

Software Engineering

Data Capture

Data capture

Data entry :- Direct input output of data in the appropriate data fields of a database through the use of human data input device such as keyboard mouse or touch screen

Data collection :- the collection of data from surveys or from independent or network location via data capture , data entry or data logging

Data capture :- (Automatic identification) input of data , not as a direct result of data entry but instead as a result of performing a different but related activity

E.g. capturing inventory related data while recording a sale

It is a identification and extraction of data from a scanned document

Data input :- Transfer information from forms ,such as surveys ,questionnaires ,suggestion card and application

Database :- a collection of data fundamental to a system

What is data capture?

- It cover all stages from recording of basic data to the feeding of this data into computer for processing
- The basic steps are
 - Original recording
 - Data transmission
 - Data preparation
 - Data verification
 - Sorting
 - Control
 - Computer input

Data capture objectives

- Reduction in volume of input output to the extent possible
- Lesser manual preparation
- An input design will ease the work of the person engaged in input preparation
- Minimize the number of steps practice able in data capturing process

Basic steps in data capturing

- Original recording :- collection of data at its source this involves clerical preparation of source document including manual checks
E.g. preparing an examination mark list,
Giving job application form
- Data transmission :- the data moves from the point of origin to the data processing center
e.g. group of related mark list are bunched into batches and sent to data processing center, main processor has many terminals the transfer of data from a particular terminal to the main processor can be termed as transmission

- Data preparation :- the transcription of source document on to a input media such as magnetic tape, magnetic disk, magnetic drum is data preparation
- Verification :- it is to verify that the transcription has been done correctly this is vital because it can result in wrong output
- Sorting :- process of arranging data in some desired sequence sorting may be done manually or mechanically e.g. punched card have to be arranged in a logical order for production of particular input or output

- Control ;- throughout all stages listed above it is essential that checking , verifying and validity controls are maintained this is to ensure that all data collected , transmitted and input are correct
- Computer input :- the data is read by input device like magnetic disk drive and transferred to the internal store where it goes internal checks , invalid data will go through entry stage again.

Data validation

- The objective of data validation is to detect errors at the earliest possible stage before costly activities are performed on invalid data
- Data validation is done by way of manual verification in data capture stage itself
- In spite of this still there may be incorrect batches of input data ,missing data, duplicated data and incorrect file records

- Data vet or data validation :- the objective of data validation is to allow only valid data to be written on the media which will be used in subsequent processing
- Invalid data is also identified and recorded separately
- This invalid data is checked manually for low casting errors after correcting these errors the data is again subjected to above data validation process of accurate input.

Validation checks

- Field check
 - Limit check: applied to each field (data item) of a record to ensure that its content lie within predefined range
 - Picture check :may be applied to each field to detect entry of incorrect characters in the field
 - Valid code check : to validate input against predefined transaction codes these predefined may either be embedded in the program or in stored files
 - Check digit : detect transposition error when recording key field
 - Arithmetic check : are use to ensure validity of the result by performing arithmetic operations in different ways
 - Cross check : may be applied to verify fields appearing in different files to verify the result fully

- Transaction check
 - Sequence check: are applied to detect any missing transaction (e.g. off serially numbered vouchers, any emp being missed out from payment or salary)
 - Format completeness : used to check presence and position of all fields in a transaction
 - Redundant data check: are employed to check the validity code with reference to description
 - Combination check : may be applied on various fields of a file (e.g. exercised on various fields of a file to check amt written in figures and in words)
 - Probability check : are used to avoid unnecessary rejection of data
 - Password check : entry of data by authorized user in online system
 - Checks: may be incorporated to ensure that transaction pertain to the current periods

- Batch total : transaction have been transcribed correctly a total of some common component of a batch of data so as to enable control to be maintained over the validity of data
- Hash total : a control total – sum of values in a particular field or recorded area of a file to ensure that transaction have been transmitted currently

Types of errors

- Transcription error : we transcript word differently (54786 as 54789)
- Transposition error: here position of digits is changed(54786 as 54768)