

Lecture 4: Data and Knowledge Management

Data Management

- An effective IS provides users with **accurate, timely and relevant information**
 - Accurate information is free of error
 - Info is timely when it is available to decisions makers when it is needed
 - Info is relevant when it is useful and appropriate fo the types of work and decision that require it.
- Some orgs. Don't have timely, accurate or relevant info because their data have been poorly organised and managed.
- **Difficulties of managing data**
 - Amount of data is increasing exponentially
 - Data are scattered throughout orgs. And collected by many individuals using various methods and devices.
 - Data comes from many sources
 - Data are subject to **data rot** (outdated, destroyed storage media)
 - Inconsistent, conflicting data due to non - integrated information systems.
- **Therefore MANAGING DATA is important!!!!**
- **How do orgs. Manage the availability, usability, integrity and security of company data?**
 - **Establish Data Governance:** data governance is all about to ensure data standards and polices are defined and enforced across the entire organisation.
 - **Key Data Governance Principles**
 - **Sponsorship:** Role of sponsorship is providing overall leadership and funding for DG initiatives
 - **DG council:** leadership to drive adoption and compliance to establish controls
 - **Data standards:** standards and policies are clearly defined
 - **Data Owners:** the role of data owners is to provide the execution of data governance principles.
 - **Monitoring:** Establish process to track, audit and report compliance data standards and polices
 - **Change Control:** Establish change control process to evaluate, approve and communicate changes to data standards and polices

- **Key Governance Process:**
 - **Establish** DG Org.
 - Identify DG program exec sponsor
 - Establish council
 - Engage key FG stakeholders (e.g. DG owners)
 - **Define** data standard and polices
 - **Enforce** Data Standards and Polices
 - **Audit Monitor and Control**
 - Perform regular audit and control activities
 - Monitor and measure performance on DG activities.
- **Problems with traditional file environment:**
 - **Data Redundancy:** same data are stored in many places
 - **Data Inconsistency:** various copies of the data do not agree
 - **Poor security:** there is little control or management of data, access and dissemination of info may be out of control
 - **Poor data sharing and availability:** pieces of data are in diff files and diff parts of the org cannot be related to each other

Database

- **Database is a collection of data organised to serve many applications and controlling redundant data.**
- Rather than storing data in separate files for each application, data appears to users as being stored **in only one location.**
- **Database Management System (DBMS):** is a software that permits an org to centralise data, manage them efficiently and provide access to stored data by application program.
 - **E.g. Microsoft Access,** a popular DBMS for personal computers enable users to interact with data in database, add, delete and modify data stored in the database.
- **The Data Hierarchy**



- A **bit** represents the smallest unit of data a computer can process (0 or 1)

- A **byte** represent a single character which can be can be a letter, a number, a symbol
- **Field:** Logical grouping of characters into a word, a small group of words or an identification number. E.g. a student's name in a uni computer files would appear in the "name filed and his or her date of birth would appear in the "date of birth" field.
- **Record:** A logical grouping of related fields such as the student's name, course and grade.
- **File:** A logical grouping of related records.
- **Database is a collection of related attributes about an entity**
 - An entity is something you collect data about, such as people or classes
 - Each record typically contains many attributes e.g. a name
- **Advantages of Database approach:**

| Advantages | Description |
|---|---|
| Minimal data redundancy | A single copy of data ensures that storage requirements are minimized. |
| Improved data consistency | Eliminating redundancy greatly reduces the possibilities of inconsistency. |
| Increased security | A centralized system makes it easier to enforce access restrictions. |
| Improved data quality | Centralized control, minimized redundancy, and improved data consistency help to enhance the quality of data. |
| Improved data accessibility and sharing | A centralized system makes it easier to deploy and control access for personnel within or outside organizational boundaries. |
| Enforcement of standards | A centralized system makes it much easier to enforce standards and rules for data creation, modification, naming, and deletion. |