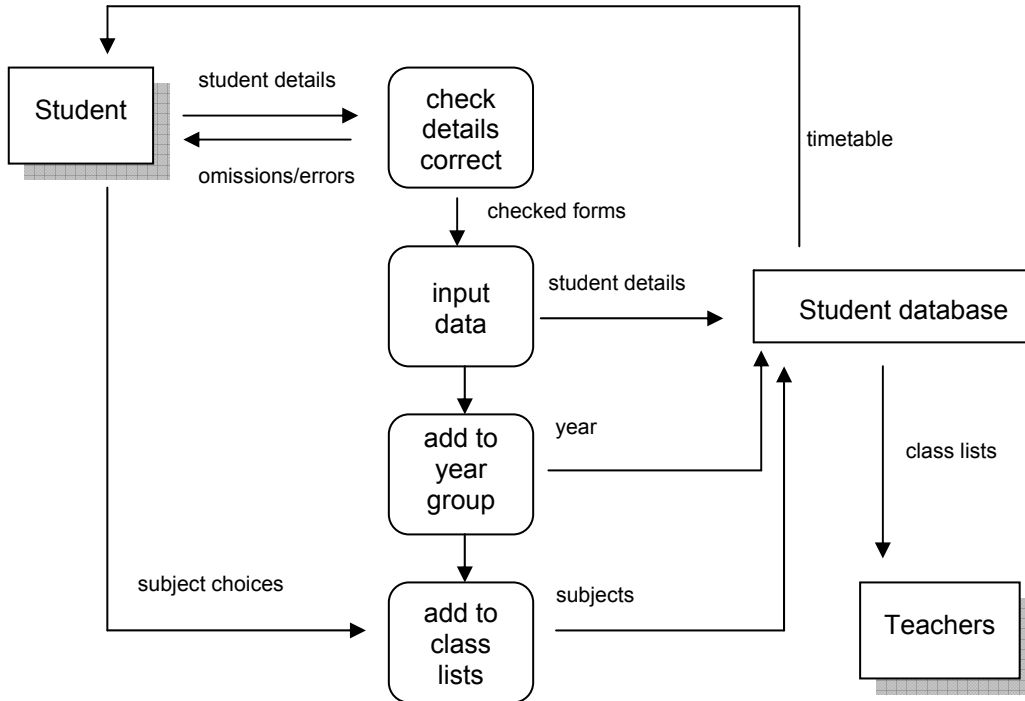


## Data flow diagrams

A data flow diagram (DFD), as the name suggests, shows how data flows through an information system. It is different from a conceptual schema diagram in that it concentrates on how the data is processed (changed) rather than the objects and relations in the system. By looking at the flow of data we not only see the information system from a different perspective, it will also give us a good idea of the inputs and outputs to and from the system.

A DFD shows how data enters and leaves a system, how it moves from one process to another, and where it is stored. Here is a DFD for a student enrolling at a school:



Enrollment DFD

Here the student to be enrolled is represented as a rectangle. The student provides details such as name, age, gender, etc., here shown by an arrow. This data must be checked to see if it is complete or contains any obvious errors. There is an arrow showing the flow of data back to the student if there are any problems.

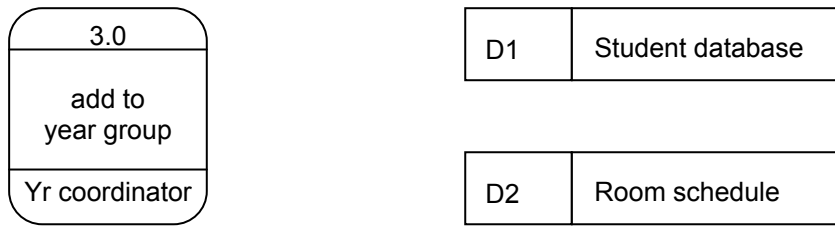
Once checked the data can then be input into the student database. The student must also be added to a year group and must be added to class lists. As part of this last process the student will get to choose which subjects he or she would like to do.

Once all the data has been added to the student database it can then be used to generate information such as class lists and timetables.

In a DFD sources or destinations of data such as students or teachers are represented as rectangles. The flow of data is shown by arrows. Processes, such as *add to year group*, are represented as rounded rectangles. A *process* is where data is taken as input, something is done to it, and then an output is made. Each process must therefore have at least one input and one output arrow.

Finally data stores are shown by open ended rectangles.

If you wish to add more detail to a DFD the processes and data stores can have identifying information added.

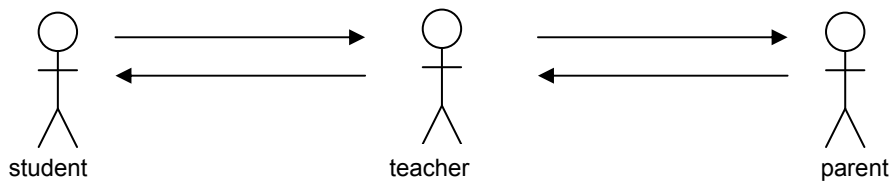


The process can be numbered to indicate its place in a sequence, with sub-numbering (3.1, 3.2, etc.) if required. The person or device making the change to the data is recorded at the bottom. If there are a number of data stores they can also be identified by a code.

The best way to develop a DFD for a given scenario is to use either a whiteboard or a large sheet of paper. The situation should be talked through, preferably by more than one person, and the inputs, outputs, processes and data stores identified. Changes should be made as necessary until a representative diagram of the flow of data is achieved.

### Activity 9.1 – Flow of data

1. Write a paragraph describing the flow of data through the school enrolment scenario described above. Assume a typical student and describe the data collected, where it goes, and what processing it undergoes.
2. Identify the sorts of data that flows between student and teacher, and teacher and parent. Draw this on a diagram like the following:



3. Explain what each of the following is and draw the symbol used to represent it.
  - a An external source or destination of data.
  - b A flow of data.
  - c A process.
  - d A data store.
4. The diagram on the next page shows the flow of data for the *Repairs* database used earlier. Owners of computer and other equipment to be repaired will leave the device with the receptionist who will tell them when it is ready to be collected. The receptionist will record the details on the computer. From the computer the foreman will allocate jobs to be performed by the technicians and himself. Represent the information from the diagram as a DFD.

