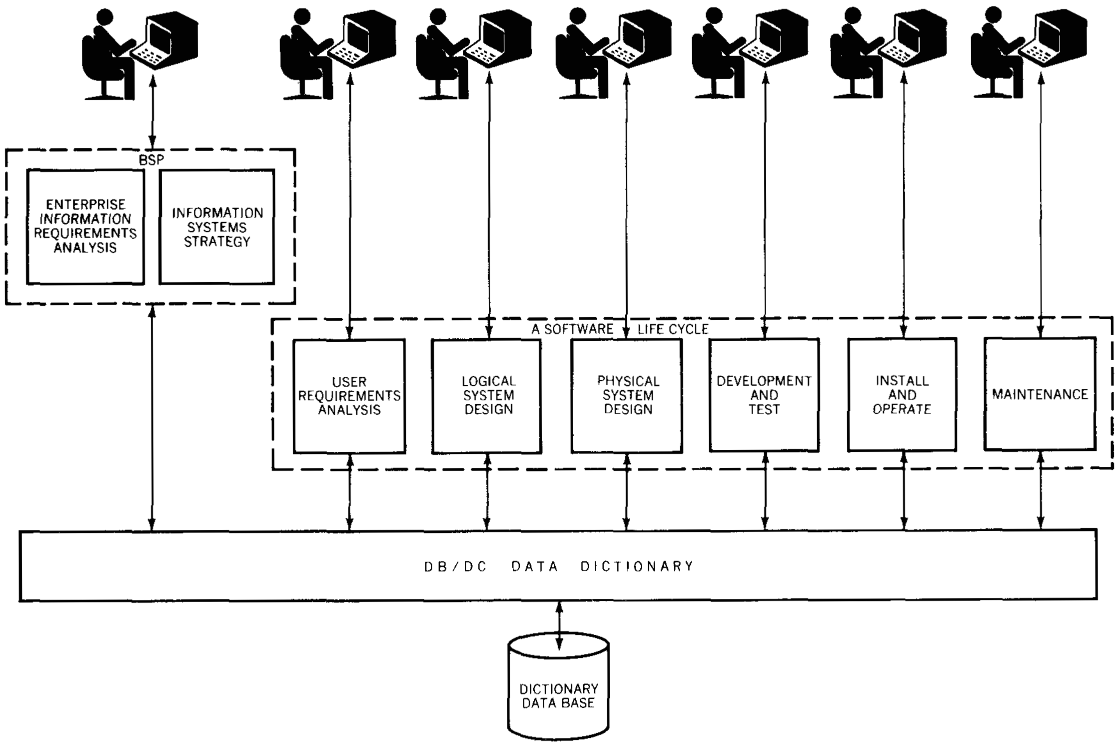
Traditionally, Business Systems Planning (BSP) studies have been conducted using manual techniques. This paper describes one approach for computer assistance to such a study. The Extensibility facility of the IBM DB/DC Data Dictionary is shown to satisfy the requirements for the capturing of and subsequent reporting on BSP study data. The possibility of extending this approach to follow-on software development activities is discussed. General overviews of the Business Systems Planning methodology and the IBM DB/DC Data Dictionary are also provided.

Supporting Business Systems Planning studies with the DB/DC Data Dictionary

by J. G. Sakamoto and F. W. Ball

The Business Systems Planning (BSP) methodology came into being as a result of an IBM effort to plan for its own information systems in the 1960s. BSP is a structured way of assisting a business to establish an information systems plan that satisfies both near- and long-term information needs.[1] BSP techniques include top-down analysis and planning, bottom-up information system implementation, managing data as a corporate resource, and orientation around business processes. The BSP methodology centers around a study team which collects and investigates data that are required to run the business. These data are then categorized into data classes; the categorization leads to the definition of information systems to support the business processes and business objectives. The study team starts by collecting facts about the business that are usually available in documented form throughout an organization. These facts are organized, abstracted, and analyzed by the study team and enhanced by members of top management, who explain the business and add those points usually not documented. The study progresses to an identification of the major activities and decision processes in the business. Management is then asked to further validate and enlarge upon the facts that have been gathered and analyzed. The analysis ends with the consolidation and comparison of the information from all sources. The study team then presents its findings, recommendations, and action plan. More insight into the BSP study approach can be obtained from Reference 1.

Figure 1 An integrated approach to managing information systems activities



Traditionally, BSP studies have used manual techniques for data storage and analysis. But, because the study data can be voluminous, and because they must be easily updated and modified, there are benefits to be gained from some form of automated assistance.[2] Automated assistance is also potentially beneficial in the integrated management of subsequent software development activities. This concept is shown in the example of Figure 1, which depicts one of the multiple possible software life cycles that might result from a BSP study. Not shown are the actual application-oriented files or data bases.

We begin by describing the potential benefits and the minimum requirements for a computer-assisted tool to support BSP studies. An overview of the IBM DB/DC Data Dictionary follows which shows how it could be used to satisfy these requirements. The creation and usage of a BSP data model in the DB/DC Data Dictionary are then described. Next, the requirements for a reporting facility and some operational issues are discussed. We conclude by discussing the concept of integrating this approach with follow-on software development activities.

A tool to support BSP studies

Tasks such as organizing, storing, manipulating, retrieving, and analyzing the data gathered during the course of a BSP study are currently performed manually. We believe there are potential benefits to be gained by the use of a computer-assisted, data-base-oriented tool in supporting these tasks. Some of these benefits are

- Large quantities of data can be stored, and the manipulation/analysis of complex relationships becomes manageable. This ability could potentially improve the quality of the study and/or reduce the duration of a study by allowing a study team to concentrate on the analysis of problems. It also could enhance the probability that the study recommendations, action plan, and follow-on activities will be implemented.
- Although the gathered BSP data are interrelated, the data base approach allows each piece of data to be stored only once. This approach helps to maintain the consistency of the analysis data and reduces possible errors.
- Information is less likely to be lost with an automated approach. Important not just for the obvious reasons, this benefit is also important for historical research into the original justification of given implementations.
- A data base approach can be a bridge between the end of a BSP study and subsequent information systems (I/S) and data management activities. The BSP data base can be used to identify implementable systems, support system design and data modeling activities, and generate test plans to help ensure that the systems developed do indeed satisfy the original requirements. This area is a major one for future research. However, we feel that the establishment of a BSP data base is a good start in this direction.
- A data base approach facilitates ongoing BSP-like studies. The plan that results from a BSP study represents the best thinking at a specific time. However, as the business situations and associated information requirements change, the data base can be kept current with only incremental changes. Analysis of information requirements and information availability in the changed environment can be made quickly.

We think that the minimum requirements for such a tool are

- A facility for describing various information-related enterprise phenomena, such as relationships between organizational units, business functions, and information being created and used.
- A facility for storing these descriptions in a data base in a computer-processable form.
- A facility for adding, deleting, and changing such descriptive data in the data base. This facility should also check for correctness and consistency of updated data so that the accuracy and integrity of study data in the data base can be maintained.

• A facility for inquiry/analysis of the study data in the data base.

Although many data-base-oriented systems facilitate these requirements,[3–5] we used Release 4 of the IBM DB/DC Data Dictionary[6] since it satisfied the majority of the requirements and was also in more widespread use than other alternatives.

In the following sections we show how the Data Dictionary can be used in BSP studies. We start by briefly describing the Data Dictionary.

What is the Data Dictionary?

The Data Dictionary we refer to is Release 4 of the IBM DB/DC Data Dictionary program product, called here simply the Dictionary. For our purpose, the Dictionary can be thought of as an automated central collection of descriptive information about data resources and their relationships. The types of information maintained include, for example, the formats and characteristics of data, where they are stored and how they are used. The Dictionary provides a set of computer-based facilities for describing the information, entering it into its data base, modifying it, and obtaining reports.

The data resources are described in terms of *subjects*, their *relationships* to other subjects, and any necessary *attributes*. An illustration of the application of this concept comes from the BSP Information Systems Planning Guide,[1] which will be referred to henceforth as the BSP Guide. In BSP, a matrix is used to show data *creation* and data *usage* relationships between processes and data classes. Figure 2 shows an example of such a matrix, taken from the BSP Guide, for a hypothetical medium-size manufacturing firm. The row and column names of the matrix such as "customer," "order," and "sales territory" are considered to be *subjects* in Dictionary terminology.

Subjects are grouped into *subject categories*. The previous subjects, for example, belong to a subject category called "DATACLSS" (data class). Another example of a subject category is "PROCESS," containing such subjects as "territory management," "selling," "sales administration," and "order servicing."

The *relationships* between the subjects are described by establishing relationship names between the subject categories. For our example, the relationship names between the "PROCESS" and "DATACLSS" categories are USES and CREATES (and conversely USED_BY and CREATED_BY):

"PROCESS" USES "DATACLSS" or
"DATACLSS" is USED_BY "PROCESS" and
"PROCESS" CREATES "DATACLSS" or
"DATACLSS" is CREATED_BY "PROCESS."

Figure 2 Process versus data class matrix

DATA CLASS / PROCESS	CUSTOMER	ORDER	VENDOR	PRODUCT	ROUTINGS	BILL OF MATERIAL	COST	PARTS MASTER	RAW MATERIAL INVENTORY	FIN. GOODS INVENTORY	EMPLOYEE	SALES TERRITORY	FINANCIAL	PLANNING	WORK IN PROCESS	FACILITIES	OPEN REQUIREMENTS	MACHINE LOAD
BUSINESS PLANNING							U						U	C				
ORGANIZATION ANALYSIS														U				
REVIEW AND CONTROL													U	U				
FINANCIAL PLANNING											U		U	C	U			
CAPITAL ACQUISITION													C					
RESEARCH				U								U						
FORECASTING	U			U								U		U				
DESIGN AND DEVELOPMENT	U			C		U		C										
PRODUCT SPECIFICATION MAINTENANCE			U	U		C		C										
PURCHASING			C				U											
RECEIVING			U						U									
INVENTORY CONTROL									C	C					U			
WORKFLOW LAYOUT				U	U											C		
SCHEDULING			U	U											C	U		U
CAPACITY PLANNING			U		U											U	U	C
MATERIAL REQUIREMENTS			U	U		U											C	
OPERATIONS					C										U		U	U
TERRITORY MANAGEMENT	C	U		U														
SELLING	U	U		U								C						
SALES ADMINISTRATION		U										U						
ORDER SERVICING	U	C		U														
SHIPPING		U		U						U								
GENERAL ACCOUNTING	U		U								U		U					
COST PLANNING		U	U				C											
BUDGET ACCOUNTING							U				U		U	U	U			
PERSONNEL PLANNING											C		U					
RECRUITING/DEVELOPMENT											U							
COMPENSATION											U		U					

LEGEND: U = USES
C = CREATES

SOURCE: BSP INFORMATION SYSTEMS PLANNING GUIDE

Having established the relationships between the subject categories, we can now describe the relationships between the individual subjects in these categories, as the following examples show:

"territory management" USES "order"
"selling" CREATES "sales territory"
"product" is USED_BY "order servicing"

The *attributes* of subjects are the information that is needed to describe the properties or characteristics of the subjects. For example, if one wants to describe the locations in which various "PROCESSes" are performed, the LOCATION attribute may be established as

"territory management" in LOCATION='LOS ANGELES,'
"order servicing" in LOCATION='NEW YORK' and
"sales administration" in LOCATION='SAN FRANCISCO.'

Attributes can also be used to describe such relationships as

"order servicing" USES "product" FREQ='DAILY'

where FREQ (for frequency) is an attribute of the relationship USES and describes the fact that order servicing uses the product information on a daily basis.

Attribute description is not necessary for analyzing data class and process relationships to determine an information architecture. However, attributes such as LOCATION and FREQ may be useful information to maintain for distributed processing considerations. Other examples of attributes are discussed in subsequent sections.

Figure 3 illustrates the concepts of subject categories, subjects, relationships, and attributes, using some of the previous examples.

The Dictionary product is distributed with eleven predefined subject categories that are mainly oriented toward describing the units of data within a data base system and components of the computer programs that use the data. Some examples of these are DATABASE, SEGMENT, ELEMENT, PROGRAM, MODULE, and TRANSACTION. The individual subjects for these categories may be interactively entered into the data base and modified or "browsed" by using the Interactive Display Forms facility (IDFF) of the Dictionary. A REPORT facility is provided to produce reports on category contents, details of subjects, and lists of related subjects, with various levels of detail.

two major facilities

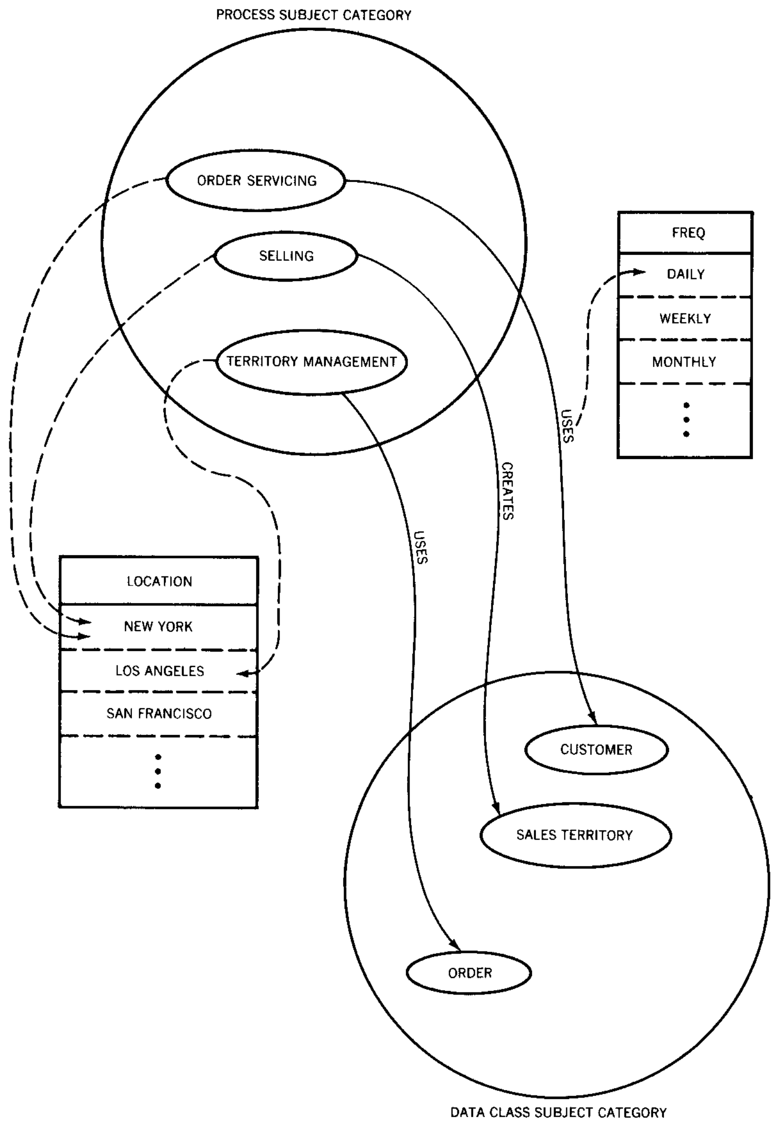
Two major facilities of the Dictionary, introduced with Release 3, make it a very attractive tool for supporting BSP studies. They are

- The Extensibility facility
- The Program Access facility

The *Extensibility facility* allows an installation to define additional subject categories plus any associated relationships and attributes. It is this capability that allows us to create new subject categories to describe the BSP study data. Examples are organizational units, business functions, data classes being created and used, current I/S support, and informational problems being experienced. We have previously discussed how the processes and data classes can be described as "PROCESS" and "DATACLSS" subject categories. We will define more subject categories to describe other BSP data definitions in the following section.

The *Program Access facility* (PAF) allows an installation to write various analysis programs which use the descriptive information stored in the Dictionary data base. Access to the BSP study data via the PAF allows the creation of programs that can produce reports in any desired format. We call these programs PAF programs.

Figure 3 Subject-relationship-attribute concept



The BSP Data Model

The BSP data model described here is a modification of that by Sakamoto.[7] The model represents categories of BSP study data and the relationships between these categories. In our discussion, we often refer to BSP subjects or simply to subjects. These subjects are instances of BSP study data that belong to the categories. Example: PROCESS is a category that describes business processes. Subjects of that category might be such business processes as "financial planning," "forecasting," or "capital acquisition."

Figure 4 BSP data model

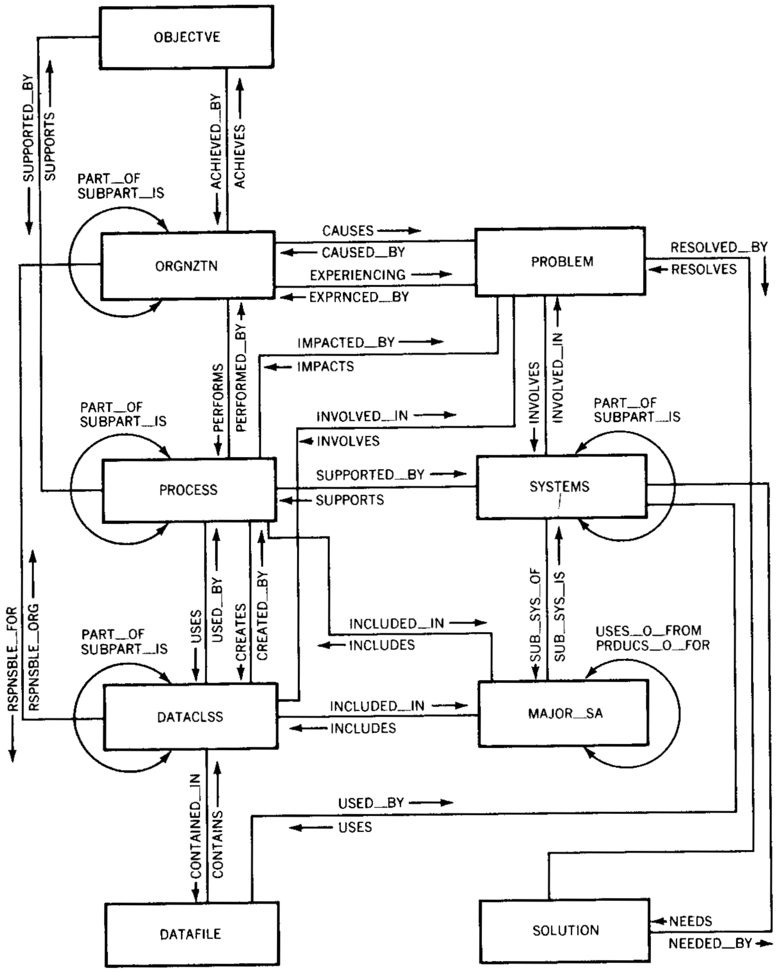


Figure 4 shows the model schematically. Nine rectangles indicate the categories of the BSP study data, and the connecting lines indicate the relationships between the categories. The figure shows both of the complementary relationship names. For example, ACHIEVED_BY complements ACHIEVES, USED_BY complements USES.

The nine categories are

1. DATACLSS—logically related information in the enterprise
2. DATAFILE—"containers" of DATACLSS
3. MAJOR_SA—major systems areas of an information architecture
4. OBJECTVE—objectives and goals of the enterprise
5. ORGNZTN—organizational units of the enterprise

Table 1 Category relationship

Category	*Relationship*	*Category*	*Category*	*Relationship*	*Category*
DATACLSS	CONTAINED_IN	DATAFILE	PROBLEM	CAUSED_BY	ORGNZTN
	CREATED_BY	PROCESS		EXPRNCED_BY	ORGNZTN
	INCLUDED_IN	MAJOR_SA		IMPACTS	PROCESS
	INVOLVED_IN	PROBLEM		INVOLVES	DATACLSS
	PART_OF	DATACLSS		INVOLVES	SYSTEMS
	RSPNSBLE_ORG	ORGNZTN		RESOLVED_BY	SOLUTION
	SUBPART_IS	DATACLSS	PROCESS	CREATES	DATACLSS
	USED_BY	PROCESS		IMPACTED_BY	PROBLEM
DATAFILE	CONTAINS	DATACLSS		INCLUDED_IN	MAJOR_SA
	USED_BY	SYSTEMS		PART_OF	PROCESS
MAJOR_SA	INCLUDES	DATACLSS		PERFORMED_BY	ORGNZTN
	INCLUDES	PROCESS		SUBPART_IS	PROCESS
	PRDUCS_O_FOR	MAJOR_SA		SUPPORTED_BY	SYSTEMS
	SUB_SYS_IS	SYSTEMS		SUPPORTS	OBJECTVE
	USES_O_FROM	MAJOR_SA		USES	DATACLSS
OBJECTVE	ACHIEVED_BY	ORGNZTN	SOLUTION	NEEDS	SYSTEMS
	SUPPORTED_BY	PROCESS		RESOLVES	PROBLEM
ORGNZTN	ACHIEVES	OBJECTVE	SYSTEMS	INVOLVED_IN	PROBLEM
	CAUSES	PROBLEM		NEEDED_BY	SOLUTION
	EXPERIENCING	PROBLEM		PART_OF	SYSTEMS
	PART_OF	ORGNZTN		SUBPART_IS	SYSTEMS
	PERFORMS	PROCESS		SUB_SYS_OF	MAJOR_SA
	RSPNSBLE_FOR	DATACLSS		SUPPORTS	PROCESS
	SUBPART_IS	ORGNZTN		USES	DATAFILE

6. PROBLEM—informational problems discovered
7. PROCESS—business processes of the enterprise
8. SOLUTION—suggested solution of PROBLEM
9. SYSTEMS—computer systems in support of business processes

Table 1 summarizes the possible category and relationship names of the BSP model. For example: DATACLSS CREATED_BY PROCESS.

Table 2 shows the names of subject attributes and their values. We apply four basic attributes to each subject in each category:

1. ANALYST_NAME—identifies the team member who collects and maintains the study data. Used primarily in management of the study.
2. DATE—date of update of a BSP subject.
3. ORDER—values that arrange subjects in ascending or descending order. Used in analysis reports for logical grouping.
4. SELECTION—allows for selective retrieval of BSP subjects. Allowing subjects to have multiple selection attributes permits a computer program to retrieve selected subjects based on this attribute. Later sections describe such retrieval and analysis.

Table 2 Attributes of subjects

Category	*Attribute*	*Values of attribute*
DATACLSS	ANALYST_NAME	Name of responsible study team member
	DATE	6-digit numeral (YYMMDD)
	ORDER	1-9999 numeral
	SELECTION	1-13 characters of alphanumeric data
	TYPE_OF_DATA	Allowable values are PLANNING, INVENTORY, TRANSACTION, and SUMMARY
DATAFILE	ANALYST_NAME	Name of responsible study team member
	DATE	6-digit numeral (YYMMDD)
	DF_TYPE	1-13 character description of file type
	ORDER	1-9999 numerals
	SELECTION	1-13 characters of alphanumeric data
MAJOR_SA	ANALYST_NAME	Name of responsible study team member
	DATE	6-digit numeral (YYMMDD)
	ORDER	1-9999 numerals
	SELECTION	1-13 characters of alphanumeric data
OBJECTVE	ANALYST_NAME	Name of responsible study team member
	DATE	6-digit numeral (YYMMDD)
	ORDER	1-9999 numerals
	SELECTION	1-13 characters of alphanumeric data
ORGNZTN	ANALYST_NAME	Name of responsible study team member
	DATE	6-digit numeral (YYMMDD)
	ORDER	1-9999 numerals
	SELECTION	1-13 characters of alphanumeric data
PROBLEM	ANALYST_NAME	Name of responsible study team member
	DATE	6-digit numeral (YYMMDD)
	ORDER	1-9999 numerals
	PROBLEM_TYPE	1-13 character description of problem type
	SELECTION	1-13 characters of alphanumeric data
	STATUS	1-13 character description of status
PROCESS	ANALYST_NAME	Name of responsible study team member
	DATE	6-digit numeral (YYMMDD)
	ORDER	1-9999 numerals
	SELECTION	1-13 characters of alphanumeric data
SOLUTION	ANALYST_NAME	Name of responsible study team member
	DATE	6-digit numeral (YYMMDD)
	ORDER	1-9999 numerals
	SELECTION	1-13 characters of alphanumeric data
	STATUS	1-13 character description of status
SYSTEMS	ANALYST_NAME	Name of responsible study team member
	DATE	6-digit numeral (YYMMDD)
	ORDER	1-9999 numerals
	SELECTION	1-13 characters of alphanumeric data
	STATUS	1-13 character description of status

Table 2 also shows some less frequently used attributes for selected subjects.

Table 3 gives an example of attributes that qualify relationships. For example:

Table 3 Attributes of relationships

Relationship	Attribute	Values of attribute
IMPACTS	CRITICALITY	1-13 character description of criticality of the problem
PERFORMS	INVOLVEMENT	Allowable values are DECISION_MAKER, MAJOR_INVOLVEMENT and MINOR_INVOLVEMENT
SUPPORTS	STATUS	1-13 character description of the degree of support

"inaccurate_data" IMPACTS "budgeting" (CRITICALITY='severely')
"controller" PERFORMS "budgeting" (INVOLVEMENT='major_involvement')
"system_a" SUPPORTS "budgeting" (STATUS='planned_for')

Categories

We now describe the BSP data model in terms of categories and their relationships to other categories (see Figure 4).

DATACLSS (data class) category

The DATACLSS category is used to describe classes of logically related information created and used in an enterprise. Figure 2 shows examples, and the BSP Guide provides approaches to identifying data classes.

There are relationships between DATACLSS and some of the other categories. The CREATED_BY and USED_BY relationships with PROCESS are the basis for developing an information architecture. RSPNSBLE_ORG relates the organizational unit to the data class. For example, RSPNSBLE_ORG for "employee" information might be "personnel director." PART_OF (SUBPART_IS) describes the hierarchical structure of data classes and is essential for subsequent data modeling since, eventually, it would be necessary to identify relationships among all data entities, from BSP data classes to records, segments, and fields. INVOLVED_IN can indicate that a data class is experiencing a PROBLEM in some organizational unit. For example, a DATACLSS subject "sales forecast" may be INVOLVED_IN an "untimely data" PROBLEM. INCLUDED_IN as a relationship with MAJOR_SA (major systems areas) describes the data part of the information architecture. CONTAINED_IN describes where data are maintained in relation to DATAFILE.

DATAFILE (data file) category

The DATAFILE category is used to describe the "containers" of data. These may be automated, as data bases in a computer system, or manual, as the physical drawers of a filing cabinet. Sometimes the definition is expanded to include documents. The attribute DF_TYPE indicates the type of such "containers." The USED_BY relationship with SYSTEMS indicates those data files that support business processes. The "data file CONTAINS data class" relationship is the complement of "data class CONTAINED_IN data file."

MAJOR_SA (major systems areas) category

The MAJOR_SA category is used to describe major systems areas of an information architecture. Major systems areas are logical groupings of business processes and data classes at the enterprise level. They should not be confused with computer-based systems that are defined as SYSTEMS in this model. The INCLUDES relationship with the PROCESS and DATACLSS subjects is used to indicate the business processes supported and the data classes created and used. The SUB_SYS_IS (subsystem is) relationship with SYSTEMS subjects is used to describe the existing and/or proposed systems that fall under the major systems areas of the information architecture. USES_O_FROM (uses output from) and PRDUCS_O_FOR (produces output for) are used to describe the information flow among the major systems areas.

OBJECTVE (objective) category

The OBJECTVE category is used to describe the objectives and goals of an enterprise. They are related to organizational units and business processes by

OBJECTVE is ACHIEVED_BY ORGNZTN
OBJECTVE is SUPPORTED_BY PROCESS

This information is useful in establishing priorities in I/S activities and in identifying potential benefits of the I/S.

ORGNZTN (organizational unit) category

The ORGNZTN category is used to describe the organizational units of the enterprise. Relationships are: PART_OF (SUBPART_IS), describing the hierarchy of organizational units; PERFORMS (with PROCESS), relating the unit to its function, along with the INVOLVEMENT attribute that indicates the degree of each unit's involvement in the process; EXPERIENCING (with PROBLEM subjects), relating the unit to problems, discovered during interviews, that may be informational or associated with other aspects of the business; CAUSES (with PROBLEM subjects), describes the interviewee's perception of which organizational unit causes the problem; and RSPNSBLE_FOR (with DATACLSS subjects), indicating which unit is responsible for which data classes. Examples of these relationships are: The "Plant Operations Director" and "Production Planning Director" are PART_OF the "Vice President of Production." A "personnel director" might be RSPNSBLE_FOR the "employee" information.

PROBLEM (problem) category

The PROBLEM category is used to describe problems discovered in a BSP study, especially during interviews. They need not be informational problems. Some relationships of PROBLEM subjects to other subjects are

PROBLEM is CAUSED_BY ORGNZTN
PROBLEM is EXPRNCED_BY ORGNZTN
PROBLEM INVOLVES DATACLSS

RESOLVED_BY (with SOLUTION) documents solutions suggested by the interviewees. INVOLVES (with SYSTEMS subjects) identifies problem areas in automated support of business processes. PROBLEM_TYPE is an attribute used to clarify the problem, such as "inaccuracy" or "untimeliness."

PROCESS (process) category

The PROCESS category is used to describe the business processes of an enterprise. The PERFORMS relationship and its complement, PERFORMED_BY, express the degree of involvement that organizational units have with processes. SUPPORTS relates objectives and goals (OBJECTVE) to their specific support by a process. The concept of PART_OF (SUBPART_IS) describes the hierarchical structure of processes. This concept allows for the systematic decomposition of business processes and is useful in system design activities. IMPACTED_BY (with PROBLEM subjects) indicates how a process is affected by an informational problem. For example, "Capacity planning is severely IMPACTED_BY late sales forecast data." In this case, "capacity planning" is a subject in the PROCESS category and "late sales forecast data" is a subject in the PROBLEM category. The degree, "severely," is expressed by the CRITICALITY attribute. SUPPORTED_BY (with SYSTEMS subjects) indicates any automated support for the process, whether in effect, proposed, or planned for. The STATUS attribute expresses the distinction. For example: "Funds management" is SUPPORTED_BY "general ledger" STATUS='planned_new.' INCLUDED_IN (with MAJOR_SA subjects) describes that part of the information architecture where support for the process is considered. The USES and CREATES relationships (with DATACLSS) describe data that the process creates and uses.

SOLUTION (solution) category

The SOLUTION category is used to describe possible solutions to PROBLEMs. The solutions may be an interviewee's opinions, or they may be formal proposals made by the BSP study team. RESOLVES relates a PROBLEM to its SOLUTION. DATE and STATUS are attributes that keep track of actions taken. NEEDS (with SYSTEMS subjects) describes I/S needs that interviewees have expressed.

SYSTEMS (computer systems) category

The SYSTEMS category is used to describe computer systems that provide automated support of business processes. The STATUS attribute of the SUPPORTS relationship specifies whether the system is proposed, in development, or operational. The SUB_SYS_OF relationship indicates that the computer-based system is a subsystem of a MAJOR_SA. The USES relationship describes data files needed by the system, and INVOLVED_IN indicates the involvement of the system in informational problems. NEEDED_BY describes an interviewee's expression of I/S needs. PART_OF (SUBPART_IS) describes the hierarchical structure of the computer system.

The Data Dictionary allows for the definition of other characteristics of the data model. For example, the Data Dictionary can validate BSP data during input. Table 2 shows possible specifications of maximum length and allowable characters for names as well as allowable values for subject attributes. The validation can be accomplished by either the standard Dictionary facility or user-written routines. Also, clarifying text can be included with the model definition to explain BSP subject categories and to serve as reminders to the study team members.

Versions of the BSP data model described here have been installed within IBM as well as at several other locations, including Standard Oil of California.[8]

Installation of the BSP data model in the Dictionary

The BSP data model discussed in the previous section should be adequate to describe the BSP study data. However, each study team is encouraged to review and evaluate this model in light of its own situations and areas of emphasis. Some subject categories may require additional relationships and/or attributes. Also, additional subject categories may be required to keep track of certain information for the follow-on I/S management activities.

Next, the BSP data model must be *installed* in the Dictionary, using the Extensibility facility. As stated before, the Dictionary comes with predefined subject categories such as DATABASE and SEGMENT. The subject categories such as ORGNZTN and PROBLEM will be new entries in the Dictionary.

The process of incorporating these nonstandard (user-defined) data models into the Dictionary involves the use of various Dictionary commands such as ADD, ADD_RELATIONSHIP, and INSTALL. These commands can be entered either on line through the Interactive Display Forms facility (IDFF) or in batch mode. We recommend that this task be performed by a data base administrator or a systems programmer with a working knowledge of the Dictionary. Our experience shows that such a person can define and install this BSP data model in a few days.

Following successful installation of the BSP data model in the Dictionary, a report named GUIDE (no connection to the IBM users group of the same name) should be generated. This report serves as a "user manual" and contains all of the BSP subject categories, relationships, and attributes with their associated information.

Storing BSP study data in the Dictionary

After the installation of the BSP model, BSP data can be entered into the Dictionary. As was previously mentioned, the Dictionary has two operating modes. The IDFF is an on-line capability that displays forms with labeled fields and columns that guide the user's entries for specific data. This method of entering data relaxes the syntax rules and helps avoid omissions of parameters, specification of contradictory information, and misspelling of keywords. The IDFF can also increase productivity, because processing of a single form can result in many items of subject and relationship data being entered into the Dictionary. The other mode of operation of the Dictionary is batch. This mode entails the use of a command language that allows users to enter data and to change, delete, and copy stored data, all accomplished by creating batch input job streams. This operation is less efficient than the on-line approach since errors are not discovered at the time of entry but rather later in an output listing. Which mode of operation is chosen depends on the operational environment. However, for the BSP application, we recommend the use of the IDFF.

Reports for BSP

The Dictionary reporting facility

BSP study data stored in the Dictionary data base must be retrieved, analyzed, manipulated, and presented in the forms most convenient for analysis by BSP study team members. The Dictionary provides SCAN and REPORT commands. The SCAN command searches subject names, relationships, or descriptions in a designated subject category for a specified character string or strings. Output can list the subject names under which the string or strings were found. The REPORT command can produce a report on subjects, their attributes, and relationships, with various levels of detail. The REPORT command can also produce a glossary report: a group of summary reports on a set of subjects in the Dictionary data base. Additionally, the IDFF screens can be used to "browse" through the subjects and their relationships in the Dictionary.

Additional reports

The standard Dictionary reporting facility is of an "inquiry" nature. Although it is an important facility, additional "analysis"-oriented reports, such as the ones provided by the Problem Statement Language/Problem Statement Analyzer (PSL/PSA),[3] are required for BSP studies. The BSP methodology, for example, uses matrices extensively for representing relationships between entities (Figure 2). We have created a prototype of such a matrix report since it is not available with the standard SCAN and REPORT commands of the Dictionary. The five useful reports we have identified are (1) MATRIX report, (2) NETWORK report, (3) STRUCTURE report, (4) SUMMARY report, and (5) SELECT report.

MATRIX report

The MATRIX report presents relationships between two categories of entities in a matrix format. That is, it shows how the entities (subjects) of two categories are related. The most important matrix

Figure 5 Process versus data class matrix with an information architecture

PROCESS \ DATA CLASS	PLANNING	FINANCIAL	PRODUCT	PARTS MASTER	BILL OF MATERIAL	VENDOR	RAW MATERIAL INVENTORY	FIN. GOODS INVENTORY	FACILITIES	WORK IN PROCESS	MACHINE LOAD	OPEN REQUIREMENTS	ROUTINGS	CUSTOMER	SALES TERRITORY	ORDER	COST	EMPLOYEE
BUSINESS PLANNING	C	U	U						U					U			U	U
ORGANIZATION ANALYSIS	U																	
REVIEW AND CONTROL	U	U																
FINANCIAL PLANNING	C	U								U								U
CAPITAL ACQUISITION		C																
RESEARCH			U												U			
FORECASTING	U		U											U	U			
DESIGN AND DEVELOPMENT			C	C	U									U				
PRODUCT SPECIFICATION MAINTENANCE			U	C	C	U												
PURCHASING						C											U	
RECEIVING						U	U											
INVENTORY CONTROL							C	C		U								
WORKFLOW LAYOUT			U						C				U					
SCHEDULING			U			U			U	C	U							
CAPACITY PLANNING						U			U		C	U	U					
MATERIAL REQUIREMENTS			U		U	U						C						
OPERATIONS										U	U	U	C					
TERRITORY MANAGEMENT			U											C		U		
SELLING			U											U	C	U		
SALES ADMINISTRATION															U	U		
ORDER SERVICING			U											U		C		
SHIPPING			U					U								U		
GENERAL ACCOUNTING		U				U								U				U
COST PLANNING						U											C	
BUDGET ACCOUNTING	U	U								U							U	U
PERSONNEL PLANNING		U																C
RECRUITING/DEVELOPMENT																		U
COMPENSATION		U																U

LEGEND: U = USES
C = CREATES

SOURCE: BSP INFORMATION SYSTEMS PLANNING GUIDE

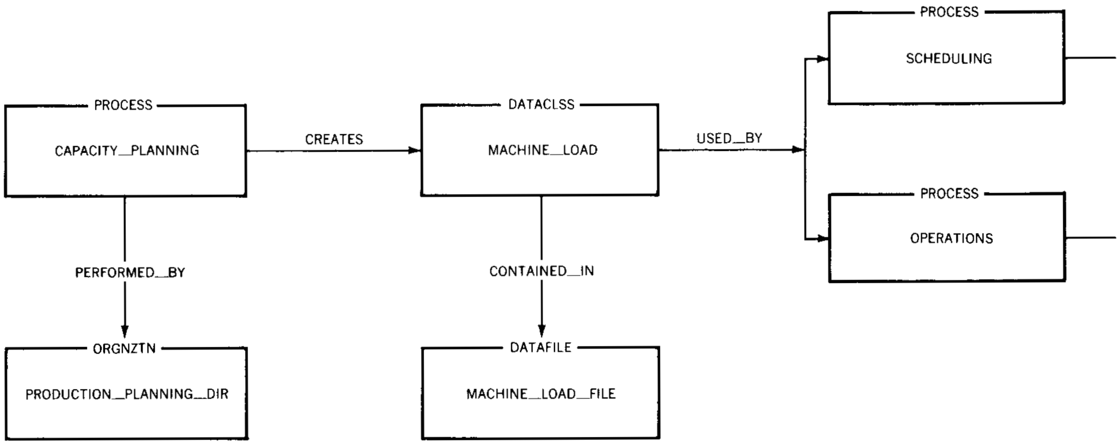
report in a BSP study would be a process versus data class report. This report would show, in a concise matrix form, which processes "create" which data classes and which processes "use" which data classes. Other examples of matrix reports are

- Data class versus data file (DATACLSS versus DATAFILE)
- Organization versus process (ORGNZTN versus PROCESS)
- Organization versus data class (ORGNZTN versus DATACLSS)
- Problem versus data class (PROBLEM versus DATACLSS)
- System versus organization (SYSTEMS versus ORGNZTN)
- System versus process (SYSTEMS versus PROCESS)
- System versus data file (SYSTEMS versus DATAFILE)
- Problem versus process (PROBLEM versus PROCESS)

Some of the functional requirements for MATRIX reports follow:

- In addition to a regular process versus data class matrix report, such as the one shown in Figure 2, a special matrix showing an information architecture, such as the one shown in Figure 5, can

Figure 6 Example of a NETWORK report



be generated. The BSP Guide offers instructions for generating an Information Architecture Diagram which can be used for grouping processes and data classes to identify major systems areas (MAJOR_SA in our BSP data model). Ewusi-Mensah[9] has also developed a technique for grouping processes and data classes, based on some concepts from graph theory. Both techniques produce similar results.

- An information architecture matrix may at times need to be modified. For this reason, the order of row names and column names should be specifiable. This requirement is useful for other matrix reports where logical groupings of subjects are desired.
- Relationships of interest should be selectively specifiable. For example, our BSP data model allows two relationships (ignoring the complement of each) between the PROBLEM and ORGNZTN categories:

  PROBLEM is CAUSED_BY ORGNZTN
  PROBLEM is EXPRNCED_BY ORGNZTN

  If we are only interested in which organizations are experiencing informational problems, we would specify that only the EXPRNCED_BY relationship be shown on the matrix report.
- Multiple relationships should be able to be shown at a single matrix position. If, for example, process-a USES and CREATES data class-x, we should be able to specify a symbol that represents both of these relationships.

NETWORK report

The NETWORK report presents a network diagram of related subjects of interest. This graphical report is intended to foster understanding of various informational entities and their relationships in an enterprise. It would be especially useful in tracing the cause and effect of informational problems and in identifying dependencies of processes and data classes.

Figure 7 Example of a STRUCTURE report

```
1. BUSINESS_PLANNING
    2. CAPACITY_PLANNING
    2. CAPITAL_PLANNING
        3. FINANCIAL_ANALYSIS
    2. PERSONNEL_PLANNING
            •
            •
            •
1. MARKET_FORECASTING
    2. CONSUMER_SURVEY
        3. INTERVIEWS
    2. MARKET_SHARE_ANALYSIS
        3. COMPETITION_ANALYSIS
        3. TREND_ANALYSIS
```

Some of the functional requirements for a NETWORK report are

- Categories and relationships of interest can be selectively specified.
- The starting subject in a NETWORK diagram can be specified.
- Names of subjects should be enclosed (and the categories to which they belong identified) and lines should be used to identify relationships.

An example of a NETWORK report is shown in Figure 6.

STRUCTURE report

The STRUCTURE report presents a hierarchical structure of the subjects in a given category. The structure is implied by the PART_OF (SUBPART_IS) relationship in our particular BSP data model. This report would become especially useful in the subsequent design and development of systems and in the data base administration area. For example:

BUSINESS_PLANNING SUBPART_IS CAPITAL_PLANNING
FINANCIAL_ANALYSIS is PART_OF CAPITAL_PLANNING

An example of a STRUCTURE report is shown in Figure 7. Some of the functional requirements for a STRUCTURE report are

Figure 8 Example of a SUMMARY report for BUSINESS_PLANNING

```
BUSINESS_PLANNING (PROCESS)

  DESCRIPTION: ESTABLISH CORPORATE POLICIES AND LONG-TERM PRODUCT
               STRATEGIES.

  DATE:              05/08/81
  ANALYST_NAME:      JOHN_SMITH
  SELECTION (1):     MANAGEMENT
  ORDER:             123

      PERFORMED_BY:     DEPT_A                                    (ORGNZTN)

      SUPPORTS:         BUSINESS_OBJECTIVE_A                      (OBJECTVE)

      PART_OF:          LONG_RANGE_PLANNING                       (PROCESS)

      SUBPART_IS:       PERSON_POWER_PLANNING                     (PROCESS)
                        CAPACITY_PLANNING                         (PROCESS)
                        CAPITAL_PLANNING                          (PROCESS)

      USES:             ***DATA DOES NOT EXIST IN DATABASE***

      CREATES:          PLANNING_DOCUMENT                         (DATACLSS)

      SUPPORTED_BY:     ***DATA DOES NOT EXIST IN DATABASE***

      IMPACTED_BY:      LATE_SALES_FORECAST                       (PROBLEM)
                        INACCURATE_SALES_DATA                     (PROBLEM)

      INCLUDED_IN:      ***DATA DOES NOT EXIST IN DATABASE***

      SUPPORTED_BY:     TREND_ANALYSIS_PROGRAM                    (SYSTEMS)
```

- Relationship names that imply hierarchies can be specified. For our BSP data model, we have chosen PART_OF (SUBPART_IS). However, to be general, this information should not be contained in the STRUCTURE report program. The design of such a program should be open-ended, and relationship names should be specifiable to the program when it is invoked.
- Selectivity should be allowed. In other words, if one is not interested in seeing the entire structure, it should be possible to specify the "top" of the structure and the number of levels that follow.

SUMMARY report

The SUMMARY report presents, for each subject specified, attribute data, relationship data, and text data (Figure 8). Note that the USES relationship of BUSINESS_PLANNING is not connected to another subject. This incompleteness illustrates the usefulness of this report for completeness checks as well as for presenting concise, summary information about a subject.

Some of the functional requirements for a SUMMARY report are

- Selectivity should be allowed. In other words, one should be able to specify only attribute data, or only relationship data, or only text data as well as some combinations of each type of data.
- The category name of the related subject name should be indicated.
- If no related subject has been specified, it should be so indicated.

SELECT report

The SELECT report presents a list of names of those subjects from the Dictionary data base that satisfy given selection criteria. The intent of this report is *not* to provide a full query facility for the Dictionary, but to provide selective retrieval of subject names. For example, business processes are described in the PROCESS category. In order to distinguish between processes in different functional areas such as manufacturing, marketing, and administration, we would use a SELECTION attribute such as

SELECTION='manufacturing'
SELECTION='market'
SELECTION='admin'

where 'manufacturing,' 'market,' and 'admin' are the values of the SELECTION attribute. By specifying SELECTION='market,' one could SELECT all marketing processes. Multiple values should also be specifiable for this attribute.

Some of the functional requirements for a SELECT report are that

- The selection criteria should allow for relationships or attributes as well as some combinations of both.
- The selection criteria should also allow some comparison operators, such as "greater than," "less than," and "not equal to," to be used with the values of attributes.
- The selected subject names should be ordered before output according to the user's specification, such as in ascending order or descending order.
- Subject names selected should be usable as input to other report programs. This condition allows for selectivity of subjects being analyzed, say, in a matrix report. For example, in a PROCESS versus DATACLSS matrix, one might be interested in only the marketing processes.

As an alternative, the functions of the SELECT report could be incorporated into other reports so that selective analysis of subjects would be possible.

Writing report programs

After the types of additional inquiry and analysis reports have been decided, detailed functional specifications (user interfaces) for the reports must be developed. The functional requirements suggested previously should be considered. Additionally, the detailed specifications of these report programs must be documented, including input parameters and output formats. BSP study team members, who are the end users in this case, should work closely with systems programmers, who are familiar with the Dictionary and its Program Access facility (PAF), in specifying the functional requirements of the reports. The systems programmers would be able to advise on the feasibility of certain requirements and on performance issues (for example, the advisability of default limits and LIMIT parameters for each report to address both performance concerns and output volume control).

Next, the computer programs must be written for generating these reports, using PAF. These PAF programs may be written in ASSEMBLER, COBOL, or PL/I. A sample PAF program for each of these languages is distributed with the Dictionary. They are helpful in understanding how PAF is used to retrieve the necessary BSP data from the Dictionary and to provide the desired results. Each PAF program written is installed (or "link-edited") in the Dictionary and becomes a part of the Dictionary. It can then be executed by specifying its name in the Dictionary EXECUTE command.

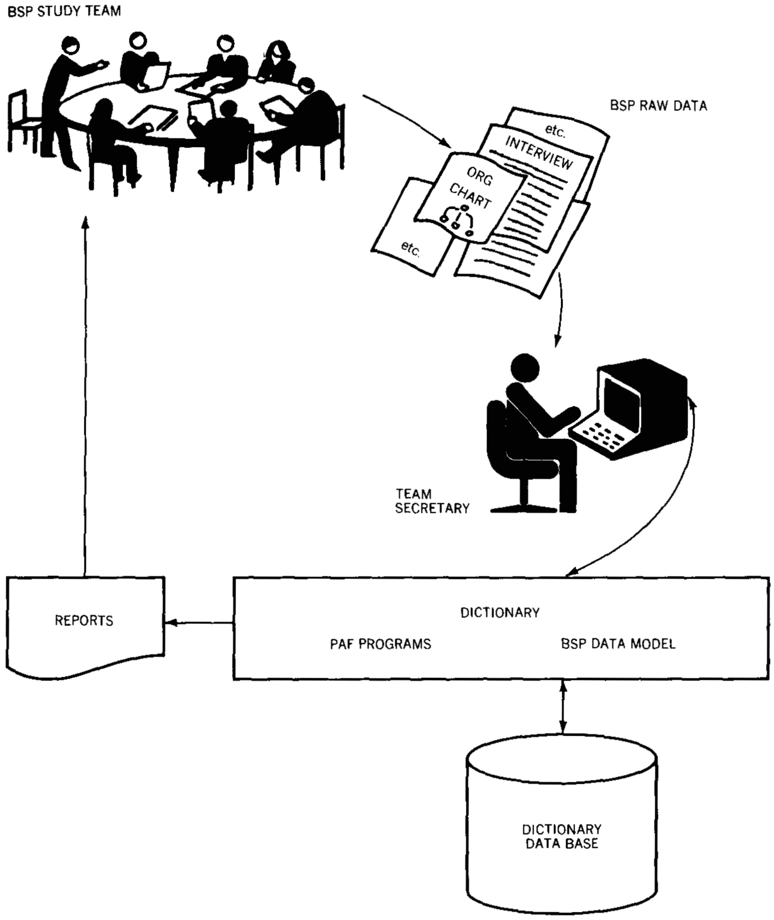
Obviously, the development of PAF programs requires a great amount of coordination between the BSP study team and the PAF programmers. This is the most time-consuming aspect of the approach we are advocating. Therefore, the program specifications, designs, and programming tasks must be scheduled to ensure that the report programs are ready in time for the initial BSP study.

A question may be raised as to the feasibility of this approach if the PAF program development requires several weeks (or even several months) for a two-month BSP study. We feel it is a worthwhile investment since the Dictionary approach can facilitate ongoing BSP-like studies. We may refer to these studies as "mini" BSP studies. After the establishment of a BSP data base with the initial BSP study, it can be updated with the follow-on "mini" study data as situations change. These PAF programs can then be executed to obtain various reports for monitoring, analysis, and planning purposes. We feel that data base administration (DBA), if such a function exists, would be the most suitable organization to carry on this responsibility. These PAF report programs will also be useful in subsequent I/S and data management activities, as discussed in the section on future research.

Preparing for the BSP study and some operational issues

Prior to the recommended two-week preparation phase for a BSP study, the Dictionary must be available and must have the BSP data model installed and the necessary report programs written, debugged, and included in the Dictionary.

Figure 9 A BSP study operational environment



The BSP education plan should include an overview discussion of the Dictionary, the BSP data model, and the planned uses of the report programs so that the team members are aware of what to expect from this tool. It is particularly important to note the benefits of the approach to the enterprise as well as to the individual study team members.

Not every team member is expected to interface with the Dictionary. Perhaps the team secretary would be the most appropriate person for this job. In this case, the secretary should be trained to enter BSP data into the Dictionary, as well as modify the data and request up-to-date reports from the Dictionary. The secretary can perform these activities interactively through the Interactive Display Forms facility (IDFF) or in batch mode. We recommend using the IDFF because of its ease of use. Figure 9 illustrates an operational environment, with the team secretary interfacing with the Dictionary. To facilitate

Figure 10 An example of a PROCESS form

Name of PROCESS: ______________________ DATE: __________

Description: ______________________

ORDER value: ________ ANALYST_NAME: ______________

SELECTION values: __________, __________, __________
__________, __________, __________

PERFORMED_BY (organization names): ______________

SUPPORTS (objective names): ______________

PART_OF (process name): ______________

SUBPART_IS (process names): ______________

USES (data class names): ______________

CREATES (data class names): ______________

SUPPORTED_BY (system names): ______________

IMPACTED_BY (problem names): ______________

INCLUDED_IN (name of major systems areas): ______________

Notes: ______________

entering of the data, forms for the subject categories could be developed. The team members would fill in these forms in the course of the interviews. Figures 10 and 11 show examples of such forms for the PROCESS and DATACLSS categories.

Each BSP data subject must be given a name, such as capital planning, budgeting, and forecasting. It is important that a naming convention be established so that BSP subject names are meaningful to all parties involved. Such a convention should include a short description of each subject name. When it is difficult to resolve naming problems, the alias feature of the Dictionary may be used. For example, people in manufacturing and those in marketing may, in the course of conducting daily business, refer to the same informational entity with different names. In such a case, aliases allow reference to the identical entity by different names. Another concern is that some of the BSP data collected could be of a sensitive nature. Operational standards and procedures should be established to ensure both the privacy and security of data. The Dictionary provides

Figure 11 An example of a DATACLSS form

Name of DATACLSS: ______________________ DATE: __________

Description: ______________________

ORDER value: ______ ANALYST_NAME: ______________________

SELECTION values: ________ , ________ , ________ ; ________ , ________ , ________

TYPE_OF_DATA: ______________

CONTAINED_IN (data file name): ______________________

PART_OF (data class name): ______________________

SUBPART_IS (data class names): ______________________

USED_BY (process names): ______________________

CREATED_BY (process names): ______________________

INCLUDED_IN (name of major systems areas): ______________________

INVOLVED_IN (problem names): ______________________

RSPNSBLE_ORG (name of responsible organization): ______________________

Notes: ______________________

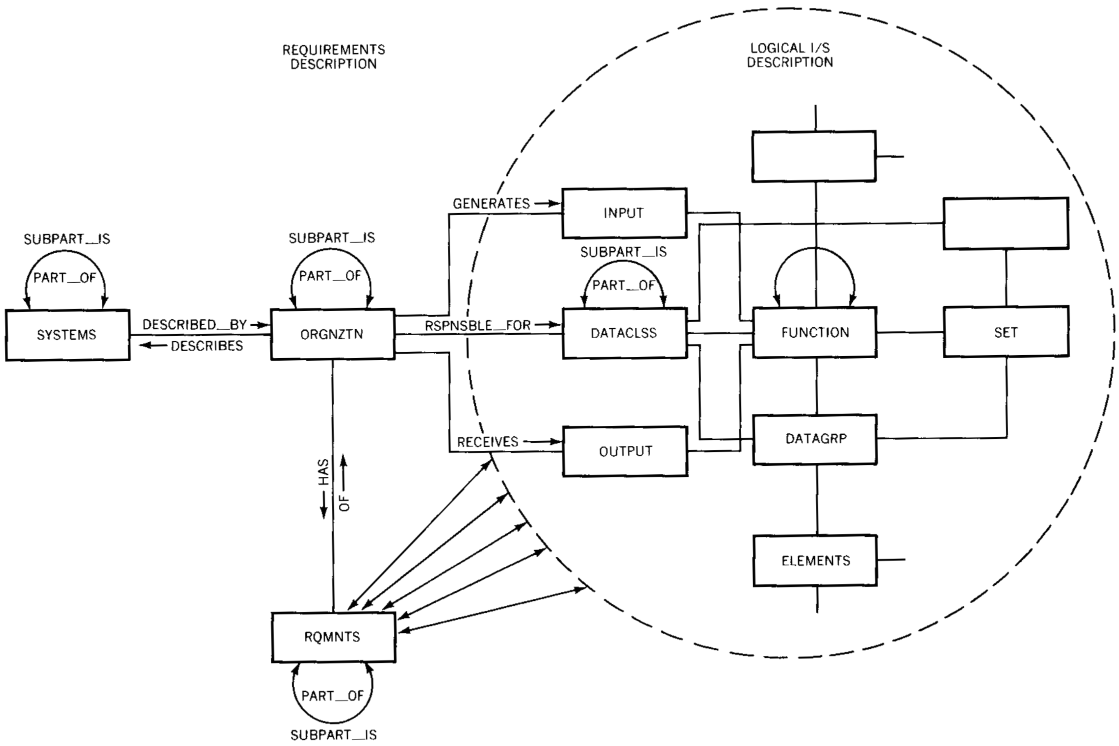
a security facility for controlling access to Dictionary definitions. Its use requires the creation of "access profiles" which can limit individual users' access to subsets of the data model.

Future research

The Dictionary has more potential than its use solely as a tool for BSP studies or a tool for data base administration. As stated previously, one greater value lies in its ability to offer an integrated approach to the management of activities from the BSP study on through requirements definition, data base administration, systems design, development, and maintenance (see Figure 1).

The rectangular groupings in the information architecture shown in Figure 5 depict major systems areas (which are subjects in the MAJOR_SA category shown in Figure 4). The BSP Guide offers guidelines in determining which major systems areas and which

Figure 12 Integration of BSP follow-on activities



systems (SYSTEMS category) should be implemented first. Once this decision has been made, the next step will be to gather, define, and analyze the user requirements for the system in question. Figure 12 depicts a conceptual extension of the BSP data model into the realms of requirements and logical I/S description. The only categories shown from Figure 4 are SYSTEMS, ORGNZTN, and DATACLSS since they will now require some additional relationships. Functional decomposition (SUBPART_IS) would be applied to the subjects in the ORGNZTN category until these subjects (departments, users, customers, etc.) were at the appropriate level to describe the user requirements for the system to be implemented. These user requirements would be captured as subjects in the RQMNTS category. Each of these subjects would be connected to one or more subjects in the ORGNZTN category by a relationship (shown here as OF (HAS) for simplicity).

Logical I/S description should delineate any interactions between the system and the environment in which it exists. This concept is achieved, in this example, by the relationships between the Logical I/S Description data model (shown in skeletal form) and the ORGNZTN category.

As the user requirements are refined (SUBPART_IS), they will be assigned to the appropriate category in the Logical I/S Description data model. This concept is illustrated by the relationships (arrows) on the RQMNTS category. This refinement of requirements could conceivably continue through the physical system design phase (Figure 1) with additional relationships between the RQMNTS category and a Physical System Design data model (not shown).

An approach of this nature would enhance the ability to determine if the original user requirements were being satisfied, since one could trace a given development specification back to the original requirements (and their organizational source) with PAF programs. Also, this approach would facilitate analysis in the areas of completeness and consistency checks, change impact/management, and project management.

Future areas of research include

- Investigation into the appropriate definition of subjects, relationships, attributes, and naming conventions so as to facilitate multiple requirements description disciplines and techniques.
- Development of data models for subsequent phases. All such models need to be logically related.
- Development of techniques for iterating through the different phases. Perhaps the notion of the traditional system life cycle would change considerably in this environment.
- Identification and development of various inquiry and analysis reports to assist all phases of the system life cycle.
- The study of management and personnel issues in this new environment so that the productivity of the information-providing services of an enterprise can be improved.

Concluding remarks

We have shown an approach in which the IBM DB/DC Data Dictionary can be used as a tool to support Business Systems Planning studies. We have defined a BSP data model in terms of Dictionary subject categories, relationships, and attributes, and identified some of the useful reports that can be implemented as PAF programs. We have also discussed some of the operational issues involved in implementing this approach.

Although this approach is only one of several possible, we believe it is a viable one which can contribute greatly to BSP studies. However, it has a greater value in its potential to offer an integrated approach for managing activities from the BSP study on through requirements definition, data base administration, systems design, development, and maintenance. The approach shown here, we feel, is the logical beginning of such an integrated approach to managing I/S activities.

ACKNOWLEDGMENTS

We are grateful to C. Coray, R. Donovan, P. Fehder, M. Fleisher, C. Mapes, R. Schouw, R. C. Summers, B. P. Whipple, J. Wright, and J. A. Zachman for their valuable comments and suggestions on this work and earlier related efforts. This paper is based on one by Sakamoto published in *The Economics of Information Processing*, Vol. 1, John Wiley & Sons, Inc. (1981).

CITED REFERENCES

1. *Business Systems Planning—Information Systems Planning Guide*, GE20-0527-3, IBM Corporation (July 1981); available through IBM branch offices.
2. C. E. Peterson, "Software support for maintaining and analyzing BSP output," Discussion Paper, SHARE Data Administration Project, SHARE 53.5 (November 12, 1979).
3. D. Teichroew and E. A. Hershey III, "PSL/PSA: A computer-aided technique for structured documentation and analysis of information processing systems," *IEEE Transactions on Software Engineering* **SE-3**, No. 1, 41–48 (January 1977).
4. *SQL/Data System: General Information*, GH24-5012 (Program Product Number 5748-XXJ), IBM Corporation (January 1981); available through IBM branch offices.
5. *Query-by-Example: Program Description/Operations Manual*, SH20-2077 (Installed User Program Number 5796-PKT), IBM Corporation (September 1978); available through IBM branch offices.
6. *DB/DC Data Dictionary—General Information Manual*, GH20-9104, IBM Corporation (March 1980); available through IBM branch offices.
7. J. G. Sakamoto, *Use of DB/DC Data Dictionary to Support Business Systems Planning: An Approach*, G320-2705, IBM Corporation, Los Angeles Scientific Center (July 1980); available through IBM branch offices.
8. C. Estep, "DB/DC Data Dictionary—A Business System Planning model," presentation by Standard Oil of California at GUIDE 52, Session No. SD-33, Atlanta, GA (May 1981).
9. K. Ewusi-Mensah, *Criteria for Decomposing an Information System into Its Subsystems for Business Systems Planning*, G320-2702, IBM Corporation, Los Angeles Scientific Center (March 1980); available through IBM branch offices.

GENERAL REFERENCES

The following references are available through IBM branch offices.

Business Systems Planning—Planning for Distributed Information Systems, GE20-0655-1, IBM Corporation (January 1980).

K. Ewusi-Mensah, *Computer-Aided Modeling and Analysis Techniques for Determining Management Information Systems Requirements*, G320-2703, IBM Corporation, Los Angeles Scientific Center (March 1980).

DB/DC Data Dictionary—Administration and Customization Guide, SH20-9174, IBM Corporation (July 1979).

DB/DC Data Dictionary—Application Guide, SH20-9173, IBM Corporation (July 1979).

DB/DC Data Dictionary—Release 3—Implementation Primer, G320-5780-0, IBM Corporation (September 1979).

DB/DC Data Dictionary—Installation Guide, SH20-9084, IBM Corporation (July 1979).

DB/DC Data Dictionary—Terminal User's Guide and Command Reference, SH20-9083, IBM Corporation (July 1979).

The authors are located at the IBM Scientific Center, P.O. Box 45013, Los Angeles, CA 90045.

Reprint Order No. G321-5161.