

ICT4 13.4 Data

GIGO: Remember that garbage in = garbage out.

If data is entered incorrectly then the information output will be incorrect, this is referred to as GIGO - Garbage In, Garbage Out.

Keying in Data

The drawback of manually typing in data is that it is slow, it inevitably leads to errors and there is the risk to health of people typing at a VDU all day. Data verification systems such as batch processing can be used to improve accuracy. Errors can occur at all stages of processing, especially during data entry and whenever there is human intervention. A VDU operator taking data from a piece of paper and typing it into a database is an example of data being entered indirectly. Indirect entry is more likely to lead to mistakes, due to human error.

"Transpositional" errors involve getting numbers the wrong way round e.g. keying in 58762 instead of 57862.

"Transcribing" means copying down data from a human source.

"Verification" means checking to see that something is accurate, e.g. entering a password twice and comparing the two versions to spot errors. Verification involves double-checking something.

"Data Validation" involves checks being made to prevent errors. Validation checks are usually built into the system.

Example: A web site asks you to enter your e-mail address to register. You realise that you can enter anything@anything.com and the site accepts this as a valid e-mail address simply because it contains an @ symbol. However, the site then tells you that your password has been e-mailed to you. This is a problem because you gave a false e-mail address and, therefore, you will not receive the password. This web site is performing both a validation check and then also a verification check.

Data Capture

Wherever possible, data should be captured directly in order to avoid mistakes being made in entering the data. Bar code readers and magnetic stripe readers are good examples of direct data capture methods.

Aggregated and Disaggregated Data

Disaggregated data are single pieces of data e.g.

- a person's age, gender or level of education
- a registration number of a single vehicle
- one candidate's exam result

Aggregated data is a group of entries e.g.

- The average age of employees in a store
- The total number of vehicles going down a road in a week
- The overall pass rate for an exam

Source of Data

- There are direct and indirect sources of data.
- An indirect source is where data is used for a purpose other than that for which it was originally collected
- Examples:
 - A "clocking on" system records who was at work and at what times. This data is used to generate information e.g. wage slips, staff attendance records and so on. This is the primary purpose for which the data is collected, therefore, it is a direct source of data.
 - A library might have an IT system, which collects data about which books are taken out, when and by whom. This is the primary purpose for which the data is collected.
 - When a credit card is scanned at a POS (point of sale) terminal, data is collected about the card number, the amount, the place of purchase, the till number and so on. This is the primary purpose for which the data is collected (this data is necessary to process the transaction). However, there are also secondary purposes for collecting records of credit card purchases. Can you think what these might be?

Bar Codes

Although there is an initial investment in hardware that has to be made, bar codes have such great benefits that they are increasingly being very widely used. It is estimated that a 2% investment leads to a 6% saving in costs.

Bar code readers are very accurate. As little as one mistake per 100,000 transactions.

Information can be passed to a MIS from barcode transactions, enabling them to make judgements about the effect of repositioning items in the store, identify fast or slow selling items, historical data showing seasonal fluctuations in trade in certain items, etc.

Bar coding is used in supermarkets but increasingly it is also being used in many other areas, such as:

- Warehousing (robotic cranes which stack shelves know exactly where to place goods by reading the barcode)
- Transport (packages are bar coded so that the precise location of any package can be known at any time)
- Manufacturing (data can be obtained on work in progress to improve production efficiency)
- Marketing (bar coded multiple choice questionnaires)
- Medical (bar codes are used to identify blood and other samples)
- Libraries (The bar code updates the computer system which can say whether or not a book is available)

In supermarkets, bar codes pass information to a computer system. If the customer has a "loyalty card", the supermarket can make a profile of which items a shopper buys and when. In some supermarkets, there is "self-scanning". This can enable the company to know at

exactly what time each item was put in the trolley. It can also detect which items were picked up and then put back.

EDI - Electronic Data Interchange

Business data e.g. orders and invoices can be sent from one firm's computer system to that of another firm. This means that data does not need to be keyed in twice. The transmission is virtually instantaneous and, therefore, the supplier's system can check for availability and respond with confirmation. When students sit exams, the results can now be sent directly to the schools rather than being posted.

Smart Cards

Cash loaded smart cards, such as Mondex, could replace cash. The card could be used as a form of ID. It could also be used to monitor where the holder goes, everything s/he buys and when. This could have privacy implications.

Acknowledgements

This document, by D. Yates, first appeared on the KJS website. It has been added to, but credit is given to how it appeared in the original format.

www.thekjs.essex.sch.uk/yates/