



Fig 1: Components of a Form

Normally a data capture form would be used together with a mouse and keyboard. However, in some situations, a different input device might be used. Perhaps the most obvious would be a hand-held bar code scanner. This can be plugged into a USB socket or keyboard input socket and supplies data to the computer as though it had come from a keyboard typing in the digits. The user would position the cursor within the appropriate text box on the form and then point the bar code reader at the bar code and press a switch. The reader would transmit the data to the computer and the bar code digits would appear in the text box. This

type of solution would be appropriate for, say, a small corner shop which used a personal computer to handle sales and stock control.

Other Input Methods

A data capture form cannot be used to capture sound (audio), video and graphic data and specialist equipment must be used. A personal computer will usually have a sound card with a microphone input socket. Software can be used to record sound input directly. Alternately there is a wide range of digital recording devices many of which have USB connections that allow sound files to be transferred directly from the device to a computer.

Pictorial based data can input using a scanner or, in the case of photographs, downloaded from a digital camera.

Video data can be downloaded from digital video recorders and, as is the case with sound and pictorial data, specialised software is available to edit and manipulate the files.

Audio, picture and video data all require a large amount of storage space on disk and within the computer's memory when being processed. The actual amount of data depends on the length of time the recording lasts or the size of the picture. It also depends on the quality of the recording. Increasing quality, time and size result in increased storage requirements and this type of data is often stored in compressed formats which may result in some loss of quality.

Validation and Verification

Two methods, validation and verification, are used to try to ensure that any data processed is correct. It is important to realize that, no matter what checks are in place, it is impossible to ensure that any data in a computer system is 100% correct.

When data is transcribed (copied) from one medium to another there is always a danger that errors will be introduced. The copy will not then be the same as the original. This is particularly true if the transcription is being done manually as is the case when a human operator reads a source document and uses a keyboard to enter the data into a computer system as might be the case with a data entry form.

Verification is designed to trap a transcription error; it should not be confused with validation where unreasonable data that has been correctly copied from the source is being identified and rejected. Verification will only help to identify errors made by the copying process.

Full verification on data input by a keyboard operator is done by having the data re-entered by a different operator. The computer system automatically checks for differences as the second copy is entered. Any differences must be the result of a transcription error by one of the two operators. In this case the transcribed versions can be checked manually against the source document and the error corrected. If the two operators make the same error then it will not be spotted.

Using a different operator to enter all the data a second time is costly and time consuming. If only small amounts of data are being entered – for example a customer order being copied from an order form – then a simple manual verification check would be done by proof reading the entry. The operator, having entered all the details, would then manually check what was displayed on the screen against the original order form. Of course, the data would also be validated as it was entered using one or more of the techniques described below, and any validation errors reported to the operator.

Validation is a check that is performed by the computer system to make sure that the data value is sensible and possible. Unfortunately simply because the data is reasonable and possible does not mean that it is correct. Invalid data must contain an error; valid data may contain an error. A number of possible validations checks are described below. Applications software will usually allow these to be set up easily. Bear in mind that the type of check will depend on the particular circumstances of the application and also that for some data more than one validation check may be appropriate.

Range Check

A range check can be used to validate any data value that has to lie between two fixed values. Range checks can be used on non-numeric data as long as there is natural sequence of values. For example dates of birth of students in a school year group could be range checked since they should all lay between two known dates (last day of September in one year and 31 August in the following year).

The two values may be absolute as in the case of percentage exam mark which cannot be larger than 100 or less than 0 or, where there is no absolute value, some sensible value may be chosen for one or both of the limits. An example of this might be an electric meter reading. In this case the lower limit is absolute since it must be greater or equal to the previous reading. The upper limit for the range would however have to be chosen as some value that represented more electricity than would be possible to use. Setting this limit too low would risk rejecting correct data as invalid while setting it too high would risk accepting incorrect data as valid.

Data Type Check

A data type check validates data as being of the correct data type for the field it is going to be stored in. Basic data type types checks identify that data is numeric, Boolean, a date/time or an integer as required.

This type of validation is usually automatically set up particularly by database software as long as the appropriate data types have been declared for each field (see page 98).

Check Digit

Check digits are used to validate items of data such as account numbers, ISBN numbers or bar code numbers. The data that is entered consists of the actual data plus a check digit, which is usually placed at the end of the data itself. An example is the ISBN number 0-471-25311-1 where the final one is the check digit. The value of the check digit is based on both the other digits and their positions in the number so that,