

#19 Data Management

The notes are 14 very dense pages filled with paragraphs. Here is a simple pdf summarizing everything as much as possible.

ATB! - Raghav, Computing Rep

Data

- ▼ Data
 - ▼ Information encoded and structured for subsequent processing
 - ▼ Contains multiple symbols
- ▼ Information
 - ▼ Data that has been interpreted
- ▼ Data Life Cycle
 - ▼ Creation ⇒ Storage ⇒ Usage ⇒ Archival ⇒ Destruction

Data Privacy

- ▼ Data Privacy
 - ▼ The requirement for data to be accessed by or disclosed to authorized persons only
- ▼ Examples for Data Privacy
 - ▼ Websites tracking shopping preferences
 - ▼ Identity theft using ID cards
 - ▼ China's facial recognition based credit score system
- ▼ Data Security
 - ▼ Ways to protect data from malicious threats

Data Integrity

▼ Data Integrity

- ▼ A process that ensures the accuracy, completeness, consistency, and validity of data

- ▼ Covers data in storage, during processing, and while in transit

▼ Compromised Integrity

- ▼ When data is replicated or transferred, it should be intact and unaltered between updates.

- ▼ Error-checking methods and validation methods are used to ensure data integrity

Storage Threats

▼ Behavioral Lapses

- ▼ Human stupidity, users using weak passwords or being careless when dealing with data

- ▼ \$227.8m lost to top 10 scams in first half of 2022

▼ Mismanagement by Multiple Users

- ▼ Multiple users working on the same file may overwrite each others' data

▼ Natural Disasters

- ▼ Natural disasters can destroy the computer system and/or storage device physically

- ▼ 50% of data centers can't withstand natural disasters without failures, Zenium Tech Partners

▼ Unauthorized intrusion into the system

- ▼ If the system is vulnerable, data is susceptible to corruption and lost

- ▼ Singhealth data breach, 1.5m users' data compromised

Data Protection

▼ Data Backup

- ▼ Creating copies of data and storing them separately
- ▼ Safety precaution for unexpected loss of data due to unintended data corruption or errors.
- ▼ A backup is typically a short-term solution and may be performed on a regular basis
- ▼ Data Archive
 - ▼ Storing data that is not actively used but kept for historical references or auditing purposes
 - ▼ It ensures that important records remain available years after they were created
 - ▼ In most cases, the purpose of data archiving is to meet legal and compliance requirements
- ▼ Data Backup VS Data Archive
 - ▼ Comparison Chart

	Backup	Archive
Storage Method	Backups are copies stored in a different location from Original	Archived Data is Original, moved to an archive storage location
Data State	Constantly Changing	Once archived, do not modify
Retention Policy	Periodically overwrite	Store for the long term
Storage Type	Hot Cloud & Local Storage	Cold Cloud & Tape Archives
Data Scope	All data	Specific Files

- ▼ Restriction of Access
 - ▼ Passwords, 2FA, Biometrics
 - ▼ IAM (Identity and Access Management) Solution
 - ▼ RBAC (Role Based Access Control) Solution
- ▼ Version Control (Git)

- ▼ Modified (Changed file, not committed)
- ▼ Staged (File marked to be committed)
- ▼ Committed (Data safely updated in DB)
- ▼ File Naming Convention
 - ▼ Consistent naming
 - ▼ <25 characters
 - ▼ Avoid spaces or special chars
 - ▼ Use capitals and underscores
 - ▼ Date Format ISO 8601 (YYYYMMDD)
 - ▼ Version Number
- ▼ Validation and Verification
 - ▼ Validation checks if data is as per expected input
 - ▼ Verification checks if data is what was meant to be entered
- ▼ Disaster Recovery
 - ▼ Set up a remote disaster recovery data site
- ▼ Data Encryption
 - ▼ Alters data content according to an algorithm that can only be reversed with the right key
 - ▼ Protects your data from unauthorized access even if data is stolen by making it unreadable
- ▼ Data Erasure
 - ▼ More secure than standard data wiping, completely overwrites data on any storage device
 - ▼ Limits liability by deleting data that is no longer needed
 - ▼ Usually a compliance regulation

PDPA

▼ What is PDPA

- ▼ Data protection law comprising various rules that govern the collection, use, disclosure, and care of personal data.
- ▼ Recognizes rights of individuals to protect, modify or access their personal data
- ▼ Recognizes rights of organizations to collect, use or disclose personal data
- ▼ Accounts for Consent, Notification, Appropriateness, Accountability

▼ Do Not Call Registry

- ▼ Registering a phone number in the DNC registry stops all marketing callers from calling
- ▼ Types are No Voice Call Register, No Text Message Register and No Fax Message Register

▼ National Registration Identification Card

- ▼ Unique identifications for Citizens and PRs, containing sensitive information like address
- ▼ Organizations are not allowed to collect, use, or disclose NRIC numbers unless required by law or necessary for identity verification
- ▼ Partial NRIC or other identifiers like mobile numbers can be used for verification instead

▼ 9 Data Obligations

- ▼ Consent Obligation → Only collect, use or disclose with CONSENT
- ▼ Purpose Limitation Obligation → Only collect what is appropriate
- ▼ Notification Obligation → Inform users before collecting, using or disclosing
- ▼ Access and Correction Obligation → If a user asks for his data, give it within a year unless it is harmful
- ▼ Accuracy Obligation → Ensure all information collected is accurate and complete

- ▼ Protection Obligation → Make sure the data is securely protected
- ▼ Retention Limitation Obligation → Reduce a user's tie to their data so it can be easily deleted if needed
- ▼ Transfer Limited Obligation → When internationally transferring data, ensure it is still legal abroad
- ▼ Accountability Obligation → Make information about data protection policies available on request