

Information Processes and Technology

Information Systems and Databases

Information Systems

Information systems perform a set of information processes requiring participants, data/information, and information technology. They are created to provide access to information for an organisation or individual.

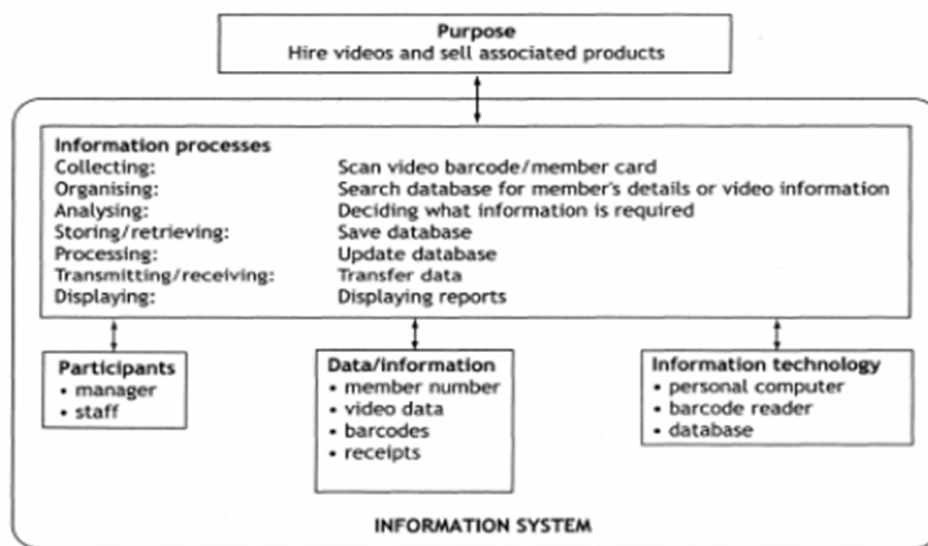
Characteristics of an information system:

- Organisation of data into meaningful information; includes data dictionaries
- Analysing information to gain knowledge; includes tables, queries and reports

Types and purposes of information systems:

- *Transaction processing systems (TPS):*
 - Collects, stores and modifies transactions of an organisation
 - Purpose: Provides data to other information systems about a transaction
 - E.g. credit card systems, ATM's and EFTPOS
- *Management Information Systems (MIS):*
 - Takes data and organises it into information reports (usually from TPS)
 - Purpose: Provides information on the performance of an organisation
- *Decision Support Systems (DSS):*
 - Takes data (usually from TPS, MIS & external sources) to assist in decision making
 - Purpose: Provides analysis tools, information and models to help decision making
- *Office Automation System (OAS):*
 - Improves efficiency and effectiveness in completing administrative tasks
 - Purpose: Manages vast quantities of data within an organisation
 - OAS tools include: word processors, DBMS software, voice mail, scheduling software, and spreadsheet software.

An ISC diagram as a systems representation tool:



E.g. Video store information system:

Database Information Systems (examples)

School databases:

- **Environment:** School community (students, staff, parents), authorities (BOS, government), any organisation, business or individual that receives information generated by the school or supplies data for the system.
- **Purpose:** To maintain the efficient operation of the school, personnel, resources and administration tasks. To provide information on student enrolments, subject selections and assessments for both school staff and educational authorities.
- **Data/Information:** Information on people and resources, library loans, staff information, student and teacher timetables, student marks and grades.
- **Participants:** Office staff, teachers and library staff who enter data
- **Information technology:**
 - **Hardware:** Computers, barcode readers, scanners
 - **Software:** Operating system, spreadsheet and word processor applications, database
- **Information processes:** Cover the seven processes; collecting, organising, analysing, storing and retrieving, processing, transmitting and receiving, displaying.

Roads and Traffic Authority (RTA):

- **Environment:** New South Wales; vehicle owners, government (NSW department of transport), police and courts.
- **Purpose:** To manage registration of all drivers and motor vehicles in NSW. To provide information to drivers/applicants on matters such as licensing, vehicle registration, etc. as well as statistical information to government and other authorities.
- **Data/information:** Drivers' details (contact, payment, driving history, license information), vehicle details (registration numbers, make of car, etc.)

- *Participants:* Data entry operators, administration staff at offices, police department
- *Information technology:*
 - Hardware: Computers, cameras, barcode scanners, networking between computers
 - Software: Operating system, database software
- *Information processes:* Cover the seven processes; collecting, organising, analysing, storing and retrieving, processing, transmitting and receiving, displaying.

Video stores:

- *Environment:* Staff, customers around the area, suppliers, other chain stores, head office.
- *Purpose:* Allowing customers to hire videos for a period of time, keeping an accurate record of rentals and stock.
- *Data/Information:* Customer details, current loans and overdue items, rental records (in stock/on loan), rental fees, video details, barcodes, receipts
- *Participants:* Staff, head management
- *Information technology:*
 - Hardware: Computers, barcode scanners
 - Software: Operating system, database software
- *Information processes:* Cover the seven processes; collecting, organising, analysing, storing and retrieving, processing, transmitting and receiving, displaying

Organisation

Non-computer methods of organising:

Examples:

- Telephone books
- Card based applications; e.g. old libraries
- Filing cabinet; e.g. small businesses
- Reference books; e.g. encyclopaedias, dictionaries

Advantages:

- It is cheaper as it doesn't require computer hardware and software
- Highly portable
- No specific skills or training required

Disadvantages:

- Requires larger amounts of physical storage space
- One person only can access it at one time
- Takes longer to organise, analyse, manipulate and retrieve the data

Computer-based methods of organising:

Examples:

- Flat file systems
- Database management systems (DBMS)
- Hypermedia

Advantages:

- Data is easily exchanged between applications and over networks
- Much faster processing
- Can be easily edited
- Large storage of data is possible
- Can be accessed by many at one time if required

Disadvantages:

- Expensive hardware and software have to be purchased
- Vulnerable to viruses and hackers
- Staff need to be trained to use it

Logical organisation of flat file databases:

Flat file database:

- Single table of data stored as a single file
- All rows (records) are composed of the same sequence of fields (attributes)

Files: Block of data which is divided into a set of related records

Records: Collection of facts about one specific entry in a database; organised in rows

Fields: Specific category of data in a database; organised in columns

Key fields (primary key): A character or group of characters used to uniquely identify a record in a database

Characters: Single unit of data used in a database; such as a letter or number

Logical organisation of relational databases:

Relational databases:

- Organises data using a series of related tables, linked by common key fields from one or more additional tables/files

Schemas: An organised plan of the entire database showing data relationships

- Entities: A specific title about which an information system collects information; i.e. a category or grouping such as Customers, Orders, Rentals, Videos, etc. (table name)
- Attributes: Defined property of an entity; e.g. CustomerID, FirstName, etc. (field name)

Relationships:

- One-to-one: One record in a table relates to one record in another table. Another determining factor is both fields in the relationship are key fields.
- One-to-many: One record in a table is related to many records in another table.
- Many-to-many: Relating many records in a first table with many records in a second table and vice-versa.

Tables consist of:

- Attributes (fields - columns)
- Records (rows)

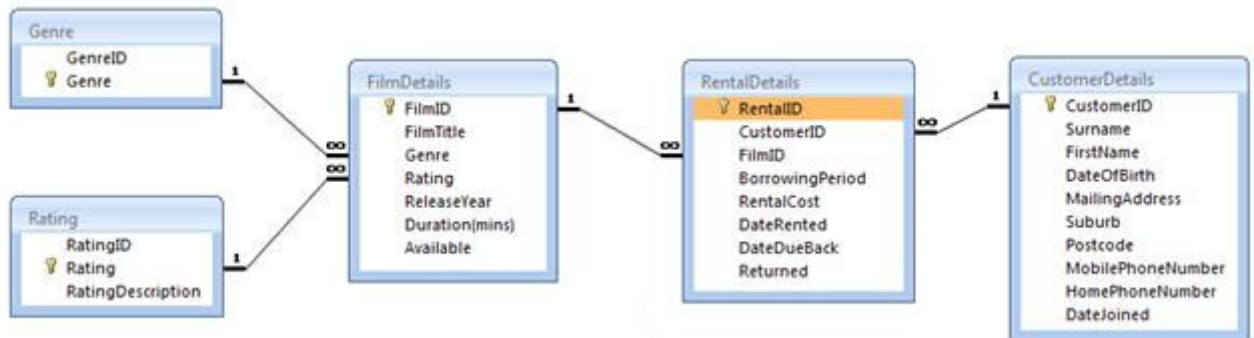
Linking tables using:

- Primary key: uniquely identifies a record
- Foreign key: Attribute (field) of a table that is a primary key of another table

Schematic diagram:

A schema shows the organisational structure of a database. It should show the entities (the tables in the database), and their characteristics (i.e. their properties/attributes/fields). It should clearly identify the primary key in each table and the links and relationships between tables.

E.g. A video store:



Data modelling tools for organising databases:

Data dictionaries: Describes characteristics of data (metadata)

- Field name, data type, data format, field size, description, example

Schematic diagrams:

- Shows the relationships between entities in a database (as above)

Normalisation:

- Reduces data redundancy by removing repeating fields or reorganising data where needed, hence improving performance

Logical organisation of hypermedia:

Hypermedia: Hypermedia is an extension to hypertext that supports linking graphics, sound and video elements in addition to text elements.

Nodes and links:

- A node is a computer that has the destination of the link

Uniform resource locator (URL):

- Address of computer where file/resource is stored on the internet
- E.g. <http://www.boardofstudies.nsw.edu.au/index.htm>
 - Protocol: allows access to web pages based on hypertext – ‘http’
 - Domain name: address of specific computer where resource is located – ‘www.boardofstudies.nsw.edu.au’
 - File path: path followed to the file being retrieved – ‘index.htm’

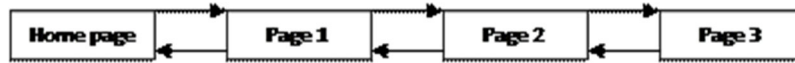
HTML tags (hypertext markup language):

- Metadata – data that describes other data
- Indicates how parts of a document are displayed and navigated
- Viewed as pages in web browsers

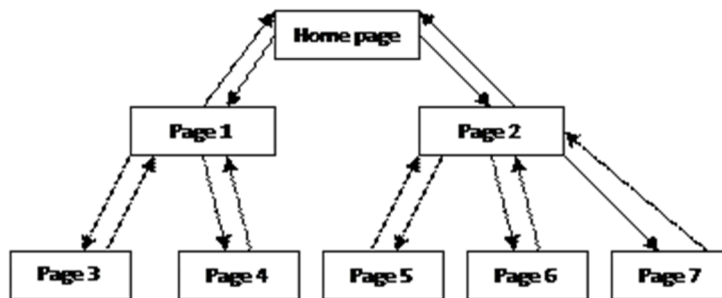
Tools for organising hypermedia:

Storyboards:

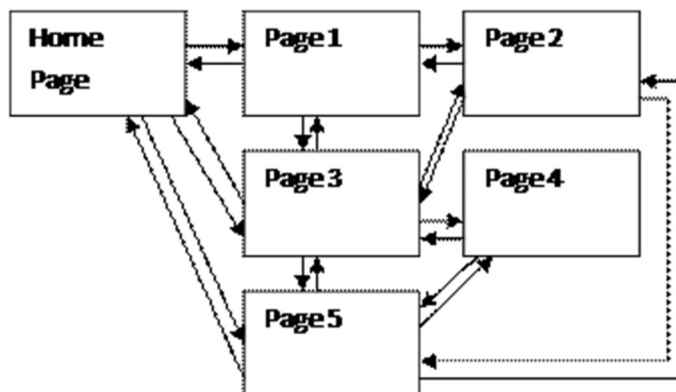
- Series of frames, each representing a different action or screen image
- Consists of navigation paths, information and graphics
- Linear: All pages are linked in a simple line



- *Advantage – very simple to create and maintain the data is presented in a set order.*
 - *Disadvantage – searching is slow, you cannot go directly to a particular page.*
- Hierarchical: choices branch off into further choices, based on navigation



- *Advantage – locating a particular page takes fewer links and is faster.*
 - *Disadvantage – more complicated to create and maintain*
- Non-linear: no structure - free navigation



- *Advantage – the user has more freedom to explore.*
- *Disadvantage – very difficult to create and maintain.*

- Composite: mixture of the above

Storage and Retrieval

Database management systems (DBMS):

- Software application that allows the entry, manipulation & storage of data in a database
- Handles access to the database- establishes & maintains data security by restricting access to authorised personnel
- Carries out validation & verification of data
- Independence of data from the DBMS
- Functions include; sorting, searching, storing, creating, editing and accessing data.

Data access methods:

Sequential:

- Each record must be accessed in a linear progression; i.e. from first to last
- E.g. accessing data on a magnetic tape

Direct:

- Data is accessed in any order; i.e. without accessing previous data items
- E.g. storing data on disc such as DVD or CD, hard drive.

On-line and off-line storage:

On-line:

- Storage device permanently positioned
- Available immediately to connected computers
- E.g. hard drive

Off-line:

- External devices used to store data
- Cannot be accessed until storage media is mounted into a drive
- E.g. USB, CD, DVD, memory cards, magnetic tape

Centralised and distributed databases:

Centralised:

- Accessed through a single DBMS server; a single database
- All users connect directly to DBMS

Distributed:

- Set of connected databases stored on multiple computers
- Appears as one large database to users

Storage media:

Hard discs:

- Stores data magnetically on precision aluminium or glass platters; Direct access

CD-ROMs:

- Data is read and written using laser technology; Direct access

Cartridges and tapes:

- Magnetic tape encased in a cartridge
- Stores large amounts of data inexpensively, erasable, reusable.
- Used for backup
- Sequential access

Calculation of storage

Images:

- Divide by 8 for bytes. Divide by 1024 for kilobytes, then 1024 again for Megabytes
- E.g. the size of an image with 1024 x 768 resolutions at 24 bit depth.
- $1024 \times 768 \times 24 = 18874368$ bits

Audio:

- Sample rate x sample size x number of channels
- E.g. for 1 minute stereo song with a sampling rate of 44.1khz with a sample size of 20bits
- $44100 \times 60 \times 2 \times 20 = 105840000$ bits

Video:

- Frame rate x resolution x bit depth x running time (seconds)
- Therefore the size of a 1 hour movie at 25fps, bit depth of 32 and with a resolution of 1024x 768:
- $25 \times 1024 \times 768 \times 32 \times (60 \times 60) =$ answer

Database:

- Total field size x amount of records

Encryption and decryption:

Encryption:

- Process of encoding data; maintains confidentiality and security
- Algorithm or key is required to encode the data
- Involves manipulations of bit patterns
- In asymmetric encryption, two encryption keys are set up. One key is called a public key and is available on a wide scale. The other key is called private key and is restricted to those authorised to have the key. The authorised user of the data must have both keys to decode the data.
- Symmetric encryption uses one key for both encryption and decryption. This key is possessed by both the sender and the receiver of the data and is only accessible to those authorised to use the data.

Decryption:

- Process of decoding data; receiver can translate to its original state
- Reverse algorithm or key is required to decode the data

Backup and security procedures:

Backup:

- Storing/copying data to another permanent storage device
- Full backup: All the data in the system is backed up
Advantages: data is quickly recovered, all data is saved.
Disadvantages: backup time is longer than other methods.
- Differential: Only data changed since the last FULL back up is saved

Advantages: backup time is rapid, minimal space required

Disadvantages: recovery requires a combination of full and one differential back up.

- Incremental: Only data changed since the last back up (full/differential/incremental) is saved

Advantages: fastest back up time, minimal space required

Disadvantages: recovery takes time as it requires a combination of full back up and any number of incremental back up since the last full back up.

Security:

- Physical security measures - locking rooms/buildings
- Usernames and passwords
- Encryption and decryption
- Restricted access
- Record/field locking
- Firewall, antivirus, etc.

Tools for database storage and retrieval:

Searching and sorting:

- Carried out on a field alphabetically, numerically or chronologically; relevant information is extracted

Query By Example (QBE):

- Operator enters criteria against a field
- Records meeting specific conditions are displayed

Structured Query Language (SQL):

- Performs complex searches of the database.
- Commands include:
 - SELECT (field to be displayed)
 - FROM (table where the field is to come from)
 - WHERE (the search criteria)
 - ORDER BY (order in which the results are sorted)
- Relational operators: =, <>, <, >, <=, >=
- Logical operators: AND, OR, NOT
- E.g.

```
SELECT *
FROM Persons
WHERE Persons.FirstName='Simon' OR Persons.FirstName='Daisy'
ORDER BY LastName DESC;
```

Tools for hypermedia search and retrieval:

Free text searching:

- Technique for searching a computer based document or database for characters or words
- Search engine searches all words and tries to match search words supplied by the user

Search engines:

- Indexing - databases of indexed websites that can be searched using keywords
- Search robots - programs that facilitate indexing by accessing websites and gathering information
- Metadata – search engines look for metadata (keywords, titles, etc.)

Reporting on data found in hypermedia systems:

- Web browsers
- Stand-alone applications
- Media players

Other Information Processes for Databases

Displaying:

Reports:

- Formatted and organised presentation of information
- E.g. mailing labels, invoices, sales summaries and telephone lists

Constructing different views of a database for different purposes:

- Forms - used to enter, view and edit data easily

Issues Related to Information Systems and Databases

Acknowledgement of data sources:

- Requires permission from source to use data before publication
- Legal requirements - copyright
- Ensures justification of outputs; e.g. results from surveys
- Provides a mechanism for tracking and checking data; determines accuracy of data

Access to data, ownership and control of data:

- Limited access to authorised personnel only; control over data
- Prevents unauthorised access, alteration/destruction of data or programs, unauthorised use of resources and release of information

Freedom of Information Act:

- The rights of individuals to privacy and control of information concerning their personal details need to be considered. However the existence of a store of information concerning individuals can, in some cases, be beneficial to society in general.
 - Unwanted names on mailing list
 - The sale of mailing lists
 - Loss of control of information

- Surveillance information held on computers controlled by government agencies
- Tracing missing persons
- Information on individuals needs to be accurate

Privacy principles:

- Database systems usually have security measures in place to prevent both unauthorised access to the system and corruption of the data stored. This introduces the concern about security.
- Data may get into the wrong hands; may become damaged, deleted accidentally or intentionally, be taken or distributed
 - Restrictions on the level of access – password protection, biometric devices.
 - Data encryption, firewall, virus protection
 - Data backups

Accuracy of data and the reliability of data sources:

- Data accuracy is a major issue with data retrieved from the internet
- Can minimise the issue by checking the source of the data and cross-checking the data against a known reputable source.
- Reliable data sources commonly have a domain name with .edu or .gov, whereas most unreliable data sources would have .com

Quality of data

- Data quality/integrity can be maintained by ensuring that the database is able to resist user mistakes, system malfunctions, deliberate or accidental alteration.
- Data validation including; entering data within a range and with a certain format, radio buttons and drop down lists to decrease user input errors.
- Reliable data should be up-to-date

Current and emerging trends in the organisation, processing, storage and retrieval of data:

Data warehousing:

- Electronic storage centre of data from a variety of sources for future analysis
- Data may be available for sale to interested parties; privacy issue

Data mining:

- Sifting through data to detect patterns; e.g. analysing shopping habits of individuals to identify frequently purchased products to target advertising; can lead to spamming issue. Issues of ownership of data and privacy are also a concern.

Online analytical processing (OLAP):

- Quickly answer multi-dimensional analytical queries
- Provides statistical evidence to corporations; assists in decision making

Online transaction processing (OLTP):

Facilitate and manage transaction applications typically for data entry and retrieval e.g. bank AT