༄ འབྲུག་གཞུང་འཕྲུལ་རིག་ལས་སྡེ།
GovTech
Bhutan

DATA MANAGEMENT GUIDE

2023

Introduction

The rapid development of technology on the one hand and the demand for online services from the citizens, on the other hand, require the need for collecting various data types from different sources. It ranges from public to personal to confidential data and as such Government agencies are concerned about security and ownership when requested to share data. Accessibility and availability of data through e-services heighten the information security risk and it necessitates safeguarding these information assets in governmental agencies. The step towards safeguarding and protecting data is to consider factors such as what kind of data can and cannot be shared with whom and in which format besides its criticality and sensitivity.

Objective

The objective of these guidelines is to:

a. Minimize potential errors and risks by establishing data handling processes and policies;
b. Uphold service users' trust in handling their sensitive information by complying with legislative standards;
c. Optimize the management of data for the agency's needs (e.g. to identify trends for service planning, enhance the quality of service delivery, monitor service users' journeys, etc.)

Scope

The scope of these Guidelines shall apply to

a. Data that is generally collected and created online and offline, generated, collected, and stored information data.
b. Every government agency collects and stores\ information assets.
c. Online and offline data

Legal Basis

These Guidelines are issued in the exercise of the power conferred Information Communications and Media Act of Bhutan 2018[1] to govern the electronic data collection, protection, storage, and disclosure.

[1] https://www.dit.gov.bt/sites/default/files/attachments/ICM%20Act%202018.pdf

Title

These Guidelines shall be called the Guidelines on Data Management and shall come into force on .... corresponding to the .... day of the ...... month of the Bhutanese .... Year.

Definitions

a. *Data* means a representation of information, knowledge, facts, concepts, or instructions that are being prepared or have been prepared in a formalized manner, and which is intended to be processed, is being processed, has been processed, or is capable of being processed in a computer system or computer network, and may be in any form including computer printouts, magnetic or optical storage media, punched cards, or punched tapes or stored internally in the memory of a computer, computer system or computer network;

b. Government means The Royal Government of Bhutan

c. Person means any individual

d. Sensitive Personal Data or Information means password;

   i. financial information such as bank account or credit card or debit card details, etc;
   ii. physical, physiological, and mental health conditions; sexual orientation;
   iii. medical records and history;
   iv. biometric information; and
   v. other information that may be legally deemed to be private.

e. Provided that, any information that is freely available or accessible in the public domain or available under any other existing national laws shall not be regarded as sensitive information.

Guide Layout

These guidelines are organized in five stages of the data management lifecycle:

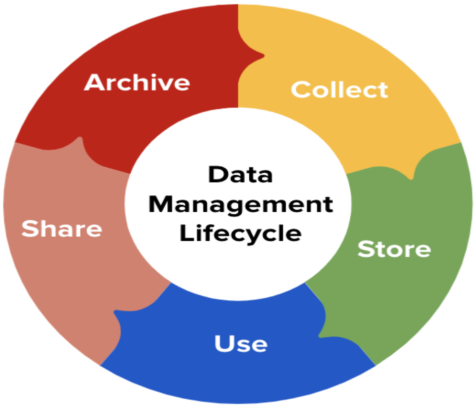


Figure 1: Stages of Data Management

Chapter 1: Data Collection

Data collection refers to the process of acquiring information, and data. Agencies collect data for various purposes including service delivery, analysis, and evidence-based decision-making. An effective process of any data collection includes clarity on the following questions:

a. Why do agencies need to collect data?
    a. Define the primary objective of data collection

b. What are the types of data to be collected?
    a. Identify the type of data needed to meet the objective [qualitative or quantitative, primary or secondary]

c. How to collect the required data?
    a. Establish a suitable method of data collection [interviews, surveys, observations]

d. How to control the quality of the collected data?
    a. Apply quality control measures [data validation, cleaning, and verification before, during, and/or after data collection to ensure data accuracy and reliability

Chapter Overview

1.1. Determine data collection objective

It is crucial for agencies to clearly establish the main objective for data collection, and accordingly ensure that the data collected is used only for the intended purpose. In addition, agencies should inform individuals and other data sources regarding the primary objective before or during the data collection process.

a. *Establish the purpose of Data Collection:* Agencies /individuals should establish the purpose of data collection.
   For example, for delivering citizen services an agency is required to collect citizen data or to enhance an existing system an individual or an agency may conduct a feedback survey or customer satisfaction survey, or a group of volunteers to drive a blood donation campaign may collect individual information to record the donor's blood details and also to contact donors in emergencies.
b. *Determine data sources and data collection methods:* Agencies shall determine the sources and methods that will be used to obtain data.
c. *Establish data analysis and usage:* The plan on how to analyze and draw conclusions from the collected data should be clearly established.

Some common examples of data collection objectives include:

[Objective] *The agency aims to effectively deliver citizen services* **[Data requirements]** Individual Information; Required Services
[Objective] *The agency plans to enhance an existing system* **[Data requirements]** User feedback/satisfaction; User behavior
[Objective] *The agency wants to study the impact of a service/program* **[Data requirements]** Tracer; Impact assessment

Agencies are also recommended to develop internal data policies for internal and external stakeholders which should document the Do's and Don'ts when handling data and who should have access to data.

1.2. Types of data to be collected

Data is broadly categorized into two types:

Quantitative Data

- □ *numerical and measurable*
- □ *provide standardized information that can be easily summarized and tested*
- □ *difficult to capture complexity of the issue*

Examples: counts, percentages, or scores

Qualitative Data

- □ *non-numerical and descriptive*
- □ *provide rich and detailed insights into the phenomena of interest*
- □ *often difficult to generalize and compare*

Examples: words, images, or opinions

Based on agencies' objectives and sources, the types of data collected include:

Personal Data:

- □ *any information which relates to a person and can be used to identify an individual*
- □ *often used for administrative purposes*

Examples: Name; CID number; Contact details

System Data:

- □ *any information collected, processed or generated by a computer system or network for and to provide services by the agency*
- □ *used to manage and monitor service delivery*

Examples: Log data; Performance metrics; Error reports

Research Data:

- □ *any information collected in the course of conducting research*
- □ *used to support and/or validate prior research findings*

Examples: Surveys; Observations; Data compilation from various sources

Program data:

- □ *any information that measures the inputs, outputs, and outcomes of an agency's program or intervention*
- □ *used to guide decisions, show effectiveness, or justify funding*

Examples: Indicators; Results/Outcomes; User satisfactions

Case Data:

- □ *any information collected that is related to an agency's specific case*
- □ *used to track progress and evaluate performance*

Examples: Case number; Referrals; Status

These data can be in form of any one of the following:

St data:

- □ organized and formatted in a specific way, making it easy to search and analyze

Examples: spreadsheets, databases, and financial records

Unstructured data:

- □ not organized in a specific format and is more difficult to search and analyze

Examples: emails, social media posts, and video or audio recordings

Semi-structured data:

- □ a mix of structured and unstructured data

Examples: XML files and JSON data

Different data types require different storage and processing techniques, and may require different levels of security and privacy protections depending on the type of data and its intended use.

1.3. Methods of data collection

There are various methods to collect data that agencies can use to gather information for a specific purpose. When choosing a data collection method, agencies must consider clear objectives, feasibility, and resource optimization. Agencies should choose the data collection method that is most appropriate for their needs and budget, allowing them to gather the information they need reliably and accurately.

Some examples of common data collection methods include:

QUANTITATIVE DATA

Surveys:

- □ involve a series of questions to a group of people to gather information
- □ can be conducted in person, by phone, or online

Observations:

- □ involve watching and recording the actions or behaviors of individuals or groups
- □ can be conducted in person or remotely (through video recordings)

Experiment:

- □ involve manipulating one or more variables to study their effects on a particular outcome
- □ commonly used in scientific research

QUALITATIVE DATA

Interviews:

- □ involve one-on-one conversations between a researcher and a participant.
- □ conducted in person, by phone, or online.

Focus Group:

- □ bring together a small of people to discuss views and opinions on a specific topic or issue
- □ moderated by a trained facilitator to guide the discussion and asks open-ended questions

Records and Documents:

- □ can take many forms, such as written reports, emails, photographs, videos, and audio recordings
- □ may be stored in physical or electronic formats

Considerations for data collection under the ICM Act 2018

PERSONAL DATA

Personal Data refers to *Personally Identifiable Information* which is any information that can be used to identify a specific individual. Examples of Personal Data include a person's name, date of birth, CID number, address, email address, and phone number.

Personal data management should comply with Chapter 21 of the ICM ACT of Bhutan 2018 on data protection. International best practices on personal data protection can also help agencies gather, handle, and disclose personal data fairly and responsibly and provide individuals control over their data.

CONSENT

An agency shall obtain the express written permission of the subject for the collection, collation, or processing of any personal information unless permitted or required to do so by law

DISCLOSURE NOTIFICATION

Agency shall not disclose any of the personal information held by it to a third party unless required or permitted by law or specifically authorized to do so in writing by the concerned person.

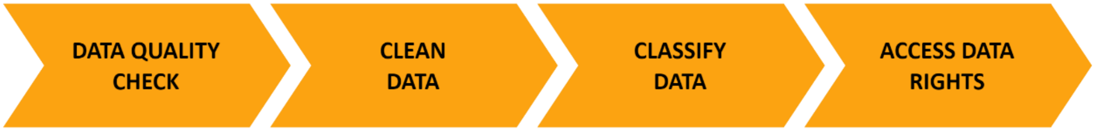
DATA ARCHIVAL OR DESTRUCTION AND ANONYMIZATION

The agency possessing, dealing with, or handling any personal data, including sensitive personal data or information, shall delete or destroy all personal information that has become obsolete.

Provided that the person may use that personal information for statistical purposes as long as the profiles or statistical data cannot be linked to any person by a third party.

1.4. Data quality control

Agencies must ensure that the collected data is of high quality with proper validation, organization, classification, and assignment of access rights before storing, using, and sharing the collected data.



1.4.1. Check Data Quality

Checking data quality involves checking and ensuring that the data is accurate, complete, consistent, relevant, and timely.

Some components of checking and verifying the quality of data include:

1. Accuracy

- Verify that the data is correct with minimal or no errors

Example: If the agency collects the CID number of their service users, to ensure that the CID is accurate, ensure that the CID provided by the user is accurate by cross-checking with the Census information.

2. Validity

a. Mandatory fields:

- Ensure all mandatory fields are filled

Example: Name and contact details of users are not left blank for any online services for

b. Cross-field validation:

- Check for any data that may affect other data fields.

Example: Counter-check user’s inputs on age through different in years of data collection and birth year

c. Unique identifiers:

- Ensure the requirement of unique conditions for data.

Example: National Citizen Identity number, Business License number, Driving License number

d. Constraints:

- Check if data is within the restricted set

Example 1: The number of days in February is either 28 or 29
Example 2: The date of graduating grade 10 is around the time the user was 10 years old

e. Normal Structure:

- Ensure data is of specific format

Example 1: Contact numbers recorded as 17XXXXXX or 0232XXXX

Example 2: Email addresses should be in the format ‘xxxx@xxx.com’

f. Conformity:

- Compare data to predefined rules or standards to ensure conformity

Example 1: Using Kilogram(s) for Weight; Number of completed years for Age

Example 2: Names (and spellings) of Dzongkhags, Gewogs, and Villages consistent with Census standards

3. Completeness

- Ensure the required data is complete.

Example 1: The date of birth is not complete if either date, month or year is missing:

⊗ 12th April
✓ 12th April 1996

Example 2: For occupations, avoid users from providing abbreviations:

⊗ AFO, DTO
Assistant Finance Officer, District Trade Officer

4. Consistency

- Define rules or restrictions to ensure data consistency by specifying which combinations of values are valid

Example 1: If you use the DD/MM/YYYY format for Date of Birth, continue using it in other date-related fields.

Example 2: Suppose there are two field types for the employment section: employment type and employment sector. If the employment type is “farmer” but the sector is “Government agencies,” this indicates data inconsistency.

5. Timeliness

- Ensure there is no wide gap between data collection and data availability for the predefined objective

Example 1: Collect and distribute hospital data with healthcare providers quickly to help them make patient care decisions.

6. Data auditing

- Review the data quality processes and outcomes to assess their effectiveness and compliance

Example: Collect, store, and manage data processing; Update incorrect records or add data validation rules to fix data errors; Update data gathering methods or add new data management systems to increase data quality.

Resources to assess data quality

The Bhutan Statistics Quality Assurance Framework (BSQAF) 2020, based on the United Nations National Quality Assurance Frameworks Manual for Official Statistics, outlines how agencies can maintain official statistics throughout the entire Bhutan Statistical System[2] (BSS) at the highest level. The Framework also guides coordination between BSS, and data providers, and statistics producers outside of the BSS to generate quality official statistics.

Agencies may refer to the Guideline on Assessing Quality of Administrative Data for Producing Official Statistic 2022 along with the mapping tool and questionnaire for guidance on assessing and improving the quality of their existing administrative data.

[2] BSS includes the National Statistics Bureau (NSB) and other producers of official statistics.

Additional resources from NSB to conform and standardize data across various agencies include:

1. Standardization of Measurement of Unit 2002,
2. Bhutan Standard Statistical Code (BSSC) 2020,
3. Bhutan Standard Statistical Geographic Code (BSSGC) 2020,
4. Bhutan Standard Industrial Classification (BSCI) 2020, and
5. Bhutan Standard Classification of Occupation (BSCO) 2021.

1.4.2. Clean Data

Data cleaning process will help agencies to address and resolve issues with data that is incorrect, corrupt, improperly formatted, duplicated, or incomplete within a dataset.

Some data cleaning techniques that may be used include

Clean irrelevant data:

□ removing any observations that is not relevant after confirming that they are not needed
□ checking for and correcting any errors - incorrect or inconsistent values to ensure the accuracy of the analysis
□ identifying and handling outliers to avoid any biases or skewing

Remove duplicates:

□ remove any duplicate observations after confirming that they are duplicates

Check for missing values

□ if the actual values are not available, consider either replacing them with default values, such as “0”, “-99” for numerical data, or “MISSING”, “UNKNOWN”, “NA”, “99” for categorical data

Categorize field types:

□ Group the data into categories or classes based on similar characteristics to simplify the data for further analysis[3]

[3] Agencies can refer and adopt the standards specified by the NSB in their available resources for ease of intra-agency and inter-agency data sharing.

1.4.3. Classify Data

Data classification agencies to classify data and define the principle for data classification such that data is maintained and shared, internally and externally, without violating any legal obligations or bringing any harm to the proper function of the agency or the welfare and safety of its staff or third parties upon disclosure.

An example of standard data classification is discussed below:

Public data:

- □ Applies to data with minimal or no foreseeable risks from public disclosure
- □ Minimum requirement of security by the agency to protect the integrity and availability of this data

Examples: Data routinely distributed to the public regardless of whether the agency has received a public records request - Annual reports, publicly accessible web pages, marketing materials, and press statements.

Internal data:

- □ Applies to the data created and to be consumed by the agencies concerned
- □ This data usually requires permission from the Head of the agencies before sharing
- □ Its compromise may bring inconvenience to the agency, but is unlikely to result in a breach of confidentiality, loss of value, or serious damage to integrity.

Examples: Data that have not been finalized or adopted, Internal Memos and notifications, personal communications (including email, Minutes of meetings, and file notes of agencies)

Confidential data:

- □ Applies to data or material whose unauthorized disclosure could reasonably be expected to cause damage to the work of the agency or person
- □ Most restricted distribution must be protected at all times as its compromise could seriously damage the safety or integrity of an agency, its staff or the users.

Example: Personal case files such as benefits, program files, or personnel files; Tax returns or financial health of the agency; Personal health information of individuals; and Salary information

Strictly confidential data:

- □ Applies to data whose unauthorized disclosure could be detrimental to the sovereignty and security of the State
- □ Mandatory level of protection as is prudent or as required by law

Examples: Susceptible information affecting the interest of the sovereignty, security, unity, peace, or foreign relations of Bhutan

1.4.4. Assign Access Rights

Assigning access rights is the process of granting or restricting permissions to different users or groups to view, modify, delete, or share data. Assigning data access rights is crucial to reduce information security risks by protecting the confidentiality, integrity, and availability of data, as well as ensuring compliance with relevant laws and regulations.

Access rights can be based on various criteria, such as the roles, responsibilities, or levels of trust of the users, and some key components in the process include:

Identifying data owners and stakeholders:

Data owners are responsible for defining the purpose, scope, and classification of data. Stakeholders are the users who need access to data for their roles and tasks. Data owners and stakeholders should collaborate to establish access rights policies and procedures.

Defining access levels:

Access levels are the permission levels that determine what actions users can perform on data. Access levels and roles should be aligned with the data classification, sensitivity, and type of users.

Implementing access control mechanisms:

Access control mechanisms are the tools and methods that enforce the access rights policies and procedures. They can be based on different models, such as discretionary, mandatory, or role-based access control. Access control mechanisms should be consistent, secure, and auditable across all platforms and devices that store or process the data.

Data access control should apply for both electronic and hardcopy data. For online data, system administrators can regulate data access depending on user roles. For hardcopy data, physical security measures such as locked cabinets or rooms, and establishing protocols for handling and storing sensitive documents are recommended. To ensure only authorized users have access, agencies can mechanisms to track the distribution and disposal of hardcopy data.

Regularly reviewing and updating access rights:

Access rights should be regularly reviewed and updated to ensure validity and relevancy. Periodic audits and evaluations of access rights and usage should be carried out along with updating or revoking access rights when there are changes in the data or the users.

Considerations for personal data protection in data collection

When collecting personal data, agencies can consider the following:

1. Personal data is collected and processed fairly, where the users or individuals are informed about the objective of the data collection, and their consent is obtained wherever necessary.

2. Users or individuals are aware of how and who will use their personal data.
3. Data collection is limited to ensure only the personal data needed for the intended purpose is collected.
4. Personal data is not stored longer than needed. Data should be securely removed or anonymized after use.
5. Protect personal data from unwanted access, loss, and misuse.
6. People can access, correct, delete, and object to the use of their personal data.

Chapter 2: Data Storage

Data storage refers to storing collected data in a physical medium, thereby allowing users to access, modify, and share their data. With the increasing number of data collected and stored by agencies, effective data storage - storing information in a safe, structured, and organized manner – has become a critical component of data management.

Efficient data storage allows agencies to manage data efficiently, readily access and retrieve information, and make decisions based on accurate and relevant data. It also ensures sensitive and confidential information is protected from security breaches, cyberattacks, and data theft as data is stored in a secure and compliant manner.

There are several types of data storage depending on the capacity, performance, dependability, and accessibility of the medium. Different types of data storage have different advantages and disadvantages in terms of capacity, speed, reliability, and cost.

Chapter Overview

2.1. Data storage types

Data storage is a critical component of data management. To manage and use data effectively, it must be stored in a way that is secure, reliable, and accessible. Data storage types can be broadly categorized into two – (1) On-premise (government) data storage, and (2) Commercial cloud data storage – with their respective pros and cons.

Agencies should choose the storage type that meets their specific needs and requirements. In general, commercial cloud data storage is typically known for being cost-effective and flexible, while on-premise data storage offers greater control and security, managing and storing sensitive government information.

On-premise (Government) Data Storage

This storage type is owned, operated, and managed by the agency for its own purposes.

Pros:

□ *High level of control over data and infrastructure*
□ *Stronger data security and compliance with regulations*
□ *Greater flexibility to customize to specific business needs*
□ *Reduced latency and improved performance*

Cons:

□ *High upfront costs for hardware, software, and maintenance*
□ *Limited scalability and agility*
□ *Increased complexity with management and upgrading*
□ *Potential for power and cooling inefficiencies leading to higher utility bills*

Commercial Cloud Data Storage

This data storage type is owned and operated by third-party providers to store and process data over the internet.

Pros:

□ *Cost-effective and flexible scaling*
□ *Increased reliability and uptime due to redundancy and failover capabilities*
□ *Access to a wide range of computing resources and services*
□ *Reduced management and maintenance requirements*

Cons:

□ *Limited control over infrastructure, data management, and security policies*
□ *Higher risk of security breaches and data loss, especially for less established providers*
□ *Depen internet connectivity leading to latency and network issues*
□ *Subscription-based pricing model can be costly for businesses with consistent, high use.*

Examples: Google Cloud; Amazon Web Services

Considerations for Choosing a data center

Choosing a type of data storage is an important decision that can have a significant impact on the reliability, security, and performance of your agency's IT infrastructure. Here are some factors to consider when selecting a data storage type:

DATA VOLUME

When planning for data storage, the quantity of data generated by the agency should be the primary consideration. Understanding the current and projected data volume will help determine the necessary storage capacity.

DATA ACCESS NEEDS

It is essential to evaluate the frequency and type of data access required to effectively meet the agency's data collection objectives. This helps determine the proper storage infrastructure, including the required storage type, such as Solid-State Drive (SSD), Hard Disk Drive (HDD), and Network Attached Storage (NAS).

SCALABILITY

As data volume increases, storage needs will also shift. Taking scalability into account during the planning phase can help in ensuring that the storage infrastructure can accommodate future demands without experiencing delays or data loss.

SECURITY AND PRIVACY

Prioritizing on security and privacy when planning for data storage can help reduce the risk of unauthorized access to sensitive and confidential data. Consideration should be given to the legal and regulatory requirements for data security and privacy when determining the location of data storage types.

CONTINUITY AND DISASTER RECOVERY

Planning for data storage should take into account requirements for continuity and disaster recovery. This involves ensuring that the data storage type supports high availability and data protection.

COMPLIANCE

It is important to understand the specific compliance needs and requirements that apply to an agency in order to ensure that its storage infrastructure complies with these regulations. By accounting for compliance while planning for data storage, agencies can ensure that they comply with all legal and regulatory requirements, minimizing their risk exposure.

Choosing a data center that meets the agency's needs in the above areas will ensure that the IT infrastructure is reliable, secure, and high-performing.

Some good practices for email, websites, and web application security:

1. *Implement a strong password policy:* Use strong passwords for email, websites, and web applications and change them regularly. Avoid using easily guessable passwords such as your name, date of birth, or a dictionary word.

2. *Implement two-factor authentication:* Implement two-factor authentication to add an extra layer of security to your email, website, and web application logins.

3. *Keep software up-to-date:* Keep all software, including operating systems, web servers, and web applications up-to-date with the latest security patches and updates. Use genuine software.

4. *Use encryption:* Use encryption for sensitive data, such as credit card information, passwords, and personal data. Implement TLS encryption (SSL Certificate) on your website to protect data in transit and Secure File Transfer Protocol.

5. *Secure email communication:* Use end-to-end encryption for sensitive email communication and avoid sending sensitive information via email whenever possible.

6. *Implement input data validation and output encoding in web applications:* To ensure that any inputs are valid and do not contain malicious codes.

2.2. Stored data protection

The protection of stored data requires a multilayered strategy combining technical and administrative measures. By using encryption, access controls, network security, backups, updates, physical security, and proper retention and disposal practices, agencies can protect their stored data from unauthorized access, loss, and theft.

Based on the level of security needed and the specific requirements of the data, here are some ways to protect stored data:

1. Encryption:

One of the most effective ways to protect stored data is to use encryption. This involves converting the data into a code that can only be read by someone with the key to decrypt it. Encryption can be applied to individual files or folders, or to entire drives or databases.

2. Access controls:

Access controls can be used to restrict access to stored data to only authorized users. This may involve using passwords, PINs, or biometric authentication, as well as implementing role-based access controls that limit access to certain data based on job function.

3. Firewalls and network security:

Firewalls and other network security measures can be used to prevent unauthorized access to stored data from outside the agency. This may involve using intrusion detection and prevention systems, as well as network segmentation to isolate sensitive data from other parts of the network.

4. Backup and recovery:

Regularly backing up stored data is important to protect against data loss in the event of a hardware failure, natural disaster, or other catastrophic event. Backup data should be stored in a secure location that is separate from the primary storage location.

5. Patching and updates:

Keeping software and operating systems up-to-date with the latest security patches and updates can help protect against known vulnerabilities that could be exploited to gain unauthorized access to stored data.

Considerations for personal data protection when storing data

When storing personal data, agencies can consider the following:

1. *Obtain consent:* Agencies should obtain the individual's consent before collecting, using, or disclosing their personal data. Consent should be informed, meaning the individual understands the purpose of the data collection and how their data will be used.
2. *Limit collection:* Agencies should collect only the personal data that is necessary for the stated purpose, and should not collect data that is not relevant or required.
3. *Protect personal data:* Agencies should take appropriate security measures to protect personal data from unauthorized access, use, or disclosure. This may include physical, technical, and administrative safeguards.
4. *Be transparent:* Agencies should be transparent in their data handling practices and should provide individuals with information about their rights and how to exercise them.
5. *Implement accountability measures:* Agencies should have policies and procedures in place to ensure compliance with the Personal Data Protection, and should be accountable for any breaches of personal data protection.
6. *Provide access and correction:* Agencies should allow individuals to access their pers onal data and request corrections to it if it is inaccurate.
7. *Retain data only as long as necessary:* Agencies should not retain personal data for longer than necessary to fulfill the stated purpose.

2.3. Data backup

In order for the agencies to minimize the risk of data loss or damage and effectively manage cyber threats, it is recommended that the agencies:

- develop a data backup process, and
- develop an incident response and reporting plan.

2.3.1 Data backup process

Processes to routinely backup data enable agencies to protect data against data loss due to system malfunction, cyberattacks, human error, or natural disasters. Here are some components of a data backup process:

1. Determine data backup scope:

☐ Identify data that needs to be backed up

2. Set backup frequency:

☐ Consider the amount of data generated, the frequency of data updates, and recovery point objectives (RPOs)[4].

3. Select backup storage locations:

☐ Backup storage should be done in secure locations, either on-premises or in the cloud[5]

☐ Require sufficient storage capacity to meet the agency's needs

4. Select a backup method:

☐ Choose a backup method that best suits the agency's needs and budget:

   a. Full back-up: This method copies everything you want to back up;
   b. Incremental backup: This method only backs up changes made since the last backup;
   c. Differential backup: This method backs up all changes since the last full backup.
   a. Mirror backup: This method creates an exact copy of the original data

5. Set backup window:

☐ Allocate sufficient time to complete the backup process when determining backup frequency and method with no or minimal interference to regular work

[4] RPOs to identify how frequently backup data should be captured in order to meet the agency's goals and ensure continuity of operations in the case of data loss.

[5] In the case of the cloud option, it is important that agencies have a thorough understanding of the cloud-based backup services offered by the Cloud Service Provider (CSP).

6. Assign backup personnel roles:

☐ Define the roles and responsibilities of the backup personnel, including backup administration, testing, and monitoring

7. Backup process testing:

☐ Regular testing of the backup process to ensure that it is working effectively, and that data can be restored easily and/or recovered quickly in case of data loss

Agencies are recommended to regularly review and update their data backup processes to ensure that they meet changing requirements and objectives, technology advancements, and regulatory compliance requirements.

Incident response and reporting plan

As agencies host and manage several data and systems, developing an incident response and reporting plan is critical. Having an incident response and reporting plan in place ensures that incidents are detected, managed, and resolved in a timely and effective manner, minimizing the impact of incidents on the agency's day-to-day operations, protecting against data loss, and maintaining the trust of customers and stakeholders.

Here are some key components of an incident response and reporting plan:

1. Define and classify incidents:

- Clearly define what constitutes a reportable incident, such as malware attacks, unauthorized access, or denial of service
- Establish a classification system defining the severity of incidents based on impact, privacy protections, or regulatory compliance

2. Establish incident response procedures:

- Establish a clear and concise reporting process, including how incidents should be reported, who should be notified, and response procedures.

2.1 Incident resolution and recovery:

- ○ Define how incidents will be resolved - restore service, recover data, and implement any necessary changes - to prevent similar incidents from occurring in the future

2.4 Develop communication plan:

- ○ Develop a communication plan, including internal notifications and external notification requirements to government agencies or affected parties

2.3 Review and update incident response and reporting plan:

- Periodically review and update the incident reporting policy, considering changing threats, business needs, or regulatory requirements
- Regularly test the incident response plan to ensure that it is effective

2.5 Train incident response team members:

- Ensure that incident response team members are trained on incident response procedures

In the absence of an incident response plan and incident response team if an incident related to data occurs, the agencies (Government and Non-Government) or individuals could report to Bhutan Computer Incident Response Team (BtCIRT) through their website : www.btcirt.bt or via emailing at cirt@btcirt.bt.

Chapter 3: Data Usage

Effective data usage requires collecting and maintaining accurate and reliable data, analyzing and interpreting the data for insights or trends, and using the data to improve service delivery and drive informed decision-making.

Chapter Overview

3.1. Verify data usage objective

Data can be used in a variety of ways, depending on the specific context and the type of data that is being used. Before using data, your agency should verify the purpose for data usage:

- Determine the objective of data usage (Refer to Chapter 1, 1.1. on setting clear objectives during the initial data collection phase)
- Verify that the agency has the required data for the intended objective[6]
- Ensure accuracy, completeness, and timeliness of the data
- Ensure data usage is compliant with the agency's internal data policies

General purposes of data include:

1. Analysis:

Data can be analyzed to identify patterns, trends, and insights.

2. Decision-making:

Data can be used to inform decision-making processes. For example, a business might use sales data to make decisions about inventory management or pricing.

3. Predictive modeling:

Data can be used to build predictive models, which can be used to forecast future outcomes or behaviors.

4. Personalization:

Data can be used to personalize experiences for individual users. For example, a website might use data on a user's browsing history to provide personalized recommendations.

5. Optimization:

Data can be used to optimize processes or systems. For example, a manufacturing company might use data to optimize its supply chain or production processes.

6. Research:

Data can be used for research purposes, such as studying social trends, analyzing public opinion, or investigating scientific phenomena.

7. Reporting:

[6] Since various agencies are collecting information, before deciding to collect, store and use data, agencies can identify means to coordinate, collaborate data using and data sharing in order to avoid duplication of data collection.

Data can be used to create reports and visualizations that communicate information to stakeholders. For example, a company might use data to create a sales report or a dashboard that shows key performance indicators.

By understanding the ways in which data can be used, agencies can make better use of the data that is available to them, leading to improved decision-making and better outcomes.

Considerations for data usage

DATA QUALITY
Data quality is crucial when using data to make strategic decisions. It is essential to ensure that data is accurate, complete, and up-to-date to avoid making decisions based on flawed or incorrect data.

DATA INTEGRATION
Agencies often collect data from multiple sources, and it is important to integrate these sources to ensure that the data is complete and accurate.

DATA ANALYSIS
Data analysis involves reviewing and interpreting the data and identifying trends, patterns, or relationships that can be used to inform business decisions.

DATA SECURITY
It is essential to ensure that the data used in decision-making is secure and protected from unauthorized access, theft, or modification.

DATA PRIVACY
It is crucial that agencies ensure that personal or sensitive data is handled in compliance with applicable data protection and privacy laws to maintain users' trust and confidence.

DATA-DRIVEN DECISION-MAKING
Agencies can leverage data analytics to drive their strategies by making informed decisions based on relevant data insights.

3.2. Data usage methods

There are several types of data analysis methods, and the choice of method will depend on the type of data being analyzed and the specific objective of the analysis. In general, data can be used as follows, but not limited to:

- Improving data-driven services through data analytics
- Improving agency's operational efficiency with timely data
- Evaluating and assessing the impact of services/programs

Improve data-driven services through data analytics

Types of data analysis methods include:

1. Descriptive analysis:

Descriptive analysis involves summarizing and describing data using basic statistical measures such as mean, median, and mode. This type of analysis can be useful for providing an overview of the data and identifying patterns and trends.

Examples:

- Measures of central tendency: Mean, median, mode
- Measures of variation: Range, standard deviation, variance
- Frequency distributions: Histograms, frequency tables, bar charts, pie charts

2. Diagnostic analysis:

Diagnostic analysis involves exploring the relationships between different variables to understand the causes of specific outcomes. This type of analysis can be useful for identifying the root causes of problems or inefficiencies.

Examples:

- Correlation analysis: Measures the strength and direction of the relationship between two variables
- Regression analysis: Examines the relationship between one dependent variable and one or more independent variables
- Multivariate analysis: Looks at the relationships between multiple variables at once

3. Predictive analysis:

Predictive analysis involves using data to make predictions about future outcomes. This type of analysis can be useful for forecasting trends, identifying risks, and making informed decisions.

Examples:

- Machine learning: Uses algorithms to identify patterns and make predictions based on data
- Time-series analysis: Examines patterns over time and makes predictions about future trends
- Decision trees: Analyzes a decision-making process by identifying possible outcomes and assigning probabilities to each outcome

4. Prescriptive analysis:

Prescriptive analysis involves using data to identify the best course of action to take in a given situation. This type of analysis can be useful for making decisions about resource allocation, process optimization, and other complex problems.

Examples:

- Optimization: Finds the best solution to a problem based on specific constraints and goals
- Simulation: Models a real-world system or process and tests different scenarios to identify the best course of action
- Decision analysis: Helps decision-makers choose the best course of action by considering all possible outcomes and their probabilities.

General benefits of data analysis to an agency include

IMPROVED DECISION-MAKING
By analyzing data, agencies can gain insights that inform better decision-making. This can lead to better outcomes and more efficient use of resources.

ENHANCED EFFICIENCY
Data analysis can help agencies identify inefficiencies and areas for improvement. By optimizing processes and systems, agencies can become more efficient and effective.

IMPROVED PERFORMANCE
By monitoring and analyzing data, agencies can track their performance over time and make adjustments as needed to achieve better results.

BETTER COMMUNICATION
Data analysis can help agencies communicate their results and findings more effectively to stakeholders, such as funders, policymakers, and the public.

Overall, data analysis can be a powerful tool for agencies to improve their performance and achieve better outcomes. By choosing the right type of analysis[7], agencies can gain insights that lead to better decision-making, improved efficiency, and more effective communication.

Anonymization of personal data for use

If the data to be used contains personal information, your agency must conduct a risk impact assessment to ensure that personal data is processed in accordance with the ICM Act (Chapter 21 Data Protection). One of the measures is to use the data in an anonymized form.

[7] When choosing the type of data analysis, it is important to consider the objective of data analysis, the type of data to be analyzed and availability of resources.

Data anonymization refers to the process of converting personal data into data that cannot be used to identify any particular individual in order to protect privacy and confidentiality when using data. Here are some key considerations for anonymizing personal data for use:

OBJECTIVE OF ANONYMIZATION

It is important to have a clear understanding of why the data is being anonymized and what risks are being mitigated. This can help inform the specific anonymization techniques that are used.

TYPE OF DATA

The type of personal data being anonymized can impact the effectiveness of anonymization techniques. For example, sensitive data such as medical records or financial information may require more robust anonymization techniques than less sensitive data such as demographic information.

RISK OF RE-IDENTIFICATION

Anonymization techniques should be chosen with the risk of re-identification in mind. The goal is to ensure that the data cannot be re-identified, even if it is combined with other datasets.

DATA QUALITY

Anonymization techniques should not compromise the quality of the data. The data should still be useful for the intended purpose after it has been anonymized.

LEGAL AND ETHICAL CONSIDERATIONS

Anonymization should be done in accordance with legal and ethical guidelines. For example, certain types of personal data may be subject to specific privacy regulations that need to be taken into account.

TRANSPARENCY

It Is important to be transparent about the anonymization process and to clearly communicate any limitations or risks associated with using anonymized data.

Overall, anonymization of personal data is a critical consideration for protecting privacy and confidentiality when using data. By carefully considering the purpose, type of data, risk of re-identification, quality of data, legal and ethical considerations, and transparency, agencies can ensure that they are using anonymized data in a responsible and effective way.

DATA ANONYMIZATION TECHNIQUES

1. *Masking or redaction:* This technique involves replacing sensitive data with a placeholder or symbol. For example, masking credit card numbers or social security numbers by replacing them with "X" or "*".

2. *Generalization:* This technique involves replacing specific data values with a general range or category. For example, replacing the exact date of birth with the age or age range of the individual.
3. *Aggregation:* This technique involves combining data into groups or categories to conceal individual data points. For example, combining the ages of individuals into age brackets or ranges.
4. *Perturbation:* This technique involves adding random noise to the data to make it more difficult to re-identify individuals. For example, adding small amounts of random noise to the exact location of individuals in a geographic dataset.
5. *Cryptographic techniques:* This technique involves using encryption or hashing to protect personal data. For example, encryption or hashing a person's name or address to make it more difficult to re-identify the individual.
6. *Differential privacy:* This technique involves adding random noise to the data to protect the privacy of individuals while still allowing useful insights to be obtained. This technique has gained popularity in recent years, especially in the context of data sharing.

The choice of anonymization technique will depend on the type of data being anonymized, the level of protection needed, and the intended use of the data. It's important to carefully consider the trade-offs between privacy protection and data utility to ensure that the anonymization techniques used are effective and appropriate for the intended use of the data.

Improve Agency's Operational Efficiency with timely data

Maintaining current and accurate information is crucial in enhancing the agencies' operational efficiency. It is important for agencies to perform periodic data updates to maintain a single, reliable source of information thereby enhancing their efficiency. There are several ways to maintain current and accurate information including:

- Reducing the time and effort needed to verify users' information when receiving applications
- Ensuring consistency in staff assessments of user eligibility for services that have predetermined criteria
- Demonstrating accountability to stakeholders by meeting legal requirements for data management
- Efficient resource planning and development of data collection mechanisms.

Considerations for improving access to timely data in an agency

When using personal data, agencies can consider the following:

1. Regular data updates: Ensure that data is updated on a regular basis. This can be done manually or automatically depending on the type of data and its source.
2. Real-time data feeds: Consider using real-time data feeds to provide the most current information to decision-makers. This can be particularly important in industries where real-time information is critical, such as finance or healthcare.
3. Automated data collection: Automate data collection processes to reduce the lag time between data creation and analysis. This can involve using technologies such as IoT sensors, web scraping, or automated data extraction tools.
4. Data governance: Implement data governance processes to ensure that data is managed effectively and that data quality is maintained over time. This can involve establishing data standards, implementing data quality checks, and assigning ownership and accountability for data.

Data integration from different sources *and real-time data analytics* are recommended to create a more comprehensive and up-to-date view of the agency's data.

3.3. Internal protocols and procedures for data usage

Developing internal policies for data usage is an important aspect of ensuring that data is used in a responsible, ethical, and secure manner. Here are some key considerations to keep in mind when developing internal policies for data usage:

a. Purpose:

Clearly define the purpose for which data is collected and used, and ensure that it aligns with the agency's mission and values.

b. Data classification:

Policies should define the categories of data that are collected, processed, and used by the agency. These categories should reflect the sensitivity of the data and the potential impact on the agency and its customers if the data is lost, stolen, or disclosed inappropriately.

c. Compliance:

Ensure that the policies are compliant with all relevant laws, regulations, and industry standards.

d. Transparency:

Be transparent about data collection and usage practices, and provide individuals with clear information about their rights to access, correct, or delete their data.

e. Access controls:

Develop procedures for granting and revoking access to data, and ensure that access controls are appropriate for the sensitivity of the data being accessed.

f. Data retention:

Establish retention periods for different types of data, and develop procedures for securely deleting data when it is no longer needed.

g. Data protection:

Ensure that personal data is collected, used, and disclosed in accordance with applicable data protection laws (ICM Act). This includes the obligation to obtain explicit consent from data subjects for the collection, use, and disclosure of their personal data.

h. Data security:

Develop security measures to protect data from unauthorized access, loss, or theft, such as encryption, firewalls, and regular backups.

i. Communication and training:

Provide training to all employees on the agency's data usage policies and procedures, and ensure that they understand the importance of protecting data.

j. Audit:

Establish procedures for regularly auditing data usage practices, identifying areas of risk, and taking corrective action where necessary.

k. Accountability:

Define roles and responsibilities for data usage and ensure that individuals and teams are accountable for complying with the policies and procedures.

l. Data catalog:

Develop a centralized inventory or directory of an agency's data assets, which includes metadata such as data source, data format, data owner, etc.

Considerations for personal data protection in data usage

When using personal data, agencies can consider the following:

1. Obtain the consent of the individuals before collecting, using, or disclosing their personal data, and ensure that the consent obtained is valid, specific, and informed.
2. Purpose limitation to ensure personal data is collected and used only for the intended purposes, and provide prior information about the purpose to the individuals.

3. Notification to the individuals about the purposes for the use of data, on or before such use.
4. Access and correction mechanisms wherein agencies will need an internal policy in place to facilitate access and correction requests to personal data used or shared.
5. Accuracy of data before using it to make decisions affecting individuals.
6. Protection of personal data by anonymizing data as much as possible before analysis, or releasing only aggregated results to ensure individuals cannot be re-identified from anonymized data.

Chapter 4: Data Sharing

Data sharing is the process of facilitating data transfer across agencies. By granting agencies and individuals authorized access, data sharing can improve collaboration, operational efficiencies and service delivery, enabling data-driven decision-making. Simultaneously, agencies must implement policies and systems to ensure data-sharing practices are secure and ethical.

Chapter Overview

4.1. The objective of data sharing

Identifying clear data-sharing objectives allow agencies to share information with each other for a variety of reasons. The benefits of data sharing can include:

1. *Collaboration:* Data sharing allows multiple parties to work together on a common goal, which can lead to better outcomes and solutions.
2. *Efficiency:* By sharing data, agencies can avoid duplicating efforts and reduce the time and resources required to collect data.
3. *Transparency:* Data sharing can promote transparency and accountability by making information available to stakeholders and the public.
4. *Innovation:* Sharing data can lead to new insights and discoveries that would not have been possible if data were siloed.
5. *Better decision-making:* Access to more data can help individuals and agencies make more informed decisions.
6. *Cost savings:* Sharing data can reduce the costs associated with data collection, storage, and analysis.

However, it is important to note that data sharing must be done in a responsible and ethical manner, taking into account issues such as privacy, security, and intellectual property rights.

Considerations for personal data protection when sharing data

There are international best practices that are designed to ensure that personal data is collected, used, and disclosed in a fair and responsible manner and that individuals have control over their personal data. By following these practices, agencies can build trust with their customers and stakeholders, and avoid potential legal and reputational risks associated with non-compliance with data protection laws.

Agencies can consider the following best practices along with the ICM ACT of Bhutan 2018 on personal data protection:

1. *Purpose limitation:* Personal data should only be used for the specific purpose for which it was shared and not for any other purpose without the data subject's consent.
2. *Data minimization:* Only collect and share the minimum amount of personal data necessary for the intended purpose.
3. *Lawfulness, fairness, and transparency:* Personal data should be collected, processed, and shared in a lawful, fair, and transparent manner, with appropriate notice given to data subjects.
4. *Data accuracy:* Ensure that personal data is accurate, complete, and up-to-date.
5. *Data security:* Implement appropriate technical and agency-level measures to protect personal data against unauthorized or unlawful processing, accidental loss, destruction, or damage.
6. *Data retention:* Personal data should not be retained for longer than necessary, and should be securely deleted when no longer needed.

7. *Data subject rights:* Data subjects have the right to access, correct, and delete their personal data, as well as to object to processing, restrict processing, and request data portability.
8. *International data transfers:* When sharing personal data across borders, ensure that appropriate safeguards are in place to protect the data, such as standard contractual clauses or other approved transfer mechanisms.
9. *Data breach notification:* In the event of a data breach, promptly notify affected data subjects and relevant authorities.
10. *Accountability:* Maintain documentation and records to demonstrate compliance with personal data protection obligations.

By highlighting these personal data protection obligations, agencies can help ensure that personal data is shared and processed in a responsible and compliant manner, protecting the privacy and rights of data subjects.

4.2. Assess data for sharing

Before sharing data, it is important to assess the data to ensure that it is appropriate for sharing and that the risks associated with sharing are understood and managed appropriately. Here are some factors to consider when assessing data for sharing:

1. *Data quality:* The quality of the data must be assessed to ensure that it is accurate, complete, and relevant to the intended purpose of the data sharing.
2. *Data sensitivity:* The sensitivity of the data must be evaluated to determine if any special protections or controls are necessary to prevent unauthorized access, disclosure, or use.
3. *Legal and regulatory requirements:* The legal and regulatory requirements that apply to the data must be understood to ensure that the data can be shared in compliance with applicable laws and regulations.
4. *Security and confidentiality:* The security and confidentiality of the data must be assessed to determine if additional security measures or controls are necessary to protect the data from unauthorized access, disclosure, or use.
5. *Privacy considerations:* The privacy considerations associated with the data must be evaluated to determine if any privacy risks exist and if any privacy controls or protections are necessary to protect the privacy of individuals whose data is being shared.
6. *Ethical considerations:* The ethical considerations associated with the data must be evaluated to determine if any ethical issues exist and if any additional safeguards or protections are necessary to ensure that the data is shared in an ethical and responsible manner.
7. *Data use:* The intended use of the data must be evaluated to ensure that the data is being shared for a legitimate purpose and that the data will be used in accordance with any restrictions or conditions that may apply.

By assessing data for sharing, agencies can ensure that the data is being shared in a responsible and appropriate manner and that the risks associated with sharing are identified and managed effectively.

Considerations for Developing internal data protection policy

Developing an internal data protection policy is essential for any agency that collects, stores, and uses personal data. Here are some considerations for developing an internal data protection policy:

1. *Legal and regulatory requirements:* Ensure that the policy is compliant with all applicable laws and regulations prevalent in the country.
2. *Scope:* Determine the scope of the policy, including the types of data covered, the purposes for which the data will be used, and who will be responsible for implementing and enforcing the policy.
3. *Data security:* Develop appropriate security measures to protect personal data against unauthorized access, disclosure, alteration, or destruction.
4. *Data retention and disposal:* Define how long personal data will be retained, and specify how it will be disposed of when it is no longer needed.
5. *Data accuracy:* Ensure that personal data is accurate, complete, and up-to-date, and develop procedures for correcting or updating data as necessary.
6. *Data subject rights:* Develop procedures for responding to data subject requests, such as requests to access, correct, or delete personal data.
7. *Data breach response:* Develop a plan for responding to data breaches, including reporting the breach to the appropriate authorities and notifying affected individuals.
8. *Staff training:* Ensure that all staff members who handle personal data are trained on the policy and their responsibilities in relation to data protection.
9. *Privacy impact assessments:* Conduct privacy impact assessments to identify and mitigate any privacy risks associated with the collection, storage, and use of personal data.
10. *Review and update:* Regularly review and update the policy to ensure that it remains relevant and effective in protecting personal data.

By considering these factors when developing a data protection policy, agencies can help protect the privacy and security of personal data and comply with applicable laws and regulations.

Additional references include **GI Policy 2018**, **e-Governance Policy 2019**, and **ICM Act of Bhutan 2018**.

4.3. Establish a data-sharing agreement

A data-sharing agreement is a legal contract that defines the terms and conditions under which data will be shared between two or more parties. This agreement sets out the obligations, responsibilities, and limitations of each party involved in the data-sharing process, and ensures that the data is used only for the purposes specified in the agreement.

A data-sharing agreement typically includes the following elements:

1. *Purpose:* The agreement should clearly state the purpose for which the data will be shared and the intended use of the data.
2. *Data types:* The types of data to be shared should be specified, including any restrictions on the use of sensitive or confidential information.
3. *Roles and Responsibilities:* The roles and responsibilities of the parties involved should be clearly stated.
4. *Consent:* The parties involved should seek prior consent of the data owner for sharing data to any third parties
5. *Data ownership:* The agreement should specify who owns the data and who has the right to access and use it.
6. *Confidentiality:* The agreement should include provisions to protect the confidentiality of the data, such as limiting access to authorized individuals and requiring the parties to maintain appropriate security measures.
7. *Data retention:* The agreement should specify how long the data will be retained and how it will be disposed of when no longer needed.
8. *Data quality:* The agreement should include provisions to ensure the accuracy and completeness of the data, as well as the methods for resolving any disputes over data quality.
9. *Liability:* The agreement should specify the liability of each party for any loss or damage resulting from the use of the data.
10. *Data sharing frequency:* Agencies should determine the frequency of data sharing as per the requirements of the data consumers, i.e., whether data sharing is ad-hoc, one-off, or periodic, etc.

Data-sharing agreements are important to ensure that data is shared in a responsible, timely, and secure manner and that the rights and interests of all parties involved are protected.

4.4. Ways to share data

Here are some common ways data sharing is carried out across agencies:

1. *File transfer:* Data can be shared by transferring files from one location to another, such as through email, file-sharing services, FTP (File Transfer Protocol) and offline data transfer using portable drives, cloud sharing;

2. *APIs:* Application Programming Interfaces (APIs) allow software applications to communicate with each other, and data can be shared between applications using APIs.
3. *Data Sharing Portal:* Data can be shared through web-based data sharing platforms.
4. *Data warehouses:* Data can be stored in a centralized data warehouse, where it can be accessed and analyzed by multiple users or applications.
5. *Data sharing platforms:* There are many data sharing platforms available that allow researchers and agencies to share data and collaborate with others.

When choosing a method to share data, it is important to consider factors such as the sensitivity of the data, the size of the data, the technical expertise required to share the data, and the security and privacy implications of the sharing method. Agencies should also ensure that they have the appropriate legal and regulatory permissions to share the data in the chosen manner.

Chapter 5: Data Archival and Destruction

Data that is no longer actively used but still has value should be archived, while data that are identified to be no longer required should be destroyed to prevent unauthorized access. Archiving and destroying data also require relevant guidelines and protocols in place to ensure data is archived or destroyed in a secure manner. Effectively archiving and destroying data is an important part of an effective data management lifecycle, reducing storage costs and protecting data from unwarranted access.

Chapter Overview

5.1. Benefits of archiving data

Archiving data refers to the process of moving data that is no longer actively used or needed from a primary storage system to a secondary storage system. Here are some of the benefits of archiving data:

1. *Recovery:* Archiving data can provide an additional level of protection in the event of a disaster by allowing agencies to recover archived data from secondary storage systems.
2. *Cost savings:* Archiving data can help agencies reduce storage costs by moving less frequently accessed data to less expensive storage media, such as tape or cloud storage.
3. *Improved performance:* Archiving can help improve the performance of primary storage systems by reducing the amount of data that needs to be managed and accessed on a regular basis.
4. *Regulatory compliance:* Archiving data can help agencies meet regulatory requirements for internal data retention policy;
5. *Data preservation:* Archiving data ensures that important data is preserved for long-term retention, allowing agencies to access historical data for analysis or reference purposes.

Considerations for data collection under the ICM Act 2018

DATA ARCHIVAL OR DESTRUCTION AND ANONYMIZATION

The agency possessing, dealing with, or handling any personal data, including sensitive personal data or information, shall delete or destroy all personal information that has become obsolete. Provided that the person may use that personal information for statistical purposes as long as the profiles or statistical data cannot be linked to any person by a third party.

Data Archival on Clouds

Data archival on clouds refers to the process of storing data that is no longer needed for immediate access in an online cloud-based storage system. Here are some considerations for data archival on clouds:

1. *Data retention policies:* As with any data archival process, it is important to have a clear data retention policy that outlines how long data should be retained and under what circumstances it should be archived. This policy should be based on legal and regulatory requirements, as well as business needs.
2. *Cost:* Archiving data on the cloud can be cost-effective, but it is important to consider the costs associated with storage, access, and retrieval. Cloud storage providers typically

charge based on the amount of data stored, so it is important to estimate the amount of data that needs to be archived and to choose a storage plan that meets your needs.

3. *Security and privacy:* Cloud storage providers typically have robust security measures in place to protect data, but it is important to ensure that the provider you choose meets your security and privacy requirements. This includes compliance with applicable data protection regulations and adherence to best practices for data security.
4. *Data format:* It is important to consider the format of the data being archived. Cloud storage systems typically support a wide range of file formats, but it is important to ensure that the format is compatible with the storage system you choose.
5. *Access and retrieval:* While archived data may not be needed for immediate access, it is important to ensure that the data can be retrieved quickly and easily if required. This may involve implementing a system for indexing and searching archived data.
6. *Disaster recovery:* Cloud storage providers typically have disaster recovery plans in place, but it is important to understand the provider's plan and to have a contingency plan in place in case of data loss or disruption.

Data archival on clouds can be a cost-effective and efficient way to store data that is no longer needed for immediate access. However, it is important to consider data retention policies, cost, security and privacy, data format, access and retrieval, and disaster recovery when choosing a cloud storage provider for archival purposes.

5.2. Assess data for archival or destruction

When determining whether to archive or destroy data, agencies can take into account the following key areas:

A. Data Archival

1. Document retention requirements: Agencies are required to retain certain types of data for specific periods of time to comply with legal and regulatory requirements. Failure to comply with these requirements can result in legal or financial penalties.
2. Business needs: It's important to consider the value of the data to the agency. Some data may be critical for ongoing operations or may have future value for analysis or reference purposes.
3. Storage costs: Storing data indefinitely can be expensive, especially if the data is stored on high-performance storage systems. Archiving data can help reduce storage costs by moving less frequently accessed data to lower-cost storage media.
4. Security and privacy: Data that contains sensitive or confidential information should be securely stored or destroyed in accordance with the internal data policy to prevent unauthorized access or breach.
5. Disaster recovery: It's important to consider the potential impact of data loss in the event of a disaster. Archiving data can provide an additional level of protection against data loss.

6. Data management: It's important to have a clear understanding of the data that is being stored and to have a well-defined data management strategy to ensure that data is stored, archived, and destroyed in accordance with policies and regulations.

B. Data Destruction Profile

1. Redundant: Data has met its purpose of collection, use, and disclosure. Data is duplicated within the same system or across multiple systems. In such scenarios, your agency should only keep one source of reference.
2. Trivial: Data does not serve as evidence of business activity or historical value. Data does not provide corporate relevant knowledge and business insights.
3. Obsolete: Data has been superseded by other information and is inaccurate or invalid.

Document retention requirements

Document retention requirements refer to the legally mandated period of time that certain types of documents must be kept and maintained by individuals and agencies. These requirements are intended to ensure that important information is preserved and can be accessed if needed for legal, regulatory, or other purposes.

The specific document retention requirements vary depending on the type of document and the jurisdiction in which the individual or agency operates. Some common types of documents that are subject to retention requirements include

1. *Financial records:* These include bank statements, tax returns, receipts, and invoices.
2. *Employment records:* These include employee contracts, performance evaluations, and payroll records.
3. *Legal documents:* These include contracts, agreements, and legal correspondence.
4. *Medical records:* These include patient medical histories, diagnoses, and treatment plans.

The retention period for these documents can range from a few years to several decades depending on the type of document and the jurisdiction.

It is important for individuals and agencies to understand the specific document retention requirements that apply to them to ensure that they comply with legal and regulatory obligations. Failure to comply with document retention requirements can result in legal and financial penalties.

5.3. Archive or destroy data in a secure manner

Archiving and destroying data in a secure manner is important to ensure that sensitive information is not compromised. Here are some steps agencies can take to securely archive and destroy data respectively:

A. Data Archival

1. Develop internal data retention policy: An internal data retention policy outlines what data is to be retained, how long it is to be kept, and how it should be disposed of. It is important to ensure that the policy complies with legal and regulatory requirements.
2. Securely store archived data: Archived data should be stored in a secure, offsite location. It should be protected with appropriate physical and logical security measures, such as access controls, encryption, and backup systems.
3. Regularly review archived data: Regularly reviewing archived data can help ensure that data is still needed and that it is being stored in accordance with the data retention policy.
4. Maintain an audit trail: It is important to maintain an audit trail of all data that is archived and destroyed. This can help ensure that data is being managed in accordance with the data retention policy and can also be used to demonstrate compliance in the event of an audit or legal dispute.
5. Train employees on data security: All employees should be trained on the agency's data retention policy and procedures for securely archiving and destroying data. This can help ensure that employees understand the importance of data security and follow established procedures.

B. Data Destruction

1. Use secure methods to destroy data: Data should be destroyed in a manner that prevents recovery. This can be accomplished by using data destruction software or physical destruction methods, such as shredding or incineration.

Securely archiving and destroying data requires a well-defined data retention policy, secure storage and destruction methods, regular reviews of archived data, an audit trail, and employee training on data security. By following these steps, agencies can help ensure that sensitive information is protected and that they remain compliant with legal and regulatory requirements.

Retention of Personal Data

The retention of personal data refers to the period of time that an agency retains personal data after it has been collected. Personal data retention policies are typically based on legal and regulatory requirements, as well as business needs. Here are some considerations for the retention of personal data:

1. *Legal and regulatory requirements:* Agencies must comply with applicable data protection regulations. These regulations specify the requirements for the retention of personal data, and agencies must ensure that they comply with these requirements.

2. *Business needs:* agencies should also consider their business needs when determining the retention period for personal data. For example, some data may need to be retained for a longer period of time for business continuity or to meet contractual obligations.
3. *Data minimization:* Agencies should only retain personal data that is necessary for their business purposes. Unnecessary personal data should be deleted or destroyed in a secure manner.
4. *Data security:* Personal data must be protected with appropriate technical and agency-level measures to prevent unauthorized access, accidental loss, or destruction. This applies to both retained and destroyed personal data.
5. *Data subject rights:* Individuals have the right to request access to their personal data, request corrections or erasure of their personal data, and object to the processing of their personal data. Agencies must have procedures in place to respond to these requests.
6. *Data protection impact assessments:* Agencies should conduct data protection impact assessments to identify and mitigate any potential risks to individuals' privacy rights that may arise from the retention of personal data.

Considerations for Personal Data Protection for Archiving and Destroying Personal Data

1. *Data minimization:* Only archive or destroy personal data that is necessary and delete or destroy any unnecessary personal data in a secure manner.
2. *Lawful basis:* Ensure a lawful basis for processing personal data, such as consent or legitimate interests for both archiving and destruction of personal data.
3. *Privacy notices:* Inform individuals about the processing of their personal data, including archiving and destruction processes, through clear and concise privacy notices.
4. *Data retention:* Retain personal data for no longer than necessary and in accordance with applicable data retention policies and legal requirements.
5. *Security measures:* Protect personal data with appropriate technical and agency-level measures to prevent unauthorized access, accidental loss, or destruction, including personal data that are archived and destroyed.
6. *Data subject rights:* Ensure procedures in place wherein individuals have the right to request access to their personal data, request corrections or erasure of their personal data, and object to the processing of their personal data.
7. *Data protection impact assessments:* Conduct data protection impact assessments to identify and mitigate any potential risks to individuals' privacy rights that may arise from archiving or destroying personal data.

8. *Data portability:* Ensure procedures are in place to facilitate the transfer of personal data to individuals or other agencies in a structured, commonly used, and machine-readable format, if requested