MODELS



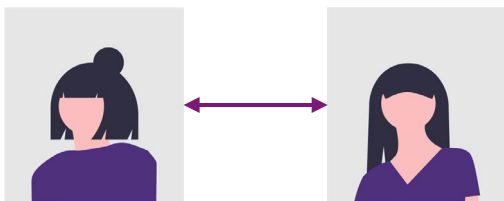
Define data sharing model



Common data sharing models

Data sharing can occur directly between two or more parties or via service providers such as data exchanges or other data sharing facilitation mechanisms. Listed below are some examples of data sharing models.

1 Direct sharing



Bank/s

Business partner/s*

Direct connection between Data Provider and Data Consumer. This can be either a bilateral or multilateral mechanism for data transfer, commonly made via APIs or flat files.

Example: Bancassurance

A local bank and insurance company partner to offer integrated products to their customers. To enhance their product and marketing strategy, they share aggregated insights on their customers.

2 Data exchange / Data-as-a-Service

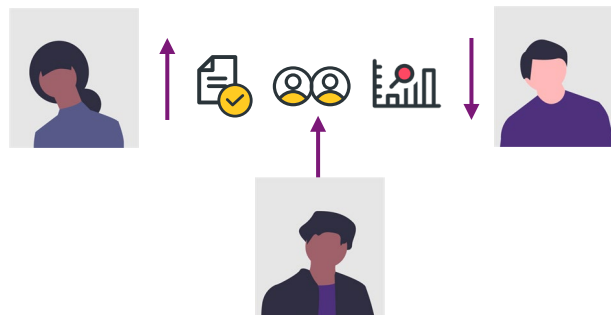


Typically a platform that enables sale of data assets to a broad and not necessarily specific set of data consumers.

Example: Macroeconomic data

A leading market research firm publishes monthly macroeconomic data. A bank subscribes to this data to assist their strategic decision-making and planning.

3 Facilitated sharing



Data Service Provider providing infrastructure and value-add services that bring together Data Providers, Data Consumers and potentially additional Data Service Providers.

Example: Fraud detection algorithm

On a secure data platform, two or more banks pool anonymised data on fraudulent transactions. A data scientist analyses the data to identify trends in customer behaviour/profiles that could predict or flag fraud before it happens, and shares the insights with the banks for mutual benefit.

At a glance

- There is no single strict data sharing model. What is important is that considerations around data sensitivity and risk management are considered in selecting the data sharing model.
- Existing platforms e.g. marketplaces, service providers may speed up and ease the process of data sharing as often they have existing legal and security protocols in place.

Common data transfer methods at banks

Web interface

Combines a number of different protocols for authentication, authorisation and data transfer. Commonly used as a means to upload or download data from a platform or marketplace via a browser.

SFTP

Secure File Transfer Protocol. Provides file access, file transfer and file management. Encrypts both authentication information and data files being transferred.

Cloud storage bucket

Data provider uploads data to a cloud storage bucket. Data Consumer accesses the bucket and downloads data directly to their cloud storage or local storage.

Email

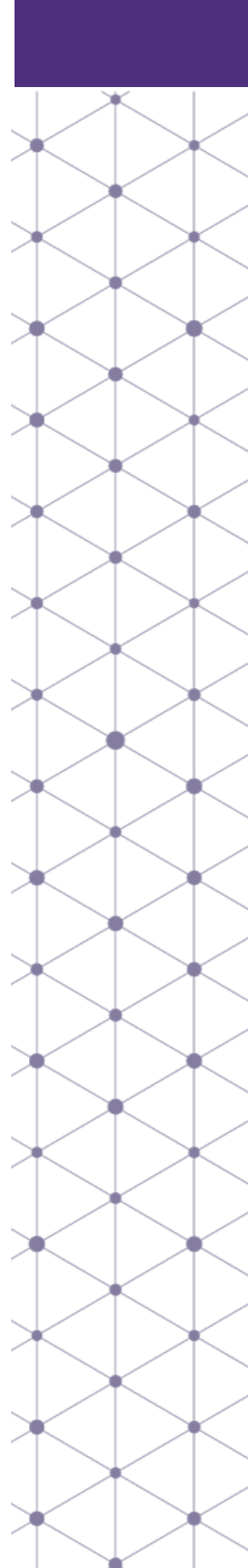
Electronic mail. Not recommended for sensitive data as unsecure plain text can be sent and is at risk of interception by a third party. To be avoided unless sharing open, public data sets.

API

Application programming interface. Often used as a method for transferring data directly between two applications.

Portable storage devices

Devices such as external hard drives, solid state disks, and universal serial bus ("USB") devices that can store and be used to upload to data to new systems, subject to endpoint security considerations and requirements.





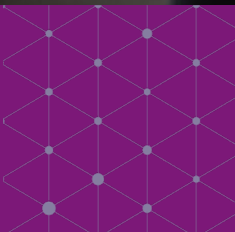
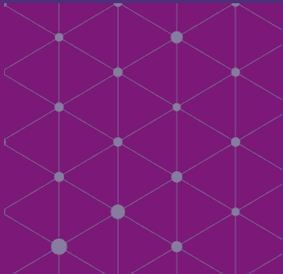
“ We are firm believers that data sharing can create a data network effect, which brings about value for our customers and partners. This handbook clearly lays out the data sharing principles and discusses key steps of the data sharing process: from ideation, to evaluation, and then finally contracting and implementation. Key considerations are introduced, including privacy preserving techniques like Differential Privacy, and certifications like IMDA’s Data Protection Trustmark. These emphasize the importance of trust and security, which are central to our principles as an exchange and market infrastructure provider. This is a good resource for financial institutions that are embarking on their data sharing journey.

Aaron Tan

Group Head of Data & Analytics |
Singapore Exchange

Ng Kin Yee

Head of Market Data and Connectivity |
Singapore Exchange



LEGAL & REGULATORY CONSIDERATIONS



Identify legal & regulatory considerations



Legal & regulatory considerations

Banks operate in a regulated environment within Singapore. Regulations cover permissible activities, reporting and control expectations. Financial regulations for banks are set by the Monetary Authority of Singapore (MAS). Besides these, banks are also subject to other laws and regulations of Singapore (e.g. Personal Data Protection Act) and the applicable laws of other jurisdictions that they operate in.

This chapter outlines the common regulatory and legal considerations that should be taken into account when entering into data sharing agreements. However and importantly the Handbook is only intended as a guide and **should not be construed as being comprehensive, prescriptive or constituting legal advice.**

Banks looking to enter into data sharing arrangements should continue to conduct their own due diligence on the third-party recipients and the specific circumstances of the arrangements to ensure that safeguards concerning the protection and specified purposes/uses of the data are in place. This is more acute in situations where the third party is processing data on behalf of the bank.

At a glance

- The **Banking Act** regulates the permissible activities of a bank and the obligations around the privacy of customer information. This provides context within which a bank can operate.
- MAS Notices and Guidelines provide information on obligations for Financial Institutions. With respect to data sharing and topics such as cybersecurity, technology risk and third-party handling of data relevant Notices and Guidelines include those on: **Technology Risk Management and Outsourcing.**¹
- **Data sharing agreements** are a vital part of defining what is required and boundaries for safe data sharing.
- The level and detail of the agreement will often be a function of the complexity and nature of data being shared.
- Contractual provisions do not diminish individual participants obligations under law or the Banking Act.

¹MAS has worked on industry programs for data sharing on topics such as SGFinDex. For more information on SGFinDex, visit: <https://www.mas.gov.sg/development/fintech/sgfindex>

Common regulatory & legal questions

Below are some common questions that banks and ecosystem partners should consider when embarking on their data sharing journey.

How can I share data?

With explicit or deemed consent (as permitted), exemption or exception;
If required by law;
Or if data is no longer referable to the customer or individual.

What kind of rules and restrictions apply when it comes to data sharing?

Applicable requirements from regulatory, technical and operational controls as well as legal obligations may apply.

These can include:

- Laws and regulations (e.g. Personal Data Protection Act (PDPA), Banking Act)
- Notices and guidelines (e.g. Monetary Authority of Singapore (MAS) Guidelines on Outsourcing and Technology Risk Management, Notice on Cyberhygiene)
- Contractual obligations
- Intellectual property rights (e.g. copyright, database rights)

How can I ensure that data is no longer referable to a customer or individual?

Organisations may consider adopting a set of risk management controls such as anonymisation to ensure that data being shared cannot be used to identify any particular customer or individual, and to reduce the risk of re-identification.

Is there published guidance on data sharing that I can refer to?

It is included within this document and in the sources referenced in the following page.

For more information

Please refer to the following:

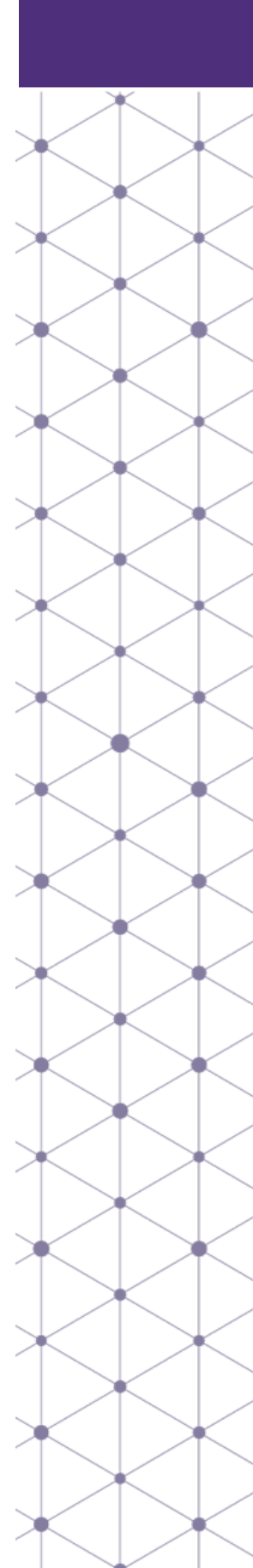
Personal Data Protection Act (PDPA)
<https://sso.agc.gov.sg/Act/PDPA2012>

Banking Act
<https://sso.agc.gov.sg/Act/BA1970>

MAS Guidelines on Outsourcing
<https://www.mas.gov.sg/regulation/guidelines/guidelines-on-outsourcing>

MAS Guidelines on Technology Risk Management
<https://www.mas.gov.sg/regulation/guidelines/technology-risk-management-guidelines>

MAS Notice 655 Cyber Hygiene
<https://www.mas.gov.sg/regulation/notices/notice-655>



Contracting guidance

A data sharing agreement sets out the important terms that should be agreed between parties to the data sharing partnership. There is no prescribed format for a data sharing agreement as it differs based on the nature of the partnership. The sections below illustrate:

- Common contractual components that are essential for trusted and safe data sharing.
- Types of agreements based on complexity and scale of the partnership
- Legal framework that can be adopted for multilateral data sharing via a data service provider

For sample Data Sharing Agreements from IMDA, please refer to:
www.go.gov.sg/data-innovation

Practical considerations required for contractual data sharing agreements

1 Categories of Agreements

As the data sharing relationship evolves, parties should ensure a commensurate level of protection when choosing the type of agreement required.

This section details different types of agreements and scenarios where they are used.

2 Components of Data Sharing Agreements

Details the components of creating an agreement in order to establish clear boundaries in a data sharing partnership, and protect each parties' respective positions.

3 Legal framework for multilateral data sharing

In the case of multilateral data sharing via a data service provider, there may be different legal agreements with participants of the shared platform to ensure trusted and transparent data sharing.

Contracting guidance

1 Categories of Agreements

The level of detail and the extent of parties' rights in the above categories should take into account factors such as the nature of the data sharing relationship (e.g. stage of the relationship, sensitivity of data shared), the use of the data including the extent of reliance on the data shared and associated liabilities, the parties involved (e.g. whether subject to industry regulation) and risk posture of the party (e.g. relating to cybersecurity risk). As the data sharing relationship evolves, parties should ensure a commensurate level of protection when choosing the type of agreement required. Broadly, parties may rely on the following categories of agreements:

Non-disclosure Agreement (NDA)

Definition: The non-disclosure agreement is typically used at the initial (i.e. Plan) phase of the Data Sharing Journey. The focus is to facilitate open discussions on potential data sharing arrangements with emphasis on the confidentiality protections, but does not contemplate further use of the information.

Under an NDA, typically data is not shared unless it is thoroughly anonymised & aggregated. However, metadata can be shared to facilitate mutual understanding and it is often valuable to share early in the data sharing process. The emphasis is on the confidentiality protections regarding the nature of data sharing, types of data and outcomes. Issues such as data sharing model, data access, warranties and licensing would not generally feature at this stage.

Pilot Agreement

Definition: At the phase where parties look to evaluate the real-world applicability of the use case (i.e. Implement), the type of data shared could be expanded (e.g. live or production data) but used in a limited and controlled context. The pilot agreement would also feature issues around data access, licence, data sharing model, warranties and IT security.

Proof of Concept (PoC) Agreement

Definition: As parties enter into the Build phase and the use case develops, a PoC agreement enables parties to assess the viability of the use case and identify the types of data and technical/operational considerations around data sharing.

At this stage, the type and extent of data shared may remain limited (e.g. non-production/artificial data sets), with issues of data access and data sharing model featuring more prominently in the agreement.

Data Sharing Agreement

Definition: This agreement sets out the full rights and obligations of the parties in an established data sharing arrangement (i.e. Implement). In addition to the issues addressed in the preceding agreements, due to the ongoing nature of the relationship, the Data Sharing Agreement will also address issues around ongoing monitoring and termination of the relationship. There should also be clear positions on the extent to which data shared can be relied upon for its intended use or purpose, and the associated liability in the event of losses arising from such reliance (e.g. where the data provided is inaccurate or incomplete). Components of Data Sharing Agreements are elaborated in the next page.

Contracting guidance

2 Components of Data Sharing Agreements

To establish clear boundaries in a data sharing relationship, parties should protect their respective positions through a legally enforceable agreement. In addition to standard contract clauses, data sharing agreements should cover these topics:

Component	Description
Purpose	Description of use case
Roles	Data recipient, data provider, data service provider (if applicable)
Data to be Shared	Categories of data, data elements
Use of Data	Permissible uses of the data and/or specific restrictions
Data Sharing Model	Technical requirements, specific responsibilities of the parties and dependencies
Data Access	Scope, frequency, access rights
Data Loss & Liability	Safeguards and liability around data loss prevention such as data protection, IT security, storage, segregation, transmission
Intellectual Property Ownership	Retention of background IP, ownership of newly developed/derivative IP
Grant of License	Rights to use the data for intended purpose, distribution or sub-licensing, authorised users
License Restriction	Territorial/time limitation, exclusivity or commercialisation rights
Governance, Controls and Ongoing Monitoring	Audit, incident notification and escalation, issue and error management, internal controls and quality assurance, relationship management
Warranties	Data quality, integrity, fitness for purpose, reliance, free of data defects
Compliance with laws	Complying with and enabling the other party to comply with applicable laws, not causing other party to breach laws, liability for breaches (including indemnities)
Confidentiality	Protections in relation to the disclosure of confidential information
Termination	Notice period, data return and disposal
Data Subject Consent	Disclosure to and consent from data subjects, consent database

Contracting methods

There are two distinct ways to contract, which will depend on data sharing model utilised.

Point to Point contracts

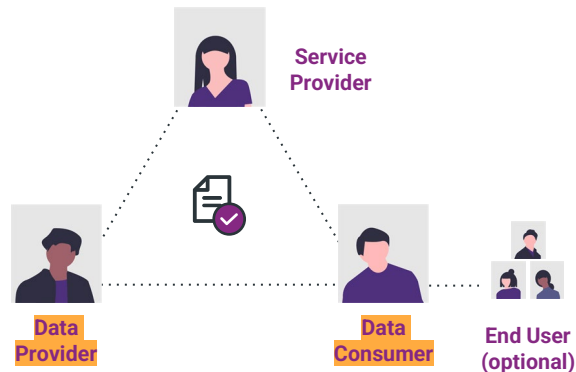


Bilateral data sharing agreement

Observed model: Direct sharing

- Agreements between Data Provider and Data Consumer
- Sets out the rights, responsibilities and obligations of both parties

Service Provider contracts



Multilateral data sharing

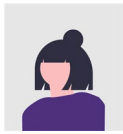
Observed model: Facilitated sharing

- Agreements between Data Provider(s), Data Consumer(s) and Service Provider(s)
- Designed for scale - large number of parties using consistent terms
- Sets out the rights, responsibilities and obligations of all parties
- Additional agreements between Data Consumer(s) and optional End Users

Requirements for participants of data sharing parties

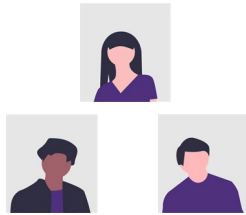
The participants involved in multi-party data sharing may have different requirements.

1 Participant Requirements



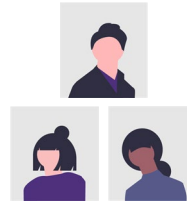
DATA PROVIDERS

Require governed and secure access to the data sharing service provider, a legal framework for multi-party data sharing and an approved permitted use of data with any **Data Consumers**.



DATA CONSUMERS

Require governed and secure access to the data sharing service provider, a legal framework for multi-party data sharing, an approved permitted use of data with any **Data Providers** and an agreement in place with any **End Users**.



END USERS (optional)

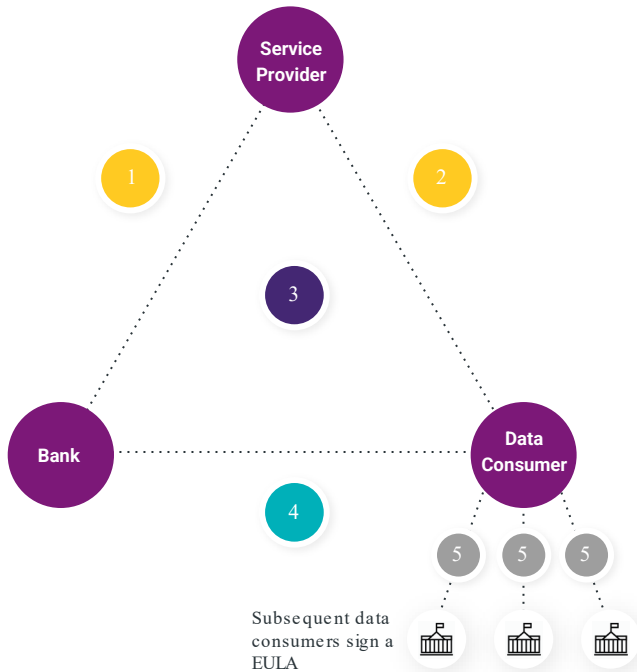
Required to agree to the terms set by **Data Consumers** in order to protect the interests of the **Data Providers**.

2 Use Case of Agreements

Agreements		Data Provider	Data Consumer	End User
Software	Software Agreement	✓	✓	
Legal	Framework for data sharing (e.g. Service Provider Common Legal Framework)	✓	✓	
	Use case terms that define permitted use of data	(multiple)	(multiple)	✓
	End User Agreement			(multiple)

Data sharing via service provider example

This example showcases contracting in data sharing via a service provider.



1 Platform Agreement
Document to govern Platform usage between a bank and a service provider.

This is executed once offline by the bank.

2 Guest Agreement
Document to govern Platform usage between a data consumer and a service provider.

This is executed once offline by a data provider.

3

Service Provider Common Legal Framework

Framework for multi-party data sharing. Defines the role and responsibilities of participants, IP ownership, confidentiality, privacy, data destruction, etc. This is executed once offline by a bank and data consumer..

4

Data Licence

Specifies the use case specific terms, including the permitted use of data. References the Common Legal Framework and data product EULAs. This is an online form executed by a bank and data consumer.

5

End User Licence Agreement

Protects the interests of data providers, to the extent that their IP still exists within. This is executed **once** offline by each subsequent data consumer that subscribes.

Case Study 3: Contracting

To build real economic resilience, you need real information.

Assist with economic decision-making.

What was the problem?

A consultancy specialising in demographics, economics and spatial planning, wanted additional data sources to enable economic decision-making.

It required access to transactional data from a bank, in which to apply their own IP to clean, transform and aggregate the data into meaningful insights.

What was the solution?

Under an agreed legal framework, the consultancy receives license-based access to de-identified transactional data on a recurring basis from a major bank. Access is provided via a secure on-demand analytical environment.

The bank ensures the process is secure and complies with regulation and policy through clearly defining the permitted use of data and utilising software which provides well-governed repeatable workflows to access data and a full audit trail.

Aggregate data insights are built within the analytical environment and extracted as input into a data product.

What was the outcome?

The data product provides insight and knowledge, benefiting stakeholders through better decision-making.

Stakeholders received real-time, up-to-date information about the market giving them an edge in planning and negotiations.

Case Study 3: Contracting

To build real economic resilience, you need real information.

Linking contracting to data sharing principles

Data Sharing is Purpose Driven

- Use case purpose is to assist in economic decision-making by incorporating additional data sources. This benefited both data sharing participants and ultimately society via more efficient allocation of resources.

Data Shared is Proportionate

- No identifiable data is required; use case utilises anonymised transactional data that is accessed via a data sharing service provider where aggregate insights are built and extracted as input into data product
- In this way, the data shared with end users is minimised

Data sharing participants have necessary skills & authority

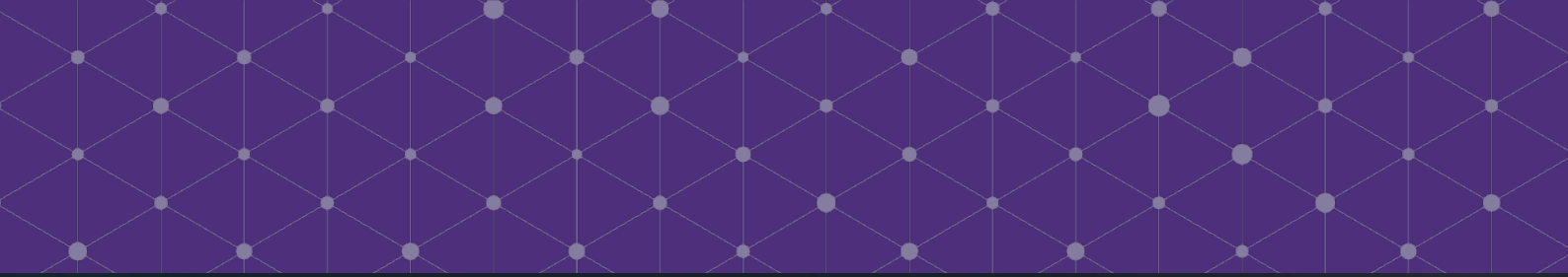
- The bank provides licensed-based data access to specific users from the data consumer who are granted appropriate authority to use the de-identified data in the agreed way over a specific timeframe

Data Protection is aligned with data sensitivity

- Data is de-identified by the bank and aggregated by the data consumer prior to input into data product to reduce the risk of customer re-identification
- The bank ensures the process is secure by utilising software which provides well-governed repeatable workflows to access data and a full audit trail

Data Sharing has a legal basis

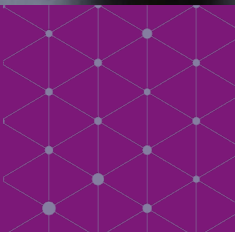
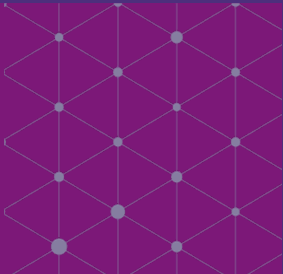
- An agreed legal framework is in place for both the bank and the data consumer that defines their roles and responsibilities, IP ownership, confidentiality and privacy obligations
- The bank and data consumer also have a project-specific legal agreement in place that specifies the use case specific terms including the permitted use of data
- Data sharing has a legal basis and is compliant with applicable laws and regulations



“ *In order for Banks to harness the potential of data; we all need a framework and a mechanism to do so in a safe way – the Data Sharing Handbook provides clarity and guidance on how banks can achieve that.*

Mohammed Rahim

Global Head Data Management Risk | Standard Chartered Bank (Singapore)



IMPLEMENT: PUTTING IT ALL TOGETHER



Implement



Putting it all together



A successful data sharing partnership ultimately provides benefits to both the data provider and consumer. For those benefits to hold true both parties need to be conscious of the ongoing obligations they have. The success of a data sharing initiative is closely tied to strong alignment in purpose and collective understanding of benefits, risks, and regulatory expectations between the parties involved. Importantly, participants should continually revisit their data sharing principles and monitor for risks posed throughout the data sharing arrangement.

As part of the Data Sharing Journey, three steps are identified in the Implement stage. These are further explained below with additional considerations for data sharing participants.

Putting it all together

1. Check operational readiness and go live

- Define the specific use cases and their associated proportionate data
- Plan and align on activities, roles and responsibilities
- Establish data governance, protection, security and quality protocols, frameworks and controls
- Establish data lifecycle management requirement for on-boarding & off-boarding data
- Perform relevant control due diligence
- Prepare, build, and test technical environments, data, algorithms
- Share and integrate data, models
- Establish communications strategy to manage platform, customer, user perceptions
- Fulfil contractual obligations

Subject to the sensitivity of data shared, careful consideration is required of not just technical requirements but also how to demonstrate other requirements, these include:

- Need for appropriate skills and authority by parties receiving the data
- Protocols for how any incidents are managed, data destroyed and retained
- Controls required to demonstrate any oversight required by the data consumer
- Mechanisms to ensure any data security breaches are reported and acted upon within required timelines
- Mechanisms for data retrieval, destruction where data sharing agreements expire or are terminated. The extent of such mechanisms will be heavily dependent on the sensitivity of the data.

Note that many of these considerations are covered under various MAS Notices and Guidelines which are referred to in the “Legal and Regulatory Considerations” chapter.

Putting it all together

2. Monitor data sharing

- Monitor risks and manage defects on an ongoing basis
- Report and resolve incidents and breaches as required
- Execute roll-back or contingency plans where necessary
- Monitor and manage consent, where required, to ensure data usage is consistent with authorised purposes

Data sharing should, when aligned with the principles, be something that drives value to end customers and the broader ecosystem. To ensure that the value is not eroded, ongoing monitoring is critical, particularly when sensitive data is being shared. The extent to which this is needed is subject to the data's sensitivity. However, even for less sensitive data it is important to monitor to manage reputation.

Key activities would typically include:

- Regular checkpoints on progress and outcomes
- Active risk and issue tracking and escalation protocols
- Potentially audit / ongoing review of controls to ensure they remain in place through the duration of the agreement
- Defined protocols and measures for data retention / destruction
- Training and evaluation of roles & responsibilities to ensure employees have sufficient knowledge, skill and authority in data sharing.

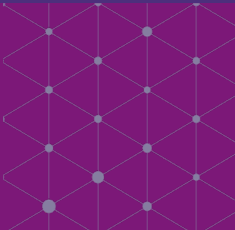
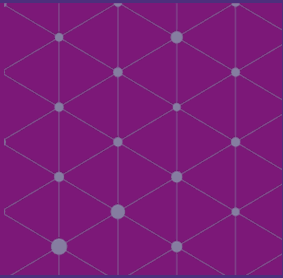
Putting it all together

3. Build for the future

- Analyse ways to improve and enhance data sharing models
- Explore additional data sharing opportunities
- Work with regulators, industry bodies and broader ecosystem to identify high value sharing opportunities
- Showcase success stories for new value creations

Data sharing is an emerging topic, but has significant potential to benefit consumers and society as a whole. Data sharing should be encouraged in line with data sharing principles as well as laws and regulations, and it is important that organisations continually look for additional opportunities to engage in data sharing. Ultimately data sharing is most valuable when done safely and securely with a clear purpose. Working with regulators, industry bodies and the broader ecosystem is an important part of enabling future data sharing and its corresponding benefits.





The Association of Banks in Singapore (ABS)
#12-08, MAS Building 10 Shenton Way,
Singapore 079117 Tel: (65) 6224 4300 / Fax:
(65) 6224 1785 Email: banks@abs.org.sg

Published : 30 August 2021