one of the most important still to be defined in the context of a wholly online system is the place of the operator in the functioning of the system.

5. User relationships

Sections 2 and 3 each identified one area in which the relationship between the analyst/designer and the user became immediate. These were the areas of business analysis and man/machine interaction.

The Cambridge project generated a further instance. Any sub-system of the total integrated system required a sequence of occurrences of one or more applications for its successful completion. For example, the admission of an emergency patient required:

1. Searching of the patient master index to ascertain whether the patient was already registered.
2. If not, then registration of the patient.
3. Admission of the patient.

Each application was a sequence of transactions which were processed serially. A transaction consisted of a message input, a process which accessed the data base and some message output.

As a design aid and implementation tool the project developed its own system for designing transactions through use of one of the interactive terminals. There was also a system for the routine implementation of those transactions. But the design aid was so simple that there was no inherent reason why the user should not use it directly. In practice the user, now that he was in direct contact with the computer system, was much more aware of what was possible and what was desirable and was only too eager to suggest additional information which might be generated or additional facilities which might be created. With the coming of individual microprocessors on the practitioners desk, which ultimately will contain plug-in software, the possibility of the user being able to ‘do his own thing’ becomes much closer. The consequence for the analyst/designer is that he must now regard the user as a completely integral part of the analysis/design process.

In an integrated system there is not one user but many. These are in many departments and at many levels. The clerk who will operate the terminal is at least, if not more, important than the line manager or the departmental manager. As stated in the previous paragraph all of these must now become integrated into the analysis/design process. The total group of ‘chief designer team’ plus the relevant users became a design team closely akin to the ‘system design team’ described by Mumford, Land, and Hawgood (1978). The only, and marginal, differences between the Mumford approach and that which took place at Cambridge were:

1. Many staff were involved in the design team for whom the impact on them of the system was minimal, e.g. the doctor who now received the results of laboratory tests in printed form rather than handwritten although the details were identical.
2. All staff were involved directly (i.e. in person) in any decision which affected them and never through a third party (say a trade union official).

6. Conclusions

As a consequence of implementing an integrated online computer system it was realised that new technology would bring about significant changes in the way that the analyst/designer went about his job. These could be summarised as:

1. The analyst/designer could no longer work in isolation, he had to become an integral part of the business in which he was working.
2. The analyst/designer would need to acquire a large number of complex new skills.
3. All analysis and design would in future be undertaken by a design team.

It is surely significant that these conclusions, although arising from an entirely different starting point, are so akin to those presented in the other papers.

References


The data base approach

P. Prowse

Esso Europe Inc., 5 Hanover Square, London W1

1. Background

The data base approach enables computer systems to be much more responsive and more intimate. The term ‘user friendly’ has been coined to describe a system which helps a person to do a job in a natural way that is easy to understand and remember. Data base systems are becoming user friendly.

The justification for a data base is to support online interactive systems for the majority of operational systems in a corporation and many of the planning systems also. The term ‘operational systems’ is used to describe those systems that support the day-to-day business of the company (Anthony and Dearden, 1976). The information required for planning and management control systems is derived from the operational systems but need not form part of those systems. In Fig. 1 we show the basic business functions supporting management control and planning and the top level decision support systems which help determine the strategic thinking of the company. The upper levels derive data from the lower levels but also require more external information from the outside world.

There is not just one data base supporting all activities of a company but rather a series of data bases which are loosely connected. These data bases support particular functional areas; for example, in an oil company the refinery planning systems incorporate a special data base to support large linear
programming systems.

There is definitely a need for one source of basic data and one carefully controlled means of updating that basic data. The basic data may be replicated many times for differing applications, both for centralised and distributed systems, but there has to be a single source. In Fig. 2 the types of data that are involved in the various levels of the pyramid of information systems are shown alongside the pyramid. Basic operations involve transaction data, whereas planning systems need case data and management control requires aggregation of data.

The concept of ownership of basic data is important, but it must be ownership on behalf of the corporation, not for a particular group or department. Ownership is important because it induces a sense of responsibility and the right environment for 'caring' for data. The interest shown by users when this concept of ownership on behalf of the corporation is explained has been considerable. Identification of the 'best' owner requires some consultation and the decision should not be made lightly; politics are involved and a consensus agreement must be reached.

2. Benefits of a database approach

The benefits of a database approach accrue to users as well as the data processing organisation. They can be summarised as follows:

1. **User benefits**
   (a) visibility
   (b) consistency
   (c) clear responsibility and authority at correct point in organisation
   (d) high availability and good security.

2. **Data processing benefits**
   (a) one source of updates for each logical collection of data
   (b) one definition—single dictionary entry
   (c) one control point—accounting controls
   (d) increased user support through appreciation of the value of the data base approach
   (e) integrity—adequate set of tools
      —effective log tape for recovery
   (f) security— one point to control access
   (g) generalised data base software to aid development.

3. **Data base policy**

A policy has been established to promote the use of database technology when justified in the affiliates (Esso Europe, 1978). It is believed that significant benefits can be achieved by implementing database systems, particularly those which share data between several functional areas. These benefits are in the form of better information and cost reduction to the users. Better information results from data which are more consistent, more current and better validated. Cost reductions result from reduced manpower for resolving data inconsistencies, lower system life cycle costs and reduced redundant processing.

The policy states that the short term benefits of data base should come from:
(a) online entry and validation of data
(b) online enquiry facilities
(c) fast response reporting.

These facilities demand the data structuring, integrity and security features of a data base management system to be effective, provide high availability and good response time.

The long term benefits due to reduced data redundancy, more consistent data and lower maintenance costs are much more difficult to quantify. Thus the implementation of a DBMS without the potential for incorporating online facilities is hard to justify.

The planning of the future data base environment should be based on a sound overall concept which addresses the following main tasks:

1. Define the scope of future data base development work.
   The use of data base technology is particularly attractive for systems which share data between several functional users.

2. Identify the impact of distributed processing on the future data base environment.

3. Establish a long term plan for the development of 'business subject' data bases which will replace the current application oriented data files. Prerequisite for this development is a detailed knowledge of the information elements of the company and a long term migration plan for the applications concerned.

4. Select and optimise the use of data base software and related software components. The selection will impact on the necessary systems design standards.
4. Systems analysis
With that background to what is seen as a data base approach let us consider what systems analysts have to do to support that view.

First, it is necessary to agree that applications can no longer be considered in isolation. A high proportion of the basic data is common. Of course some of the data appears to be required for a single purpose but, as government demands for information grow, it is getting obvious that the majority of data will be common to more than one application in the future. Each application may need its particular view of data; for example, daily operational type systems may require subsets of data that can be processed very efficiently.

Second, it is essential to understand which type of information system is being investigated. It is important to deal with only one type of system at a time, i.e. basic business systems should be kept separate from planning systems, etc. This is because the nature of the systems probably demands different solutions and it is much simpler to keep them apart. Planning and control systems require sets of data which can be extracted, modified for a particular purpose and then held until the planning cycle has been completed. In that case processing efficiency may not be a very significant factor.

Third, analysts should recognise the value and opportunity provided by data administration. If it is agreed that the majority of data is common to more than one application then some analysts need to concentrate on data rather than on applications. These are the analysts who look across a company for the varied uses of data and support the application analysts. These data analysts are termed ‘data administrators’ for they do truly administer data. They are concerned with common definitions, naming conventions, coding systems and regulations for usage. To support data administrators a highly qualified group of data base administrators are required. These people are primarily concerned with the integrity, security and efficient storage of data. They will probably be responsible for data base design and the day-to-day use of a data dictionary that serves as a focal point for all data definitions.

The data administration role is only beginning to be understood. Efforts to define the role are generally inadequate to convince top management of the need to allocate sufficient resources to the task, but a start has been made. As success is achieved in convincing users to become owners of data on behalf of the corporation, then the need for co-ordination of data will become more obvious and resources will be reallocated to specific data administration roles.

It is important to talk about the reallocation of resources because many of the data administration tasks are buried in existing applications and among the manpower which is in place at present it is just not recognised as being data administration. There are some new data co-ordination tasks but they are likely to be less than the effort currently wasted in reconciling inconsistent figures or in duplicated updating systems. The total effort which is expended when all applications are permitted to own their own data has not yet been measured and compared against the effort required for a similar system under good data administration. Instead it is possible to take a ‘subject’ approach to data (Martin, 1976) and justify improved data administration for each subject at a time.

5. The changing role of application analysts
What are the changes necessary in application analysis? At the outset it must be understood that a model of the data required for the application should be established at a very early stage in the analysis process. This model should be produced in association with data administration. Sometimes the model is built by data administration and at other times by the application depending upon training and the availability of resources.

Data administrators can often provide valuable input and probably have a greater understanding of the data analysis process. However, the application analysts may be in a better position to carry out the work. They are more likely to have the resources, and in any case the data analysis is fundamental to understanding the problem which the application is trying to solve.

Most large applications are developed these days under some form of project management system. Applications are phased with decision points at the end of each phase. Typical phases are scoping, exploration, functional specification, systems design, development, start up and audit. An outline of the data model is required at the exploration stage and a full logical model of data is essential at the functional specification stage. A physical model of data for a particular data base management system is a prerequisite to the development stage.

What are these phases in a little more detail?

1. Scoping is setting the boundaries, confirming that there is a problem, estimating the size of the task and indicating possible savings that might result.
2. Exploration is confirmation of the boundary, a clearer statement of the problem and an outline of some alternate solutions with likely costs and benefits.
3. Functional specification is a complete definition of user interfaces and the functions that have to be carried out.
4. Systems design is how to do it.
5. Development is coding and testing.
6. Start up is going operational.
7. Audit is appraising what has been achieved.

A new technique of structured systems analysis is now in use at the scoping and exploration stages. This is an extension of program structure technology into the systems area. It is a technique that is just being introduced in European affiliates and therefore it is not included in this paper. Another technique which could be helpful at the systems design stage is ‘life history analysis’ (Robinson, 1979) but experience is limited and it is also too early to report on the practical use of the technique.

In contrast, the data base approach has been in use for a number of years and data base topics can be related to the project management phases. Data requirements, by subject, should be identified at the scoping stage and confirmed at the exploration stage. The task of building an outline data model should also be tackled at the exploration stage. It is a new requirement at a very early stage in an application.

The definition of a full logical model of data at the functional specification stage is essential. Analysts need to be trained to produce these models. Analysts also need to understand the reasons why these models are required. One reason that could gain great acceptance amongst analysts is the better estimation of likely development costs that could result from an improved understanding of detailed data requirements.

It is frequently the case that costs escalate as the development phase is reached. That is because of the greater appreciation of the details essential to the client specifications which have been agreed. Now these details can be determined at the time the interfaces are proposed, rather than later, and the estimates can be much more reliable.

Business factors which deal with the rules of the business should be identified at the function specification stage.

A brief explanation of the logical data model, subject data bases and business factors follows.

6. Logical model of data
What is a logical model of data? One concept is that it is a set of ‘relations’ plus ‘constraints’. There are other candidates such as the entity relationship model (Chen, 1976) and the network
model (CODASYL, 1971). Experience over the past four years indicates that the relational approach (Codd, 1970; Date, 1977) works well for the definition of a data model. It also enables fairly easy transition to a physical model which can run on existing data base management systems. The notion of adding constraints to the relational model is now well accepted (Benci et al., 1976). Some more definitive information on how to develop a logical data model is given in a companion paper by Prowse and Johnson.

Analysts in most Exxon affiliates are taught logical data analysis on a one week course. This course was developed in 1975, it includes lots of practical material, as well as some relational theory. To date, well over 1,000 analysts have been trained. The training has improved the understanding of data and its role in the definition of systems; it has also led to easier communication on data requirements between analysts and with users. To talk of data being in 'third normal form' is now acceptable to the majority of analysts.

The basis for logical data analysis can be expressed in terms of three environments:

1. Business environment—assists with the analysis of the business information requirements.
2. Technological environment—represents the only constant in an environment of changing technology.
3. Processing environment—applicable to both a centralised and distributed application.

7. Subject data analysis

Another stage of data analysis that is in development is the recognition of the value of categorising data into a limited set of 'subjects'. The use of these subjects helps enormously in communication with users. Computer people argue about the definition of a subject but to users it seems a natural way of organising data. A preliminary subject data base approach has been defined as 'A corporate approach to the filing of data which addresses clearly defined "subjects" irrespective of business functions'. Subjects are the broad classifications of data that exist in most corporations. Examples are customers, products, locations and orders, etc.

Initially a particular subject definition is very loose; just an outline concept. Then a data analyst starts to define the subject and investigates it across the company. As his understanding grows the definition becomes clearer; finally the analyst is able to provide a fairly watertight definition. This is then published for general information and agreement.

Once the basic structure of the subject data base has been established then it is frozen. Small additions may be made later but no structural changes are permitted. This is because of the limitations of existing data base management systems.

8. Business factors

It is extremely helpful to recognise the commonality of many business factors across systems. Business factors are the rules of the business. They consist of tables and lists which describe how the business is run. They are updated fairly infrequently and are different from the variety of master data that is involved in subject data bases. Examples of business factors are:

(a) sales areas
(b) price zones
(c) VAT rates, etc.

In the past this type of data has been embedded in programs or application tables. Now it is clear that it is necessary to use one set of tables which can be updated centrally.

To get a business factors concept off the ground it is a good idea to appoint a business analyst with wide company experience to be the data administrator for this type of data. Good supporting software is vital.

Business factors should be held in basic form in one set of tables and subset tables produced automatically for each application. In that way efficiency can be maintained.

9. Conclusion

It has been shown that the systems analyst task has changed considerably. Analysts need to learn about logical data models, subject analysis and business factor determination. They must be willing to co-operate with and possibly work with data administration staff. They can no longer work on an application in isolation. They should be looking out for new techniques like structured systems analysis and life history analysis.

Analysts should phase their work and recognise that the scoping, exploration, functional specification and systems design stages of a project require different levels of analysis and design work.

Acknowledgement

The author wishes to thank Esso Europe Inc. for permission to publish the paper.

References

CODASYL DBTG. (1971). Data Base Task Group of CODASYL Programming Language Committee Report (April 1971) and DDL.