

Database Documentation

Database documentation includes the recording of the procedures, standards, guidelines and database descriptions necessary for the proper, efficient and continuing use of the database.

The documentation must be specially prepared for and selectively distributed to

- end users;
- DBA function itself;
- computer operations function;
- applications development function.

The DBA has the responsibility for providing and maintaining adequate documentation for these recipients. This chapter discusses the types of documentation that are required. The list given is not necessarily exhaustive.

This chapter covers the following topics:

- Standards
 - Description of the Database
 - Data Dictionary, Function and Use
 - Predict : The Adabas Data Dictionary
 - Applications Using the Database
 - Description of Data Sources
 - Data Access and Manipulation Procedures
 - Passwords and User Identification
 - Back-Up Procedures
 - Restart and Recovery Procedures
 - DBMS Performance and Measurement
-

Standards

Establish and maintain a consistent set of database controls, particularly in the following areas:

- Data definition: a uniform methodology should be adopted;
- Data usage;

- Ownership of, and responsibility for, identifiable portions of the database;
- Data access and manipulation: Standards for the way data is accessed and updated in the database are crucial to ensure data integrity. Standard procedures for coding requests reduce the possibility of errors. For example, the updating of key values should be strictly controlled;
- Data edit and validation: The DBA must establish procedures to ensure compliance with the rules and maintain consistent levels of data quality in the database. The DBA should, therefore, become involved in systems and acceptance testing of new applications that use the DBMS;
- Computer operations: The DBA is responsible for ensuring that standard procedures are used by computer operations personnel when they deal with the database. This includes standard backup procedures, restart and recovery procedures, and other operations-related activities.

Some personnel will resist the establishment of standards for the database environment. Give careful consideration to the status of a standard and of areas where standards should be established. In general, a new standard will require negotiation, arbitration, and compromise before all the parties concerned will accept it even as a proposed standard. The only way to determine whether a standard is practical is to implement it.

Standards are subject to change. The process of changing an existing standard must be as tightly controlled as that of installing a new one. Changes should be formally proposed and communicated to all the affected users. After a trial period, review the proposed change with users to decide whether it should become a standard.

Periodically review the database standards to evaluate their effectiveness and to ensure that they are being followed. Corrective action may need to follow such a review. The DBA also has the principal responsibility for ensuring that all personnel who work in the database environment are aware of, and adequately trained in, the use of the standards.

Because each site has its own procedures and requirements, this documentation does not suggest a specific set of standards.

Description of the Database

The database description should cover the following main areas:

- Conceptual database: a formal description of the data to be stored and a definition of its inherent logical data structures. The DBA should explicitly define data structures which reflect the DBA's knowledge of foreseeable developments and include the needs of other known or potential users.
- Use of the database: information should be recorded at both the application level and the individual data item level. This documentation will be of immense value when new systems are being developed. The information that should be recorded is discussed in the next subsection.
- Implementation: how the data is currently stored in the database. This is the documentation of the physical storage structure of the database and it should include (among other items):
 - Implemented data structures (logical data formats and relationships). These are normally a subset of the conceptual database;

- Storage structure (physical data formats and relationships), in terms of Adabas, this means the contents of the Field Definition Tables, coupling relationships and other inherent relationships;
- Volume of data: number and average size of records in each file;
- Anticipated growth: by file and field size (and, therefore, record size);
- Additions and deletions of records: number in an average maintenance run. It is also useful to keep a note of the distribution of additions and deletions (i.e., are they random across the file or restricted to a small portion of it), as well as the manner (user program or utility) in which these additions and deletions are performed.

In addition, the reasons why this particular implementation has been chosen should be recorded. This information will later prove useful when maintenance or new application design is undertaken.

The DBA is responsible for formally describing the database in the manner discussed above and maintaining this description on a data dictionary (whether this process is automated or not). The project team will be required to provide the DBA with all the information the DBA needs to perform this task.

Data Dictionary, Function and Use

A data dictionary contains information about the definition, structure and use of data. It does not store the actual data itself, but rather data about data. Simply stated, the data dictionary contains the name of each data type (element), its definition (size and type), where and how it is used and its relationship to other data elements.

A data dictionary enables the DBA to exercise better management and control over the organization's data resources. Advanced users of data dictionaries have also found them to be valuable tools in project management and systems design.

The data dictionary will enable the DBA independently to manage actual data items and the programs that manipulate and access them. This independence of control results in the substantially enhanced usefulness of the data. The data dictionary serves to collect the information needed in order to make the data more useful.

Containing all of the definitions of the data, the dictionary becomes the information repository for the data's attributes, characteristics, sources, use, and interrelationships with other data. The data dictionary should provide the following information:

- The kind of validity tests which have been applied to this data type;
- What modules, programs, systems and reports use this data type?
- The level of security which has been applied; Who is allowed to access the data type? Who is authorized to update this data type?
- By what other names is the data type known in various application environments?
- What is the input source for this data type?

- Textual description of the data type.

Predict : The Adabas Data Dictionary

Predict, the Adabas data dictionary system, is used to establish and maintain an online data dictionary.

Database information may be entered into the dictionary in online or batch mode. The description of the data in the Adabas dictionary includes information about files, the fields defined for each file and the relationship between files. The description of use includes information about the owners and users of the data in addition to the systems, programs, modules and reports that use the data. Dictionary entries are provided for information about

- network structures
- Adabas databases
- files, fields, and relationships
- owners and users
- systems, programs, modules and reports
- field verification (processing rules)

Standard data dictionary reports may be used to

- display the entire contents of the data dictionary
- print field, file, and relationship information
- print field information by file

In addition, the dictionary data can also be accessed directly from Natural since it is stored in a standard Adabas file.

Refer to the Predict documentation for more detailed information about the Adabas data dictionary system.

Applications Using the Database

For each application using the database, the following information should be recorded:

- Application characteristics:
 - Description of the function of the application;
 - Mode of use: batch/online, single or multiuser;
 - Frequency of use: standard scheduled times;

- Type and volume of transactions;
- Performance considerations: minimum acceptable response time;
- File requirements:
 - Which files are accessed;
 - How files are accessed (the use of descriptors);
 - Specific data items (fields, subfields, superfields, and so on) used;
- Security requirements:
 - Ciphering and/or password usage (that is, cipher keys and/or passwords are not to be made generally available);
 - Authorized users of this system/program/enquiry;
 - Back-up requirements: frequency and content of file backups;
 - Restart requirements.

Since any database is only a partial implementation of the conceptual database (see previous section) and user's requirements change with time, new applications of the database will be found as time passes. Some of these applications may be developed into new systems or additions to existing systems, but they first arise as a simple user requirement .

Establish procedures for recording unplanned applications of the database; if one becomes relatively frequent or important, you can often gain a processing advantage by redesigning or reorganizing the database or files within it. For example, assume that a file was initially loaded in Customer Number order. Subsequently, applications that process the file by Salesman Number assume greater significance. You can unload the file and reload it in Salesman Number sequence without affecting the logical operation of most existing applications, thus achieving an overall reduction in the processing time needed by all the applications that use the file.

Records of this unplanned use of the database or shift of emphasis in processing priorities, can be made when a user makes an interactive request for information, whether the user does this in his or her own department or through the DBA. These records should be regularly reviewed by the DBA and discussed with the affected users.

Description of Data Sources

For any new application, the data dictionary is the first reference document for determining the potential sources of information. The description of data sources will be derived during the systems analysis and design phases of a new project.

Record and retain the following information about each system:

- The present form and location of data: forms, files, computer storage media;

- Access techniques to be used to acquire the data;
- The intended use of the data in relation to its present accuracy, completeness, and timeliness, including necessary validation or editing;
- The need for modification of the data before it is stored on the database;
- The authorized agent for the use of the data;
- The cost of acquiring the data.

Data Access and Manipulation Procedures

The DBA must have administrative control over all access to and updating of the data in the database. Unless this is so, there can be little meaningful control or protection exercised over it. The lack of such control can result in serious security and integrity problems.

Because authority and responsibility for the database cross organizational boundaries, a corporate policy covering database usage by and among operating units should be published. Such policy statements can enhance the administrative control of the DBA and help to promote clear understanding of database procedures among users and data processing personnel.

Part of this policy will include statements on

- the use of DBMS commands-who can and (more important) cannot use the various facilities provided by the DBMS?
- database usage-is this to be achieved by user programs? Are standard interfaces such as Natural or SQL to be used? What error handling procedures should be observed? To whom should difficulties be reported, and how (for example, a trouble report)?
- maintenance and update procedures-who will be responsible?

This policy statement should be proposed and drafted by the DBA, and then reviewed and agreed upon by all the affected parties.

Passwords and User Identification

User ID and password information needs to be stored securely by the DBA, as only the DBA and the affected users should have access to it. This documentation will include

- assignment procedures for cipher keys, passwords and user identification;
- actual assignments: cipher keys (where it is necessary for the DBA to know this), passwords and user identifications;
- terminal and data access procedures;
- access authorities must be established for each data entity. These should define:

- Who has the right and/or need to know the content of the data, as well as of its existence;
- Who can read the data from the database, add new occurrences of the data, update existing values of the data, and/or delete the data from the database.

Once this authority has been established, it is important to set up proper control procedures in order to ensure that violations of database security do not occur.

- database security procedures. The physical protection of the data in the database should be recorded, detailing
 - positive human control over the database (communications rooms, access to the computer and terminals, storage of database backup and log tapes);
 - physical separation of data entities (separate files, separate databases, use of partial mount and file cluster facilities);
 - secure areas for terminals (lockable terminals, keyholders, leased or dial-up lines, taken offline when not in use);
 - persons authorized to receive published information about the database.

The DBA uses the Adabas Security utility to implement and control password security (see the Adabas Security documentation for more information). The DBA must be the only person at an installation who is permitted to use this utility.

The DBA must implement procedures to physically secure the security utility itself, and all documentation concerning security.

Back-Up Procedures

The content of back-up files, as either database or file copies (taken by the Adabas ADASAV utility) should be recorded together with the following information:

- What state (or point in time) the backup data refers to;
- Identification of the data which can be backed up;
- The volume of data that is involved;
- The backup facilities that should be used in order to reestablish the database to that state;
- The frequency and schedule of database backup operations.

The DBA should help computer operations personnel to develop procedures for carrying out the database backup task. Database backup is an essential step in ensuring that the database can be restored to its proper state in the event of destruction or damage. The decision will have to be taken (for every application) as to whether the entire database is to be backed up or whether dumping and restoring of specific files is more appropriate.

Information about developing backup procedures for a particular application is included in the Adabas Operations documentation.

Restart and Recovery Procedures

The DBA is responsible for formulating and supervising procedures for

- restarting the DBMS after failure;
- recovering the database to a recent checkpoint (if necessary), thus removing the need to repeat database maintenance work;
- controlling the priority and sequence of database restoration.

Restart and recovery is an important database protection consideration. The DBA must develop standards, procedures and rules to provide such a capability. The DBA must be certain that the standards and rules are being adhered to and enforced. Restart and recovery must be planned for and designed in conjunction with the implementation of the DBMS. It should not be added as an afterthought.

Detailed information about Adabas restart and recovery is included in the Adabas Operations documentation.

DBMS Performance and Measurement

The DBA has a continuing role in maintaining and improving the performance of the database system.

To do this, the DBA must monitor the performance of the system and try alternative design strategies to improve it. As work patterns change in the company, both the volume and relative proportion of types of transactions may change. This may affect performance and design changes may be necessary to counteract it.

In the longer term it may be possible to predict changes in workload, and plan how to meet them by redesign or equipment enhancement.

The effect of new hardware or software should also be evaluated, and possible changes should be cost-justified and incorporated into the long term strategy.

Keeping track of (and measuring) the performance of the DBMS is therefore an important part of the DBA's function. The DBA should establish and maintain records of

- the computing resources used, including frequency of use, by each application area;
- the users who are serviced by a particular application; and
- DBMS effectiveness with respect to response time and cost.

The DBA will also need to establish and document procedures for

- monitoring the frequency of DBMS usage; and
- DBMS performance management.

It is the responsibility of the DBA to monitor the database environment on a continuing basis, in order to ensure that an efficient level of service is provided while database integrity is maintained. This responsibility for monitoring takes the form of a variety of activities and procedures, of which performance management is but one.

The Adabas Online System provides the DBA with a powerful tool for monitoring the database. See the Adabas Online Systems documentation for more information.