ST JOHN'S RC HIGH SCHOOL

COMPUTING SCIENCE DEPARTMENT

HIGHER COMPUTING SCIENCE

INFORMATION SYSTEM DESIGN & DEVELOPMENT

CONTENTS:

DATABASE STRUCTURE

What is a database?

A database is an organised and structured collection of related data which you can search through. Databases can be:

- Stored manually (in a filing cabinet or on index cards).
- Storage could be in book form, (eg a telephone directory or dictionary).
- Stored electronically using computer hardware and database software.
- Stored electronically using computer hardware and web authoring software.

Why do we need to structure and organise data?

It is necessary for any large amounts of data to be stored in some kind of order, **data is generally meaningless on its own. Once data has been organised and structured it turns into information which we need to use for every aspect of our lives.**

Once information is stored in an electronic database (including web pages), the information can be accessed easily and quickly – allowing updates to be done quickly so that the information is always accurate and up-to-date. Electronic databases also allow for **single entry – multiple use**, ie the data is entered by one person but can be accessed by many different people, which helps to minimise errors in the data being held.

You probably use a filing system at home, eg:

- Store your music CD's or Computer Games in a storage rack.
- Keep your books on shelves.

Databases stored electronically using a computer system, can be:

- Flat
- Relational
- Web-based

FLAT FILE DATBASES

A **flat file database** has only **one table and holds all the records.**

Advantages of flat file databases include:

- All the records are stored in one place.
- Easy to sort and search.
- No need for database software to be used – the data could alternatively be stored using spreadsheet software.

Disadvantages of flat file databases include:

- **Duplication** – with the potential for thousands of records to be stored in the one place it can be difficult to spot duplicated records.
- **Inefficient –** you may have records which many of the fields have no data entered.
- **Poor security –** you may have to limit access to the database, because some of the fields contain information which should be kept confidential.

RELATIONAL DATABASES

A more efficient way of storing data is to use a **relational database**.

A **relational database** is one in which all data is stored in **relations** which (to the user), are **tables with rows and columns**. Each **table (entity)**, is composed of **records** (called Tuples) and each **record is identified by a field (attribute)** containing a value. **Every table (entity), shares at least one field (attribute), with another table in 'one to one (1:1),' 'one to many (1:M),' or 'many to many (M:M)' relationships.**

These relationships allow the database user to access the data in almost an unlimited number of ways, and to combine the tables as building blocks to create complex and very large databases.

The **advantages of using a relational database** include:

- Minimises data duplication.
- Easier to maintain security.
- Records can be added, removed or updated easily.
- The data format, ie the adding, removing or data type of fields can be changed easily.

WRITTEN TASK

1 **Give two advantages** and **two disadvantages** of using a flat file database.

2 In a relational database **what** is an:

(a) An entity.
(b) An attribute.

3 **List** the **three types of relations** which are available in a relational database.

4 **Suggest two advantages** of using a relational database *rather* than a flat file database.

CREATING A DATABASE

To create a database **table (entity),** you must first design it - this mean that you must decide:

- What it is you want to store **– Records**
- What information you want in the records – **Fields (Attributes)**
- What information will be unique to each record **(Primary Key)**
- What the **format of the fields** will be (eg text, number, alphanumeric, date etc)

TABLE (Entity) – a collection (or list) of related information held in columns and rows.

Each **ROW** in the table is called a **RECORD**

(Example: your information in a telephone book is a **RECORD**)

Each **column** in the table is a category or **FIELD (Attribute)**

(Example: a column of phone numbers in a telephone book would be considered as the Phone Number **FIELD**)

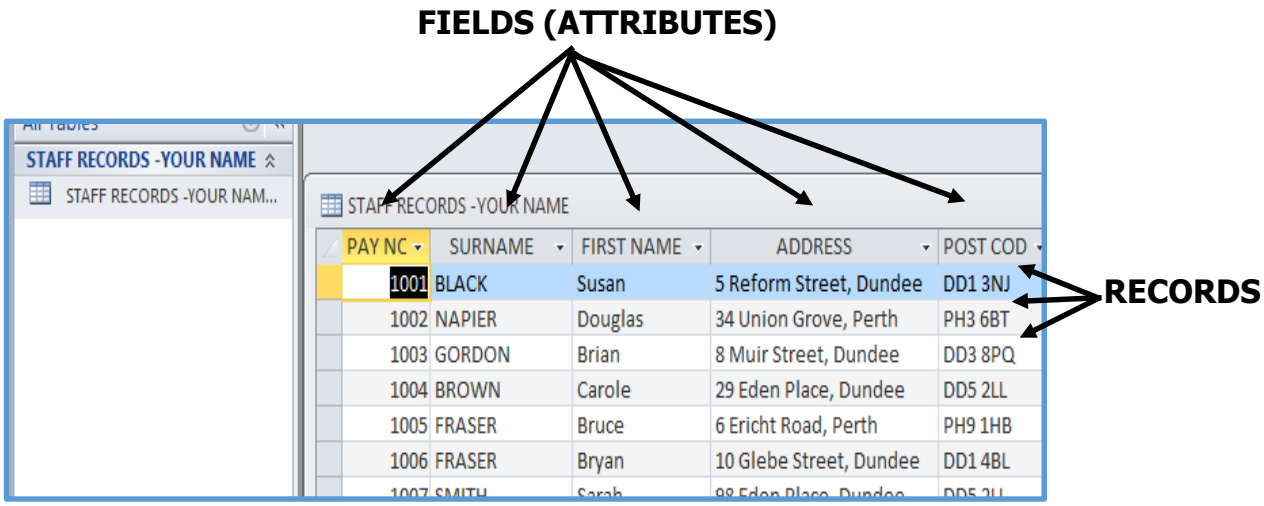
Each piece of data is a **DATA VALUE**

(Example: your phone number in a telephone book is a **DATA VALUE**)

WHAT DOES A DATABASE TABLE (ENTITY) LOOK LIKE?

A database has columns which are called **fields (attributes)** and rows which are called **records.** Individual pieces of data or information are stored in **fields** and all the data or information relating to someone or something is stored in a **record.**

The screenshot below shows a **DATABASE TABLE (ENTITY)** in **list view** (all the records showing):

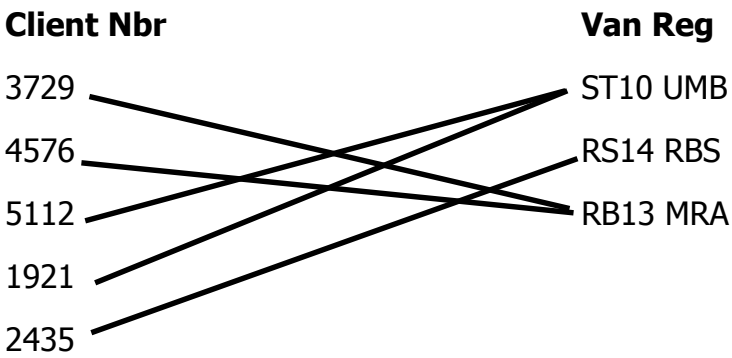


ENTITY-RELATIONSHIP OCCURRENCE MODELLING

Once the decision has been made for the need for a **relational database management system (RDBMS)** to be created, the **system analyst** examines any entities and relationships that exist in the current system.

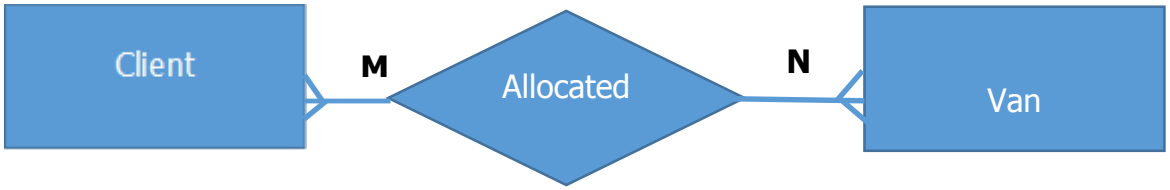
The **first step** towards creating an **entity relationship diagram** is identifying the relationships between **two entities (tables).**

The **second step** is to identify a primary key for each **entity** (table). **The primary key for each entity (table) should be unique).** Once the **primary keys** have been identified a list of values for each key should be completed, by creating an **entity occurrence diagram**, for example:



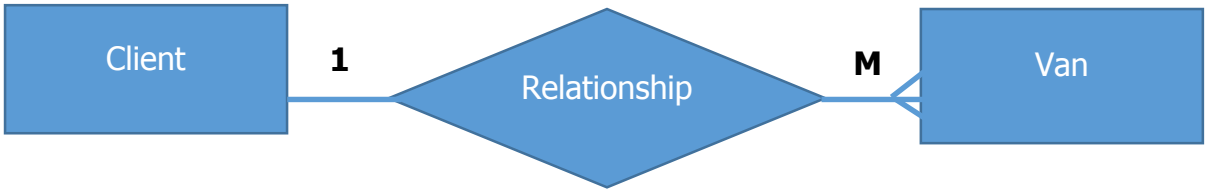
Each line indicated a link between Client Nbr entity set and the Van Reg entity set.

Once all the entities and relationships have been identified an **entity relationship diagram** is used to illustrate the logical structure of the database. For example:



The **entity relationship diagram** shown above, shows that **many** clients can be allocated **many** vans.

The other **ERD connectors** possible are:

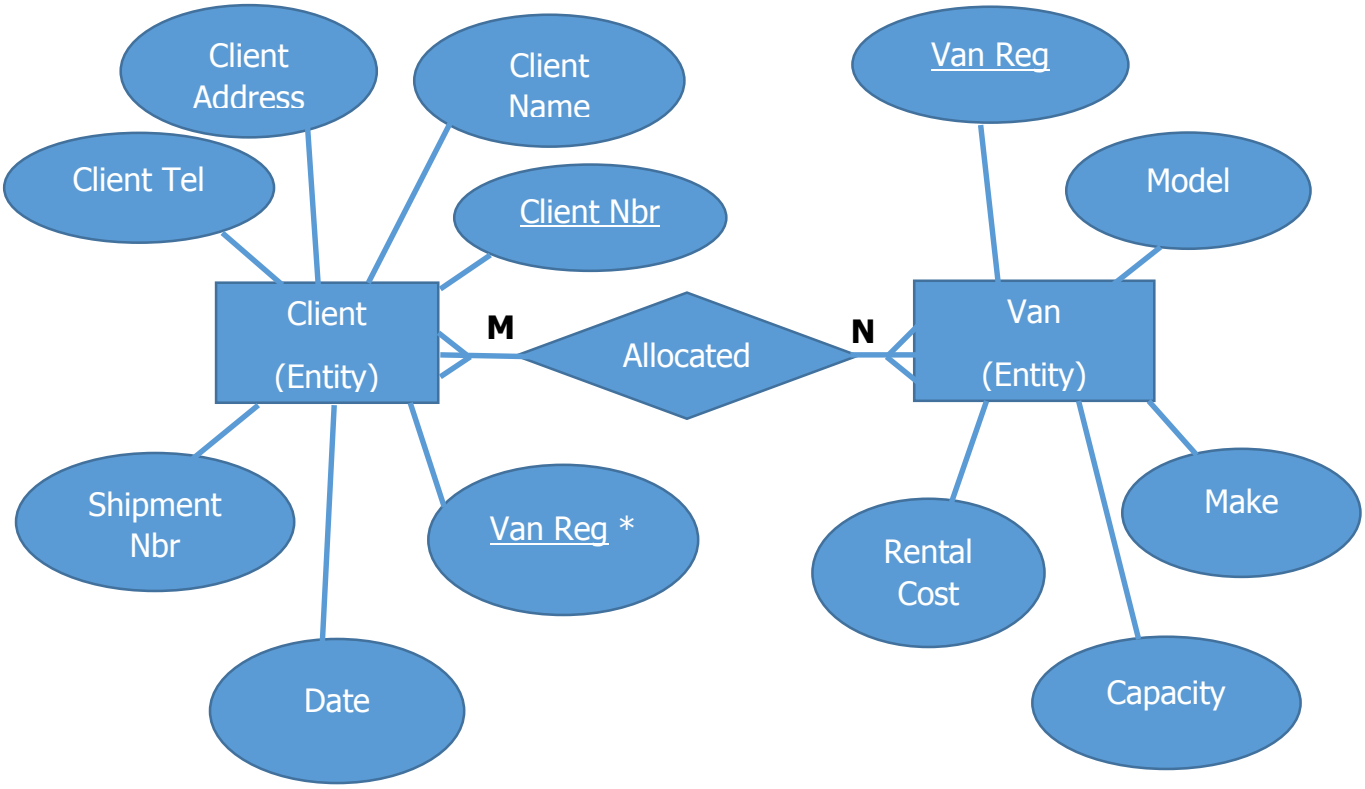


The **entity relationship diagram** shown above, shows that **one** client can be allocated **many** vans.

The **entity relationship diagram shown above**, shows that **one** client can be allocated **one** van.

ADDING ATTRIBUTES (FIELDS)

Attributes (fields) can be added to each of the **entities (tables)** in **the entity relationship diagram (ERD)** to complete the entity model. An example of a completed **ERD** – including attributes, is shown below:



NOTE: Primary Key in an entity is denoted by being **underlined**

Foreign Key in an entity is denoted by being **underlined*** with a star

NORMALISATION

In relational database design, the process of organizing data to minimize **data duplication** (redundancy) is called **normalization. Normalisation** usually involves dividing a database into two or more **tables (entities)** and defining relationships between the tables. The objective is to isolate data so that additions, deletions, and modifications of any **field (attribute),** can be made in just one **table (entity),** and then propagated through the rest of the database via the defined relationships.

There are three main normal forms, each with increasing levels of normalization:

First Normal Form (1NF): Each **field (attribute)** in a **table (entity),** contains different information. For example, in an employee list, each table would contain only one birthdate field.

Second Normal Form (2NF): Each **field (attribute)** in a **table (entity)** that is not a determiner of the contents of another field must itself be a function of the other fields in the table.

Third Normal Form (3NF): No duplicate information is permitted. So, for example, if two **tables (entities**), both require a birthdate **field (attribute),** the birthdate information would be separated into a separate table and the two other tables would then access the birthdate information via an index field in the birthdate table. Any change to a birthdate would automatically be reflect in all tables that link to the birthdate table.

An example of normalization to 3NF is shown below:

List all attributes (fields) in UNF(Normalised Form):

client_nbr
client_name
client_address
client_telephone
client_fax
contact_person
shipment_nbr
insurance_ref
from
to
date
time
van_reg
van_reg
model
make
capacity
fuel_type
rental_cost

UNF – Remove duplicate items:

client_nbr
client_name
client_address
client_telephone
contact_person
shipment_nbr
insurance_ref
from
to
date
time
~~**van_reg**~~
model
make
capacity
fuel_type
rental_cost

client_nbr
client_name
client_address
client_telephone
contact_person
shipment_nbr
insurance
from
to
date
time
van_reg
model
make
capacity
fuel_type
rental_cost

To transform UNF into 1NF we have to:

Choose a **start key** from a **field (attribute) which contains unique data**. **Indent repeating data** to show the **repeating group for the primary key selected. Complete UNF give the relation a name and enclose it in brackets:**

client (**client_nbr**
client_name
client_address
client_telephone
contact_person
shipment_nbr
insurance_ref
from
to
date
time
van_reg
model
make
capacity
fuel_type
rental_cost)

Making **1NF by removing and naming repeating group** and creating the **relationship** by adding the **foreign key:**

client (**client_nbr**
client_name
client_address
client_telephone
client_fax
contact_person)

removal (**shipment_nbr**
insurance_ref
from
to
date
time
van_reg
model
make
capacity
fuel_type
rental_cost
client_nbr*)

There are still some problems associated with 1NF:

- Insertion Anomalies
- Modification Anomalies
- Deletion Anomalies
- **Data redundancy** - which refers to the amount of data in a database that is not required. The aim of good relational database design is to **reduce data redundancy by:**
  - Ensuring that each item of data is stored only once within the database and
  - Ensuring that only data that cannot be calculated from other data held in the database is stored.

Making **2NF by removing partial dependencies – ie, non-key attributes dependent on part of the key:**

client (**client_nbr**
client_name
client_address
client_telephone
client_fax

contact_person)

shipment (**shipment_nbr**
insurance_ref)

removal (*<u>shipment_nbr</u>
from
to
date
time
van_reg
model
make
capacity
fuel_type
rental_cost
***<u>client_nbr</u>**)

There are still some problems associated with **2NF** as there is still **redundant data in the relations/tables and possible insertion, deletion and modification problems.**

We can solve these problems by taking the design of our database to **3NF**:

2NF

client (**client_nbr**
client_name
client_address
client_telephone
client_fax
contact_person)

shipment (**shipment_nbr**
insurance_ref)

removal (***shipment_nbr**
from
to
date
time
van_reg
model
make
capacity
fuel_type
rental_cost
***client_nbr**)

3NF

client (**client_nbr**
client_name
client_address
client_telephone
client_fax
contact_person)

shipment (**shipment_nbr**
insurance_ref)

removal (***shipment_nbr**
from
to
date
time
***van_reg**
***client_nbr**)

van (**van_reg**
model
make
capacity
fuel_type
rental_cost)

WRITTEN TASK

1 In a relational database **what** is an:
(a) An entity.
(b) An attribute.

2 **List** the **three types of relations** which are available in a relational database..

3 **What** should the contents of a primary key always be?

4 **Explain** the difference between a primary key and a foreign key.

5 **Why** do we *normalise* data when designing a relational database?

6 **Suggest two** ways in which we can minimise data redundancy.

7 **Explain** why storing the address as a *single attribute* is not good database design.

DATA DICTIONARY

A **data dictionary** is a tool used by a **systems analyst** to document a **relational database management system (RDBMS).** It contains **data about data** because it describes the domain constraints that apply to each **entity (table)** and **attribute (field).** Because it contains data about data, a **data dictionary is often called metadata.**

When the system analyst creates a **data dictionary** he is really creating the design of the database because of the description of the **entities (tables)** and their **attributes (fields)** that are in the relational database system. Each row in a data dictionary shows the constraints for each **entity (field)** within the database system and each column holds specific data relating to a particular domain.

A data dictionary contains the following columns:

Table (Entity): When a database is created using a RDBMS the **entity set name is the name of the table.**

Field (Attribute): This column in the data dictionary is where the names given to the **fields (attributes)** in each **table (entity)** are listed. It is important that the names given to **fields (attributes) are meaningful.**

PK/FK (Primary Key/Foreign Key): This column is used to identify **the primary and foreign keys** in each table (entity).

Date Type/Size: This column is used to record the **data type** to be held in each **field (attribute),** for examp**le – text, integer, real, object, boolean, date etc. This column will also give the** size which is the maximum number of characters permitted to be entered in that field (attribute), for example a **field (attribute)** with a text data type, to be used to store a first name might have a maximum size

of 20 characters. The maximum size for any **field (attribute)** must be specified to ensure storage space can be reserved in computer memory.

Unique: this column allows the system analyst to declare whether the data entered into this **field (attribute)** will be unique or note and is usually represented in the data dictionary by **Y or N**.

Required: This column allows the system analyst to declare whether or not this **field (attribute)** can be left blank in any record. Again this is usually represented in the data dictionary by **Y or N.** This is an example of a **presence check** (validation).

Validation: Each field (attribute) is constrained by a **range of acceptable values**. These constraints are one of four types:

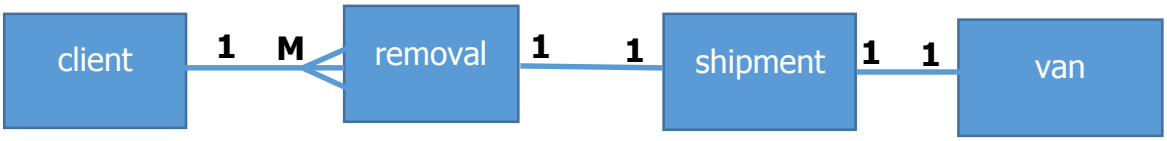
- **Presence** (are null values allowed or not)?
- **Range** (above, below or within certain values).
- **Permitted values** (value must be chosen from a selection of allowed values).
- **Format** (the data entered must conform to a specific pattern and/or number
- of characters).

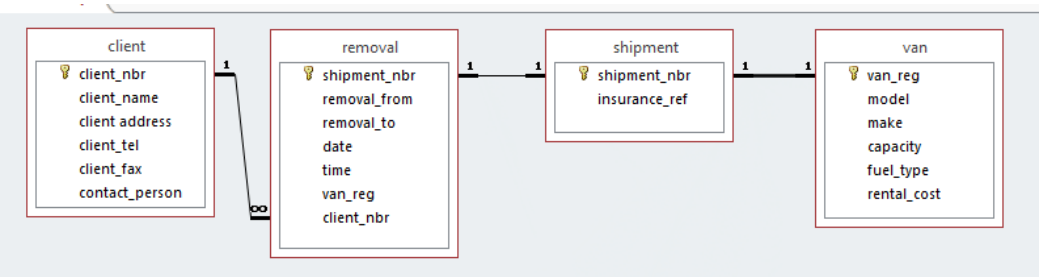
The Data Dictionary for the **tables (entities)** in the **Removal** example used above to show **normalisation**, would look like this:

Table	Field Name	PK/ FK	Data Type/ Size	Unique	Required	Validation	Format
client	client_nbr	PK	Text (10)	Y	Y		C1
	client_name		Text (20)	N	Y		
	client_address		Text (40)	N	Y		
	client_tel		Text (15)	N	Y		
	client_fax		Text (15)	N	N		
	contact_person		Text (20)	N	Y		
shipment	shipment_nbr	PK	Text (10)	Y	Y		S1
	insurance_ref		Text (10)	Y	N		
removal	*shipment_nbr	FK	Text (10)	Y	Y	Lookup from shipment table	S1
	removal_from		Text (40)	N	Y		
	removal_to		Text (40)	N	Y		
	date		Date/Time	N	Y		Short Date
	time		Date/time	N	Y		Medium Time

	*van_reg	FK	Text (10)	Y	Y	Lookup from van table	
	*client_nbr	FK	Text (10)	Y	Y	Lookup from client table	
van	van_reg	PK	Text (10)	Y	Y		
	model		Text (15)	N	Y		
	make		Text (15)	N	Y		
	capacity		Text (15)	N	Y		
	fuel_type		Text (10)	N	Y	Restricted choice of Petrol or Diesel	
	rental_cost		Number (currency)	N	Y	>=50 and <=200	

The **entity relationship diagram** of the relationships between the four tables (entities) would look like this:





In the **relationship diagrams** shown above, **the relationships are as follows:**

1 client can have **many** removals

1 removal can have **1** shipment

1 shipment can have **1** van

KEY FIELDS IN DATABASES

Linking database tables (entities), saves duplication of data and input time.
Linking tables also helps to reduce input errors as, if data has only to be entered once, there is less chance of mistakes.

Tables are linked using **Primary and Foreign Keys.**

A **primary key** is assigned to a **field which is used to uniquely identify a particular piece of information on the database** so that you can find a particular record in a file as quickly as possible, we also call this a **Key Field**.

The data held in the **primary key field in a table where is unique**. For example your:

- SQA (Scottish Qualifications Authority) Registration Number.
- National Insurance Number
- Network User ID

In the example given above the Scottish Qualifications Authority (SQA) will have a database of all the students who are sitting National or Higher exams. The **primary key field in the database would be the candidate number as each number can only belong to one person - therefore it is unique to them.**

A **foreign key** is field in one table, **which is the primary key in another and is used to link the tables together.**

A database can have many tables, linked together using primary and foreign keys, to avoid duplication of data.

Linking tables also helps to **reduce input errors** as, if data has only to be entered once, there is less chance of mistakes.

COMPOUND KEY

Sometimes, when we look at the data we want to store in a database, there is no obvious field to use as the primary key. When this happens, a **compound key** can **be created from more than one field (attribute**) to uniquely identify an **entity** occurrence.

DATA INTEGRITY

An added advantage of using a relational database is that it enforces **data integrity** – it forces records to be consistent. It does this in two ways:

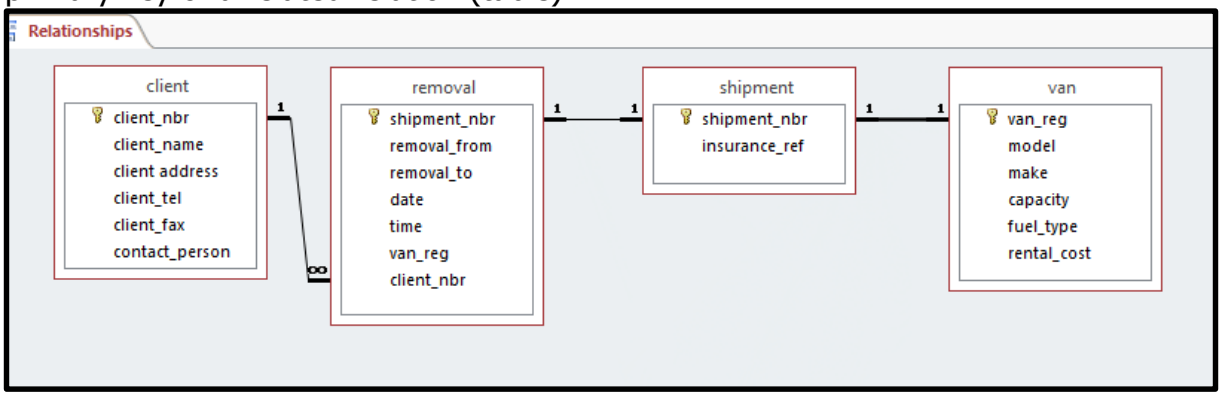
ENTITY INTEGRITY

Entity integrity relates to primary keys. This rule states that every relation must have a primary key and the column or columns selected for the primary key should be unique and not null.

REFERENTIAL INTEGRITY

Referential integrity enforces the rule that the contents of the field which has been set as the **primary key** cannot be duplicated in a table. It also ensures that whenever the contents of a primary key are changed, it will automatically update the contents when this key is used as a **foreign key** in other tables.

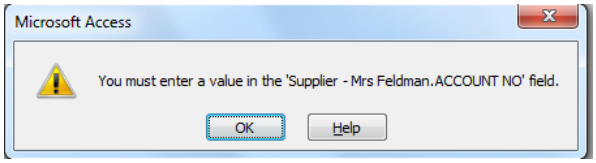
The **referential integrity** rule states that foreign key should be linked to the primary key of a related relation (table).



VALIDATION

Data is validated to check that it is **sensible and within allowable limits**. For example, validation checks would not allow:

- A numeric field to contain text.
- A required field to be left blank.

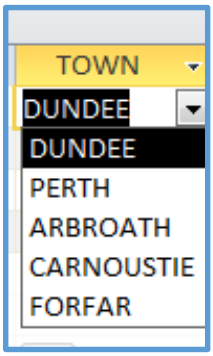


OR

Some other validation checks include:

Range Checks – data must be within a certain range, eg a date cannot have more than 31 days or 12 months in it.

Restricted Choice – data must be selected from a drop-down menu.



Presence Check - this field in any record cannot be left blank, some data must be entered. (Primary Keys will be validated by a presence check).

VERIFICATION

It is essential that the information you hold in a database **is accurate.**

To verify data, is to check that it has been entered correctly. Sometimes this is done by the application asking you to enter the data twice, other times the user is asked if what has been typed in is correct.

For Example, when you set up or change a password, you are usually asked to **enter it twice to verify it.**

ADDING AND ALTERING RECORDS IN A DATABASE

Once you have created the basic record structure (by creating the fields) you can begin adding information to each record.

Most packages allow you to alter the record format by:

- adding, deleting or changing the name of a field
- changing field lengths or field types (eg a number
- field may be changed to a text field
- by altering the way the screen looks
- by rearranging the position of the fields
- by using only selected fields in the printout

WRITTEN TASK

1 **Explain** the purpose of the **data dictionary.**

2 **Draw** a diagram of the **three relationships** which can be represented using an **ERD (Entity Relationship Diagram).**

3 **Why** might you choose to create a **compound key**?

4 **Explain** what **data integrity** is and the **two** ways in which this can be achieved.

5 **What** is **metadata?**

6 **Explain** the difference between **data** and **information.**

6 **Suggest** why **validation checks** are needed and give **three** types of validation checks and give an example for each one given.

7 **Explain** what **verifying data** means and give **one** suggestion as to how data might be verified.

SEARCHING A DATABASE

A database can be **queried (searched)** for information in two ways:

A **simple search** – which is searching on one field or a **complex search** searching on more than one field.

The **advantage of using complex searches** is that it narrows down the information you receive, making it easier to find the precise information you are looking for.

Linking Search Options: You can link the conditions of a search in a field - by using **AND** or **OR** - at the end of the search and all the records which match the set of conditions (criteria), you have selected will be displayed, eg if we searched a database for pupils whose "subject" is computing the result could be 3 pupils - Hamilton, Easton and Adie - but if we add the condition that the "grade" is 1 - this time the result is Adie.

	SURNAME	FIRST NAME	CLASS	SUBJECT	GRADE
2	ADIE	Jean	3S2	Computing	1
3	NAPIER	Alan	3S5	History	3
5	DUNCAN	Tom	3S7	English	1
6	EASTON	Robert	3S5	Computing	2
8	WILSON	Judith	3S4	Art	1
10	ROBERTS	Margaret	3S2	English	2
12	GORDON	Roy	3S1	Modern Studie	3
13	HAMILTON	Graham	3S1	Computing	3

KEYWORDS

A **keyword** is a text or code that is stored in a key field and is used to conduct searches and sorts. Some systems use the term string instead of keyword. For example each record in a file containing data about bank accounts would have a field named Account Number - this would contain a string which is unique to that record such as 714568.

By choosing your keywords carefully you can save a lot of time when searching databases.

SORTING A DATABASE

A sort allows the user to put records into order according to one of the fields. This allows you to arrange the records in a database file in alphabetic or numeric and ascending or descending order. Ascending numeric order would be: 1,2,3,4…, while descending alphabetic order would be, Z,Y,X,W…

To start **a simple sort**, you must choose a field on which to **sort** the database, or the order will stay in the order you typed them in.

COMPLEX SORTING A DATABASE

A **complex sort** allows the user to put records into order **according to more than one of the fields.** This allows you to arrange the records in a database file in alphabetic or numeric and ascending or descending order. Ascending numeric order would be: 1,2,3,4…, while descending alphabetic order would be, Z,Y,X,W…

For Example: You might put a staff database in alphabetical order of Surname and then by Initial – both alphabetically ascending and then by Salary Numerically descending – the result of this sort is that staff would be listed firstly alphabetically and then by their salary from highest to lowest.

REPORTS

A formatted and organized presentation of data is a **report**. When you **select information from a database**, eg by searching and sorting to find the records you want and getting them into the correct order, you are setting up a **report definition.**

You could of course print out the whole database - but databases can be very large and you will seldom want to see all the information held there.

FORMS

A **form** in database is an interface (window), which allows a user to have access to the data stored. It makes it possible for manipulations such as data entry, modification and even retrieval**. Forms** have direct access to table used to store information. As with reports, **forms** can be set up which will only show selected fields in the database.

COMPUTED FIELDS

Databases can have **computed field(s**) **or (calculated field(s)** which will carry out **calculations on another field or fields within the database** and display the answer (like a cell in a spreadsheet). Eg the average field shown in the diagram is a **calculated field:**

SURNAME:	Bruce
FORENAME:	Robert
CLASS:	2G
Test 1:	25
Test 2:	31
AVERAGE:	28

METADATA

Metadata is "data about data". Structural metadata is about the design and specification of data structures and is more properly called "data about the containers of data"; **descriptive metadata**, on the other hand, is about individual instances of application data, the data content.

As information has become increasingly digital, **metadat**a are used to describe digital data using metadata standards specific to a particular discipline. By describing the contents and context of data files, the quality of the original data/files is greatly increased. For example, a webpage may include **metadata** specifying what language it is written in, what tools were used to create it, and where to go for more on the subject, allowing browsers to automatically improve the experience of users.

STORAGE SPACE FOR DATABASE FILES

Database files have to be stored - either in main memory or on disc. Main memory is limited in size. Other storage mediums such as hard discs, magnetic tape, CD & DVD R & RWs. For the computer to manage the file storage efficiently we must decide on:

- The storage space needed for each record.
- The total number of records in the file.

From the values of the storage space needed for field names, records and total number of records, we can make a calculation of the total amount of storage space needed for the database file using the formula:

data for one record*number of records = total storage space required

We measure the data storage space in bytes. The number of data bytes needed for each record depends on the type of data stored under each fieldname.

EXPERT SYSTEMS

An **expert system** is a program what has access to a **huge database of knowledge about one particular subject.** It can make decisions and offer advice based on the knowledge it has. An expert system is made up from **facts and rules - known as the knowledge base.** This is like a giant database, but is more flexible as it contains rules as well as facts.

A **knowledge engineer** extracts these fact and rules from human experts and put them into a form that the computer is able to understand.

WRITTEN TASK

1 **Explain** the difference between a **simple search** and a **complex search**.

2 **What** is the advantage of using a **complex search?**

3 **Give two** terms which can be used in a complex search to link search options.

4 **Detail** the **two** ways in which a database can be *sorted into order.*

5 **What** name is given to a **formatted and organized presentation of data?**

6 **Suggest** one use of **forms** in a database.

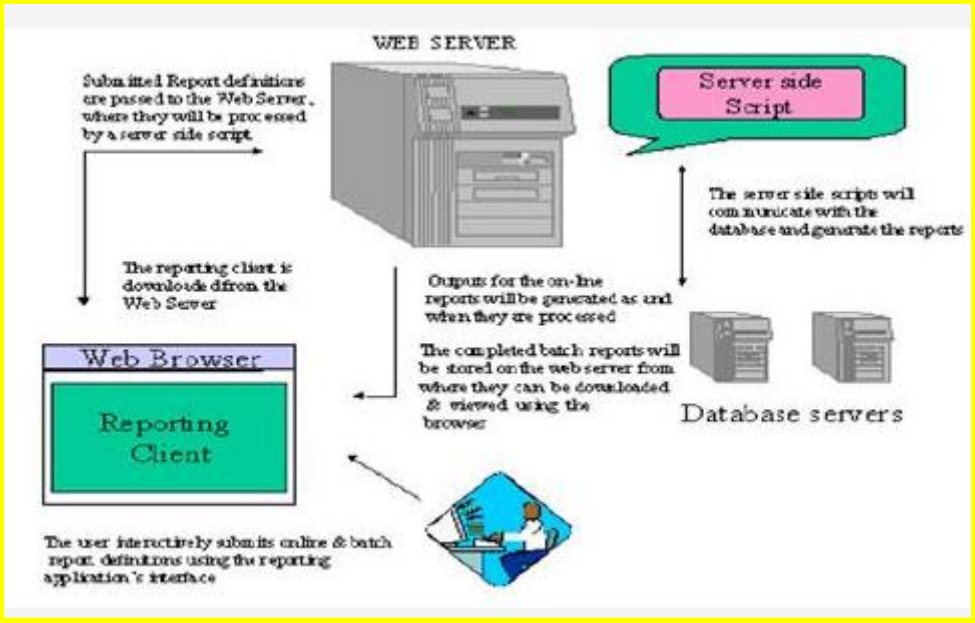
7 **What** is a **computed field?**

8 **Explain** what an **expert system** is what the role of a **knowledge engineer** in relation to an expert system is.

WEB BASED DATABASES (WBDB)

Organizing and retrieving information in a databases with a limited volume of information is fairly straightforward, but as the Internet emerged, information storage and retrieval changed radically and as a result, a new generation of databases called WBDBs have been created to meet user needs. **A web-based database (WBDB) is an organized listing of web pages.**

Most databases in use nowadays are **relational databases**, but client/server databases are the basis for **WBDBs**. They are set up to operate 24 hours a day and are used by ISPs as well as individuals The most common language used by relational databases is SQL.

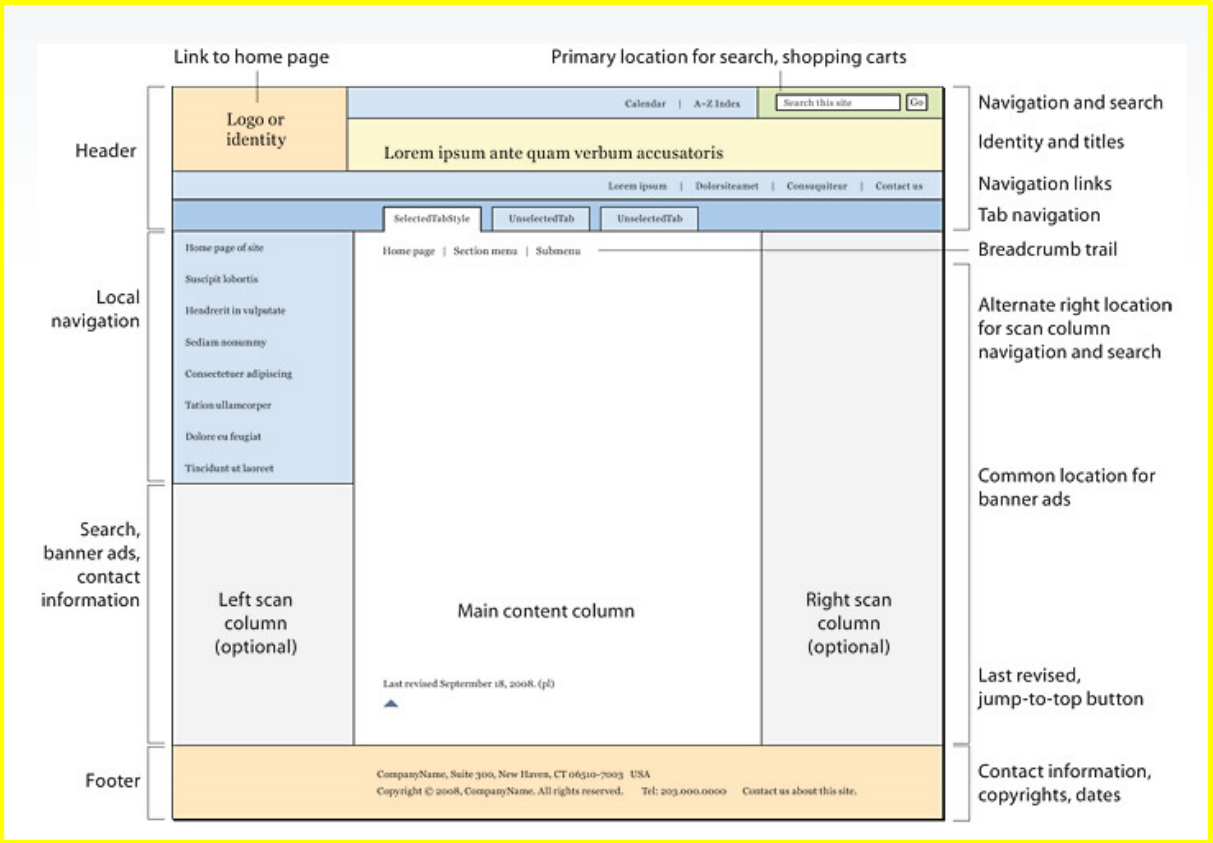


WEB STYLE SHEET

A **web style sheet** is a form of separation of presentation and content for web design in which the mark-up (i.e., HTML or XHTML) of a webpage contains the page's semantic content and structure, but does not define its visual layout (style). **Instead, the style is defined in an external style sheet file using a style sheet language such as CSS.** This design approach is identified as a "separation" because it largely supersedes the antecedent methodology in which a page's mark-up defined both style and structure.

PAGE STRUCTURE

Web "sites" are complete abstraction - they don't exist, except in our heads. When we identify a site as such, what we're really describing **is a collection of individual linked pages that share a common graphic and navigational look and feel.** What creates the illusion of continuity across a cohesive "site" is the design features that pages share. Individual html pages and how they are designed and linked are the atomic unit of web sites, and everything that characterises site structure must appear in the page templates.



CSS (CASCADING STYLE SHEETS)

Cascading Style Sheets (CSS) is a **style sheet language used for describing the look and formatting of a document written in a mark-up language**. While most often used to style web pages and interfaces written in HTML and XHTML, the language can be applied to any kind of XML document.

CSS is designed primarily to enable the separation of document content from document presentation, including elements such as the layout, colours, and fonts. This separation can improve content accessibility, provide more flexibility and control in the specification of presentation characteristics, enable multiple pages to share formatting, and reduce complexity and repetition in the structural content.

Using CSS solves the problem of formatting web pages:

HTML code isn't really intended for more than the basics of web design. Another problem is that the coding must be added to each item in each page, even though they are identical in design or style.

On the other hand, **Cascading Style Sheets** are used to format the design or presentation of your pages. **CSS code** can be added to a page or exist on a separate file, used by many web pages**. CSS** enables the separation of document content from document presentation.

Another aspect of **CSS** is that you can divide a page into divisions and set properties or styles within each division. Also, the divisions can be nested, allowing for greater flexibility in formatting a page.

CSS rules can be internal or external:

CSS rules can be embedded within the <head> of a web page HTML code. Those rules apply to the whole page. More often, an external CSS file is used, such that it applies to all pages referencing that file.

CSS rules can define the styles of the <body> and the text tags, such as <h1> and for the whole page. Class and ID rules define styles for specific situations.

Simple rule - An example of a simple rule is: p {font-size: 12px;}. Thus, the font size of all paragraph text would be 12px, unless there were some exceptions made.

Class rule - An example of a rule for a CSS class is: .note {font-size: 10px}. Applying the rule to a whole paragraph in HTML would be:

xxxx

Applying the rule to a section of text would be:

Section of text in paragraph.

ID rule - An **ID rule** can be used only one time on a page. An example of an **ID rule** in the CSS file is: #menu {font-size: 10px; font-color: green;}

Thus all the text in the menu division would be 10px and green:

Green text

Cascading rules - The "cascading" concept of a CSS means that rules stated later in the CSS file can take precedence over previously states rules. For example, you might set the general size of the tag but then later set a different size when is in the right column of your page.

Divisions used for formatting - You can divide your web page with divisions, using the tag. Styles can be assigned for all elements within that division. Also, divisions can be nested, giving more flexibility of formatting design.

For example: your web page HTML code could include:

Header content

Article content

Menu content

META TAGS

Meta tags are contained in the **HEAD** section near the top of the page. They're not displayed to the end user unless you view the source code of the page. They allow webmasters to provide information to a browser, search engine, or automated program (i.e., robot).

They are ignored by default unless the browser or search engine specifically recognizes them. There are several **meta tags**, but the most important are:

Title tag

This is the first and most important tag in **search engine placement algorithms**. Ideally, the Title tag should contain five to ten good, descriptive words and not exceed 70-80 characters. The specific keyword should be used, preferably as the first word of the title. **Remember, the title appears as the link in search engine listings, so you need to make it attractive to humans as well.**

For example: <title> Meta Tags and High Search Engine Placement </title>

Meta Description tag

This is the second important tag in search engine placement algorithms. The **Description tag** is used by many search engines as a short description of the page in search engine listings, so don't repeat the Title in this tag, and make sure it's attractive, without being misleading. It should not exceed 150-200 characters. The specific keyword should be used at least once (try twice if it fits and see how it ranks). Include common synonyms and one or two of your most important general keywords.

For example:

<meta name="description" content="How to use Meta tags to get high search engine placement.">

Meta Keywords tag

The **Keywords tag** is designed to tell the search engine what keywords are important to your page, and thereby how people should be able to find you when they search. But this meta tag has almost completely lost its relevance as far as search engine positioning is concerned - it hardly affects the rankings these days.

For example:

<meta name="keywords" content="meta tags, high search engine placement, meta expires tag, meta refresh tag">

Meta Robots tag

This tag contains instructions to web-indexing robots, as to whether or not to index this page and to follow the links on it. By default, robots both index and follow everything, unless you tell them not to. So, you generally don't require this tag.

However, you may need to prevent the search engines from spidering the pages containing Meta Refresh tags or other risky techniques that are often used for spamming, because the search engines may penalize your site for this.

For example:

<meta name="robots" content="noindex,follow">

will tell the robots that they shouldn't index this page, but they should follow the links that are on it.

Meta Revisit tag

The **Revisit tag** defines how often a search engine or robot should come to your website for re-indexing. This meta tag is supported by many search engines and is often used to control search engine placement of websites that change their content on a regular basis.

For example:

<meta name="revisit-after" content="30 days">

The following Meta tags control your visitors' browsers...

Meta Expires tag

The **Expires tag** allows you to force the browser to not cache a page, so that it loads a new copy from the server each time the page is viewed. It is primarily useful for pages with dynamic content. Just set the expiry date to 0, and caching will be disabled.

For example:

<meta http-equiv="expires" content="0">

To make a page expire at a specific time in the future, spell it out in Greenwich Mean Time…

<meta http-equiv="expires" content="Mon Dec 31 23:59:59 GMT 2007">

Meta Refresh tag

The Refresh meta tag will automatically send users to another URL, after a specified amount of time has elapsed. The syntax for the content part is - the number of seconds to delay, then a semicolon, then the new destination URL.

For example:

<meta http-equiv="refresh" content="5;URL=http://www.site.com/">

Note that the search engines don't like Meta Refresh tag and either ignore the current page and index the second page or may not index the site at all.

WRITTEN TASK

1 **What** is a **web based database?**

2 **Explain** the purpose of a **web style sheet.**

3 **Is** there actually any such things as a **website? Explain** your answer.

4 **What** is used for describing the **look and formatting of a document** written in a mark-up language?

5 **Give four tags** which may be used in web pages **and describe the purpose of each.**

DYNAMIC WEB PAGES

A **dynamic web page** is a web page managed as a single page in the web browser while the actual web content given on that page varies.

Parameters determine how the assembly of every new web page proceeds. After that, some of the objects on the dynamic web page, will be smart objects, adapting their own content without needing new page parameters. The same dynamic web page is reloaded, by the user, a computer program, scheduling, etc., to change some variable content, but it is the same page in the browser.

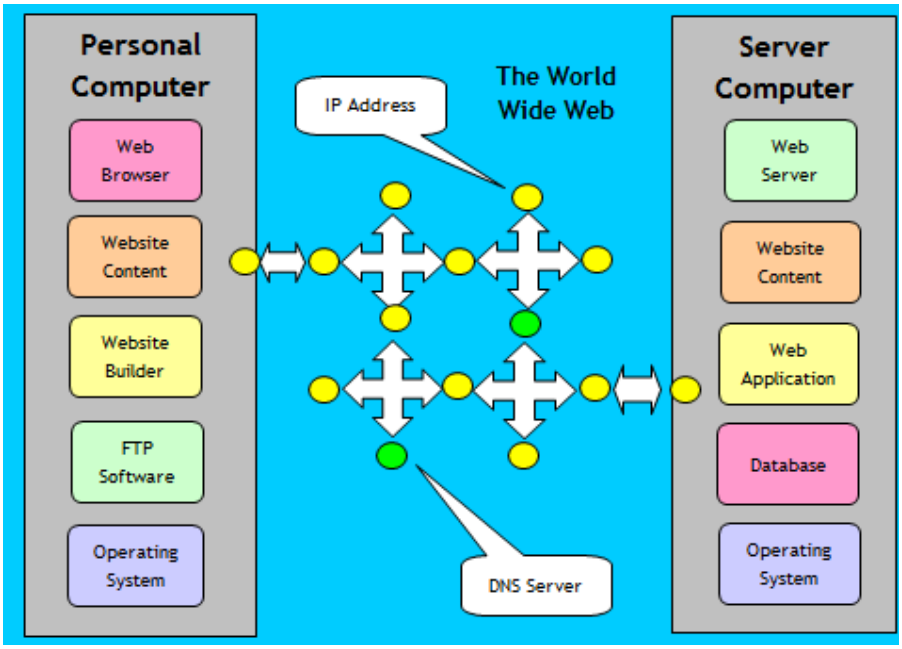
DHTML also determines the intelligence in a page, but in general it is a new page in the browser. Client-side-scripting, server-side scripting, or a combination.

DATABASE DRIVEN WEBSITE

A **database driven website** is a website that uses a database to store information that is captured during an Internet session or a site that uses data from a database to dynamically assemble web pages using live data from a database. Most, but not all database driven websites are "authenticated" sites. An authenticated site requires that the visitors identify themselves using a login ID and password. Examples of data driven sites include Amazon, Hotmail, and web banking websites.

It is much more complex to build a database driven website than a "brochure ware" website or a web-application that doesn't require a database. There are web authoring tools now available that can help you generate simple database driven sites using data from Access databases, spreadsheets, and XML data sources.

There are many different types of **database management systems (DBMS)** that can be used in a database driven website. The most widely used DBMS for database driven sites is MySQL, a relatively simple one that runs on computers that use Unix and Linux operating systems. On server computers that use Windows based operating systems, SQL Server, MS Access and XML data sources are popular.



When a request for a **data driven web page** is made by a **web browser to a web server,** the web server calls an application system (or server-side script) that in turn makes a request for data from the database. If the request is successful, the data is formatted and passed back to the web server which integrates it with other website content. **The resulting web page and related files (e.g. graphics) are then sent back to the web browser using the standard http communication protocol.**

INTERACTIVE WEB PAGE

An **interactive web page allows** the user or visitor to participate and edit the content on it. A good interactive web page is defined by its ability to engage the user in an immersive experience that delivers satisfaction. These web pages are created by integrating JavaScript in their coding.

Commonly used **interactive websites include sites such as**:

- Facebook
- ebay
- Google Maps

MULTIMEDIA APPLICATION

A **multimedia application** is an application which allows the user to **combine different media types**, ie text, graphics, sound, animation and video to present information, entertainment, training, simulations etc.

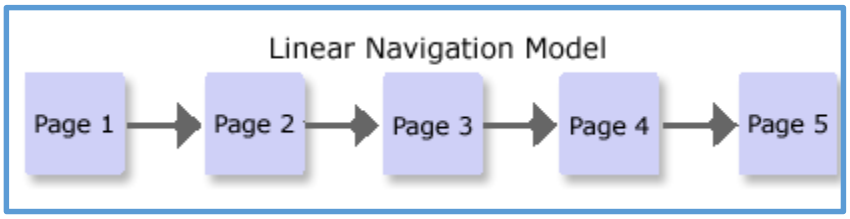
MULTI-LEVEL NAVIGATION

Any information system, including website navigation has **two main functions**:

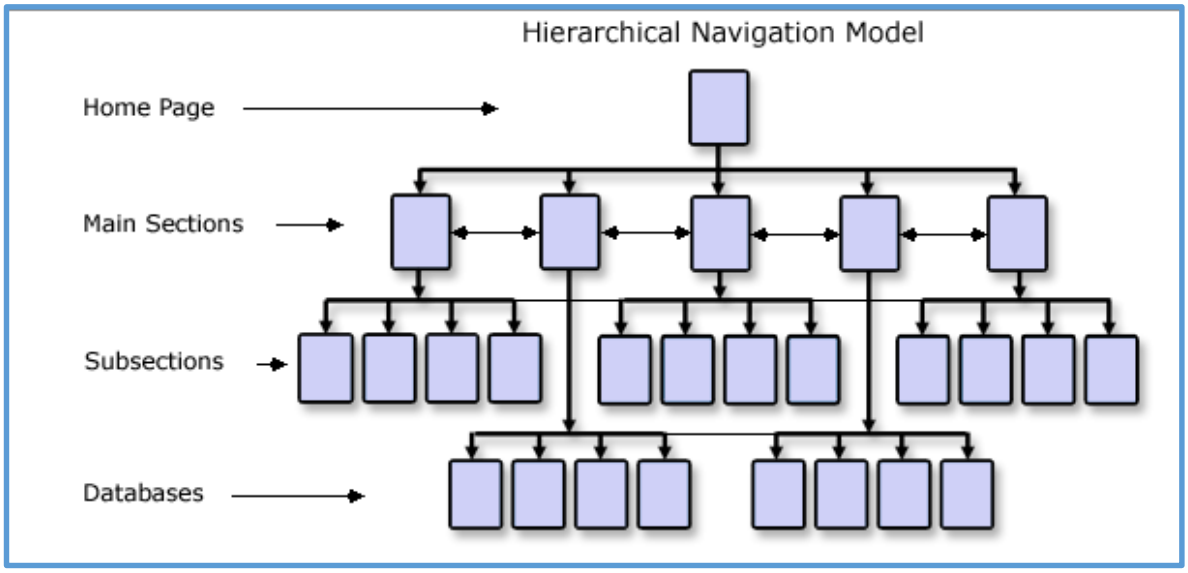
- To tell the user where they are.
- To enable the user to go somewhere else.

Navigation is the system(s) that a visitor can use to move around a website (global navigation, breadcrumb trails, related links, pagination (previous/next page), footer navigation, etc.) AND the visual indication of such systems (hyperlinked text, tabs, buttons, etc).

Linear navigation is used when you want the visitor to go from one step to another in a particular order.



Hierarchical navigation goes from the general to the specific. For example, from a homepage on a website, to main sections to subsections and databases. It is a way to tie together many areas of information into a working website structure. A visitor could easily go from the homepage to other areas of the website and back again using this pattern of website navigation. The goal of any hierarchical website navigation system is to provide an efficient means of simplifying user interaction.



A good design principle is that any page of your website is no more than two levels deep from home page (i.e. it should not take more than two clicks to go from the home page to every page of your site). Otherwise, the **search engine spiders** may not index all the pages.

WRITTEN TASK

1 **What** is a **dynamic web page**?

2 **Give two** examples of **database driven** websites.

3 **What** does an **interactive website** allow the user to do?

4 **Explain** the difference between **linear** and **hierarchical navigation**.

5 **Suggest** which method of navigation is most suited to a collection of related web pages.

MEDIA TYPES

Information systems can hold and use different types of data, such as text, graphics, sound, animation and video.

The vast majority of information systems use **standard file formats** to save their data files. Using **standard file formats** makes the data **portable** across different computing platforms and improves accessibility.

Text Files

The **standard file formats for text files** are:

- Text (**.txt**)
- Rich Text Format (**.rtf**)

A **text** (**.txt file)** is extremely **portable** as it can be opened in **any word processing or text editing application**. Another advantage of **.txt files** is that they take up less storage space than **.rtf files**. However, a disadvantage of **.txt files** is that they do not store information about how the document has been formatted, things such as alignment, style etc.

A **rich text format (.rtf file)** is also portable across most computing platforms, however an **.rtf file** does contain all the formatting information, including fonts and sizes used, styles, indentation etc.

Sound Files

The **standard file formats for sound files** are:

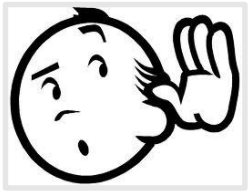
- **.wav**
- **.mp3**

Once the sound card has captured audio data and converted the samples into digital values this is known as RAW, because the sound files have not been subject to further processing. RAW sound files are very large and take up a large amount of backing storage space so some form of **compression** is needed to reduce the size of the file.

Sampling Depth refers to the number of bits used to store **each sample of sound.** The greater the **sampling depth**, the higher quality of sound. However the greater the **sampling depth**, the larger the file size will be.

SAMPLING DEPTH	NUMBER OF SOUND LEVELS
8-bit samples	256 sound levels
16-bit samples	65,536 sound levels
32-bit samples	4 billion + sound levels

This means that a sound file with a **sampling depth** of 32 bits will require four times the backing storage than the same sound with a **sampling depth** of 8 bits.



Sampling Frequency refers to the number of times per second a sample is taken. The greater the **sampling frequency**, the higher quality of sound. However the greater the **sampling frequency**, the larger the file size will be.

Sound Time simply means, the longer the sound lasts, the larger the file size will be.

Number of Channels refers to whether the sound clip has been recorded in mono or stereo.

All of the above factors, added together can lead to very large file sizes requiring compression, using either of the following standard sound file formats:

WAV file format is one step removed from **RAW** audio data. However **WAV files are encoded using lossless compression,** by storing the difference between the samples of the audio wave, instead of the actual sound samples themselves, as is done in RAW files. A **WAV file can be a quarter of the size of the same audio file stored in RAW but without any loss of sound quality.**

MP3 is an audio compression format which uses lossy compression. The **lossy technique** samples data that cuts out parts of the sound that the human ear cannot hear. Once it has cut the inaudible data it uses Huffman encoding to further compress the file size. Compressing files using **MP3 format** means that audio files can be compressed by a factor of up to 12 with no noticeable loss of quality. **MP3 files** take up less space on backing storage than other audio file formats and take less time to transmit on a network.

Graphic Files

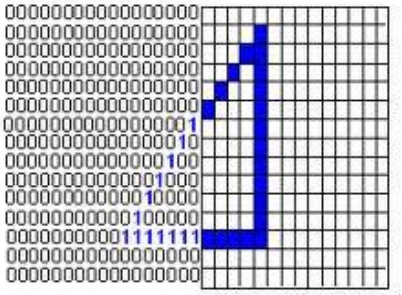
The **standard file formats for graphics files** are:

- **.bmp**
- **.jpeg**
- **.gif**
- **.png**

The **file sizes for graphic files can be very large**, which means they will take up a lot of backing storage space and will take longer to transmit across networks such as the Internet.

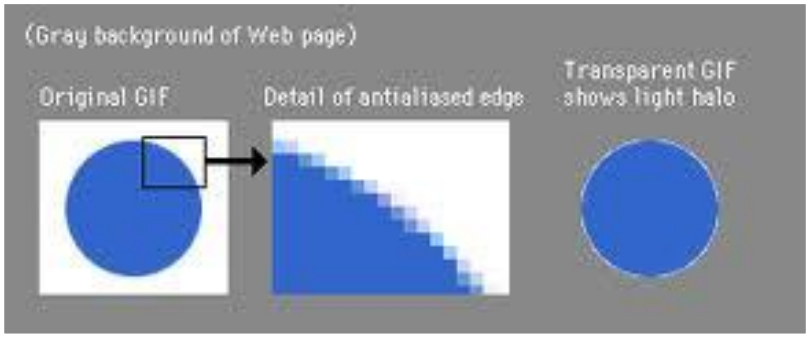
Graphic file sizes can be made smaller by varying the resolution and/or colour depth used and by using compression techniques. However, varying the resolution and/or colour depth and compressing a graphic will have an effect on the file quality and the file size.

A **bitmap file format (.bmp**), is an **uncompressed file format** where the bit-map file contains data about each pixel used to make up the image. For a black and white image 1 bit is needed to store each pixel. For an image using 256 colours 1 byte of storage is needed for every pixel and with an image using True Colour this rises to 3 bytes of storage for every pixel. **This means that an uncompressed bit-map file can take up a very large amount of storage.**



The **.jpeg file format** (**J**oint **P**hotographic **E**xperts **G**roup) - uses **lossy compression to reduce the file size.** This works by cutting out the parts of the graphics which won't be noticed by the human eye. As **jpegs use True Colour** with 16,777,216 colours available – the lossy compression cuts out some of the very subtle variations in colour. **jpegs** are the most suitable file format for storing photographs, but are not suitable for storing line drawings as the loss of data would be noticeable here. jpegs can vary the level of compression – the higher the compression, the smaller the file size but the greater loss of data.

The .gif file format use lossless compression which works by using a code to store patterns of bits that occur repeatedly throughout the graphics file. Because the **gif format** is based on an 8-bit colour code – files stored in this format are limited to 256 colours, meaning that the **gif format** is best suited for storage of drawings, charts and diagrams and is not suited for the storage of photographs. However, an advantage of **gif** is that colours can be set as transparent which is useful when displaying a **gif** image in a multimedia document, as it does not then obscure the underlying screen pattern.



The **.png file format** (**p**ortable **n**etwork **g**raphics), uses **lossless compression.** It can achieve between 5 and 25 per cent more compression that the .gif file format. It can support a bit depth of up to 48 bits per pixel – meaning it has a huge range of colours available. It also allows you to control the transparency of a graphic. **.png** was approved as a standard by the World Wide Web to replace .gif as **.png** is completely patent- and license-free. The most recent versions of web browsers support **.png.**

COLOUR DEPTH	NUMBER OF POSSIBLE COLOURS PER PIXEL
1 bit	2 – black or white
8 bits	256 colours
24 bits	16,777,216 (True Colour)

RESOLUTION (dpi)	BITS/BYTES USED TO REPRESENT EACH PIXEL	COLOURS AVAILABLE	PIXELS PER 3 INCH SQUARE GRAPHIC	FILE SIZE
300 x 300	1 bit	Black or White	90, 000	10.99 Kb
300 x 300	8 bits (1 byte)	256	720,000	87.9 Kb
300 x 300	24 bits (3 bytes)	16,772,216	2,160,000	224.26 Kb

PRACTICAL TASK

Calculate the storage requirements for the following graphics:

1 A 5 inch x 5 inch graphic with a resolution of 300 x 300 pixels using a bit depth of 8 bits.

2 A 2 inch x 2 inch graphic with a resolution of 200 x 200 pixels using a bit depth of 24 bits.

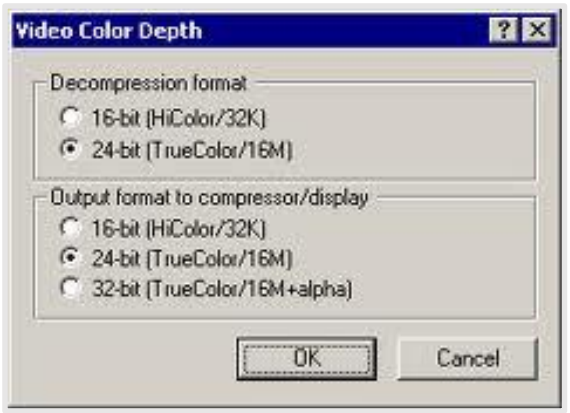
3 A 6 inch x 4 inch graphic with a resolution of 400 x 400 pixels using a bit depth of 24 bits.

4 A 3 inch x 2 inch graphic with a resolution of 200 x 200 pixels using a bit depth of 8 bits.

Video Files

The **standard file formats for video files** are:

- **.mpeg**
- **.avi**



Video is a sequence of frames (individual pictures), taken one after the other. When played back quickly the brain and eyes are tricked into seeing a moving image.

Because video files can be very large – one second of uncompressed video could demand 53 Megabytes of storage, video data is **compressed to make less demands on backing storage and more easily transmitted across networks.**

Videos are usually captured in true colour which requires 24 bits for every pixel. Reducing the colour depth reduces the file size but limits the number of colours displayed and degrades the quality of the video.

Resolution of a video clip is expressed as the number of pixels per frame – measured in height in pixels x width in pixels. Increasing the resolution improves the quality but increases the file size. Decreasing the resolution degrades the quality but decreases the file size.

Frame rate (fps) measures how many frames are processed per second. The standard rate for a video clip is 30 fps.

Reducing the frame rate (fps) to 15 fps would reduce the file size but result in a jerky, unsatisfactory playback. Cropping the video will also reduce file size, but will result in some frames being lost.

The **.avi file format** (**A**udio **V**ideo **I**nterleave), allows Video files to be stored in **uncompressed format.** This is a popular file format for Windows applications because it is compatible with the Windows Media Player.

Limitations of AVI include:

- Audio is low quality as it supports only 8-bit samples at a frequency of 11.025 KHz.
- It is not compressed – so large file size.
- Resolution is a maximum of only 320 x 240 pixels.
- Maximum file size of 2 Gigabyte – so not suitable for large projects.

.mp4 is the current standard from MPEG (**M**oving **P**ictures **E**xperts **G**roup) for compressing video data. **mp4's use lossy compression**. This works by looking the differences frame to frame in a video and cuts out the unchanged parts. Although data is lost to achieve compression, when the video is played back the human eye and brain compensate for any loss of data and fill in the gaps.

PDF File Format

Adobe Systems developed the **PDF file format** as a means of publishing and sharing documents across networks, especially the Internet. The **PDF file format is used to represent documents in a manner independent of application software, hardware, and operating systems.** PDF files can contain clickable links and buttons, form fields, video, and audio. PDFs look just like the originals and preserve all the source file information, even when text, drawings, videos, audio, 3D maps, full-colour graphics, photos, and business logic are combined in a single file.

Spreadsheet File Formats

.xls - Main spreadsheet format which holds data in worksheets, charts, and macros.

.csv - CSV files can be used with any spreadsheet program, for example, Microsoft Excel. They differ from other spreadsheet file types in that you can only have a single sheet in a file, they cannot save cell, column, or row styling, and cannot save formulas.

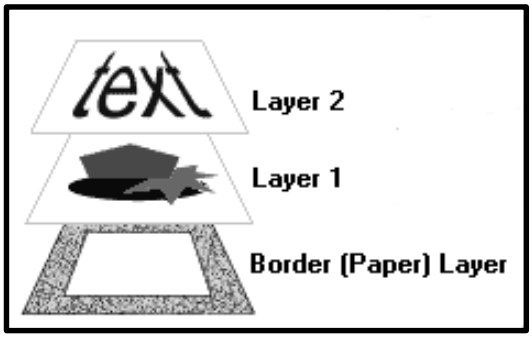
.xml - Extensible Markup Language (XML) is a markup language that defines a set of rules for encoding documents in a format which is both human-readable and machine-readable. XML-based formats have become the default for many office-productivity tools. XML has come into common use for the interchange of data over the Internet.

Vector Graphics

.svg and .vrml

The basic features of **Vector Graphics** are:

- They are scalable – resolution independent.
- Individual objects in a vector graphic can be edited.
- Graphics can be assembled by placing objects in layers.



Scalable = resolution independent.

Vector = represents objects by defining a series of attributes.

Common attributes of vector graphic objects:

- Position
- Shape
- Size
- Rotation
- Line
- Layer
- Fill

Virtual Reality Markup Language (VRML) is a standard language used to model and animate geometric shapes. It is used to define 3D environments for the WWW (World Wide Web).

Attributes of a 3D image:

- Shape
- Position
- Size
- Rotation
- Texture

Another great advantage of vector graphics is that their file size is much smaller than a bit-mapped file. This is because it is the instructions **(attributes)** needed to recreate the graphic which are stored, not the actual graphic. For example,

Common **attributes** for a circle, filled with red:

```
<svg>
<circle cx="200" cy="200" r="150" fill="red"/>
</svg>
```

WRITTEN TASK

1 **Give one advantage** a .rtf file has over a .txt file.

2 **State** which type of compression is used and the effect it has on the quality of and file size, of each of the following sound files:

(a) **.wav**

(b) **.mp3**

3 **What** effect does increasing the **sampling depth or sampling frequency** have on the quality and size of a sound file**?**

4 **What** type of compression is used for**.jpeg** and what effect does it have on the quality of the graphic file?

5 **Which** graphic file type was approved as a standard by the World Wide Web to replace .gif?

6 **Give two disadvantages** of using the **.avi** file format to store videos.

7 **Which** standard file format for videos uses **lossy compression**?

8 **Which** file format is used to represent documents in a manner independent of application software, hardware, and operating systems?

9 **Explain** what Extensible Markup Language (XML) is and the main advantage of using it.

10 **How** are **vector graphics** saved?

11 **Give two advantages** and **one disadvantage** of saving a graphic in .svg format.

SCRIPTING LANGUAGES

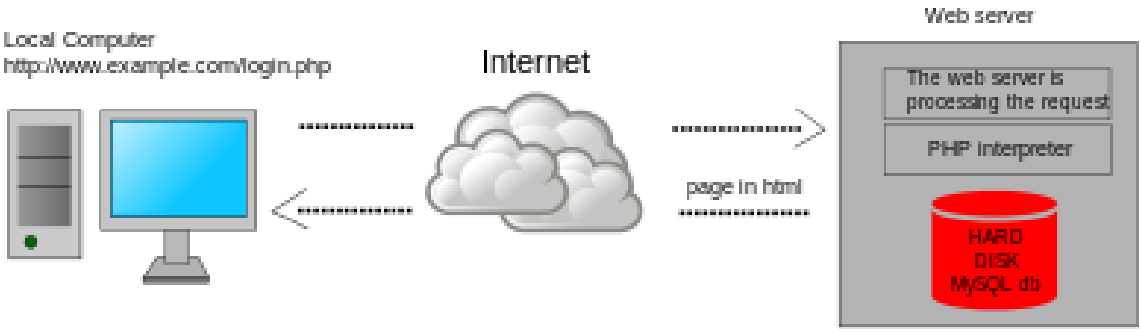
Scripting languages are designed to work alongside another program. Perhaps the best know example of this is **Javascript** which is used alongside **H**ypertext **M**ark-**U**p **L**anguage **(HTML). Scripting languages** are commonly used on the World Wide Web to increase the functionality of web pages and to make them interactive and dynamic. **PHP** is another scripting language which can be mixed with **HTML** on the World Wide Web. Another scripting language which is commonly used is **V**isual **B**asic for **A**pplications **(VBA),** which if often used to automate repeated tasks in documents created using Microsoft Office applications.

When a **Scripting language** *is run it interpreted from source code one instruction at a time.* Whereas the language which has been used for writing the entire application is usually *compiled before it is run*.

A **client-server network** has two types of computers – servers and clients. The server computers, for example the **file server, print server or web server** provide network resources and the **client (computers)** make use of the resources provided by the **server.**

It depends on whether it is the **server or the client who is running the script**, for the classification of **server-side or client-side scripting.**

SERVER-SIDE SCRIPTING



When a user requests a web page from the **serve**r, the script on the page is interpreted by the server, which creates or changes the page content to suit the user. The changed page, in its final form is then sent to the user and when this page has been sent it then cannot be changed again, in this instance, using **server-side scripting.**

The **scripting language,** such a**s Java or PHP** are executed by the server when the user requests a document. The output produced (usually HTML), is then sent to the user's computer.

SERVER-SIDE SCRIPTING is used mainly for managing a site's contents by presenting data, stored in a database, when the user requests it. It can also be used as a security measure on websites by restricting access to content.

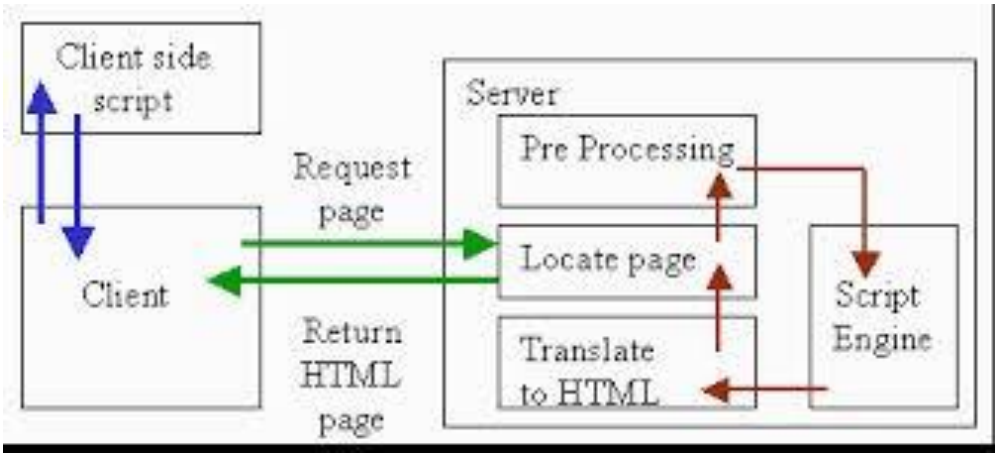
Typically, **server-side scripting** is used to permit users to have accounts which can be customised to individuals, by providing personalised data from databases. Social networking and E-commerce sites use **server-side scripting**.

Some examples of sites which use BOTH server-side and client side scripting are Amazon, Facebook and e-bay:

The **server-side scripting** deals with logging in and data which is specific to the user. Whereas the client-side scripting enables interaction on pages such as processing data which has been entered in a text box, for example when a user "updates their status" or "checks in" on Facebook.

CLIENT-SIDE SCRIPTING languages are executed on the *client computers of a network.* **Client-side scripts** may be either part of the document (**embedded script)** or be kept in a separate file **(external script),** which is referenced by the document which uses this script.

Client-side scripting is used to add functionality to websites and other applications. **HTML** is limited because its main purpose is to describe the elements of a web page such as, text, sound, graphics and video and the formatting information on the content. Therefore **scripting languages such as VBscript, Javascript** are used to add functionality to web pages and other applications. For example, client-side scripting could contain instructions to the browser to execute an action in response to a user's actions – such as selecting an item from a drop-down menu or clicking a command button. *It is important to note, that* ***client-side scripting*** *is not limited to web pages, it can be used in other applications*, for example, by creating a keyboard shortcut which, when pressed, automates an action.



WRITTEN TASK

1 **What** are **scripting languages** designed to work with?

2 **Do** you **interpret or compile** a scripting language before you run it?

3 **Wha**t is **server-side scripting** mainly used for?

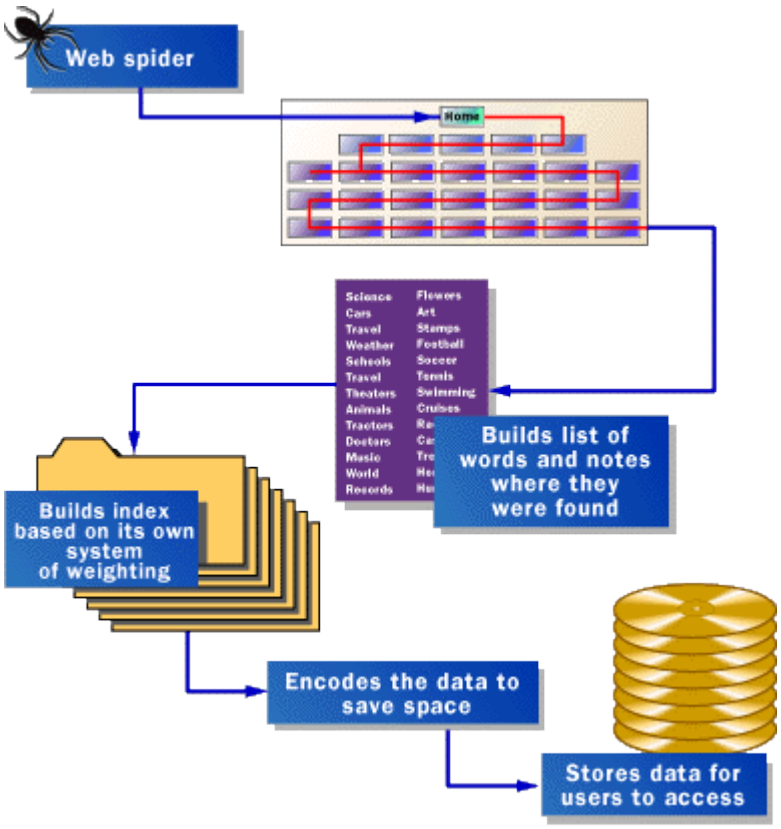
4 **What** is **client-side scripting** used for?

SEARCH ENGINES

The Internet and its most visible component, the World Wide Web, makes hundreds of millions of **web pages** available, waiting to **present information** on an amazing variety of topics. Problems arise because of the very fact that there are hundreds of millions of pages available, most of them titled according to the whim of their author, almost all of them sitting on servers with cryptic names. When you need to know about a particular subject, how do you know which pages to read? If you're like most people, you visit an **Internet search engine**.

Internet search engines are special sites on the Web that are designed to help people find information stored on other sites. There are differences in the ways various search engines work, but they all perform three basic tasks:

They search the Internet -- or select pieces of the Internet -- based on important words (called **keywords**).



Web Crawling

Most Internet users limit their searches to the Web.

Before a **search engine** can tell you where a file or document is, it must be found. To find information on the hundreds of millions of Web pages that exist, a search engine employs special software robots, called **spiders**, to build lists of the words found on Web sites. When a spider is building its lists, the process is called **Web crawling**.

How does any spider start its travels over the Web? The usual starting points are lists of heavily used servers and very popular pages. The spider will begin with a popular site, indexing the words on its pages and following every link found within the site. In this way, the spidering system quickly begins to travel, spreading out across the most widely used portions of the Web.

Meta tags allow the owner of a page to **specify key words** and concepts under which the page will be indexed. This can be helpful, especially in cases in which the words on the page might have double or triple meanings - the **meta tags can guide the search engine in choosing which of the several possible meanings for these words is correct.** There is, however, a danger in over-reliance on **meta tags**, because a careless or unscrupulous page owner might add **meta tags** that fit very popular topics but have nothing to do with the actual contents of the page. To protect against this, spiders will correlate **meta tags** with page content, rejecting the meta tags that don't match the words on the page.

Building the Index

Once the spiders have completed the task of finding information on Web pages (and we should note that this is a task that is never actually completed -- the constantly changing nature of the Web means that the spiders are always crawling), the search engine must store the information in a way that makes it useful. There are two key components involved in making the gathered data accessible to users:

- The information stored with the data.
- The method by which the information is indexed.

In the simplest case, a **search engine** could just store the word and the URL where it was found. In reality, this would make for an engine of limited use, since there would be no way of telling whether the word was used in an important or a trivial way on the page, whether the word was used once or many times or whether the page contained links to other pages containing the word. In other words, there would be no way of building the ranking list that tries to present the most useful pages at the top of the list of search results.

To make for more useful results, most **search engines** store more than just the word and URL. An engine might store the number of times that the word appears on a page. The engine might assign a weight to each entry, with increasing values assigned to words as they appear near the top of the document, in sub-headings, in links, in the **meta tags** or in the title of the page. **Each commercial search engine has a different formula for assigning weight to the words in its index. This is one of the reasons that a search for the same word on different search engines will produce different lists, with the pages presented in different orders.**

WRITTEN TASK

1 **Give two** basic tasks carried out by any **search engine**.

2 **Explain** how a **web spider** works.

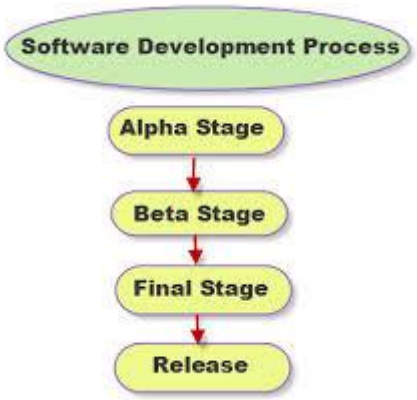
3 **Describe** the role of **meta tags** in relation to search engines.

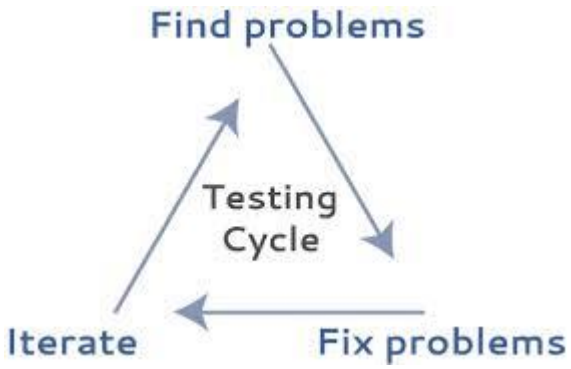
4 **What** are the **two key roles** involved in building the index of a search engine?

5 **Why** do different search engines produce different results, even when the same keyword is used in each?

TESTING

The purpose of testing an information system is to identify and correct any errors, before the product is realised for use.





ALPHA TESTING

Alpha testing is the first phase of testing and this takes place in the developer's environment. Alpha testing for an information system would include checking that the solution:

- Meets the requirements that guided its design and development.
- That all links work correctly.
- Pages look as they are intended to.
- Responds correctly to all kinds of inputs.
- Is sufficiently usable.
- Can be installed and run in its intended environments, and achieves the general result its stakeholders desire.

Any problems identified at the **alpha testing stage** which need to be rectified are the **responsibility of the developer** and this is known as **corrective maintenance.**

BETA TESTING

Beta Testing is the second phase of testing and this is when the information solution is **tested in the environment where it is intended to be used** and so that any problems not identified at the **alpha testing stage** can be identified and corrected before the information solution is released for use. **Beta testing** is therefore related to the type of information solution, for example:

- A computer game would be beta tested before it was released to obtain feedback from typical intended users as to how they feel about the game.

- A relational database would be tested to ensure that you have obtained a database system which satisfies the ACID properties (Atomicity, Consistency, Isolation and Durability).
- Usability testing covers whether the Website is designed well enough for customers to use it conveniently and without problems.
- Beta testing of websites also includes load testing, failure recovery, security, and platform and browser compatibility.

It may be that when **beta testing** takes place, the client who has commissioned the information system has some further ideas they would like implemented, which are in addition to the original specification given to the developer. If these changes are then implemented, this is called **perfective maintenance** and the cost of this is the responsibility of the client commissioning the information system.

USABILITY

Usability testing focuses on measuring a product's capacity to meet its intended purpose. Examples of products that commonly benefit from usability testing are web sites or web applications, computer interfaces, documents, and devices.

Users may be asked to perform specific tasks under controlled conditions to see how easy (or otherwise) they find it to use the software.

A **usability test** involves carefully creating a scenario, or realistic situation, wherein the person performs a list of tasks using the product being tested while observers watch and take notes. Several other test instruments such as scripted instructions, and post-test questionnaires are also used to gather feedback on the product being tested. The aim is to observe how people function in a realistic manner, so that developers can see problem areas, and what people like.

COMPATIBILITY

Compatibility testing is one type of software testing performed on a system that is built based on certain criteria and which has to perform specific functionality in an already existing setup/environment. Compatibility of a system/application being developed with, for example, other systems/applications, OS, Network, decide many things such as use of the system/application in that environment, demand of the system/application, etc.

It is a type of testing used to ensure compatibility of the system/application/website built with various other objects such as other web browsers, hardware platforms, users, operating systems, etc. This type of testing helps find out how well a system performs in a particular environment that includes hardware, network, operating system, other software, etc.

Compatibility testing includes:

Hardware: Evaluation of the performance of system/application/website on a certain hardware platform. For example: If an all-platform compatible game is developed and is being tested for hardware compatibility, the developer may choose to test it for various combinations of chipsets (such as Intel, Macintosh GForce), motherboards etc.

Browser: Evaluation of the performance of system/website/application on a certain type of browser. For example: A website is tested for compatibility with browsers like Internet Explorer, Firefox etc. (usually browser compatibility testing is also looked at as a user experience testing, as it is related to user's experience of the application/website, while using it on different browsers).

Network: Evaluation of the performance of system/application/website on network with varying parameters such as bandwidth, variance in capacity and operating speed of underlying hardware etc., which is set up to replicate the actual operating environment.

Peripherals: Evaluation of the performance of system/application in connection with various systems/peripheral devices connected directly or via network. For example: printers, fax machines, telephone lines, etc.

Compatibility between versions: Evaluation of the performance of system/application in connection with its own predecessor/successor versions (backward and forward compatibility).

WRITTEN TASK

1 **What** is the purpose of testing an information system?

2 **Suggest four** areas which would be tested during the **alpha testing cycle** of an information system.

3 **Give four** areas which would be tested during the **beta testing cycle** of an information system.

4 **Explain** the purpose of **usability testing.**

5 **Explain in detail**, **four areas** in which **compatibility testing** for an information system would be carried out.

6 **What** is meant by the term perfective maintenance?

PURPOSE OF INFORMATION SYSTEMS

It is important to know and define your reasons or motivation for having a website or any other type of information system.

Companies, organizations and individuals want websites. Each has their own reasons:

- Companies and organisations want to use a website as a form of advertising, as a means of communicating with the customer, and in some cases as a way to directly sell their products.
- Many individuals have websites to display their skills and interests, especially as a way to advance their careers.
- Finally, there are many people who have websites mainly for their own expression. Many have blogs for writing their thoughts.

Data

An information system is:

- Only as good as the data inputted.
- Poor data will lead to poor results being outputted.
- Incomplete or inaccurate information means outputted information will be incomplete or inaccurate therefore the system becomes useless.
- For the outputted data to be accurate or complete the source data needs to be of good quality.
- Data is generated by all parts of an organisation.
- Data can also be received from outside the organisation, for example sales orders.

People

People are involved in the capture, processing and the inputting of data in a organisation.

People can affect the quality of information and the information system because if the inputted data is captured in the wrong format or inputted incorrectly it can make the information become useless. This means that the system is only as good as the expert it has been collected from.

Hardware

The hardware should be:

- Flexible enough to cope with stress and strain put on it throughout the whole organisation.
- Should be able to store data in large volumes.

In large organisations the MIS is normally run on a server so that the whole organisation can use it.

Software

Software should be able to handle data quickly and efficiently, and that it can be easily searched though when needed.

Most Management Information Systems (MIS) use specialised software.

The hardware and software need to work well together.

FUCTIONALITY OF INFORMATION SYSTEMS

An information has four functions:

- Input
- Storage
- Processing
- Output

Input

Inputting information into an information system has two parts:

- Detailed data stored and processed and forms the basis for the rest of the system.
- User tells the system what sort of analysis they want from the system.

Storage

- The data should be stored efficiently with the highest level detail available.
- The IT department should take regular backups of the system and the stored data regularly, this should be kept in a different location in case of disaster.

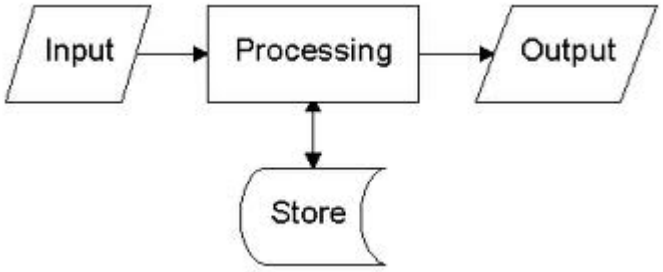
Processing

The processing of data is where the information is turned into knowledge.

Output

Outputted information can be displayed in many forms. The main three forms are:

- Graphical (e.g. Charts, graphs)
- Textual (e.g. Reports, numbers)
- Websites (multimedia – text, graphics, sound etc)



INTERACTION OF INFORMATION SYSTEMS WITH:

Novice Users - are new to the system and will need a simple and basic interface. Since they are new in the system they will expect more secure ways of doing things in the system (for example they will choose the templates or wizards to do their first steps in the system). **Novice users' interface** should provide simple ways to achieve important frequently performed tasks. When designing to novice users we should remember what the main use cases and don't overwhelm, them with unnecessary features.

Expert Users - (meaning with extended experience in the application) expects advance features and capabilities. They will want more customization options. Since they have a stable mental model of the application structure they feel free to explore the application and try new things. They will not be too worried about making mistakes since they feel secure that they will know how to bypass them.

Differing Age Ranges – it is crucial, when designing your information system, to bear in mind, the age of the audience you hope to attract. Age ranges can be anything from young children, teenagers, adults to senior citizens and the likely general level of ICT expertise for each age range.

WRITTEN TASK

1 **Give three** reasons why an individual, company or organisation may wish to have a website.

2 **Suggest** a problem which could arise if poor quality data is inputted.

3 **What** is data transformed into when it is processed?

4 **Why** is it important that regular **backups** of data are made?

5 Explain in detail, which factors have to be considered when designing an information system targeted towards:

(a) Novice Users

(b) Expert Users

(c) Differing Age Ranges

HARDWARE & SOFTWARE REQUIREMENTS

A useful computer system consists of both **hardware** and **software.**

The physical parts of the computer – that is the parts we can see and touch, are **hardware**. The actual computer consists of the systems box. The devices which are attached to the systems box **for input, output and backing storage are known as peripheral devices.**

Input Devices

Input Devices consist of **hardware peripheral devices** such as:

- Keyboard
- Mouse
- Touchpad
- Microphone
- Digital Camera
- Digital Video Camera
- Webcam
- Scanner
- Graphics Tablet
- Joystick/Joypad

Keyboard:

The most common **input device** to be used with a computer is a **keyboard**. Every key has switch beneath it, which, when pressed sends the appropriate ASCII code into the computer. **Keyboards** are used to enter text into the computer.

Mouse:

Various types of **mouse** can be used to **point and select** on the computer screen. They can be either connected to the computer with a wire or be wireless. They can also be optical – with a light under the **mouse**, where any movement is detected by a sensor, or have a ball underneath it, which, when the mouse is moved operates the mechanisms inside the **mouse**.

Touchpad:

Touchpads are usually used instead of a mouse on laptop computers. The user moves their finger over the touchpad which, (as the name suggests, is touch sensitive), to control the movement of the pointer on the monitor. Having a **touchpad** means that you do not need the extra space required for a mouse.

Microphone:

If any type of sound is required to be input to a computer system a **microphone** is needed. Sometimes a **microphone** is used along with voice recognition software so that instructions can be given to the computer or text dictated.

Digital Camera:

Digital Cameras use a lens which focuses reflected light from the subject being photographed into a sensor called a **Charged Coupled Device (CCD).** The information captured by the **CCD** is digitised and stored in the camera's memory – usually on removable memory cards.

Digital Video Camera:

A **Digital Video Camera** works in much the same way as a digital camera – using a CCD. A **digital video camera** is used to **capture moving pictures**.

Webcam:

A **Webcam** is a small **digital video camera**, which is usually built into the monitor (in laptops), or placed beside or on the monitor, of a desktop computer. The resolution of a **webcam** is not as high as a digital video camera and the webcam has to be connected to a computer system in order to capture or store images. **Webcams** can be used for video conferencing – which allows people in different locations to conduct meetings.

Graphics Tablet:

Graphics tablets enables a user to hand-draw images and graphics directly into the computer, similar to the way a person draws images with a pencil and paper. **Graphic tablets** may also be used to capture data or handwritten signatures. They also be used to trace an image from a piece of paper which is taped or otherwise secured to the surface. The device consists of a flat surface upon which the user may "draw" or trace an image using an attached stylus, a pen-like drawing apparatus. The image generally does not appear on the tablet itself but, rather, is displayed on the computer monitor.

Joystick/Joypad:

Joysticks are often used to control video games, and usually have one or more push-buttons whose state can also be read by the computer.

Joypads generally feature a set of action buttons handled with the right thumb and a direction controller handled with the left. The direction controller has traditionally been a four-way digital cross (also named a gamepad). **Joypads** are the primary means of input used on many video game consoles.

Scanner:

A **scanner** is used to capture hard copies of documents, including photographs and convert them into digital format so that they can be manipulated or stored on a computer. The **scanner** can also be used in conjunction with optical character recognition software, to scan in text based documents in a format which will allow them to be edited on a computer.

WRITTEN TASK

1 **What** is the most **common input device** used with a computer?

2 **What** is a **mouse** used for?

3 **Which** type of computer uses a **touchpad** instead of a mouse?

4 **Which** input device can be used with **voice recognition software** to input instructions or data into a computer?

5 **Give two input devices** which use a sensor called a **charged coupled device** to capture data?

6 **Suggest** a suitable input device which could be used for **video conferencing**?

7 **Which** input device used for **capturing images**, has the **lowest resolution**?

8 **Which** input device can be used to capture **handwritten signatures**?

9 **Give two input devices** which can be used to control **video games**.

10 **Explain** what a **scanner** is used for.

11 **Which type of software** is needed with a **scanner** to enable documents to be captured in a format which allows computer editing?

Output Devices

Output Devices consist of **hardware peripheral devices** such as:

- Inkjet Printers
- Laser Printers
- Liquid Crystal Display (LCD) Screens
- Thin Film Transistor (TFT) Screens
- Loudspeakers

- **Inkjet Printers:**

Inkjet Printers are **peripheral output devices** used to obtain **hard copy of output from a compute**r. Inkjet printers may produce black & white (monochrome) or colour prints. Inkjet printers are good at producing high quality photo prints. The **resolution (quality) of a printout is measured in dpi (dots per inch)** and the **speed of printing in ppm (pages per minute).** Many inkjet printers now come with a feature which allows the **flash memory card from a Digital camera to be inserted**, so that images can be printed directly, without having to be uploaded to a computer first.

Laser Printers:

Laser Printers are **peripheral output devices** used to obtain **hard copy of output from a computer**. Laser printers may produce black & white (monochrome) or colour prints. **The resolution (quality) of a printout is measured in dpi (dots per inch)** and **the speed of printing in ppm (pages per minute).**

WRITTEN TASK

1 **What** is the **resolution of a printer** measured in?

2 **How** is the **speed of printing** measured?

3 **What** is an **LCD screen**?

4 **Where** would you usually find **LCD screens** being used?

5 **What** is a **TFT screen** and what is the **main advantage** of using one?

6 What is needed to **output sound** from a computer system?

COMPUTER PROCESSORS

The **processor**, is the brain of the computer. The more powerful the processor is, the faster your PC will run. The **speed of a processor – that is the number of instructions per second that it can carry out - is measured in gigahertz (GHz).**

Also consider the number of cores. A **multi-core processor** has more than one processor on a single silicon chip, so it's better able to handle multiple tasks at once. Most modern desktops have multi-core processors.

When buying a new PC, for a basic machine you'll need a dual-core processor that runs at about 2.5GHz - this will handle everyday basic computing tasks.

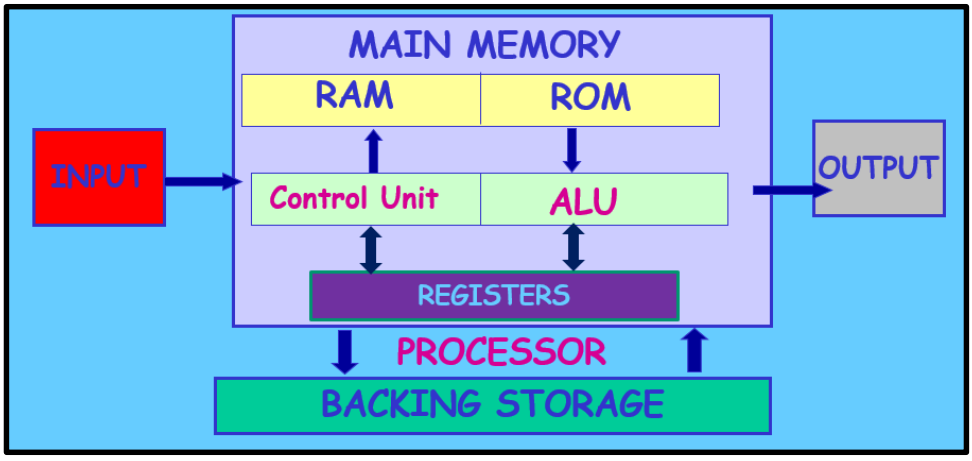
A home user, looking for amid-range PC might consider something like an i5 quadcore processor. This is the name of a range of processors from chip-maker Intel – the more expensive processors will give better performance.

Keen gamers who want all the power they can get may be tempted by the latest high-power processors such as the i7, Intel's top-of-the-range processor. However, the fastest processors don't come cheap and will push up the total cost of a PC considerably.

Most modern desktop processors will be 64-bit – this means that they can work with chunks of data made up of 64 binary digits.

MEMORY (RAM & ROM)

The **Central Processing Unit (CPU)** is where the processing takes place in the computer. It contains two parts the **PROCESSOR** and **MAIN MEMORY**.



The **Main Memory** is made up of two parts **RAM (Random Access Memory)** and **ROM (Read Only Memory).**

RAM (Random Access Memory) is used to store any program and associated data currently being run on the computer. When **RAM** is switched off all its memory contents are lost. This is why **RAM** is sometimes described as being volatile. If the user wants to make any changes to a file, it must first be loaded into **RAM** before any changes can be made.

There are two types of RAM:

- **Static RAM**
- **Dynamic RAM**

Static RAM keeps its content as long as power is maintained to the chip. However, **Dynamic RAM** needs to be refreshed very few milliseconds by rewriting its contents and of course, power also has to be maintained. The circuitry in **Dynamic RAM** is simpler than that in **Static RAM** and requires less power. As a result **Dynamic RAM** is cheaper than **Static RAM.** However the main advantage **Static RAM** has over **Dynamic RAM** is faster access.

ROM (Read Only Memory) keeps its memory contents when the computer is switched off. This is why **ROM** can sometimes be described as being non-volatile.
The contents of **ROM** cannot be altered by the user (anything stored in **ROM** must be loaded into **RAM** before any changes can be made).

The operating system program is stored in **ROM.**

Main Memory is divided into individual storage locations – with each location having a unique number or address.

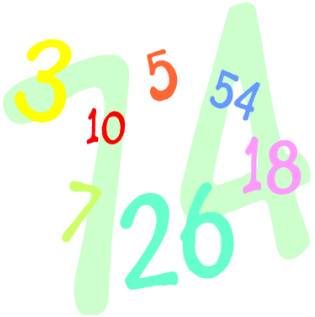
The unique address is used to find the location of data items stored in memory. The ability of the CPU to identify each storage location is called its **addressability.**

FLASH ROM is similar to **ROM,** however it can also be used to store data. **Flash ROM** is used in USB Flash Drives (Memory Sticks) and Flash Memory Cards in digital cameras, mobile phones etc. **This is called solid state storage.**

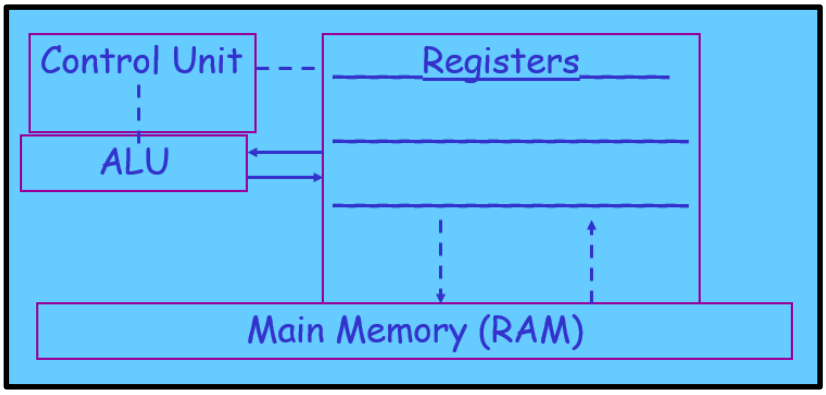
The **Control Unit** is found in the **Central Processing Unit (CPU).**

The **Control Unit** makes sure that everything happens in the correct place at the correct time. This means that it decodes instructions sent to the processor and controls peripheral devices (printers, scanners, disc drives etc), as required. The **Control Unit** also controls the flow of data around the **CPU.**

The **Arithmetic & Logic Unit (ALU)** carries out all calculations (arithmetic) and performs logical operations (it makes decisions).



Registers are individual storage locations inside the processor**. Registers** are used to hold the programs, data, instruction or the address of memory locations, while they are being processed.



The **Instruction Register** holds whichever instruction is currently being executed by the processor.

The **Accumulator** is a data register which holds the accumulated results of calculations performed by the ALU.

The **Program Counter** is a register which holds the address of the main-memory location storing the next instruction to be executed by the program.

CACHE

Cache memory is a smaller but faster compartment of memory that supports main memory. **Cache memory** is designed to accelerate the memory function. Your computer checks the cache to see if the data it needs is in storage.

Even with faster microprocessors and main memory (RAM) with greater capacities, bottlen**ecks still occur when information moves between the CPU and RAM.** Cache memory is a type of supplemental memory built for faster conveyance of information between these two components of the computer. The computer builds a library of frequently used information into the **cache memory**. The role of "caching" is to accelerate the speed of your system. **Caching** not only helps reduce logjams at the CPU and RAM, but enables users to perform tasks faster because data transmits more rapidly from cache memory to CPU.

Level 1 Cache

Typically, a high-speed cache is integrated right into the processor. This is the level 1 cache. One of the main motivations behind this concept is the "locality of reference." It means that a location just accessed by the CPU has a higher probability of being revisited in the short term. So the L1 cache holds the most recent data. When it needs the data again, the microprocessor first checks the L1

cache. Since the data is there, it eliminates the need to go back to the slower main memory. The process is usually two times faster than with the main memory.

Level 2 Cache

Level 2 cache memory, or the secondary cache, on a computer is usually located on a memory card situated close to the processor. It links directly to the central processing unit, and a circuit that is integrated into the motherboard controls it. The circuit is called the L2 controller. The level 2 cache catches recently used data not in the level 1 cache. The L2 cache on many personal computers enables the processor to get about 95 percent of the information it needs from the cache memory.

Other Cache Memory

Disk cache is another type of cache memory. It operates at a much slower speed than the other caches, because it reads data from the computer's hard drive and stores it on RAM. Software programs like DOS use it. Just as RAM is slower than L1 and L2 cache, the disk is much slower than RAM.

WRITTEN TASK

1 **Name** the **three** parts of the **CPU**.

2 **Explain** the difference between **RAM and ROM**.

3 **What** is **Flash ROM**?

4 **Explain** the purpose of the **ALU**.

5 **What** does the **Control Unit** do?

6 **What** are the **registers** used for?

7 **What** is used to **store programs and data permanently**?

8 **Explain** the term ***addressability***.

9 **What** is **cache memory** and describe the difference between level 1 and level 2 cache memory.

DEVICE TYPES

Embedded Computers:

An **embedded computer is a computer built into the system it is controlling. Embedded computers** are found inside things like cars, washing machines, games consoles etc. Embedded systems can have a processor, RAM, ROM, Input and Output interfaces all stored on a single chip. **Embedded computers process data in real-time.**

Palmtop Computers:

Palmtop computers are so called because they are small enough to be carried in one hand while being used. Sometimes they are also called PDAs (Personal Digital Assistants) or handheld computers. They come with an LCD screen for output and may have a small keyboard, however, input is usually via a touch sensitive screen using a stylus. They are battery operated.

Many mobile phones now have the same features as **palmtop computers** – these are called **smartphones.**

Smartphones:

A **smartphone** is a mobile phone built on a mobile operating system, with advanced computing capability and connectivity. **Smartphones** combine the functions of a personal digital assistant (PDA) with a mobile phone. Most **smartphones** also feature, portable media players, compact digital cameras, pocket video cameras, and GPS navigation units to form one multi-use device. Most modern **smartphones** also include high-resolution touchscreens and web browsers that display standard web pages as well as mobile-optimized sites. High-speed data access is provided by Wi-Fi and mobile broadband. In recent years, the rapid development of mobile app markets and of mobile commerce have helped drive up the use of **smartphones.**

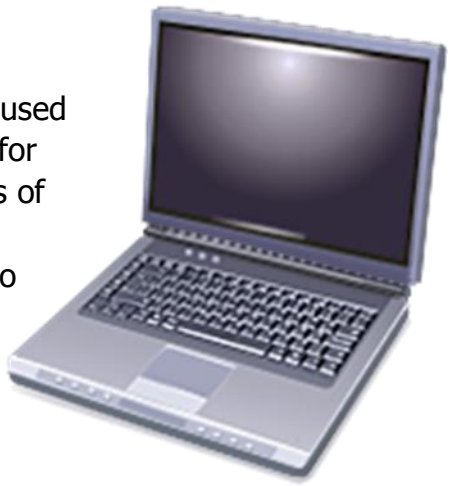
Smartphones use **mobile operating systems (OS)** such as, Google's Android, Apple's iOS, Nokia's Symbian, RIM's BlackBerry OS, Samsung's Bada, Microsoft's Windows Phone. Operating systems like these can be installed on many different phone models, and typically each device can receive multiple OS software updates over its lifetime.

Tablet computers:

Tablet computers have a large touch sensitive screen – used with a stylus for drawing and writing.

They can be used for web-browsing, e-mail, music, games etc.

Laptop (and Notebook) computers:



Laptop (and Notebook) computers are designed to be used whilst resting on your lap. The come with an LCD monitor for output and a keyboard and trackpad for input. These types of computers have a hard disk and cd/dvd drive for storage, however it is likely that solid state storage (SSD) will start to replace hard disk drives in laptops and notebooks, as they are very robust – having no moving parts. They are battery operated and come with a wireless interface card.

Desktop computers:

Desktop computers use mains electricity and as the name suggests, sit on a desktop. They are not portable. The common peripheral devices are:

- Keyboard
- Mouse
- Monitor (Liquid Crystal Display)
- Loudspeakers

Storage Devices typically include Hard Disk Drive and CD/DVD drives.

A network interface card (NIC) is also needed to connect to a local area network (LAN).

Supercomputers:

The most powerful and fastest computer is the **supercomputer**. These computers are used for intensive or demanding mathematical calculations or the production of very high quality graphics, for example the digital animation used in films.

Organisations which require large amounts of data processed at very quick speeds and multi-user access, will use **supercomputers**.

Supercomputers have very fast, multiple processors and a huge backing storage capacity.

WRITTEN TASK

1 **What** is an **embedded** computer?

2 **Explain** why it is necessary for embedded computers to process data in **real time.**

3 **What** is the **difference** between a palmtop computer and a smart phone?

4 **Suggest** a business which would use a mainframe computer and why.

5 **Wha**t is clock speed an indicator of?

6 **Wha**t is clock speed measured in?

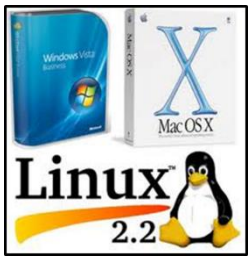
SOFTWARE

No matter what hardware a computer systems has, it must be combined with software before it can be used to do any task.

Operating Systems:

The **operating system program** must be loaded to a computer before anything else will work. **When the computer is switched on the bootstrap loader program loads the operating system.** Once the **operating system** has been

loaded it is then used to control all the application programs and peripheral devices attached to the computer.



The main functions of an **operating system** are:

- To provide a human computer interface (HCI) so that the user can interact with the operating system.
- Managing input and output to and from peripheral devices.
- Managing the loading and execution of programs.
- Managing, saving and loading of files to and from backing storage.
- Error reporting.

Current trends in Operating System Design:

Trends in **operating system design** include scaling up to utilise large address spaces, large numbers of cores, cloud computing, mobile hardware and the consequential mobile operating systems required.

The advent of **larger address spaces** presents the challenge of re-imagining how to use the additional space. Additionally, the continuous push to more and more cores brings with it the added burden of determining how best to utilise each of them most efficiently, and the creation of appropriate scheduling algorithms and systems necessary to ensure that this efficient utilization actually takes place.

Cloud computing brings with it a number of challenges. The issue of security is of paramount importance to many. When data is stored in the cloud, one has to be able to trust the security systems put in place by whoever is operating the servers where that data is housed.

There has been an explosion **of mobile hardware and the mobile OSs required for the mobile hardware.** The increasing prevalence of touch, voice, motion, and other forms of input, the internet of things (ie the trend of random things in the world around us having chips in them and being connected to the internet and reporting status / other data about themselves and their surroundings), the automation of the home, for example, the smart fridge.

Licensing

A software license is a legal instrument governing the use or redistribution of software. A typical **software license** grants an end-user permission to use one or more copies of software in ways where such a use would otherwise potentially constitute copyright infringement of the software owner's exclusive rights under copyright law.

In addition to granting rights and imposing restrictions on the use of software, software licenses typically contain provisions which allocate liability and responsibility between the parties entering into the license agreement. In enterprise and commercial software transactions these terms often include limitations of liability, warranties and warranty disclaimers, and indemnity if the software infringes intellectual property rights of others.

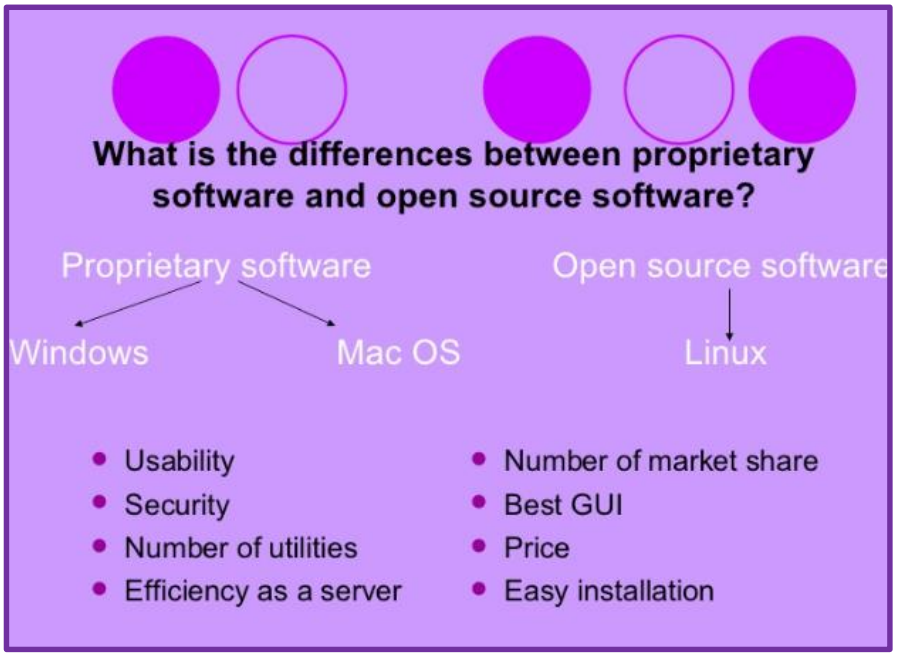
Software licenses can generally be fit into the following categories: **proprietary licenses and free and open source**. The significant feature that distinguishes them are the terms which the end-users might further distribute or copy the software.

Proprietary versus Open Source

Proprietary software is computer software licensed under exclusive legal right of the copyright holder with the intent that the licensee is given the right to use the software only under certain conditions, and restricted from other uses, such as modification, sharing, studying, redistribution, or reverse engineering. Vendors typically distribute **proprietary software** in compiled form, usually the machine language understood by the computer's central processing unit. They typically retain the source code or human-readable version of the software, written in a higher level programming language. This scheme is often referred to as closed source. By withholding source code, the software producer prevents the user from understanding how the software works and from changing how it works.

Open-source software

Open-source software is computer software with its source code made available with a license in which the copyright holder provides the rights to study, change and distribute the software to anyone and for any purpose.



Advantages and disadvantages:

1 Bug fixes - Open source applications such as Drupal have many users who can combine their expertise to fix bugs as they are discovered. With **proprietary software**, it might take longer to fix bugs as there relatively a fewer number of people who can fix it. **Proprietary or licensed software** might cost higher initially but will free you of most maintenance problems later on, saving time. On the other hand, if you want to save on cost and are technically minded, **open source programs** can offer the same kind of benefits as **proprietary software** minus the customer support.

2 Customer Support - Part of the services offered by companies selling **proprietary software** is friendly customer support, which is not normally available in **open source software**. For some open source programs with little to no paid staff, no one will be obligated to entertain your concerns. But there are online forums for many open source programs where you can address your concerns and find solutions.

3 Features & Scalability - **proprietary software programs** normally contain features that are designed for the needs of many users. But they also usually contain premium or specialized features not found yet in **open source programs**. **Open source programs**, however, may not have all the specialized features of reputable **proprietary software** but they are more scalable.

4 Product Integrity - Although software developers may not admit it, some **proprietary programs** are released to the market even when they are not yet fully tested. Purchasing premature or newly released programs may compromise your system. However, there are also many **proprietary software applications** that are reputed to be stable whenever new versions are released to the market.

Open source software programs have different versions available in the Internet and are clearly marked as stable or in a beta phase.

5 Updates - For **proprietary software**, you cannot make changes to a program. The developer takes care of its changes and updates. With **open source programs**, you can make any changes to the program because you can have access to the original source code. And depending on the **open source program** you are using, there are many other users and consultants who can advise you on how to make updates and changes.

6 Security - Both **proprietary and open source software applications** are prone to cyber-attacks or hacking. Developers of **proprietary software** are responsible for the security of their programs. They are the ones who constantly monitor and prevent potential attacks from happening to their programs. **With open source software** where the original source code can be easily accessed over the Internet, it is easier for any outdated program to be compromised, often due to missteps or oversight on the part of the user or system administrator. The **open source** community monitors its software for any security loopholes.

7 Maintenance - Updates for **proprietary software** are generally taken care of by the developer of the program, but users would have to wait for their availability. For **open source programs**, you might need to spend if you want to implement new developments or updates to your system. Additionally, if new security patches for **open source** are available, it could take a few hours of work to apply them in your system.

8 Ongoing Costs - and **proprietary software** varies from one software application to another. Some licensed software applications require a one-time payment, while others require a monthly or annual subscription. Many software applications also charge for upgrades, while updates are usually offered for free. Generally, there are no ongoing billing costs associated with **open source programs**.

9 Hidden Costs - With **proprietary software**, you purchase the license and can use it as is. You will only incur hidden costs if you need to reconfigure the software for your needs. Hidden costs are more often a part of **open source software**. *It comes from any charges you will spend to customize the software so that it can work for you.*

10 Initial Cost - Both proprietary and open source software require initial cost. For **proprietary software programs**, the initial cost is the amount you will have to pay in exchange for the software package or license. On the other hand, **open source software programs** are not really free. You might need to pay someone to implement the software to your system. For instance, Drupal open source programs may be free to acquire but you may need to tap on the expertise of someone to integrate and implement it in your system. This can be expensive, depending on the nature of work these people will have to do to make the software up and running.

Portability

Portability, in relation to software, is a measure of how easily an application can be transferred from one computer environment to another. A computer software application is considered **portable** to a new environment if the effort required to adapt it to the new environment is within reasonable limits. The meaning of the abstract term 'reasonable' depends upon the nature of the application and is often difficult to express in quantifiable units.

The phrase "to port" means to modify software and make it adaptable to work on a different computer system. For example, *to port an application to Linux* means to modify the program so that it can be run in a Linux environment.

Portability refers to the ability of an application to move across environments, not just across platforms. To clarify, a computer platform generally refers to the operating system and computer hardware only. A computer environment is much broader and may include the hardware, the operating system and the interfaces with other software, users and programmers. If software is described as **portable** *then it can be run on different operating systems* (e.g. Mac OS, Linux, PC) *with little or no change*.

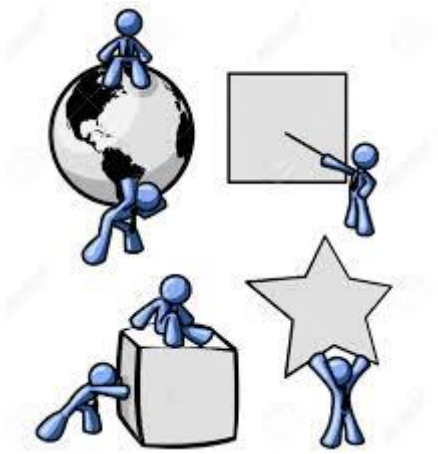
Reasons it may not be **portable** include:

- Incompatible operating systems.
- It's an older computer that isn't recognising much newer software.
- The programming language isn't available for that operating system.

Application Programs:

Application programs are pieces of software which are designed for tasks. Sometimes the software has been designed to do specific tasks – for example as a **utility program**. One example of a **utility program is anti-virus software**. This has been created specifically to protect your computer from attack by viruses and cannot be used for anything else.

General purpose packages (GPPs) include programs such as word processing, spreadsheet, database, presentation and graphics. Although each program has been designed for use on particular tasks, for example – word processing for mainly text based documents, it is possible to use software to complete tasks other than that is mas mainly designed for. **Where two or more GPPs are combined in a single program, this is called an integrated package.**



Objects and Operations:

Objects and Operations which can be carried out in a package, help the user to decide which application package is the most suitable to use for particular tasks.

As you can see from the table below, many of the same **objects and operations** can be carried out by different packages. In this case, the user should use the package which will enable him to do the task as quickly and as easily as possible.

An **object** is a piece of data in a GPP and an **operation is a process which can be carried out on that piece of data**, for example:

PACKAGE	OBJECT	OPERATION
WORD PROCESSING	Character	Enter, Insert, Edit, Delete, Copy, Paste, Format
	Word	Spellcheck, Search & Replace
	Sentence	Import/Export
	Paragraph	Import/Export
	Table	Add Row, Add Column, Split Cell, Merge Cell
	Drawing	Croup, Scale, Move, Send to Front/Back
DATABASE	Field	Enter, Insert, Edit, Delete, Copy, Paste, Format
	Record	Search, Sort, Calculate, Query
	File	Import/Export
	Layout	Report
SPREADSHEET	Cell	Enter, Insert, Edit, Delete, Copy, Paste, Set Attributes
	Row	Format, Import/Export
	Column	Format, Import/Export
	Number	Format, Import/Export
	Text	Format, Import/Export
	Formula	Calculate (sum, average, max, min, if)
	Row	Format, Import/Export
GRAPHICS	Pixel	Enter, Insert, Edit, Delete, Copy, Paste, Format
	Line	Scale, Rotate, Crop, Layer, Group, Ungroup
	Circle	Scale, Rotate, Crop, Layer, Group, Ungroup
	Rectangle	Scale, Rotate, Crop, Layer, Group, Ungroup
	Polygon	Scale, Rotate, Crop, Layer, Group, Ungroup

Web Browsers:

A **web browser** (commonly referred to as a browser), is a software application for retrieving, presenting and navigating information resources on the **World Wide Web**. An information resource is identified by a **Uniform Resource Identifier (URL)** and may be a web page, image, video or other piece of content. Hyperlinks which are present in resources, enable users easily to navigate their browsers to related resources.

Although **browsers** are primarily intended to use the World Wide Web, they can also be used to access information provided by web servers in private networks or files in file systems.

WRITTEN TASK

1 **Why** do computers need software?

2 **Explain** the difference between an **application program** and an **operating system program.**

3 **Give** an example of **utility software.**

4 **What** is an **integrated package**?

4 **State three** functions of an **operating system.**

6 **Give** the name and the primary function of **three application packages**.

7 **What** is an **object?**

8 **Suggest three operations** which could be carried out on an **object**.

9 **What** is the primary function of a **web browser**?

10 **Give two advantages** and **two disadvantages** of using **proprietary software** as opposed to **open source software**.

11 **Explain** the term **portability,** as applied to a piece of software.

12 **Suggest** a reason software may not be **portable**.

13 **Describe** a **security problem** which may occur as a result of using open source software.

STORAGE

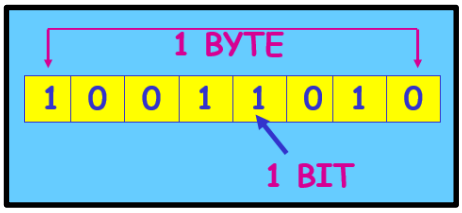
When we use computers, we need to have some kind of **backing storage** so that our program and data files are no lost when the computer is switched off and where it can be retrieved from, when it is needed again.

Backing storage may be **local, web or cloud based**, or a combination of any of these, depending on the needs of the individual or organisation.

STORAGE CAPACITY:

Computers use the binary system to store data. The binary system uses two digits 1 and 0. The number 1 represents "ON" and the number 0 represents "OFF".

A binary digit (either 1 or 0) is called a **BIT.**



Capacity measurements, small to large are:

- 1024 Bytes = 1 Kilobyte
- 1024 Kilobytes (Kb) = 1 Megabyte
- 1024 Megabytes (Mb) = 1 Gigabyte (Gb)
- 1024 Gigabytes (Gb) = 1 Terabyte (Tb)
- 1024 Terabytes (Tb) = 1 Petabyte (Pb)

Local Storage:

Local storage is where the individual or organisation owns the computers and their associated **backing storage devices**.

There are **initial and running costs,** which may be very high, associated with **local backing storage**:

- Cost of buying and maintaining the software, computers, backing storage and associated network for sharing of files.
- Technical support will need to be provided.
- File security will need to be provided.

However, organisations may find that some of the **advantages of having local backing storage** is that they can have greater control over the degree of security

they apply to their files and that, once the initial costs have been met, the running costs are lower than the ongoing costs of renting equipment and backing storage would be.

Web Storage:

Web storage and DOM storage (document object model) are web application software methods and protocols used for **storing data in a web browser**. **Web storage** supports persistent data storage, similar to cookies but with a greatly enhanced capacity and no information stored in the HTTP request header. There are two main web storage types: **local storage and session storage**, behaving similarly to persistent cookies and session cookies respectively.

Cookies:

There are **two types of cookies** – **session cookies** and **persistent cookies**:

Session cookies are temporary cookie files, which are erased when you close your browser. Webpages have no memories. A user going from page to page will be treated by the website as a completely new visitor. **Session cookies** enable the website you are visiting to keep track of your movement from page to page so you don't get asked for the same information you've already given to the site. Cookies allow you to proceed through many pages of a site quickly and easily without having to authenticate or reprocess each new area you visit. When you restart your browser and go back to the site that created the cookie, the website will not recognize you. You will have to log back in (if login is required) or select your preferences/themes again if the site uses these features. A new **session cookie** will be generated, which will store your browsing information and will be active until you leave the site and close your browser.

Persistent cookies are files which stay in one of your browser's subfolders until you delete them manually or your browser deletes them based on the duration period contained within the **persistent cookie's** file.

Web storage provides far greater storage capacity (5 MB per origin in Google Chrome, Mozilla Firefox, and Opera; 10 MB per storage area in Internet Explorer; 25MB per origin on BlackBerry 10 devices) compared to 4 kB (around 1000 times less space) available to cookies.

Web storage data is not automatically transmitted to the server in every HTTP request, and a web server can't directly write to **Web storage**. However, either of these effects can be achieved with explicit client-side scripts, which mean they are executed by the user's browser and not the server side. This allows for fine-grained tuning of the desired interaction with the server.

Web storage offers two different storage areas—local storage and session storage—which differ in scope and lifetime. **Data placed in local storage is per domain** (it's available to all scripts from the domain that originally stored the data) and persists after the browser is closed. **Session storage** is per-page-per-window and is limited to the lifetime of the window.

Cloud Storage:

Cloud storage is a networked enterprise storage where data is stored not only in the user's computer, but in **virtualized pools of storage which are generally hosted by third parties.** Hosting companies operate large data centres, and people who require their data to be hosted, **buy or lease storage capacity from them.** The data centre operators, in the background, virtualize the resources according to the requirements of the customer and expose them as storage pools, which the customers can themselves use to store files or data objects. Physically, the resource may span across multiple servers. **The safety and security of the files depends upon the hosting websites**.

Cloud storage services may be accessed through a web service application programming interface (API), a cloud storage gateway or through a Web-based user interface.

The **advantages of using cloud storage** include:

- Companies need only pay for the storage they actually use.
- Storage maintenance tasks, such as backup, data replication, and purchasing additional storage devices are offloaded to the responsibility of a service provider, allowing organizations to focus on their core business.
- Cloud storage provides users with immediate access to a broad range of resources and applications hosted in the infrastructure of another organization via a web service interface.
- Customers have access to technical support.

However, the **disadvantages of using cloud storage** include:

- On-going costs of paying for this service.
- Security concerns about possibility of hacking of important data.

DISTRIBUTED AND OFFLINE STORAGE

Distributed Storage is storage that can be shared throughout a network. For example, Windows file sharing. **Distributed file storage** security can be built in at different layers to protect files. For example - can use passwords, tickets, access control lists and encryption. **Cloud is an example of distributed storage.** Cloud storage means "the storage of data online in the cloud," wherein a company's data is stored in and accessible from multiple distributed and connected resources that comprise a cloud. Cloud storage can provide the benefits of greater accessibility and reliability; rapid deployment; **strong protection for data backup, archival and disaster recovery purposes**; and lower overall storage costs as a result of not having to purchase, manage and maintain expensive hardware. However, cloud storage does have the potential for security and compliance concerns.

Offline Storage is the term used to describe any **storage medium that must be inserted into a storage drive** by a person before it can be accessed by the computer system is considered to be a type of **off-line storage**. *Offline storage may also be written as off-line storage and is also called removable storage.*

BACKUP SYSTEMS AND STRATEGY

A **backup** is a copy of files to another medium as a precaution in case the first storage medium fails.

A **backup strategy** is developed by determining:

- Which data needs to be protected.
- Which media to use for that protection.
- How to protect the media itself.

The system administrator will create and maintain a reliable **backup system** by considering these factors. The resulting **backup system** ensures operational continuity in the event of a disaster.

The most appropriate **backup methodology** for an organisation is based upon the nature of the organisation and its processing requirements. **Backup methodologies** include:

Full – which transfers a copy of all data within the scope of the backup to the destination media, regardless of whether that data has changed since the last backup was performed.

Differential – backs up all files changed since the last full backup, regardless of whether they have been changed since the last backup operation of any kind.

Incremental – backs up only those files which have changed since the last backup operation of any kind.

Selective backup - a type of backup where only the user specified files and directories are backed up. A selective backup is commonly used for backing up files which change frequently or in situations where the space available to store backups is limited.

Hot backup - A technique used in data storage and backup that enables a system to perform a routine backup of data, even if the data is being accessed by a user. Hot backups are a popular backup solution for multi-user systems as no downtime to perform the backup is required.

Synthetic backup - a synthetic backup is identical to a regular full backup in terms of data, but it is created when data is collected from a previous, older full backup and assembled with subsequent incremental backups. The incremental backup will consist only of changed information.

Cloud backup - (cloud computer backup) refers to backing up data to a remote, cloud-based server. As a form of cloud storage, cloud backup data is stored in and accessible from multiple distributed and connected resources that comprise a cloud.

Online backup - using a high-speed Internet connection, specific files or the entire contents of a hard drive are backed up to the online storage provider's system using a Web browser interface.

Remote backup - In storage terminology, a remote backup is an online managed backup service for backing up data to a remote, cloud-based server ("cloud backup").

Snapshot backup – a virtual copy of a device or file system. Snapshots imitate the way a file or device looked at the precise time the snapshot was taken.

STORAGE DEVICES:

Backing Storage Devices include devices such as**:**

Hard Disk Drives:

Hard Disk Drives use **random/direct access** and are an example of **magnetic storage.** You get fastest access to data stored on a hard disc than any other backing storage medium. The storage capacity of a hard disk is also very large, currently at least 1 terabyte. Hard disks are usually sealed inside the hard disk drive within the systems box, however external hard disks can be plugged into the computer via a USB port – aiding **portability** by making it easy to transport large quantities of data. A hard disk can have a **data transfer rate** up to 1,030 Mbps.

Magnetic Tape Drives:

Magnetic Tape Drives use **sequential/serial access –** meaning that you have to go through the files in order to teach the one you want. They are an example of **magnetic storage.** The **storage capacity of a Magnetic Tape** can be anything from 10 to 100 Gigabytes. **Magnetic Tape** is now used mainly for **backing up data** held on computer systems – it is always important that data is backed up (extra copies kept), in case the original data is lost or damaged. **Magnetic tape drives** can have **a data transfer speed** of up to 900 Gbs an hour.

Compact Disk (CD) & Digital Versatile Disk (DVD) Drives:

Compact Disk (CD) & Digital Versatile Disk (DVD) Drives use **random/direct access and are an example of optical storage** – meaning data is written to the disk using a laser beam. The **storage capacity** of a CD-R (Recordable) & CD-RW (Rewriteable) is up to 700 Megabytes. For a DVD-R or DVD-RW it is between 4.7 and 17 Gigabytes. The cost of both is relatively cheap, they are very **portable** and the large storage capacity makes them very suitable for storing large multimedia files – such as sound, video & graphics. Different read/write speeds can be selected for CD and DVD drives with the higher the speed chosen, the higher the transfer rate. For example, a DVD with a read/write speed 0f 16x would typically transfer data at 21.13 Mbps.

Blu-ray Drives:

Blu-ray Drives use random/direct access. They are an example of **optical storage.** The **storage capacity of a Blu-ray disk** ranges from 23.3 Gigabyte (Gb) – 27 Gb for a single-layer disk to 46.6 Gb to 54 Gb for a dual-layer disk. Although a blu-ray disk is the same physical size as a DVD - the increased storage capacity on the blu-ray comes from the laser beam being used to write data being blue and having a shorter wavelength and narrower beam than the red laser used to write to DVDs. As with CD & DVDs – Blu-ray disks are **very portable**. Blu-ray Disc specification at 1x speed has a data transfer rate of 36Mbps

USB (Universal Serial Bus) Drives:

USB (Universal Serial Bus) Drives use **random/direct access**. They are an example of **solid state storage** – solid state meaning they have no moving parts, which makes them very robust. They are also **very portable** – being compatible with most computer systems. The **storage capacity of a USB Flash Drive** ranges from 1 Gigabyte upwards to 128 Gigabytes. These have a **data transfer speed** of up to 480 Mbps.

Flash Memory Cards:

Flash memory cards are used by many small devices such as smartphones, MP3 players and digital cameras to store data. A USB card reader attached to a computer allows data files to be transferred on to backing storage on a computer for editing, printing or storage. The **flash memory card** can then be wiped clean and used over and over again. **Flash memory cards use random/direct access** and are an example of **solid state storage**. Because they are very small and light, they are **extremely portable**, but can still hold many gigabytes of data. A reasonable **data transfer speed** would be around 90 Mbps.

Data Transfer Speed:

The speed with which data can be transmitted from one device to another is called the **data transfer speed**. Data rates are often measured in megabits (million bits) or megabytes (million bytes) per second. These are usually abbreviated to Mbps.

Interface Type:

An interface is how a peripheral device communicates with the processor. A USB is the most common interface would have a typical data transfer speed of around 480 Mbps upwards, depending on the model type.

Firewire is another interface often used and would have a data transfer of around 800 Mbps.

CURRENT TRENDS IN STORAGE SYSTEMS

Hybrid Clouds Become the Dominant Vision for Enterprise IT

The tension within IT about moving to the **cloud** will resolve itself as organizations recognize that a hybrid cloud model is needed to serve their application portfolio. CIOs will sort their portfolio into:

Applications that require complete IT control (using **on-premises private clouds**)

Applications that can be shared with a third party (using **enterprise public clouds**)

Applications with highly variable workloads (using **hyperscalar public clouds**)

Applications that can be addressed with Software-as-a-Service offerings

IT will act as the broker across the various cloud options, and there will be an increasing need to move application data and maintain consistent storage service levels across the different clouds.

All-Flash Startups

Enterprises will increase their adoption of **flash technology** as the leading storage companies validate this trend. The battle between mainstream players and bleeding-edge all-flash vendors will be won by those that best enable customers to deploy the right levels of performance, reliability and scalability for their specific needs and workloads. Growth in international markets will be led by mainstream players with the ability to deliver and support products globally.

Storage Virtual Machines Enable New Levels of Application Agility

Just as virtual machines enabled the movement of active applications between physical servers, storage virtual machines will liberate data from specific physical storage locations. These logical containers of data volumes simplify the migration of workloads between storage clusters and enable highly available storage clusters in metro area.

40 GbE Adoption Takes Off in the Data Centre

The next evolution of Ethernet, 40 GbE, begins seeing widespread adoption at the core of the data centre. Higher bandwidths allow larger data sets to move more quickly and easily, encouraging the growth of data.

Continued Momentum for Clustered Storage, Converged Infrastructure, Object Storage and In-Memory Databases

Important technology trends continue to grow. Clustered storage adoption accelerates. Converged infrastructure becomes the most compelling building block of data centre infrastructure. Adoption of object storage grows as applications that monetize vast capacities of data objects gather momentum. And in-memory databases, led, for example, by the popularity of SAP® HANA, enter the mainstream.

WRITTEN TASK

1 **Explain** why **backing storage** is needed.

2 **Give two costs** associated with **local storage**.

3 **Suggest two advantages** and **one disadvantage** of using **cloud storage**.

4 **What** system do computers use to **store data**?

5 **Put measurements of storage capacity** in order, **from smallest to highest.**

6 **Which backing storage device**, offers faster access to data than any other storage device?

7 **Give two backing storage devices** which use **magnetic storage**.

8 **Which backing storage devices** use **optical storage**?

9 **What type of backing storage** is a **USB flash drive**?

10 **Give two** peripheral devices which use **flash memory cards**.

11 **Explain** the term ***portability***, when applied to **backing storage devices**.

12 **What** is **data transfer speed** generally measured in?

13 **List** typical **backing storage capacities** for:

(a) Hard Disk
(b) CD-R
(c) DVD-R
(d) USB Flash Drive
(e) Flash Memory Card

NETWORKS

A **network** is formed when **two or more computers are linked together**. A computer which is **not** part of a network is called a **stand-alone computer**.

Various **types of network** can be formed, including:

- Client/server
- Peer-to-peer

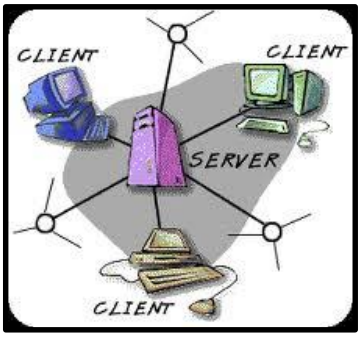
Networks can be **connected by wires, be wireless, or optically**.

Whichever type of **network** is formed, the **advantages of networking** include:

- Sharing of data.
- Sharing of expensive peripheral devices, for example coloured laser printers.
- Communication between users on the network.
- Sharing of software (if the necessary licences are in place).

Client/server Networks:

A **client-server network** is formed when the **client** (workstation) is connected to and makes use of, the resources available on the **server(s).** The **client** (workstation) requires a **NIC (Network Interface Card)** to connect it to the network.



Peer-to-Peer Networks:

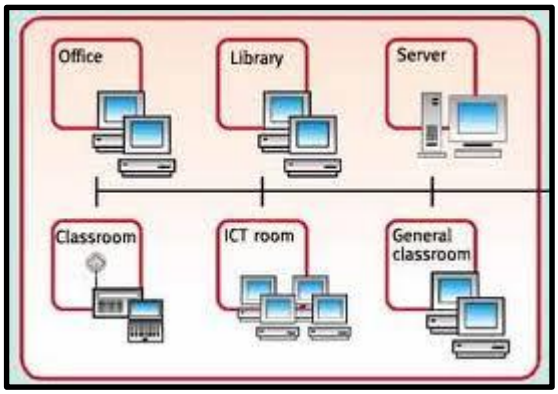
A **peer-to-peer network** is created when two or more computers are connected and share resources **without going through a separate server computer**. A **peer-to-peer network** can be as simple as a couple of computers connected via a Universal Serial Bus to transfer files, or can be a permanent infrastructure that links a half-dozen computers in a small office over copper wires.

The main **advantage of having a peer-to-peer network** is that you do not have the expense of buying a file server, as all the computers on the network have equal status. However **disadvantages** include, making backups and file security, which are more difficult to implement than on a client/server network.

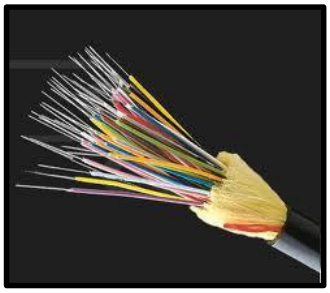
Local Area Networks (LAN):

A **Local Area Network (LAN)** covers a small geographical area such as a classroom or school building.

The **transmission media**, ie the medium which carries the data within a network can be either copper or fibre optic cable or be wireless.



Fibre optic cable or **copper cable** can be used **to connect the stations**. The signals carried by **copper cable** are subject to interference, whereas **fibre-optic** signals are not. **Fibre-optic cable** also has a higher bandwith – meaning it carries more data per second than **copper cable** does. Using **fibre optic cables** is sometimes called **optical networking** – meaning it uses pulses of light to transmit data on the network.



Fibre Optic Cable

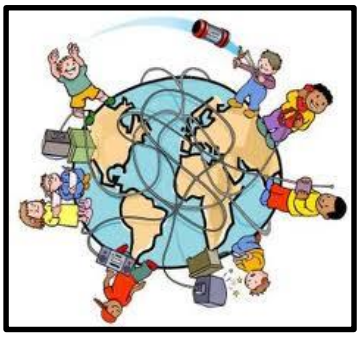
Copper Cable

Bluetooth and Wi-Fi can be used for **wireless connection in a LAN**.
Bluetooth makes short-range, temporary links between devices like mobile phones, palmtop and laptop computers.

Wi-Fi is used to **wirelessly** connect devices to one another. The range is usually up to 200 metres and is used for routers, laptop computers, games consoles etc.

Wide Area Networks (WAN):

A **Wide Area Network (WAN)** covers a larger geographical area than a LAN – typically a country or a continent.



A **Wide Area Network (WAN)** can be used for:

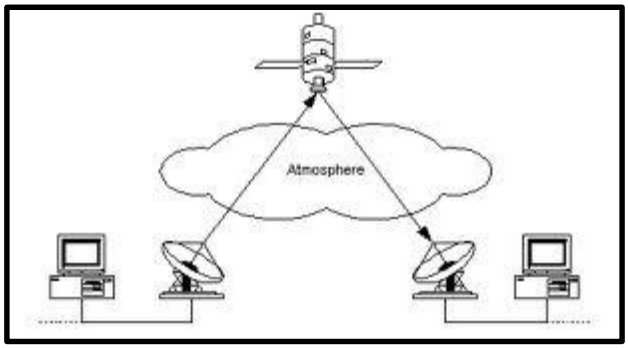
- Data sharing
- Video conferencing
- E-mail
- Internet

The Internet is an example of a **WAN** which spans the glob – the name Internet coming from Interconnecting Networks. Each network can be located anywhere in the world and include: multinational companies, governments, individuals using an **Internet Service Provider (ISP).**

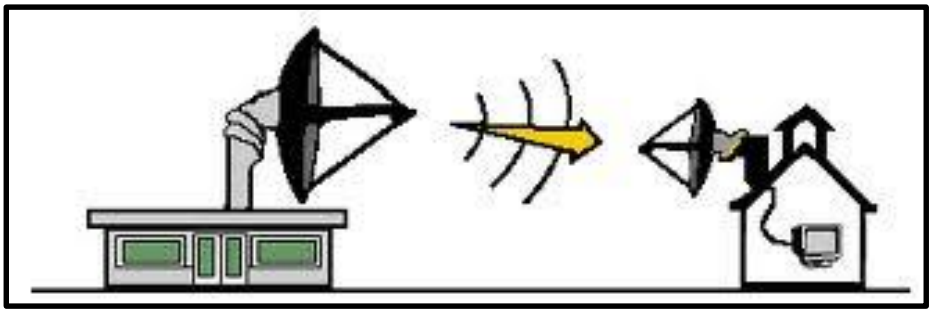
Telecommunication links are used to transmit and receive data in a **WAN**. These links include **fibre optic, satellite links and microwave transmission**.

Fibre optic cables for **WANs** are frequently laid on the seabed.

A single satellite channel is capable of carrying large numbers of separate transmissions.



Microwave systems use dish aerials to transmit data between locations.



Bandwidth:

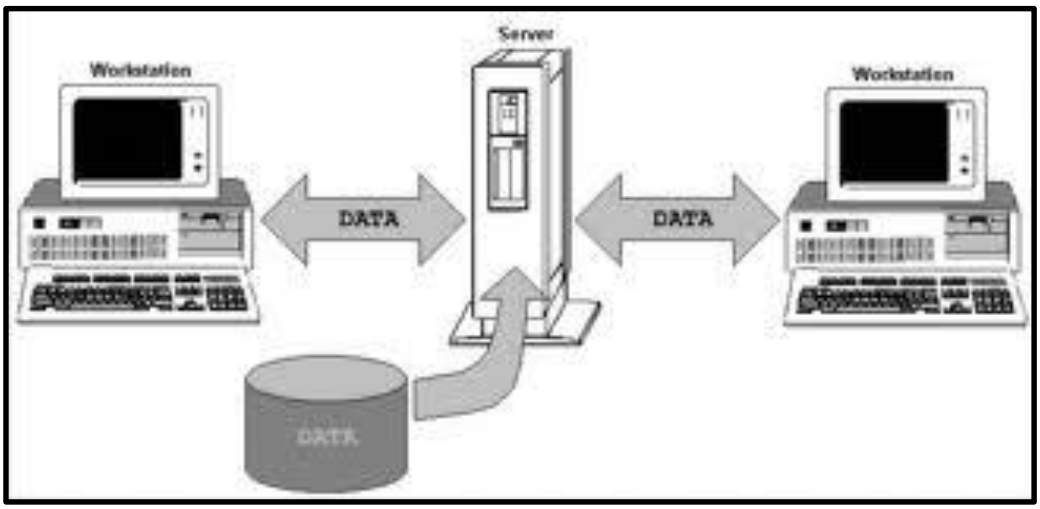
Bandwidth is the measure of the **amount of data which can be carried by the communication channel at any one time** on both **LANs and WANs.**

Bandwidth is measured in either Megabits per second (Mbps) or Gigabits per second (Gbps).

Types of Server:

A **file server** is used on a **client/server network** to provide centralised storage for user's data and programs.

Characteristics of a **file server** include: large amount of RAM, backing storage and fast multiple processors – so that users can access it quickly.

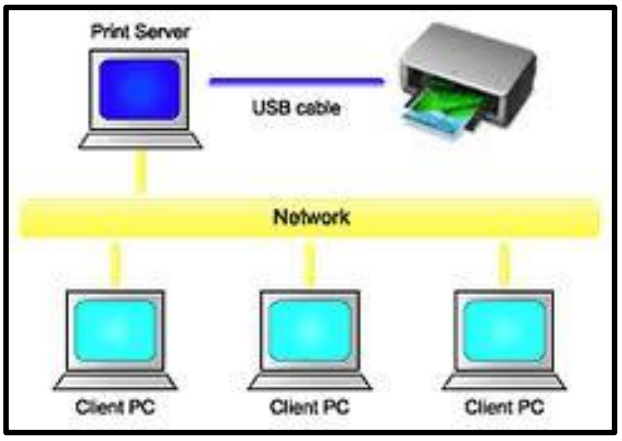


Users on a network must identify themselves to the **file server** by logging on, using their identity (username) and password. Users must always remember to log off when they have finished using the network, to keep their user account secure.

A **file server** will have a battery power supply back-up in case of power cuts and usually a **magnetic tape drive** to keep **backup** copies of users' data and programs.

A **printer server** is used to allow all the clients (workstations) on a network access to **printers controlled by the printer server**.

The **printer server** also provides a facility for queuing up printing requests so that they are printed out in turn. It will have its own processor and enough RAM to hold print jobs which are sent to it.



In a **peer-to-peer network**, the computers use a hub to connect to available printers.

Web Server:

The term **web server** can refer to either the hardware (the computer) or the software (the computer application) that helps to deliver web content that can be accessed through the Internet.

The most common use of **web servers is to host websites**, but there are other uses such as gaming, data storage or running enterprise applications.

A **web server** is used to provide **web pages** requested by clients (computers), using the **Hypertext Transfer Protocol (HTTP).** This means delivery of HTML documents and any additional content that may be included by a document, such as images, style sheets and scripts.

A **web browser or web crawler**, initiates communication by making a request for a specific resource using HTTP and the server responds with the content of that resource or an error message if unable to do so.

While the primary function is to serve content, a full implementation of HTTP also includes ways of receiving content from clients. This feature is used for submitting web forms, including uploading of files.

Web Hosting

This is a service used by individuals or organisations to make their web pages accessible using the World Wide Web. **Web hosting** is often provided free of charge by your ISP (internet Service Provider) and many of them will also offer data space in cloud storage. Large web sites, for example those used for e-commerce could require a database to install and run scripts to deliver complex functionality and to manage content.

Cloud Systems

Cloud storage systems can be **public, private or hybrid**. What they all have in common is that they are storage systems which store data on a remote computer which is accessed via the internet. The **main advantages** of this are:

- Users can access their data from anywhere, providing they have access to the internet.
- There is no data limit to storage space in cloud systems, this is specified by individual companies.

The **main disadvantage** of using cloud storage is that the speed of accessing data depends on the internet connection and it can therefore by time consuming when you require to upload or download a large amount of data.

Public Clouds

Public cloud storage may be used by people collaborating on projects because it provides a storage service that publicly hosts and manages data storage for numerous different users. Therefore, whilst it is an **advantage** to use **public cloud services for sharing resources,** it is also a **disadvantage because it is not as secure as you may wish.**

Private Clouds

Whilst **private cloud storage** works in much the same way as public cloud storage, the added advantage is that as they are privately owned by individual organisations, they can provide for a **high level of security**.

Hybrid Clouds

As the name suggest, **hybrid cloud storage** combines both public and private cloud storage.

The **main advantages** of this are, where data is not required to be kept private, organisations can use public cloud storage to reduce costs and concentrate private data, which they are required to legally keep secure and private into private cloud storage.

Current trends in Networking & Connectivity

Demand continues to increase for portable computing devices, such as smartphones, tablets etc. and the media files such as music, images and video – which have large file sizes. This has demanded a steady increase **in bandwidth requirements** and their connections. For example, already mobile network development has introduced 4G which offers 2 to 12 Mbps (Megabits per second) compared to the transfer rate for 3G which is 0.5 to 1.5 Mbps.

Optical cables are now being more widely used as **transmission media** and the use of wireless connections is increasing all the time, although this is still slower as a transmission media than wiring is.

The trend in **hardware** is also towards larger capacity main memory and backing storage devices, couples with faster processors, which have and will continue to, improve the performance of computer networks.

WRITTEN TASK

1 **What** is a **network**?

2 **What** is the **difference** between a LAN (local area network) and a WAN (wide area network)?

3 **Describe** how the Internet operates.

5 **Why** do you need an ISP (Internet Service Provider) to access the Internet?

6 **Describe two** types of transmission medium and give an **advantage** for each.

7 **Explain** what WiFi is.

8 **Explain** what Bluetooth is.

9 **What** is bandwidth?

10 **Give three advantages** of using a LAN.

11 **Which** has the **higher bandwidth** – copper cable or fibre optic cable?

12 **What** is a client-server network?

13 **What** does a client need in order to enable it to connect to a network?

14 **Give two** types of server.

15 **Describe** what can be used to aid network security.

16 **Which** type of server aids ques?

17 **How** could security be controlled on a network?

18 **What** is a file server used for?

19 In **what ways** does a file server differ from a desktop computer?

20 **Explain** how a printer server operates.

21 **What** is the primary function of a web server?

SECURITY RISKS

At the same time as the amount of computers accessing and using networks grows, so does the **security risks** associated with computers.

Security risks include:

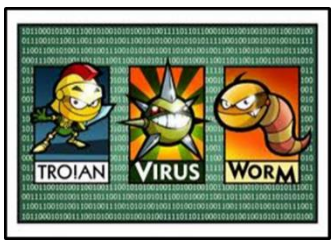
- Viruses
- Worms
- Trojans
- DOS (Denial of Service) Attacks
- Spyware
- Keylogging
- Phishing
- Identity Theft
- Online Fraud

Viruses:

A **virus** is a computer program which has been deliberately created to cause damage or disruption to a computer system. When a virus program is executed it performs the following actions:

- Replication – copies itself to other files, particularly to .com and .exe. Files and the boot sector record.
- Camouflage – attempts to disguise itself to avoid detection from ant-virus software.
- Events Watching – when the virus runs it checks for specific events, eg dates which activate the virus.
- Delivery – this is what the virus does – could be anything from mixing of letters to wiping the entire hard disk drive.

Viruses can be spread by opening e-mails or e-mail attachments. This is why you should never open e-mails from someone you do not know. Portable backing storage devices and their associated media, can also allow viruses to be spread. For example, a "homemade" CD or DVD, created on an infected computer, can allow the virus to spread to another computer, when opened there. A similar effect can happen with a USB Flash Drive. You must be very careful that you only download music, videos, software (including games), from reputable websites, as viruses can easily be downloaded over the Internet, without you realising the risk that comes from using "fun" websites.



Worm:

A **worm** is a malware computer program that replicates itself in order to spread to other computers. Often, it uses a computer network to spread itself, relying on security failures on the target computer to access it. Unlike a computer virus, it does not need to attach itself to an existing program. Worms almost always cause at least some harm to the network, even if only by consuming bandwidth, whereas viruses almost always corrupt or modify files on a targeted computer.

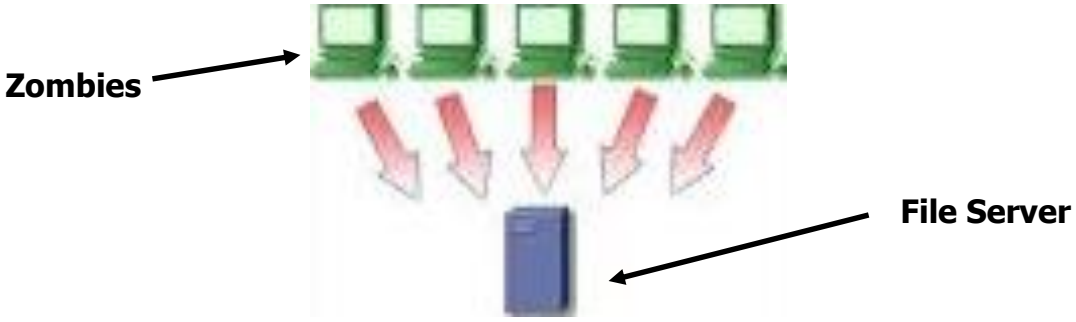
Trojan:

A **Trojan** is a program that pretends to be a helpful program. Unlike viruses, **Trojans** do not replicate themselves but they can be just as destructive. One of the most menacing types of **Trojan** is a program that claims to rid your computer of viruses but instead introduces viruses onto your computer.

DOS (Denial of Service) Attacks:

DOS Denial of Service attacks occur when networks are targeted to try and bring them to a standstill.

DOS attacks can either target a particular device on the network, for example the file server by using a **malware program such as a worm**. Or, they may make so many bogus requests to the network, using what is often termed "zombie " computers, that there is no resources left for use by genuine users.



DOS attacks can cost organisations a lot in terms of wasted manpower, loss of business and damaged reputation.

DOS attacks violate the acceptable use policies of virtually all Internet service providers (ISPs). They also usually represent breaches of the laws of individual countries.

Spyware:

Spyware is software that secretly gathers user information through the user's Internet connection without his or her knowledge, usually for advertising purposes. **Spyware** applications are typically bundled as a hidden component of freeware or shareware programs, that can be downloaded from the Internet; however, it should be noted that the majority of shareware and freeware applications do not come with spyware. Once installed, the **spyware monitors user activity on the Internet and transmits that information in the background to someone else. Spyware can also gather information about e-mail addresses and even passwords and credit card numbers.**

Keylogging:

A **keylogger** is a hardware device or a software program that records the real time activity of a computer user, including the keyboard keys they press.

Keyloggers are used by IT departments to troubleshoot technical problems with computers and business networks. **Keyloggers** can also be used by organisations to monitor the network usage of people without their direct knowledge.

Keyloggers are sometimes part of malware packages downloaded onto computers without the owners' knowledge. Malicious individuals may use **keyloggers** on public computers to steal passwords or credit card information.

Phishing:

Phishing is usually attempted by sending an e-mail to a user falsely claiming to be an established legitimate organisation, in an attempt to trick the user into surrendering private information that will be used for **identity theft**. The e-mail directs the user to visit a Web site where they are asked to update personal information, such as passwords and credit card, social security, and bank account numbers, that the legitimate organization already has. The Web site, however, is bogus and set up only to steal the user's information. **Phishing is a form of hacking used to "fish" for personal information.**

Identity Theft:

Identity theft occurs when someone pretends to be someone else, by assuming that person's identity, often in order to access resources or obtain credit and other benefits in that person's name. The **victim of identity theft (the person whose identity has been assumed by the identity thief)**, can suffer adverse consequences if they are held accountable for the thief's actions. **Identity theft** occurs when someone uses another's personally identifying information, like their name, identifying number, or credit card number, without their permission, to commit fraud or other crimes.

Online Fraud:

Online fraud is the term used to cover all crimes committed using the Internet.

Online fraud can be committed in several ways, some of which include:

- By using **hacking** into a computer system and altering or stealing data held within the system.
- By using a **keylogger** to steal passwords to online account or bank card numbers.
- **Identity Fraud**.
- **Online shopping** fraud.

The frauds listed above are just some of a very extensive lists of frauds that can be carried out over the Internet.

WRITTEN TASK

1 **What** is a **virus**?

2 **Suggest two** ways in which **viruses can be spread**.

3 **Explain** the **difference** between a **virus** and a **worm**.

4 **How** can a **Trojan** fool you into downloading it onto your computer?

5 **What** is the main purpose of a **DOS (Denial of Service) Attack**?

6 **What** would **"zombie"** computers be used for?

7 **Give** the name of the software which **secretly gathers information** through a user's Internet connection.

8 **What** does a **keylogger** do?

9 **Which** form of **hacking** is used to fish for personal information?

10 **Explain** the term ***identity theft***.

11 **Give two** types of **online fraud**.

SECURITY PRECAUTIONS

Given the numerous **security risks** that are associated with using a computer or any other device which has access to the Internet, it is important that both individuals and organisations take **security precautions to ensure the safety of their identity related data etc.**

Security precautions which can be taken include:

- Anti-virus software.
- Passwords/encryption.
- Biometrics.
- Security protocols and firewalls.
- Use of security suites.

Anti-virus software:

Every computer system should be protected by installing and using an **Anti-Virus program. Anti-Virus software** is an example of a **utility program**. Anti-Virus software is designed to **detect and remove viruses, worms and Trojans from computer systems.** It should also scan anything being downloaded to check that the item is virus free. Anti-virus software must be updated at regular intervals to ensure that it can detect and remove the new viruses that are being generated all the time.

Passwords:

The security of both standalone computers and networked computers can be improved by using **passwords**. A **password** is a secret word or string of characters used for user authentication to prove identity or access approval to gain access to a resource, which should be kept secret from those not allowed access.

The **most secure passwords** are usually formed from a combination of letters (both upper and lower case), numbers and other characters.

Security Protocols:

Network **security protocols** are used to protect computer data and communication in transit. The primary tool used to protect information as it travels across a network is cryptography. Cryptography uses algorithms to **encrypt data so that it is not readable by unauthorized user**s. Generally, cryptography works with a set of protocols that manage the exchange of data between devices and networks. Together, these cryptographic protocols enhance **secure data transfer**.

Without cryptographic **network security protocols**, Internet functions such as e-commerce would not be possible. Secure communication is necessary because hackers try to eavesdrop on communications, modify messages in transit, and hijack exchanges between systems. **Some of the tasks networks security protocols are commonly used to protect are file transfers, Web communication, and Virtual Private Networks (VPN).**

The most common method of transferring files is using **File Transfer Protocol (FTP).** A problem with **FTP** is that the files are sent in cleartext, meaning that they are sent unencrypted and therefore able to be compromised. For example, many webmasters update their sites using **FTP**; an attacker using a packet sniffer and the website's IP address can intercept all communications between the webmaster and the site's server.

As an alternative, **Secure File Transfer Protocol (SFTP)** offers a more secure way to transfer files. **SFTP** is usually built upon Secure Shell (SSH) and is able to encrypt commands and data transfers over a network, thereby reducing the likelihood of interception attacks. The SSH cryptographic protocol is also resilient to impersonation attacks because the client and server are authenticated using digital certificates.

Secure Sockets Layer/Transport Layer Security (SSL/TLS) can be used as the underlying protocol for **SFTP**. **SSL/TLS** authenticates the identity of both the server and the client, as well as encrypts communications between the two. In addition to securing **SFTP file transfers, SSL/TLS is used for securing e-mail communication.**

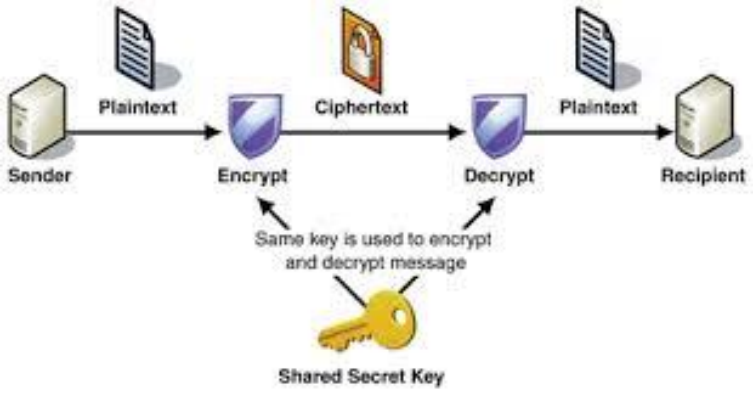
SSL is also used in combination with **Hypertext Transfer Protocol (HTTP)** to encrypt communications between a browser and a web server in the form of **HTTP over Secure Sockets Layer (HTTPS**). **HTTPS** encrypts communications and verifies the identity of a web server. When performing private transactions over the Internet, such as online banking, **it generally is good practice for a person to check the browser's address bar to make sure that the website's address begins with https:// and not just http://.**

Encryption:

When we use the Internet or any other communication network, a great deal of our time online involves sending others our own i**nformation**. Ordering something over the Internet, whether it's a book, a CD or anything else from an online vendor, or signing up for an online account, requires entering in a good deal of *sensitive personal information.* A typical transaction might include not only our names, e-mail addresses and physical address and phone number, but also passwords and personal identification numbers (PINs). All reputable organisations which use networks will have various security measures in place to stop hackers obtaining your personal information. **The most popular forms of security all rely on encryption, the process of encoding information in such a way that only the person (or computer) with the key can decode it.**

Symmetric Encryption

Symmetric encryption encrypts data by using a secret key and any encryption algorithm, which is held by both the sender and receiver and kept secure. The sender uses the **encryption algorithm** to process the data with the secret key and this produces ciphertext which is then transmitted to the receiver. The receiver then applies the secret key to the ciphertext, using the **decryption algorithm** and receives the decoded data. The advantage of **using symmetric encryption** is that it is fast and simple to use, however security issues can arise over the exchange of the secret key on what may well be insecure channels.



Asymmetric Encryption

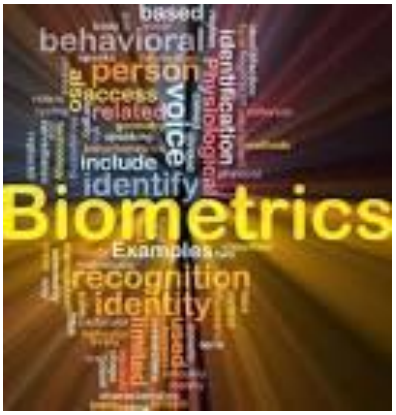
This is a more complex method of encryption which uses **both a private and public key to encrypt data.** If a public key is used to encrypt data then the ciphertext can only be decoded by using the private key and encryption algorithm. The opposite is the case if a private key is used to encrypt the data – it needs the person's public key, as a security measure to ensure the message is actually from who claims to have sent it. Because of these extra security measures, asymmetric

encryption is much more secure then symmetric encryption, but slower to implement.

Biometrics:

Some computer networks use **biometric security** to ensure only authorised users can gain access. **Biometrics** is authentication techniques that rely on measurable physical characteristics that can be automatically checked. There are several types of biometric identification schemes:

- Face - the analysis of facial characteristics.
- Fingerprint - the analysis of an individual's unique fingerprints.
- Hand geometry - the analysis of the shape of the hand and the length of the fingers.
- Retina - the analysis of the capillary vessels located at the back of the eye.
- Iris - the analysis of the coloured ring that surrounds the eye's pupil.
- Signature - the analysis of the way a person signs his name.
- Vein - the analysis of pattern of veins in the back if the hand and the wrist.
- Voice - the analysis of the tone, pitch, cadence and frequency of a person's voice.



Server-side Validation of Online Form Data

Server side validation involves passing and checking form values on the server after pressing the "Submit" button. **Server side validation** will work regardless of whether JavaScript is enabled on the client browser or not. **Server side validation** offers more security over data than client side validation.

You need to validate form data on the server (with PHP) as well as on the client (with JavaScript) for three reasons:

There are some checks you can only do on the server.

For example, suppose you have an online store selling dog toys. A user enters the number of toys he wants, if the toy is out of stock you would want to tell the user that.

Another reason to implement server-side validation would be for **security.** Every page a user sees in his or her browser is downloaded to his or her computer. That includes the JavaScript that has the validation code.

A clever hacker might be able to create a new version of your page, without the JavaScript checking. The user could then fool your server into accepting invalid data. This could mean that you get bad data in your databases. This can mess up sales, event registration, or whatever business your Website supports.

Firewalls:

Firewalls are used to stop hackers and viruses from accessing a computer.

Firewalls can be either **hardware or software-based**. A **router is a good example of a hardware device that has a built-in firewall.** Most routers can be configured to limit traffic from certain IP addresses or block requests based on other criteria. **Software programs that monitor and restrict external access to a computer or network can also serve as firewalls.** A network firewall only allows authorized traffic from the Internet to flow in and out of the network.

Security Suites:

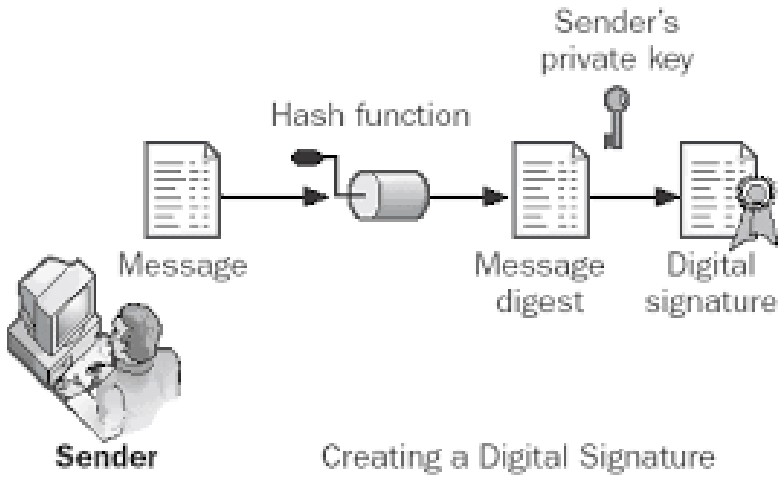
Security suites combine utility programs such as anti-virus, firewalls, anti-spyware and other malicious software, together with parental controls and password storage to protect your computer and/or network from attack.

Digital Certificates and Signatures

A **digital certificate** is a digital form of identification, like a passport. A digital certificate provides information about the identity of an entity such as the user's name and the user's public key. The entity could be a person, computer or an organisation and enables secure communication over the internet. A digital certificate is issued by a Certification Authority (CA), which acts as the middleman for sending and receiving users. Examples of trusted CA across the world are Verisign, Entrust, etc. The CA guarantees the validity of the information in the certificate.

A **digital signature** is an electronic, encrypted, stamp of authentication on digital information such as email messages, macros, or electronic documents. A signature confirms that the information originated from the signer and has not been altered.

To create a **digital signature**, you need a signing certificate, which proves identity. When you send a digitally-signed macro or document, you also send your certificate and public key. Certificates are issued by a certification authority, and like a driver's license, can be revoked. A certificate is usually valid for a year, after which, the signer must renew, or get a new, signing certificate to establish identity.



Creating a Digital Signature

WRITTEN TASK

1 **Explain** the purpose of **anti-virus software**.

2 **What** should be included in a **password** to make it very secure?

3 **Describe** what happens when data is **encrypted**.

4 **Why** is data **encrypted**?

5 **Describe three** methods of **biometric security**.

6 **What** are **network security protocols** used for?

7 **Which security protocol** authenticates the identity of **both** the server and the client?

8 **What** should the URL **start with** if it encrypts data communications and verifies the identity of a web server?

9 **Explain** what a **firewall** is.

10 **Give** an example of a piece of **hardware** which might form part of a firewall.

10 **Suggest three utility programs** what you would expect to find in a **security suite**.

COMPUTERS AND THE LAW

The Data Protection Act:

Because of the ever increasing amount of data that is held by organisations about people, the government passed The Data Protection Act to protect peoples' privacy.

The main principles of the Act include:

- Personal data shall be processed fairly and lawfully.
- Personal data shall be obtained only for one or more specified and lawful purposes.
- Personal data shall be adequate, relevant and not excessive in relation to the purpose or purposes for which they are processed.
- Personal data shall be accurate and, where necessary, kept up to date.
- Personal data processed for any purpose or purposes shall not be kept for longer than is necessary for that purpose or those purposes.
- Personal data shall be processed in accordance with the rights of data subjects (individuals).
- Appropriate technical and organisational measures shall be taken against unauthorised or unlawful processing of personal data and against accidental loss or destruction of, or damage to, personal data.
- Personal data shall not be transferred to a country or territory outside the European Economic Area unless that country or territory ensures an adequate level of protection for the rights and freedoms of data subjects in relation to the processing of personal data.

A **Data Subject** is someone about whom personal data is held.

A **Data User** is someone within an organisation which uses personal data about data subjects.

A **Data Controller** is the person within an organisation who is responsible for how that organisation collects and uses personal data about data subjects.

Organisations who wish to hold and use personal data about data subjects, must be registered on the **Data Protection Register**.

The **Data Protection Commissioner** deals with any complaints a data subject may have about how a data user or data controller is using or holding information about them.

Any data held by Customs & Excise, security forces or the police service is exempt from the Data Protection Act – which means they do not have to let a data subject see any date they hold about that person.

WHO HAS INFORMATION ABOUT ME STORED ON A DATABASE?

Some of the organisations which keep personal information of a private and confidential nature about individual people stored in computer systems are:

- Doctor's Surgery, Local Hospital
- Driver and Vehicle Licensing Centre
- Credit and Finance Companies
- Banks and Building Societies
- Police Scotland, Special Branch, MI5
- Customs & Excise, Dept of Work & Pensions
- Employers and Education Authorities etc

KEEPING DATA SECURE

Steps that can be taken to preserve data kept on computer systems from unauthorised access or deliberate damage include:

Allow only authorised personnel into the computer room by:

- using special locks on the door which require a passcode or special card to be inserted
- putting security guards on the door
- By staff wearing identification badges with photographs
- Using biometric security to allows access to the computer rooms.

Protect the computer hardware itself by:

- using a "turnkey" system which requires a special key to operate the computer

If the system is used with communication lines, take it "off-line" when communication is not essential.

Use an operating system which implements user names and passwords (this is especially important in network systems, where there may be unsupervised terminals in many accessible places.

Protect individual files stored on disc by:

- making them "read only"
- locking them for both "read" and "write" operations
- encrypting files so that they are unreadable without a special code
- storing important files on portable backing storage only, so that they can be locked away

Freedom of Information Act (Scotland)(2002)

The Freedom of Information Act 2000 provides public access to information held by public authorities. It does this in two ways:

- Public authorities are obliged to publish certain information about their activities; and
- Members of the public are entitled to request information from public authorities.

The Act covers any recorded information that is held by a public authority in England, Wales and Northern Ireland, and by UK-wide public authorities based in Scotland. Information held by Scottish public authorities is covered by Scotland's own Freedom of Information (Scotland) Act 2002.

Public authorities include government departments, local authorities, the NHS, state schools and police forces. However, the Act does not necessarily cover every organisation that receives public money. For example, it does not cover some charities that receive grants and certain private sector organisations that perform public functions.

Recorded information includes printed documents, computer files, letters, emails, photographs, and sound or video recordings.

The Act does not give people access to their own personal data (information about themselves) such as their health records or credit reference file. If a member of the public wants to see information that a public authority holds about them, they should make a subject access request under the Data Protection Act 1998.

Regulation of Investigatory Powers Act (2000)

RIPA regulates the manner in which certain public bodies may conduct surveillance and access a person's electronic communications. The Act:

- enables certain public bodies to demand that an ISP provide access to a customer's communications in secret;
- enables mass surveillance of communications in transit;
- enables certain public bodies to demand ISPs fit equipment to facilitate surveillance;

- enables certain public bodies to demand that someone hand over keys to protected information;
- allows certain public bodies to monitor people's Internet activities;
- prevents the existence of interception warrants and any data collected with them from being revealed in court.

Communications Act:

The purpose of **The Communications Act 2003** is the regulation of the provision of electronic communications networks and services.

The Act introduced new offences for 'Improper use of public electronic communications network', 'dishonestly obtaining electronic communications services' 'possession or supply of apparatus etc for contravening', and includes information disclosure offences.

Copyrights, Designs and Patents Act:

The **Copyright, Designs and Patents Act 1988**, is the current UK copyright law. It gives the creators of literary, dramatic, musical and artistic works (including software), the right to control the ways in which their material may be used. The rights cover: Broadcast and public performance, copying, adapting, issuing, renting and lending copies to the public.

Copyright arises when an individual or organisation creates a work, and if it is regarded as original, and exhibits a degree of labour, skill or judgement.

Normally the individual or collective who authored the work will exclusively own the rights. However, if a work is produced as part of employment then normally the work belongs to the person/company who employed the individual. For freelance or commissioned work, rights will usually belong to the author of the work.

The main parts of the Act which apply to computers are making it illegal to:

- Make unauthorised copies of software that is copyright protected.
- Copy and transmit software over a network without permission of the licence holder.
- To install and run unauthorised copies of software on a computer network.

It is also important to note, that you must never use photographs, drawings, music, pieces of text in any works or documents, which you will distribute for profit, without first gaining the permission of the copyright holder.

WRITTEN TASK

1 **Give** the names of **three** computer related laws.

2 **Which** law does a **hacker** break?

3 **Give three** examples of activities which may breach **copyright law**.

4 **What** is **phishing**?

5 **What** act governs the storage and **use of personal data**?

6 **What** is a person who has **data held about them** called?

7 **Give** an example of **intellectual property** and say which Act governs it.

8 **What** does a **Data Controller** do?

9 **Which** Act would you be breaking if you **deliberately spread a computer virus?**

10 **Give two rights** which employees have under **Health and Safety regulations**.

11 **Suggest** an injury which might happen to an employee who did not have a **correctly designed workstation**.

12 **Give one** example of something which would contravene the **Communications Act.**

COMPUTERS AND THE ENVIRONMENT

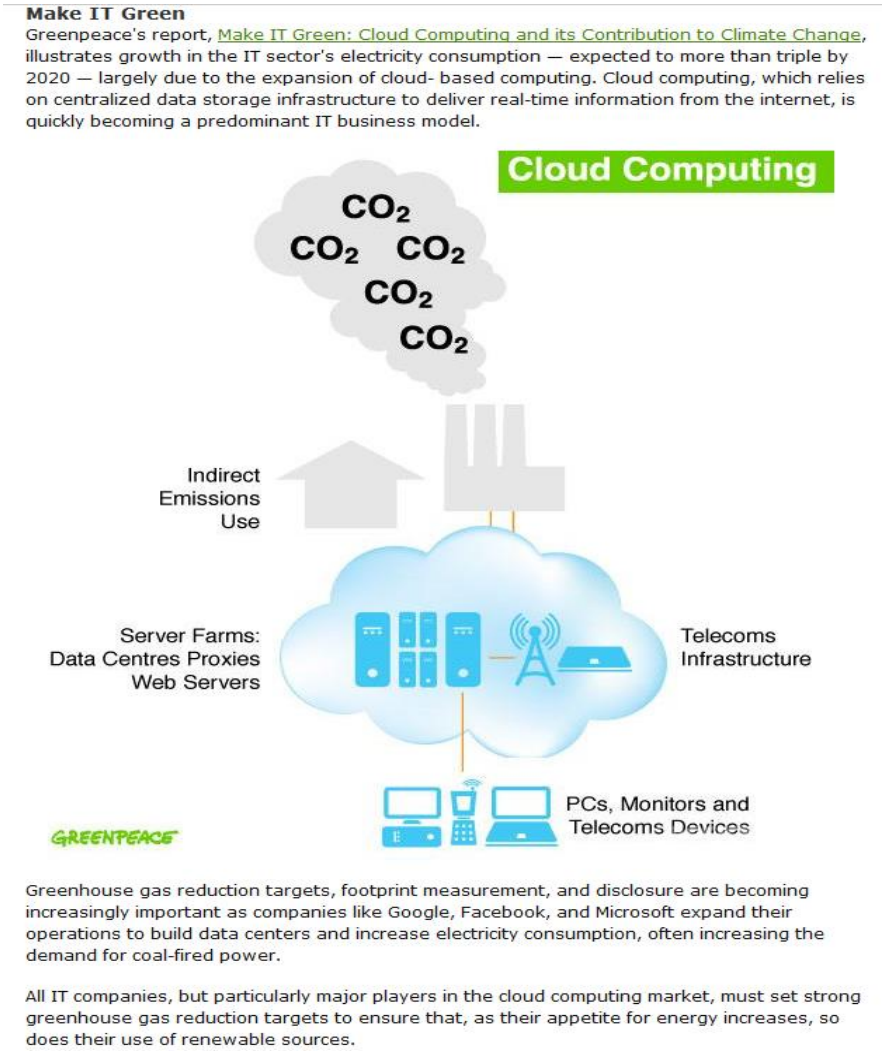
How we manufacture, use and dispose of computer equipment has an impact on the environment.

The measurement of greenhouse gasses produced from the manufacture and use of computing equipment is known as a **Carbon Footprint.**

Lifetime Carbon Footprint

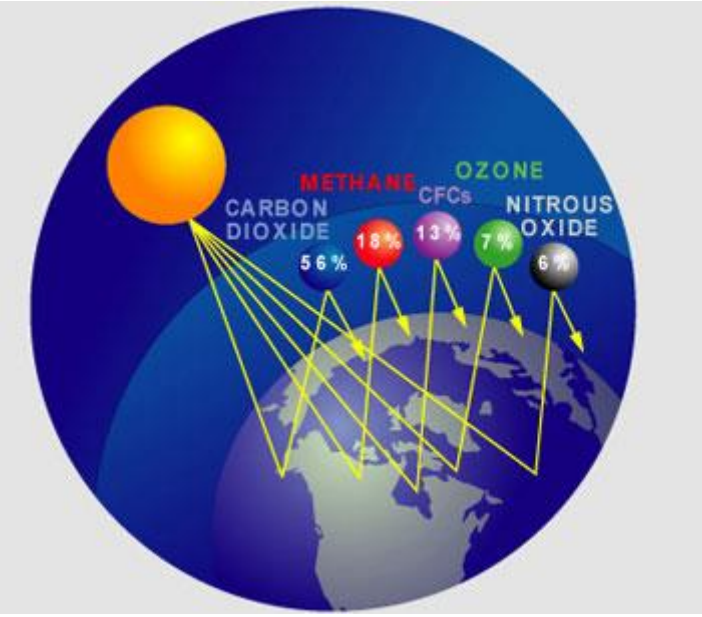
The main topics to consider in a person or companies carbon footprint are:

- Manufacture
- Use
- Disposal



A **carbon footprint** is a measurement of greenhouse gas emissions caused by an organization, event, product or person. **Greenhouse gas** is created by the production and use of computer equipment.

Greenhouse gas is a gas in the atmosphere that absorbs and emits radiation within the thermal infrared range. This process is the fundamental cause of the greenhouse effect. The main **greenhouse gases** in the Earth's atmosphere are water vapour, carbon dioxide, methane, nitrous oxide, and ozone.



Energy use:

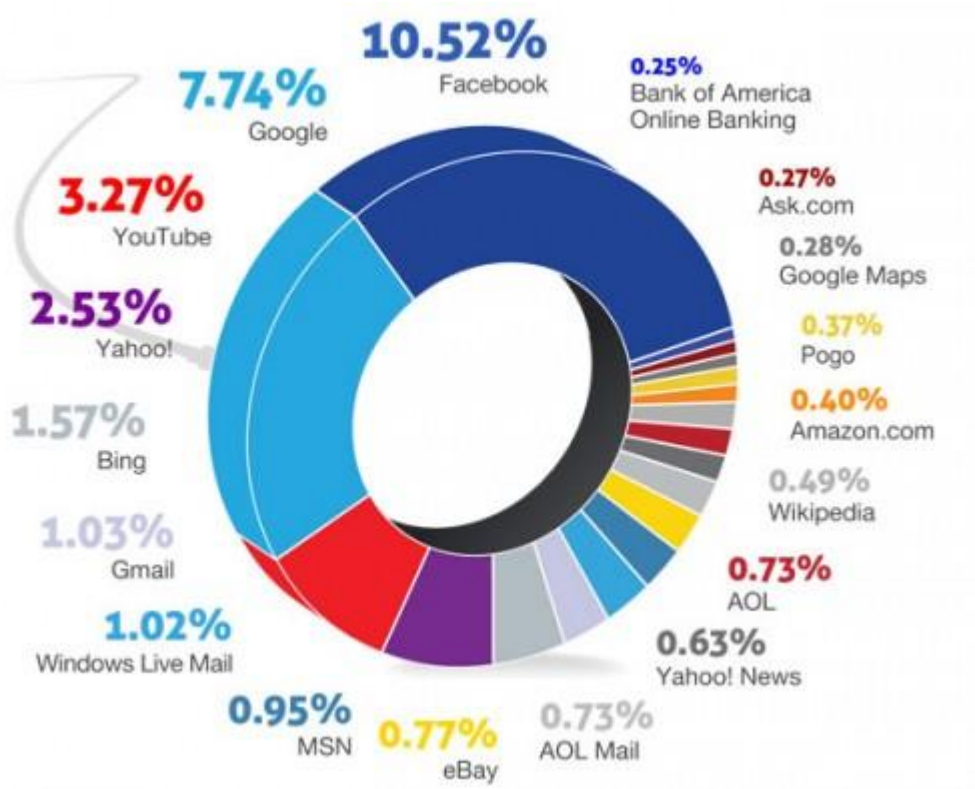
A United Nations research group has claimed that the **manufacture** of an average PC requires 10 times the weight of the product in chemicals and fossil fuels.

Many of the chemicals are toxic, while the use of fossil fuels help contribute to global warming.

Manufacturing a 24kg PC with monitor needs at least 240kg of fossil fuels to provide the energy, and 22kg of chemicals.

Using a **search engine** such Google to conduct a simple search can result in 1-10 grams of CO2 emissions.

It has been estimated that if the Internet was a country it would be the fifth biggeest energy user in the world.



Most computers create 40-80 grams of greenhouse gas emissions per hour through their **electricity use** (depending on electricity source and computer type), so the **aggregated greenhouse gas emissions just from computers is quite sizable**. Added to this is the energy used by the servers and fibre optic cabling.

Although the manufacture and use of computer equipment produces **greenhouse gases, it can also be argued that using computer equipment to enable homeworking and teleconferencing, can also contribute to the reduction of greenhouse gases** by reducing the amount of travel which people are required to do. Cutting down the amount of journeys taken, **reduces carbon emissions**.

Disposal of IT Equipment:

WEEE regulations (The **W**aste **E**lectrical and **E**lectronic **E**quipment Regulations), make it the responsibility of companies who manufacture and sell electronic equipment to take it back and **dispose of it in an environmentally friendly way**.

This means that **WEEE** must be collected, stored recycled or disposed of, separately from other rubbish.

Businesses that sell electrical or electronic equipment must now provide a service where they will take back **WEEE**. Alternatively, they must be able to tell customers where they can take their **WEEE**. Most local authorities (councils), also have facilities at waste collection centres to keep WEE separate from general household rubbish or offer a collection and disposal service for a small fee.

Hazardous Waste:

Hazardous waste from IT equipment includes things such as:

- Antimony trioxide as flame retardant.
- Polybrominated flame retardants in plastic casings, cables and circuit boards.
- Selenium in circuit boards as power supply rectifier.
- Cadmium in circuit boards and semiconductors.
- Chromium in steel as corrosion protection.
- Cobalt in steel for structure and magnetivity.
- Mercury in switches and housing.

Waste like that listed above, can lead to major health risks if dumped in landfill sites without the damaging elements being removed.

Refurbishing an old computer can often be a better option than throwing it away.

The majority of PCs routinely disposed of for recycling by businesses and consumers alike are actually far from their real end-of-life and could go on to give as much as 6,000 additional hours of use.

Identity Theft:

Whether IT equipment is to be disposed of in compliance with WEEE regulations or recycled, there are important steps that should be taken to ensure **all data which could lead to identity theft is removed from the equipment first.**

This means that all hard disk drives should be removed and destroyed, as even reformatting them does not mean that a thief will not be able to extract personal information from them.

Similar steps should be taken with removable storage media such as flash memory cards, SIM cards or external hard disk drives.

WRITTEN TASK

1 **Explain** what a **carbon footprint** is.

2 **Give three greenhouse gases** which are in the Earth's atmosphere.

3 **How many kilograms** of fossil fuels and chemicals are needed to manufacture a 24 kilogram PC?

4 **What is** the typical electricity usage of a computer, per hour?

5 **Give the name** of the regulations which govern the disposal of electronic and electrical equipment.

6 **List three** types of hazardous waste that can come from IT equipment.

7 **How many** additional hours of use could there be available from a recycled computer?

8 **Describe** what should be done to avoid **identity theft** when disposing of old computers and smartphones.

ECONOMIC & SOCIAL IMPACT

An information system is an integrated set of components for collecting, storing, and processing data and for delivering information, knowledge, and digital products. Business firms and other organizations rely on information systems to carry out and manage their operations, interact with their customers and suppliers, and compete in the marketplace. For instance, corporations use information systems to reach their potential customers with targeted messages over the Web, to process financial accounts, and to manage their human resources. **Governments** deploy information systems to provide services cost-effectively to citizens. Digital goods, such as electronic books and software, and online services, such as auctions and social networking, are delivered with information systems. **Individuals** rely on information systems, generally Internet-based, for conducting much of their personal lives: for socializing, study, shopping, banking, and entertainment.

ECONOMIC IMPACT

Network-based information systems have been a factor in the growth of international business and corporations. A relationship between the deployment of information systems and higher productivity has been shown in a number of industries when these systems complement other corporate resources.

E-commerce has moved many relationships and transactions among companies and individuals to the Internet and the Web, with the resulting expansion of possibilities and efficiencies.

The development of the Web-based ecosystem, accompanied by the **low cost of hardware and telecommunications** and the availability of **open-source software**, has led to a flowering of entrepreneurial activity and the emergence to prominence and significant market value of numerous firms based on new business models.

- Among the examples are:
- Electronic auction firms
- Search-engine firms
- Social network platforms,
- Online game companies.

Owing to the vast opportunities for moving work with data, information, and knowledge in electronic form to the most cost-effective venue, a **global redistribution of work has been taking place**. However, the wide deployment of information systems on Web platforms has not positively affected job markets.

E-commerce (electronic commerce) is the process of conducting business on-line over the Internet. It includes both **on-line banking** and **on-line shopping.**

The vast majority of people who have access to a computer at home, now use **on-line banking.**

On-line banking allows someone to access their bank account from a home computer and allows them to keep track of all the transactions on their account.

The **main advantage to the customer** is that they do not have to travel to their bank branch and they can carry out their banking business at any time , 24/7, that is convenient to them.

The **main advantage to the banks** is that it has increased the use of bank cards and decreased the amount of cash that people carry, reducing the amount of cash that banks have to handle and process. It has also meant that many banks have closed branches. The **social implications** of this are that many bank staff have either lost their jobs or have had to be retrained to do other jobs, such as manning the call-centres that most banks now use to assist customers.

A major use of the Internet is now **on-line shopping**. The process begins either with the shopper using a **search engine** to search for suppliers or, if they know which supplier they want keying in the **URL (Uniform Resource Locator)** for that company.

Most **on-line shopping websites use a shopping basket or cart**, into which the items the customer selects to buy are placed. Once the customer has finished shopping they can review their basket and then go to the checkout, at which time, if they have shopped on-line with that company before they will be asked to login, usually using their e-mail address and a password created for that supplier, and check their personal details and delivery instructions before they are asked to pay for their goods. New customers will have to create account with the supplier before following the same procedure.

Paying for the goods is usually done by bank credit or debit card or using a service such as PayPal – electronic funds transfer (EFT).

Once your payment has been accepted and the purchase is complete the website will send you an e-mail to confirm your purchase and usually another one once the goods are despatched – often giving a tracking number so that you can use the Internet to progress where in the delivery cycle your goods are.

Electronic funds transfer also takes place in shops at the point of sale (the checkout). This is called **EFTPOS** (**E**lectronic **F**unds **T**ransfer at **P**oint **O**f **S**ale). This reduces the security problems which retailers have in dealing with cash.

Advantages of companies using e-commerce to conduct the bulk (or all) of their transactions, also include:

No need to buy or rent high street premises and the associated costs, such as fittings, staff and running costs.

Retailers deal directly with customers, cutting out middlemen (wholesalers), who would take some of the profit.

As overheads are reduced, retailers can make their prices more competitive and thereby increase business.

MAINTAINABILITY

As any online information system will occasionally require maintenance, for example updating product catalogues, changing the format of a website etc. it will be important that the online

facility is not offline for any longer than necessary, as this leads to lost customers. One solution for this is to have a backup web server, which can be used to run the old information system until the new one is ready to become live. Alternatively, businesses will usually choose to take down their information system during the night, giving as much advance notice to customers as possible, banks are an example of this – as so much transactions are carried out with online banking the will put a notice on the home page of the secure server so that when customers have logged in they will get notice of when and for how long the bank's system will be down and transactions will not be possible.

SCALABILITY

Scalability in relation to an information system, is how well the system could react to increasing demand. For example, a company which uses e-commerce to sell physical goods, such as music CDs or film DVDs would have to have the ability to scale up and cope with increased demands, for example at Christmas.

However, a company which sold music and films as downloads would not face the same challenges if faced with an increased demand. An additional advantage of this is that there is virtually no delivery cost associated with digital downloads – making a business such as this, very scalable.

SOCIAL IMPACT

As the use of information systems has become pervasive in advanced economies and societies at large, several **societal and ethical issues** have moved into the forefront. **The most important are issues of individual privacy, property rights, universal access and free speech, information accuracy, and quality of life.**

Individual privacy hinges on the right to control one's personal information. While invasion of privacy is generally perceived as an undesirable loss of independence, government and business organizations do need to collect data in order to enable administration and exploit marketing opportunities. E-commerce presents a particular challenge to privacy, as personal information is routinely collected and circulated in a largely unregulated manner. The ownership of and control over, the personal profiles, contacts, and communications in social networks are one example of a privacy issue that awaits resolution through a combination of market forces,

industry self-regulation, and possibly government regulation. Preventing invasions of privacy is complicated by the lack of an international legal standard.

Intellectual property, such as software, books, music, and movies, is protected, albeit imperfectly, by patents, trade secrets, and copyrights. However, such intangible goods can be easily copied and transmitted electronically over the Web for unlawful reproduction and use. Combinations of legal statutes and technological safeguards, including antipiracy encryption and electronic watermarks, are in place, but much of the abuse prevention relies on the **ethics of the user**. The means of protection themselves, such as patents, play a great role in the information society. However, the protection of business methods (e.g., Amazon's patenting of one-click ordering) is being questioned, and the global enforcement of intellectual property protection encounters various challenges.

Access to information systems over the Web is necessary for full participation in modern society. **In particular, it is desirable to avoid the emergence of digital divides between nations or regions and between social and ethnic groups.** Open access to the Web as a medium for human communication and as a repository for shared knowledge is treasured. Many people consider free speech a universal human right and the Internet and Web the most widely accessible means to exercise this right. However, legitimate concerns arise about protecting children without resorting to censorship. Technological solutions, such as software that filters out pornography and inappropriate communications, are partially successful.

The **accuracy and security of information** contained in databases and data warehouses—whether in health and insurance data, credit bureau records, or government files—as misinformation or privileged information is a concern to everyone. Data released inappropriately can adversely affect personal safety, livelihood, and everyday life. Individuals must cooperate in reviewing and correcting their files, and organizations must ensure appropriate security, access, and use of such files.

Information systems have affected the **quality of personal and working lives**. In the workplace, information systems can be deployed to eliminate tedious tasks and give workers greater autonomy, or they can be used to thoughtlessly eliminate jobs and subject the remaining workforce to pervasive electronic surveillance. Consumers can use the Web for shopping, networking, and entertainment—but at the risk of contending with spam (unsolicited e-mail), interception of credit card numbers, and attack by computer viruses.

Information systems can **expand participation of ordinary citizens in government** through electronic elections, referendums, and polls and also can provide electronic access to government services and information—permitting, for example, electronic filing of taxes, direct deposit of government checks, and viewing of current and historical government documents. More transparent and beneficial

government operations are possible by opening the data collected by and about governments to public scrutiny in a searchable and easy-to-use form. With the Web, the public sphere of deliberation and self-organization can expand and give voice to individuals. **Information systems have also conjured images of government surveillance and business intrusion into private lives**. It remains for society to harness the power of information systems by strengthening legal, social, and technological means.

GLOBAL CITZENSHIP

Because of the advances in how we can travel, for example by aeroplane people can now travel anywhere in the world relatively easily, compared to only a century ago when people tended to live and work within the community they were born into. This, coupled with the internet which allows people who live on the opposite sides of the world to communicate, means that we are now deemed to be **global citizens**.

This increasing access to the internet means people who live on the opposite sides of the world to communicate in **online communities**. Examples of online communities include social media sites such as:

- Facebook
- Twitter
- Snapchat

Other types of online communities include **online customer communities** such as:

•Avon

•Blackbaud

•Blue Jeans Network

•DreamHost

•Hewlett-Packard

•New Holland

Specialist interest online communities also exist.

Whatever the type of online community, the purpose of it is to allow members to connect, share resources and discuss issues.

WRITTEN TASK

1 **Describe** the impact **e-commerce** has had on the retail sector.

2 **Explain** effect has the introduction of online information systems has had on the **workforce**.

3 **Suggest** a **maintenance problem** which may occur with an online information system and a strategy to overcome the problem.

4 **Explain** the term **scalability**, in relation to businesses who use online information systems for e-commerce.

5 **Briefly** describe the **social impact** online information systems may have on the following:

(a) Individual privacy

(b) Intellectual property

(c) Accuracy of information

(d) Digital divisions

(e) Quality of work

6 **Explain** how we are all now regarded as **global citizens**.

7 **Describe**, using an **example**, an **online community**.