Data Management Procedure

Document ID	DMP V1.0.1
Effective Date	2023-04-01

Tabe of Contents

1 Purpose and Applicability

1.1 The Data Management Procedure (DMP) establishes the data governance and management processes for regulatory compliance and minimizing impacts on IT and business functions

1.2 These standards and practices apply to all units of the institution engaged in any form of data processing and data management. The DMP does not supersede any regulatory requirements.

1.3 The DMP will be reviewed biennially or when significant changes occur in regulations, business models, or technologies

2 Related Policies, Guidelines, Procedures, and Applicable Laws

2.1 The AMP should be reviewed in conjunction with policies and procedures issued by both the organization and the relevant regulatory bodies, which include, but are not limited to, the following:

a. Bank Negara Malaysia (BNM) Policy Document on Risk Management in Technology (RMiT)

b. BNM Guidelines on Data Management and MIS Framework

c. Guidelines from the Department of Personal Data Protection under the Ministry of Communications and Multimedia Malaysia (KKM)

d. Circulations, notices, and memos issued by BNM

3 Glossary of Key Terms

The following are the definition of key terms in this DMP:

Critical Data refers to a data element recognized for its significant influence on a business process or analysis, requiring heightened attention to data quality.

Critical System is an application or system whose failure or disruption would substantially impact the institution's business operations, reputation, or financial standing.

Data Acquisition is the process of collecting, obtaining, or accessing data from various sources for analysis, storage, or use in specific applications.

Data Consumers are individuals, systems, or applications that use, analyse, or process data to extract insights, make decisions, or perform specific tasks.

Data Owners are individuals responsible for the overall management, integrity, and security of specific data set within the institution. They have the authority and accountability for how the data is collected, stored, accessed, and used

Data Providers are individuals, organizations, or systems responsible for supplying or making data available to others for use, analysis, or further processing

A Data Source of Truth (SoT) refers to a trusted, authoritative data repository or system that serves as the definitive source for a specific type of information within an organization. It is the single, reliable point of reference that ensures data consistency and accuracy across various systems and processes

Metadata is data that provides information about other data, offering context and details that help describe, identify, or organize the primary data. It acts as a "data about data," giving insights into its structure, meaning, and management, enabling easier discovery, usage, and organization of data.

4 Role and Responsibilities

Role	Responsibilities
Data Owner	- Accountable for data accuracy, integrity, and quality. - Defines access policies and usage rights. - Ensures data complies with legal and regulatory requirements.
Data Consumer	- Uses data for analysis, reporting, or operational purposes. - Reports data quality issues. - Follows data access policies.
Data Architect	- Designs and implements data models, structures, and systems. - Ensures data scalability and integration with other systems.
Data Analyst	- Analyses and interprets data. - Ensures data is properly processed and translated into insights. - Reports inconsistencies in data.
Chief Data Officer (CDO)	- Develops and drives overall data strategy - Ensures alignment of data management with business goals. - Oversees compliance with regulations.
Compliance Officer	- Ensures data handling complies with regulatory requirements. - Conducts audits and risk assessments. - Monitors for policy violations.
Cybersecurity	- Implements and monitors data security protocols. - Manages access control systems. - Ensures encryption and threat detection mechanisms.

5 Data Management Overview

5.1 The DMP provides guidelines on data management throughout the entire data lifecycle. The key components in data management are:

a. **Data Cataloguing**

   a. The process of creating a structured inventory of data assets within an organization. This inventory, known as a data catalogue, serves as a comprehensive repository that provides metadata, descriptions, and context about various data sets, making it easier for users to discover, understand, and utilize the data effectively.

   b. Key components of data cataloguing include:

      i. Metadata Management: Capturing essential information about data assets, such as data sources, formats, schemas, lineage, and ownership.

      ii. Data Discovery: Facilitating easy search and retrieval of data assets by users, often through a user-friendly interface or search tool.

      iii. Data Governance: Supporting data stewardship by documenting data usage policies, access controls, and compliance requirements.

      iv. Collaboration and Sharing: Allowing users to contribute annotations, ratings, or comments, fostering a community of data users and enhancing the catalogue's value.

      v. Data cataloguing helps organizations improve data accessibility, promote data-driven decision-making, and enhance data governance by providing a clear overview of their data landscape.

b. **Data Architecture**

   a. The structural design and framework that defines how data is collected, stored, organized, integrated, and utilized within an organization. It encompasses the policies, procedures, and technologies used to manage and govern data, ensuring that data assets align with business objectives and support effective decision-making.

   b. Key components of data architecture include:

      i. Data Models: Frameworks that define how data is structured, including entity relationships, data types, and schemas. This can involve conceptual, logical, and physical data models.

      ii. Data Storage: Strategies for storing data, such as databases (relational, NoSQL), data lakes, or data warehouses, and considerations for performance, scalability, and cost.

      iii. Data Integration: Methods for combining data from different sources, ensuring consistency and quality. This includes data ingestion, ETL (Extract, Transform, Load) processes, and real-time data streaming.

    iv. Data Governance: Policies and procedures that dictate how data is managed, including data quality, security, access control, and compliance with regulations.
    v. Data Flow: The pathways and processes through which data moves within an organization, including data sources, processing, and end-user access. Covers various aspects such as data storage for efficient and secure data handling, data movement including integrations, data modelling to structure and organize data, and the management of metadata, reference data, and master data.

c. **Data Quality**
   a. The overall condition and reliability of data, determined by various attributes that assess its accuracy, completeness, consistency, timeliness, and relevance for specific purposes. High-quality data is essential for effective decision-making, analysis, and operational efficiency within an organization.
   b. Key dimensions of data quality include:
      i. Accuracy: The extent to which data correctly reflects the real-world entities or events it represents. Inaccurate data can lead to misguided decisions and analysis.
      ii. Completeness: The degree to which all required data is present. Incomplete data can result in a lack of insights or incorrect conclusions.
      iii. Consistency: The uniformity of data across different data sources or systems. Inconsistent data can arise from multiple data entries or variations in data formats.
      iv. Timeliness: The relevance of data based on its age or freshness. Timely data is essential for current decision-making, while outdated data can lead to missed opportunities or errors.
      v. Relevance: The degree to which data is applicable and useful for the intended purpose or context. Irrelevant data can clutter analysis and confuse decision-making processes.
      vi. Validity: The adherence of data to predefined formats, rules, or constraints. Valid data conforms to expected standards, which enhances its reliability

d. Data Lifecycle Management (DLM)
   a. The comprehensive process of managing data from its initial creation and acquisition through its active use, storage, archiving, and eventual deletion or destruction. DLM encompasses the policies, practices, and technologies used to ensure that data is properly handled throughout its entire lifecycle, optimizing its value and ensuring compliance with regulatory and security requirements.
   b. Key stages of the data lifecycle include:

i. Data Creation/Acquisition: The initial stage where data is generated or collected from various sources, including user inputs, transactions, sensors, or external data providers.

ii. Data Storage: The process of storing data in appropriate repositories, such as databases, data lakes, or cloud storage, ensuring accessibility and security.

iii. Data Usage: The active phase where data is accessed, analysed, and utilized for various business processes, decision-making, or reporting.

iv. Data Sharing: The distribution of data across different users or systems, ensuring that the right people have access to the necessary data while adhering to privacy and security policies.

v. Data Archiving: The process of moving infrequently accessed data to lower-cost storage solutions while ensuring it remains retrievable for future reference, compliance, or auditing.

vi. Data Retention and Deletion: The implementation of policies governing how long data should be retained and the proper methods for securely deleting or destroying data that is no longer needed, in compliance with legal and regulatory requirements.

vii. Data Disposal: The final stage where data is permanently removed from storage systems, ensuring that it cannot be recovered or misused.

6 Data Lifecycle Management

The following are key guidelines for managing data throughout the entire data lifecycle:

6.1 Data Creation/Acquisition

a. All data acquired must comply with regulatory requirements; particularly PDPA

b. Data acquisition includes:

   a. Data Generation: Capturing data from various sources, such as applications, sensors, user inputs, or transactions

   b. Data Collection: Gathering data from external sources or datasets, including third-party vendors, public data, or web scraping.

   c. Data Entry: Inputting data manually or automatically into databases or systems, ensuring it meets predefined formats and standards.

6.2 Data Storage

a. All data must be stored according to its use cases while ensuring the appropriate level of data confidentiality, integrity, and availability

b. Data Classification: Categorizing data based on its type, sensitivity, and relevance to determine appropriate storage solutions.

c. Data Storage Solutions: Choosing suitable storage technologies (e.g., databases, data lakes, cloud storage) based on performance, scalability, and cost.

d. Data Backup: Implementing regular backup procedures to ensure data recovery in case of loss, corruption, or disaster.

6.3 Data Usage

a. All data usage and processing must be authorised according to the user access matrix

b. Data Access Control: Defining and enforcing user permissions to ensure that only authorized personnel can access sensitive data.

c. Data Processing: Performing operations on data to transform it into a usable format for analysis, reporting, or operational needs

6.4 Data Sharing

a. Any form of data sharing must be preauthorised with the explicit agreement of the respective Data Owner

b. Data sharing approaches:

   a. Data Integration: Combining data from multiple sources or systems to provide a unified view for users or applications.

   b. Data Distribution: Facilitating the sharing of data across departments, teams, or with external partners while ensuring compliance with data sharing policies.

   c. API Management: Creating and managing application programming interfaces (APIs) that allow systems to interact and exchange data securely.

6.5 Data Archiving

a. Data should be archived according to the usage policy defined by the Data Owners
b. Data Identification: Identifying data that is infrequently accessed but must be retained for compliance, historical reference, or auditing.
c. Data Migration: Transferring data to an archival storage solution that is cost-effective and secure.

6.6 Data Retention and Deletion

a. All data must be retained for a minimum of 7 years
b. Data Review: Periodically reviewing data to determine its relevance and necessity, identifying data that can be deleted
c. Data Deletion: Implementing secure deletion procedures to ensure data is permanently removed and cannot be recovered, including data wiping or physical destruction of storage media.

6.7 Data Disposal

a. Data Disposal Procedures: Establishing clear guidelines for the secure disposal of data and storage devices to protect sensitive information.
b. Compliance Audits: Conducting audits to ensure that data disposal practices align with legal, regulatory, and organizational policies.
c. Documentation: Maintaining records of data disposal activities to demonstrate compliance and accountability.