

A SURVEY OF THE DATA ADMINISTRATION FUNCTION
IN LARGE CANADIAN ORGANIZATIONS

by

IAN BRYCE MCCRIBICK

B.Sc., University of British Columbia, 1967

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE IN BUSINESS ADMINISTRATION

in

THE FACULTY OF GRADUATE STUDIES
(Commerce and Business Administration)

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

June 1979

© Ian Bryce McCririck , 1979

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Commerce and Business Administration

The University of British Columbia
2075 Wesbrook Place
Vancouver, Canada
V6T 1W5

Date June 30, 1979

ABSTRACT

The object of this study was to survey large Canadian organizations in order to:

- 1) determine the extent to which these organizations have established a separate Data Administration function,
- 2) empirically test Nolan's Stage Model of EDP Growth as a predictor of a separate Data Administration function, and
- 3) survey the characteristics of the Data Administration function in those organizations that have formally established such a speciality.

A survey package containing two questionnaires was sent to 555 large Canadian organizations in the private and public sectors. The "EDP Profile Questionnaire" was directed to the Manager of the EDP Activity in the surveyed organizations. This questionnaire is concerned with the EDP growth process and the existence of a Data Administrator. The "Data Administration Questionnaire" was directed to the Data Administrator in the surveyed organizations. This questionnaire is concerned with the characteristics and responsibilities of the Data Administration function. Analysis was performed on 254 EDP functions and 69 Data Administration functions.

The results obtained indicate that the Data Administration function is not prevalent in large Canadian organizations; where the function does exist its role is a fairly minor one within the EDP activity. This study found that organizations with very large EDP activities and many years of experience

with computers were more likely to have established a Data Administration function than smaller and less experienced ones. Certain organizational types (those with discretionary funds available) were more likely to have a Data Administration function than other types. The "maturity" of the organization's EDP activity was not found to be a good predictor of the existence of a Data Administrator. The sampled Data Administration functions exhibited a wide dispersion in both the activities performed and the amount of time spent on each. Few policy setting activities were performed by the Data Administrators. The Data Administration function appeared to be focused on those "data bases" using a Data Base Management System. Organizational conflicts and a general misunderstanding of the function by EDP Management have likely held back the development of the function beyond one involved primarily with the support of DBMS application systems.

Future research should be directed at understanding these conflicts and misperceptions through an analysis of the decision process involved in establishing the Data Administration function. An attempt should be made to more fully understand the data resource and how it might differ among organizational types. Before further use is made of Nolan's Stage Growth model, serious thought should be given to determining in more precise terms what the EDP growth process variables are and how they might best be measured.

TABLE OF CONTENTS

	Page
CHAPTER I. INTRODUCTION	1
Data Administration	1
Objectives	3
Scope of the Study	3
The Data Resource	4
Stage Model of EDP Growth	7
CHAPTER II. HYPOTHESES	11
The Data Administration function in large Canadian Organizations	11
Nolan's Stage Model of EDP Growth	12
Data Administration Specialization	12
Data Administration Responsibilities	14
CHAPTER III. METHOD	16
Questionnaires	16
The EDP Profile Questionnaire	17
The Data Administration Profile Questionnaire	19
The Pretest	21
The Sample	21
Survey Package	22
Follow Up	23
CHAPTER IV. ANALYSIS AND RESULTS	24
Preliminary Statistical Analysis	24
Survey Package Response	24
The Data Administration Function in Large Canadian Organizations	27
Hypothesis Tests	39
Stage Model of EDP Growth	39
Data Administration Specialization	43
Data Administration Responsibilities	50
CHAPTER V. SUMMARY, DISCUSSION AND CONCLUSIONS	55
Survey Package Response	55
The Data Administration Function in Large Canadian Organizations	57
Stage Model of EDP Growth	59
Data Administration Specialization	60
Data Administration Responsibilities	63
Role of the Data Administrator	66
Conclusions	69
REFERENCES	71
APPENDICES	74

APPENDIX A. EDP Profile Questionnaire	74
APPENDIX B. Data Administration Profile Questionnaire	81
APPENDIX C. Covering Letter	90
APPENDIX D. Follow Up Card	92

LIST OF TABLES

	Page
TABLE I. Response Tabulation	24
TABLE II. Private Sector Responses	26
TABLE III. Public Sector Responses	27
TABLE IV. Responses to Question 7 of EDPPQ	28
TABLE V. DA vs Number of EDP Employees	30
TABLE VI DA vs Number of EDP Employees-	
One-tailed T-test	30
TABLE VII. DA vs Monthly Hardware Expenditure	32
TABLE VIII. Da vs Monthly Hardware Expenditure-	
One-tailed T-test	32
TABLE IX. DA vs Experience With Computers	34
TABLE X. DA vs Experience With Computers-	
One-tailed T-test	34
TABLE XI. DA vs Organizational Type (Frequencies)	37
TABLE XII. DA vs Availability of Funds	38
TABLE XIII. EDP Employees vs Availability of Funds	38
TABLE XIV. Correlations of DA with Maturity Variables .	42
TABLE XV. Reporting Level of DA	44
TABLE XVI. Size of DA- Number of Employees	44
TABLE XVII. Experience of DA	45
TABLE XVIII. Scope of DA- Fraction of Machine	
Readable Data Under DA Responsibility	46
TABLE XIX. Size of Machine Readable Data Resource	46

TABLE XX. Sharing of Data	47
TABLE XXI. Use of DA Tools	48
TABLE XXII. Computer Applications Using DBMS	48
TABLE XXIII. DA Experience vs Use of Tools	49
TABLE XXIV. Background of Data Administrator	49
TABLE XXV. DA Experience and Time Spent on Technical and Administrative Responsibilities	51
TABLE XXVI. DA Experience and Time Spent on Technical Responsibilities- One-tailed T-test	52
TABLE XXVII. DA Experience and Time Spent on Administrative Responsibilities- One-tailed T-test	53
TABLE XXVIII. Responsibilities	54

LIST OF FIGURES

	Page
FIGURE 1. Roles of the Data Administrator	65
FIGURE 2. Relationship of Players in the Decision to Form a Data Administration Function	68

ACKNOWLEDGMENTS

This study would not have been possible without the much appreciated help, advice, and encouragement of my thesis committee: Dr. Robert C. Goldstein (chairman), Dr. Albert I. Dexter, and Dr. Craig C. Pinder. I would also like to thank my employer, B. C. Hydro and Power Authority for their help in printing the questionnaires. The financial support for the mail survey was provided from the Accounting Development Fund. Of course all the organizations that returned questionnaires deserve mention, as without their responses this study could not have been performed.

CHAPTER I

INTRODUCTION

Data Administration

In the past few years many organizations have recognized that data are a valuable corporate resource and have taken steps to manage and control this resource. For some of these organizations, action has resulted in the establishment of a separate group to manage the data resource.¹ The most important manifestation of the institutionalization of the data resource function is the appearance of the Data Base Administrator position in organizations.²

In general, the Data Administration function has been associated closely with the EDP activity. Clearly the computer has been the major impetus to specialization in the data resource.³

This relationship evolved because of the large volume of computerized, or machine readable, data that application systems used.

¹ Throughout this study the terms Data Administrator and Data Administration will be used to refer to the separate organizational unit concerned with managing the corporate data resource.

² Richard L. Nolan, ed., Managing The Data Resource Function, West Publishing, 1974, page 5.

³ loc. cit.

During the 1960's, the computerized application systems were designed with their own data files. Both the data and the responsibility for the data were fragmented - and many of those are still operating today. During the 1970's one major trend has been towards the use of the shared-data-base that serves multiple applications. Numerous benefits are being obtained from data bases, including the reduction of undesired redundancies and incompatibilities of data. But, at the same time, data bases have brought into focus the need for assigned responsibility and control of the organization's data resource. The data administrator function is being set up and given this responsibility.⁴

In short, the EDP activity has revealed to corporate management the need for better management and control of the corporate data resource.

The position of Data Administrator is a policy oriented job concerned with corporate planning and strategy. The Data Administrator is the custodian of the corporate data resource, much as the financial executive or accountant is the custodian of corporate funds; s/he has the job of setting up the infrastructure and controlling the evolution of data base systems (not necessarily limited to machine readable data) in a corporation. In summary, the Data Administrator is responsible for understanding, maintaining, developing, employing and integrating the potential in the resource data.⁵

⁴ Canning, Richard G., "Installing a Data Dictionary", EDP Analyzer, (16:1), 1978 January, page 1.

⁵ Nolan, ibid, page 24.

Objectives

The object of this study was to survey large Canadian organizations in order to:

- 1) determine the extent to which these organizations have established a separate Data Administration function,
- 2) empirically test Nolan's Stage Model⁶ of EDP Growth as a predictor of a separate Data Administration function, and
- 3) survey the characteristics of the Data Administration function in those organizations that have formally established such a speciality.

Scope of the Study

This study surveyed the 500 largest Canadian organizations in the private and public sectors. 'Largeness' was determined by sales, assets, number of employees or size of computers used.

According to Kimberly⁷ there has been little theoretical development of the concept of size in studies of size and structure. He contends that a contingency approach is required when selecting a measure of size to use in analyzing organizational structure. The optimum measure for analyzing the Data Administration function was not established a priori

⁶ The model is discussed later in Chapter I.

⁷ John R. Kimberly, "Organizational Size and the Structuralist Perspective, A Review, Critique, and Proposal", Administrative Science Quarterly, (21:4), December 1976, pages 571-579.

because a multiplicity of organizational types were being surveyed. Consequently, a number of known measurements of size were used to select respondents.

The Data Resource

Several terms are in general use when discussing the corporate data resource, for example, 'data', 'information', or 'data base'. Data may be defined as groups of non-random symbols which represent quantities, actions, things, etc. Data is formed from characters. These may be alphabetic, numeric, or special symbols such as *, \$, and %.⁸

Thus data may take the form of characters on documents, electronic impulses, a manager's knowledge, and so on. The quantity of data that exists in any organization is obviously voluminous and beyond the comprehension and processing capacity of an individual. A subset of the total organizational data resource is of interest to the manager. This subset is usually known as information, or data that has been processed into a form that is meaningful to the recipient and is of real or perceived value in current or prospective decisions.⁹ Thus the data base is all the data (or information) which is maintained

⁸ Gordon B. Davis, Management Information Systems-Conceptual Foundations, Structure, and Development, McGraw-Hill, 1974, pages 32-33 .

⁹ op. cit. ,page 32.

by an organization in order to facilitate its operations.¹⁰ Clearly, managing the data resource is an important function in any organization, as data are the raw input to managerial decisions.

In order to study the corporate data resource, it is useful to classify data according to the following dichotomies:

- 1) Public/Private,
- 2) Formal/Informal, and
- 3) Machine Readable/Non Machine Readable.

Public data refers to data that is available to anyone who is authorized to obtain the data, whereas private data is usually confined to an individual's personal collections of data in 'black books'.

Formal data refers to data that has regular routines or processes associated with its capture, storage, maintenance and dissemination. Examples of formal data are job bulletins, pay cheques and inventory records. Informal data refers to data of an 'ad hoc' nature, for example, special studies conducted to assess a product market, corporate 'politics' and 'one-shot' computer programs. The dichotomy of formal/informal data is similar to the dichotomy of programmed and nonprogrammed decisions:¹¹ programmed decisions are repeated and well defined whereas nonprogrammed decisions are occasional and unique. Machine readable data refer to data that can be processed by computers or similar machines.

¹⁰ R. C. Goldstein, in Nolan, ibid., pages 349-350.

¹¹ For a discussion of programmed and nonprogrammed decisions see Davis, ibid., pages 144-145.

In any one organization, the category of public data is generally larger than that of private data¹². More of the informal/non-machine-readable data will likely become formal/machine-readable data as computer technology and management expertise evolve to new levels.

In understanding the data resource, it is also useful to consider it from the perspective of Anthony's management decision making levels¹³ -operational, management control and strategic planning. For example, a large fraction of the machine-readable category is used for operational decision making whereas a similar proportion of the non-machine-readable category is used for strategic decision making.¹⁴

Information used in strategic planning is usually developed from 'ad hoc' studies, personal contacts and 'external data'.¹⁵ As a consequence, the importance of managing the data resource function as perceived by top management may be quite different from that perceived by operational level management.

¹² Davis, ibid., page 199.

¹³ Robert N. Anthony, Planning and Control Systems: A Framework for Analysis, Harvard Graduate School of Business Administration, Boston, 1965.

¹⁴ Richard L. Nolan, "Managing the Crises in Data Processing", Harvard Business Review, (57:2), March-April, 1979, page 119.

¹⁵ For a discussion comparing the characteristics of data used in strategic planning, management control, and operational level decision making, see G. A. Gorry and M. S. Scott Morton, "A Framework for Management Information Systems", Sloan Management Review, Fall 1971.

Historically, the EDP function has developed around the public/formal/machine-readable/operational-level data resource category. As a consequence, top and middle management are probably not well aware of the data resource concept. The privacy and security of data, however, are two Data Administration issues that should be of major concern to them. For example, the inadvertent disclosure of customer data could result in the loss of a valued client; the destruction of financial data could make an organization insolvent.

Stage Model of EDP Growth

Nolan¹⁶ contends that data is an organizational resource and, as such, should be managed as any other corporate resource (manpower, money, material etc.). In organizing to manage the data resource, a certain degree of specialization (or division of labour) will emerge. Nolan asserts that the data resource function will emerge from the EDP function, which is, in fact, the initial organizational structure established to manage the corporate data resource. He has developed a six stage model¹⁷ of the evolution of the data resource function in organizations. Several variables are used to determine where

¹⁶ Nolan, Managing The Data Resource Function, page 22.

¹⁷ Nolan originally conceptualized a four stage model but now has expanded the model to six stages. Over the last six years several articles on this model have appeared in the literature. In addition to the work referenced in footnotes above, Nolan has also published, "Thoughts about the Fifth Stage", Data Base, (7:2), 1975, pages 4-10; and "Organizational Response and Information Technology", Proceedings of the National Computer Conference 1978, pages 517-524.

an organization's EDP function is on a 'maturity' scale. Nolan contends that the appointment of a Data Administrator is characteristic of the fifth stage of maturity.

Dearden¹⁸ on the other hand, contends that centralizing the control of a company's information system is simply not feasible. He suggests that it is impossible to find an individual who combines the requisite skills in computing technology with an understanding of all the functional application areas of business and a knowledge of the management process. Thus, it is impractical to centralize control over the corporate data resource.

Dearden begins with the assumption that management information is a subject for specialization; that is, the data resource is sufficiently homogeneous for a set of principles and practices to be established for evaluating all management information needs and satisfying them. He then counters that a complete information system consists of such a huge assortment of different types of activities that no single person can possess a broad enough set of skills to apply to even a small portion of them.

Nolan's assertion is supported by Ackoff¹⁹ who suggests that managers do not really know what information they need to manage successfully. Ackoff suggests that there is a place for

¹⁸ John Dearden, "MIS is a Mirage", Harvard Business Review, (50:1), 1972 January-February, pages 90-99.

¹⁹ Russell Ackoff, "Management Misinformation Systems", Management Science, (14:4), 1967 December, pages B147-156.

the specialist to work with the manager in determining his information needs.

Nolan²⁰ states that specialization refers to the division between line and staff group responsibilities for managing the data resource. For example, the line group may perform the data entry activity but the staff group would establish rules, regulations and policies governing the entry of data to computers.

According to Nolan, the growth of the EDP activity has six stages- Initiation, Contagion, Control, Integration, Data Administration and Maturity, and four growth processes- Application Portfolio Development, Data Processing Organization, Data Processing Planning and Control, and User Awareness²¹ within each stage. A number of variables determine the stage of growth for each process. Nolan, however, has never made explicit what these variables are or how they are to be measured. His discussion of variables lacks precision and consequently replication of his studies is difficult.

The present study draws upon the work of Mantha ²² and Nolan, in building a suitable list of variables to measure the stage of development of the EDP activity. The major variables to be used as indicators of the EDP growth process are:

²⁰ Nolan, Managing the Data Resource Function, page 23.

²¹ Nolan, "Managing the Crises in Data Processing", ibid

²² Robert Mantha, "An Analysis of Skill Requirements in Data Processing Environments", unpublished master's thesis, University of British Columbia, 1978.

- 1) EDP activity size as measured by the number of EDP employees and/or monthly computer hardware expenditure,
- 2) user awareness as measured by involvement of top management and EDP users in computer system development activities,
- 3) planning and control mechanisms as measured by the existence of a chargeback process, standards for system development tasks, project reviews, a Strategic Planning Committee and a corporate data flow model,
- 4) position of EDP activity in the organizational structure as measured by the reporting relationship to top management,
- 5) age of the EDP activity as measured by the number of years since computers were first used by the organization,
- 6) applications portfolio development as measured by the fraction of the EDP budget spent on operational level type computer systems, and
- 7) EDP function performance evaluation as measured by the importance of clerical staff savings, meeting budgets, user satisfaction and contribution to organizational goals.

CHAPTER II

HYPOTHESESThe Data Administration Function In Large Canadian Organizations

According to Canning²³ a large fraction of organizations should now have established a Data Administration function because the net benefits are substantial.²⁴ This study will estimate the fraction of organizations that have established a Data Administration function, and compare these organizations with those that don't have the function, for differences in EDP activity size, experience and organizational type. Since the Data Administrator is to play a coordinating role in the management of the data resource, one might expect organizations with very large EDP activities to be more likely to have established the function than smaller ones. Experience refers to the number of years that the organization has been using computers. Various organizational types, because of the nature of their business, may assimilate computer technology differently. For example, point of sale terminals impact department stores more than they do metal producers.

²³ Canning, ibid

²⁴ For a discussion of benefits see G. K. Wiorkowski and J. J. Wiorkowski, "Does a Data Base Management System Pay Off?", Datamation, April 1978, pages 109-114; and GUIDE International Corporation, Establishing the Data Administration Function, 1977.

Consequently, the emergence of the Data Administration function may occur at different stages of EDP growth in different organizational types.

Nolan's Stage Model of EDP Function Growth

Nolan's Stage Model of EDP Function Growth posits a Data Administrator position emerging at the fifth stage; if his assumptions are correct, those organizations that score high in growth process variables should be more likely to have formally established a Data Administrator than those that score low. This study will test the hypothesis that:

- H1: More mature organizations are more likely to have established a Data Administration function than less mature organizations.

The term, mature, as applied to the EDP activity, will mean a level of sophistication with respect to Nolan's growth process variables.

Data Administration Specialization

For those organizations that have established a separate Data Administration function, the following characteristics of the function will be reported to determine the degree of specialization that has evolved in the function-organizational structure, experience, data sharing, scope, and tools. Generally, the higher the reporting position of the Data Administrator, the more specialized the function. For example, it is not unusual to see a Vice President for Finance (money resource). Experience is a measure of specialization as time

is required to learn about and apply the concepts.²⁵ Since a reduction in the amount of data redundancy is a major benefit attributable to Data Administration²⁶ the Data Administrator should attempt to integrate data where feasible and practical. As the Data Administration function becomes more specialized, more application systems should have been integrated,²⁷ and as a consequence data redundancy reduced. The evidence of integration of the data resource as manifested in the concept of shared files, is a possible measure of the degree of data redundancy. As the Data Administration function becomes more specialized, one might expect its role to expand in scope to include, for example, the non-machine-readable data resource. The application of new tools to manage the data resource is a measure of the degree of specialization in the function.²⁸ This assertion could be rephrased as the hypothesis:

H2: More experienced Data Administration functions will be more likely to use certain tools (e.g. Data Dictionary/Directory) than less experienced ones.

"Experience" is the number of years a separate Data Administration function has been established in the

²⁵ C. Babbage, in Nolan, Managing the Data Resource Function, page 7.

²⁶ Wiorkowski, ibid.

²⁷ Richard L. Nolan, "Computer Data Bases: The Future Is Now", Harvard Business Review, September-October 1973, page 113-114.

²⁸ Babbage, ibid.

organization. The study will also report the experience and background of the Data Administrators.

Data Administration Responsibilities

For those organizations that have established a separate Data Administration function, hypotheses about the responsibilities and experience of the Data Administration function will be tested. "Experience" refers to the number of years the group has been in existence.

H3: More experienced Data Administration functions will spend a larger fraction of their time on administrative responsibilities than less experienced ones.

H4: More experienced Data Administration functions will spend a smaller fraction of their time on technical responsibilities than less experienced ones.

H5: All Data Administration functions will spend a larger fraction of their time on administrative responsibilities than technical responsibilities.

The time currently spent on different responsibilities will be compared to the expected time to be spent on the same responsibilities two years hence. Administrative responsibilities are concerned with managerial or clerical tasks (policy setting or data dictionary maintenance). Technical responsibilities are concerned with computer technology (the physical storage of data or maintaining a Data Base Management System). This dichotomy of tasks is put forth

by Secrest²⁹ without any indication of the importance of each activity. If Data Administration is concerned with managing the data resource, an emphasis on administrative as opposed to technical tasks would be expected. Furthermore, if Data Administration is an evolving speciality, as suggested by Nolan, a trend should develop towards administrative responsibilities and away from technical responsibilities.

²⁹ Richard D. Secrest, "The Data Administrator", in McFarlan & Nolan, The Information Systems Handbook, Dow Jones-Irwin, 1975, pages 647-658.

CHAPTER III

METHOD

Questionnaires

This study planned to survey 500 geographically dispersed organizations. Since time and funds were at a premium, a technique that gives wide coverage at minimal expense was required. The questionnaire approach was selected as the best means for obtaining data. Several disadvantages inherent in the technique include, for example, the following possibilities:³⁰

- 1) a low response rate, which may bias the sample being examined,
- 2) questions may be misinterpreted by the respondent,
- 3) answers may be missing, and
- 4) answers may be purposefully or inadvertently distorted in a direction favourable to the sponsor of the survey or the respondent.

Several counter measures can be taken to minimize the occurrence of these events which tend to jeopardize the

³⁰ The list is not exhaustive. A more extensive discussion of the mail survey method can be found in L. Charach, "An Exploratory Investigation Toward The Development Of A Research Design For The Study of Youth Work Transition", unpublished master's thesis, University of British Columbia, 1977.

(Mantha, ibid, in a study concerning EDP managers of British Columbia based companies used Charach's methodology and achieved a response rate of 57 %).

reliability and validity of a mail survey. A well presented and professional looking questionnaire that attracts the interest of the respondent, a covering letter from a person who has credibility in the subject area, a pre-paid self-addressed return envelope and a follow up card after the initial mailing are prerequisites to obtaining a good response rate.³¹

Two questionnaires were designed, developed and pretested. The "EDP Profile Questionnaire" (Appendix A) was directed to the Manager of the EDP Activity in the surveyed organizations. This questionnaire is concerned with the EDP growth process and the existence of a Data Administrator. The "Data Administration Questionnaire" (Appendix B) was directed to the Data Administrator in the surveyed organizations. This questionnaire is concerned with the characteristics and responsibilities of the Data Administration function. Because of the possibility that the positions of EDP Manager and Data Administrator are held by different individuals, two questionnaires rather than one were developed.

The EDP Profile Questionnaire

The first questionnaire consists of a cover page followed by five pages of questions and three blank pages. The cover is purposefully pink to distinguish it from the yellow Data Administration Profile Questionnaire. The blank pages were provided to allow the respondent to express extended answers,

³¹ Charach, op. cit.

if desired.

Twenty-five (25) questions are contained in the body of the questionnaire. All but questions 1, 4, 6, and 8 are concerned with measuring the 'maturity' of the EDP activity. Question 1 is a simple graphic question on the EDP activity organizational structure designed to invoke the curiosity of the respondent. As far as Data Administration is concerned, certain organizational structures (eg. segmented by application areas or decentralized system analysis) may not be conducive to establishing a Data Administration function. In these types of structures, functional, rather than corporate perspectives are held by the EDP activity. Questions 4 and 6 are used to verify the size of the responding organization in terms of number of employees, sales and assets. Question 8 addresses organizational type.

The 'maturity' questions were originally drawn directly from Mantha³². After a pre-test, however, several of these questions were revised because the respondents found them unclear or difficult to answer without considerable research. The questions on chargeback and project development standards were expanded in scope because the researcher suspected that a wide range of actual situations could elicit positive responses. User awareness measures were considered for each phase in the system development life cycle.

³² Mantha, ibid.

The Data Administration Profile Questionnaire

This questionnaire is divided into two parts. Part 1 consists of 34 questions about the characteristics of the Data Administration function, the tools used, and the background of the Data Administrator. Part 2 consists of 38 questions about the responsibilities of the Data Administrator. Instructions are given with each question in Part 1 to ensure that the respondent knows how to reply.

Questions 1 through 11 attempt to determine where the Data Administrator reports in the organization, how long since the function was established, the number of employees in the function, and whether the organizational group responsible for the technical activities reports to the Data Administrator. Question 12 is concerned with the machine readable/non-machine-readable dichotomy. The question measures the fraction of the corporate data resource that falls under the custodianship of the Data Administrator. Question 13 measures the absolute size of the machine-readable data resource.

Question 14 measures how far the Data Administrator has progressed in reducing data redundancy. Question 15 through 30 are concerned with some of the tools used by the Data Administrator: Data Base Management System, Data Dictionary, Query Facilities and Information Requirements Analysis (IRA) Methodology.³³ For all but the IRA procedure, these questions

³³ For a discussion of various tools used by the Data Administrator see: GUIDE International, Data Base Administration Project, The Data Base Administrator, November 3, 1972, pages 2.1-2.15.

ask, how long has the tool been used and what fraction of the machine-readable resource has it been applied against?

Question 31 is concerned with the work experience and formal education of the Data Administrator.

Question 32 indicates how influential the Data Administrator is in the EDP activity. Question 33 is concerned with the existence of formal standards (a management tool), for the major processes in the life cycle of a piece of data. Question 34 is concerned with how closely the Data Administrator is involved with developing information systems and is directly related to Question 33(a) - the definition of data in the organization.

Questions 35 to 72 ask whether or not the Data Administrator is responsible for a number of particular activities, how much time is spent on those activities, and how that amount of time is expected to change two years from now. Again, each activity reflects the administrative/technical dichotomy of the Data Administration function. These activities were selected from various sources but heavy use was made of Secrest³⁴. The pre-test resulted in some minor changes to the original statements culled from the literature. The majority of these changes effectively split a question into several questions because the original statement contained more than one activity.

³⁴ Secrest, ibid

The Pretest

A pretest was conducted on three Vancouver based organizations in the sample population. It consisted of a thirty minute session with the EDP activity manager and a forty-five to sixty minute session with the Data Administrator. While the respondent completed the appropriate questionnaire the researcher sat silently and answered queries as they arose. All difficulties encountered by the pretest respondents were noted on a blank questionnaire. These sessions were invaluable in estimating the time required to complete the questionnaires; an important consideration in determining how many persons would respond. One organization in the pretest had a Data Administrator, one had recently discontinued the function because of a lack of top management interest and one had not yet established the function because of a lack of funds and top management commitment. Several changes to the draft questionnaires resulted from this pretest.

The Sample

Organizations in the sample came from both the private and public sectors. Private sector companies were selected from the Financial Post Survey of Industrials(1978) and the Financial Post Survey of the 300 Largest Canadian Companies(1978). Public sector organizations were selected from the 1978 Corpus Almanac of Canada. This book was scanned for job titles related to the EDP activity (eg. Director, Data Processing) in provincial and federal ministries. To ensure

that organizations with large computer installations were not left off the lists, the 1978 Canadian Information Processing Society's Survey of Computers was compared to the lists extracted from the aforementioned publications.

Survey Package

A survey package consisting of a covering letter (Appendix C), a self-addressed pre-paid return envelope, an EDP Profile Questionnaire and a Data Administration Profile Questionnaire was mailed to the 555 organizations selected in the sample. The survey package was directed either at a President (private sector only) or a manager in the EDP activity. Names of EDP activity managers were obtained from a Canadian Information Processing Society membership list and the 1978 Corpus Almanac of Canada (provincial and federal ministries only). Where the name of the EDP manager in the private sector was not available, the package was addressed to the President of the company. The President was selected instead of the EDP Manager for two reasons. First, EDP Managers have various titles, and in large organizations a package simply addressed to the 'EDP Manager' could be misdirected. Second, a package coming down the line from top management to the EDP manager is more likely to be receive appropriate attention.

The covering letter followed the format and content of the covering letter in Mantha's³⁵ survey. The instructions and

³⁵ Mantha, ibid

confidentiality/anonymity details were placed on the cover of the questionnaires. Essentially, the covering letter stated briefly the purpose of the survey and instructions for determining who in the organization was to complete the questionnaire.

The package was mailed on January 10, 1979, requesting that completed questionnaires be returned by February 6, 1979.

Follow Up

On January 26, 1979 a follow up card (Appendix D) was sent to all organizations, whether they had replied or not. A follow up letter was sent to 4 respondents who returned a Data Administration Profile Questionnaire without an EDP Profile Questionnaire and to 1 respondent who had returned an EDP Profile Questionnaire, indicating the existence of a Data Administrator, but failed to return a Data Administration Profile Questionnaire. One additional survey package was mailed to a respondent who received a follow up card but did not, apparently, receive a survey package.

CHAPTER IV

ANALYSIS AND RESULTS

Preliminary Statistical AnalysisSurvey Package Response

Table I gives a breakdown of the responses to the survey.

TABLE I

Response Tabulation

1. Initial mailings.....	555
2. Second mailing (by request).....	1
3. Duplicate initial mailings.....	17
4. Returned letters or telephone calls indicating that no response would be forthcoming.....	33
5. Total returned EDPPQ with or without DAPQ attached.....	262
6. Returned EDPPQ with or without DAPQ received more than 3 weeks after requested return date.....	4
7. Duplicate EDPPQs returned.....	4
8. EDPPQs used in the analysis (5-6-7).....	254
9. Returned DAPQs.....	101
10. DAPQs returned by organizations indicating that they did not have a Data Administration function...	47
11. Returned DAPQ received more than 3 weeks after requested return date.....	1

Because the survey list was compiled from several sources without an exhaustive comparison, 17 organizations were sent two survey packages. In a few instances, however, a second mailing was intentional, since more than one address was available. The Data Administration Profile Questionnaire (DAPQ) was completed by 47 respondents who stated on the EDP Profile Questionnaire (EDPPQ) that their organization did not have a separate Data Administration function. Since organizations should not be able to respond to the DAPQ if they do not have the function, the researcher examined these questionnaires more closely to determine if the respondent simply did not follow instructions or actually had some form of a Data Administration function. Note that Question 1 on the DAPQ asks where the Data Administrator reports in the organization. As a result of this inspection, the researcher recoded 16 responses on the EDPPQ to indicate the presence of a separate Data Administration function. The recoding criteria were the title of the person completing the DAPQ, the existence of a Data Base Management System and complete answers on Part 2.

A partially completed DAPQ was returned by 4 organizations that had indicated the existence of a separate Data Administration function on the EDPPQ. These organizations spent few dollars on computer hardware and used none of the Data Administration tools. Consequently, these responses were recoded to indicate the absence of a Data Administration function. These Data Administration functions were not, in the opinion of the researcher, ones that the survey was attempting

to reach.

As shown in Table I, 33 organizations responded by letter or telephone that they were not going to return completed questionnaires. The two prime reasons given were that the organization did not have the time to complete the questionnaires or the organization, as a corporate policy, did not complete questionnaires.

To verify that a reasonable cross section of the total private sector was received, the list of respondents was compared to the lists of companies used to select respondents. As shown in Table II, total responses were uniformly distributed over the sample.

TABLE II

Private Sector Responses

Out of largest 120 Canadian Industrials by sales.....	58 responses
Out of largest 6 Real Estate Companies by assets.....	5 responses
Out of largest 10 Merchandising Companies by sales.....	7 responses
Out of largest 25 Financial Companies by assets.....	12 responses
Out of largest 15 Insurance Companies by assets.....	9 responses

To verify that a reasonable cross section of the total public sector was received, a simple comparison of responses to total mailings was made. As shown in Table III, the cross section of responses was also evenly distributed over the

sample. There was only one response from the Quebec provincial government ministries. This was to be expected, as the questionnaires were not translated into French.

TABLE III

Public Sector Responses

<u>Item</u>	<u>Mailed</u>	<u>Responses</u>
Federal Government	27	17
Provincial Government	67	32
Universities	23	15
Cities	8	6

The Data Administration Function in Large Canadian Organizations

Detailed analysis of the EDPPQ was performed on 254 cases³⁶. Table IV gives a frequency distribution of responses to Question 7, on the existence of a separate Data Administration function. Only 60, or 24% of the responding organizations had positive responses to this question. Even considering the recoding of Question 7 answers, only 72, or 28% of the responding organizations indicated that they had a separate Data Administration function. Consequently, the

³⁶ All statistical analysis was performed using the software product Statistical Package for the Social Sciences (SPSS). A description of the package can be found in Norman H. Nie et al, Statistical Package for the Social Sciences, second edition, McGraw Hill Book Company, 1975.

concept cannot be considered to be pervasive in large Canadian organizations. A new variable, DA, was computed from Question 7 responses. A DA value of 1 was assigned if a separate Data Administration function exists in the organization, a value of 0, otherwise.

TABLE IV

Responses to Question 7 of EDPPQ
(Do you have a Data Administration function?)

<u>Response</u>	<u>Frequency</u>	<u>Percentage</u>
No	84	33
No, because...	22	9
No, we used to...	4	2
No, but we are thinking	67	26
Yes	56	22
Recoded from No to Yes	16	6
Recoded from Yes to No	4	2
	---	---
	253	100
Missing cases=	1	
Total analysed	254	

Note: 1. Total YES = $56 + 16 = 72$
 2. Total NO = $84 + 22 + 4 + 67 + 4 = 181$
 3. 3 organizations answering YES did not return a DAPQ.

EDP Activity Size and the Existence of a Data Administrator

To test for a possible relationship between the existence of a separate Data Administration function and the size of the EDP activity, a crosstabulation was performed between DA and the number of EDP employees (Question 2). Responses to Question 2 were divided into three groups by the number of EDP activity employees, so that the groups were about equal in size. The crosstabulation between DA and the transformed Question 2 resulted in a chi-square of 40.84966 with 2 degrees of freedom and $p < .001$ (Table V). This result suggests that organizations with a very large number of employees in the EDP activity are more likely to have a separate Data Administration function than smaller ones. The Pearson correlation coefficient between DA and number of EDP employees (Question 2) is, however, only .1890 ($n=251$, $p=.001$), suggesting that the two variables are not highly correlated. A one-tailed t-test was performed to test if the mean number of EDP employees (Question 2) for DA value 1 is greater than the mean of number of EDP employees (Question 2) for a DA value of 0. As shown in Table VI, there is a significant difference in means. Because the null hypothesis that the variances of each population are equal was rejected, a t-test based on a separate variance estimate was used rather than one based on a pooled variance estimate.³⁷ The large variances in the populations may be attributed to the wide range of responses to Question 2 (3 to

³⁷ Nie et al, ibid, page 270.

2800), with several respondents at the extreme top and bottom ends of the scale not having a Data Administrator.

TABLE V

DA vs Number of EDP Employees

EDP Function Employees
 < 28 29 - 79 > 79

DA=0	83	56	42	181
DA=1	5	24	41	70

Column Total 88 80 83 n=251

Chi-square= 40.84966 with 2 df. (p=0.0000)

Missing cases= 3

TABLE VI

DA vs Number of EDP Employees
One-tailed T-test

<u>Item</u>	<u>DA=1</u>	<u>DA=0</u>
Number of cases	70	181
Mean number of EDP employees	191.4000	83.5525
Standard deviation	376.683	183.533
F Value	4.21	
2-tail probability	0.000	
T Value	-2.29	
Degrees of freedom	81.99	
1-tail probability	.012	

As another approach to testing for a possible relationship between the existence of a separate Data Administration

function and the size of the EDP activity, a crosstabulation was performed between DA and Question 15 (monthly computer hardware expenditure). Responses to this question were divided into three groups by dollars of monthly computer hardware expenditure, so that the three groups were of about equal size. The crosstabulation between DA and the transformed Question 15 resulted in a chi-square of 30.00696 with 2 degrees of freedom and $p < .001$ (Table VII). This result suggests that organizations that have very high monthly hardware expenditures are more likely to have a separate Data Administration function than those that have low monthly hardware expenditures. The Pearson correlation coefficient for DA and Question 15 is, however, only .1722 ($n=244$, $p=.004$), suggesting that the two variables are not highly correlated. A one-tailed t-test showed a significant difference in means (Table VIII). The t-test was based on a separate variance estimate because the assumption that the population variances are equal was violated.

The results of the above analysis suggest that only large Canadian organizations with large EDP functions are likely to have formed a separate Data Administration function. But size is by no means a good predictor of the existence of a Data Administration function. Only 43% of organizations with a monthly hardware expenditure of \$100,000 or more had established a Data Administration function.

TABLE VII

DA vs Monthly Hardware Expenditure

Monthly Hardware Expenditure
(\$000's)
< 32 32-100 > 100

DA=0	79	50	46	175
DA=1	6	28	35	69

Column Total 85 78 81 n=244

Chi-square= 30.00696 with 2 df. (p=0.0000)

Missing cases=10

TABLE VIII

DA vs Monthly Hardware ExpenditureOne-tailed T-test

<u>Item</u>	<u>DA=1</u>	<u>DA=0</u>
Number of cases	69	175
Mean monthly hardware expenditure (\$ 000's)	251.749	115.513
Standard deviation (\$ 000's)	518.004	260.592
F Value		3.95
2 Tail probability		0.000
T Value		-2.08
Degrees of freedom		81.93
1 Tail probability		.02

Experience With Computers and the Existence of a Data Administrator

To test for a possible relationship between the existence of a separate Data Administration function and the experience of the organization with computers, a crosstabulation was performed between DA and Question 16 (number of years that the organization has used computers). Responses to Question 16 were divided into three groups, by years of computer usage in the organization, so that the three groups were about equal in size. The crosstabulation between DA and the transformed Question 16 resulted in a chi-square of 12.85600 with 2 degrees of freedom and $p < .002$ (Table IX). This result suggests that organizations with many years of experience with computers are more likely to have a separate Data Administration function than less experienced ones. The Pearson correlation coefficient is, however, only .2020 ($n=252$, $p=.001$), indicating a weak linear relationship. A one-tailed t-test showed a significant difference in means (Table X). The results of the above analysis suggest that organizations with many years of experience with computers are more likely to have established the Data Administration function than those with few years of experience with computers. However, even among the EDP functions with 16 or more years of experience with computers, only 30% had established a separate Data Administration group (Table IX).

TABLE IX

DA vs Experience With Computers

Number of years experience
 < 10 10-15 > 15

DA=0	68	60	53	181
DA=1	11	37	23	71

Column total 79 97 76 n=252

Chi-square= 12.85600 with 2 df. (p= 0.0016)

Missing cases= 2

TABLE X

DA vs Experience With Computers
One-tailed T-test

<u>Item</u>	<u>DA=1</u>	<u>DA=0</u>
Number of cases	71	181
Mean years of Experience	15.5634	13.2044
Standard deviation	5.477	5.041
F Value		1.18
2 Tail probability		.384
T Value		-3.26
Degrees of freedom	250	
1 Tail probability		.0005

Organizational Type and the Existence of a Data Administrator

Responses to Question 8 (organizational type) showing the frequency of occurrence of a separate Data Administration function are shown in Table XI. Responses 03, 14, 18, 19, 20, 21, and 23 have the highest proportion of organizations with a Data Administrator (Petroleum and Coal Products, Telecommunications and communications, Information processing equipment, Electric, gas or water utility, Bank, deposit or credit agency, Life insurance, and University). These organizations also happen to be ones that did not suffer adversely during the recent economic slowdown prevalent in Canada (the exception being universities which are perpetually short of funds).³⁸

To test if availability of funds was instrumental in establishing a Data Administrator, a crosstabulation was performed between DA and two groups of organizational types. The first group (with funds) consisted of the sample organizations that responded 03, 14, 18, 19, 20, 21, or 22 to Question 8. Responses of 23 were not included. Responses of 22 were included because of the large number of Data Administrators in this category and because government agencies can usually find the funds somewhere. The second group

³⁸ A comparison over the last few years of the "Earnings Adjusted to Index" indicator of the Toronto Stock Exchange '300' Stock Price Index System for these organizational types showed that these organizations did not suffer adversely over the last few years. The indices showed that the other organizational types did have reduced profit levels over the last few years.

(without funds) consisted of all other types in the sample, except "other" (response of 24). The crosstabulation resulted in a chi-square of 5.96955 with 1 degree of freedom, $p < .02$ (Table XII). This result suggests that organizations "with funds" are more likely to have a Data Administrator than those "without funds". A crosstabulation between the transformed Question 2 (number of EDP employees) and this same dichotomy of types resulted in a chi-square of 20.32210 with 2 degrees of freedom, $p < .001$ (Table XIII), suggesting that those organizations "with funds" also have more EDP employees.

TABLE XI

DA vs Organizational Type (Frequencies)

<u>Type</u>	<u>Frequency</u>	<u>DA=1 (%)</u>
1. Wood, pulp and paper	9	0 (0)
2. Primary metals and fabrication	11	2 (18)
3. Petroleum and coal products	19	7 (37)
4. Chemical and chemical products	11	3 (27)
5. Food, beverage and tobacco	18	4 (22)
6. Accommodation and food services	1	0 (0)
7. Electrical products	1	0 (0)
8. Automotive products	4	0 (0)
9. Printing and publishing	6	0 (0)
10. Miscellaneous manufacturing	1	0 (0)
11. Industrial equipment	3	1 (33)
12. Construction	3	1 (0)
13. Transportation	7	0 (0)
14. Telecommunications, communications	8	5 (63)
15. Department, supermarket, variety stores	10	1 (10)
16. Household and personal products	3	1 (33)
17. Property development, management	5	1 (20)
18. Informaticn processing equipment	2	1 (50)
19. Electric, gas or water utility	14	6 (43)
20. Bank, deposit or credit agency	12	5 (42)
21. Life insurance	13	4 (31)
22. Government department or agency	40	10 (25)
23. University	15	7 (47)
24. Other	37	12 (32)
	253	71 (28)

TABLE XII

DA vs Availability of Funds

Funds Available
Yes No

DA=0	70	87	157
DA=1	38	21	59

Column Tctal 108 108 n=216

Chi-square= 5.96955 with 1 df. (p=.0146)

Missing cases= 1

With Funds= Question 8 responses 03,14,18,19,20,21,22

Without Funds= All other responses except 24

TABLE XIII

EDP Employees vs Availability of Funds

EDP Employees
< 28 29-79 > 80

With Funds	25	31	50	106
Without Funds	49	38	21	108

Column total 74 69 71 n=214

Chi-square= 20.32210 with 2 Df. (p=0.0000)

Missing cases= 3

With Funds= Question 8 responses 03,14,18,19,20,21,22

Without Funds= All other responses except 24

Hypothesis Tests

Stage Model of EDP Growth

Hypothesis H1 :

More mature organizations are more likely to have established a Data Administration function than less mature organizations

was tested using linear regression. The dependent variable was DA with values of 0 (no Data Administrator) and 1 (a Data Administrator). The independent variables were the EDP growth variables referred to in Chapter I.

In general, multiple regression requires that the variables be measured on an interval or ratio scale.³⁹ Because the possible responses to Question 3 (EDP reporting position) and Question 9 (chargeback process) did not exhibit such characteristics, responses to these questions were transformed to "dummy" variables. The variable LEVEL with values of 0 or 1 was computed from Question 3. A value of 1 indicated an EDP activity reporting to top management (answers of 3 or 4) while a value of 0 indicated an EDP activity not reporting to top management (answers 1 or 2). Wherever a response of 5 was coded, the researcher classified the answer as 1, 2, 3, or 4 by examining the charts or descriptions supplied by the respondent. The variable CHARGE with values of 0 or 1 was computed from Question 9. A value of 1 indicated the existence of a chargeback procedure for use of EDP activity resources (answers 3 and 4) while a value of 0 indicated the nonexistence

³⁹ Nie et al, ibid, page 320.

of a chargeback procedure (answers 1 and 2). Wherever a response of 5 was coded, the researcher classified the answer as 1, 2, 3 or 4 by examining the explanation given by the respondent for answering '5'.

A new variable, SSCALE, representing standards for systems development, was computed from a composite of Question 10 responses (items 1 through 7). First, answers to these seven questions were converted to Z scores (mean of 0, variance of 1). Standardizing the raw scores for each of the 7 responses will give 'yes' answers a higher weighting if there are relatively few 'yes' answers to the question. This is acceptable, as each 'yes' on the 7 questions is seen as a measure of 'maturity'. The bias is in the proper direction. Next, a reliability test⁴⁰ was performed on the sum of these standardized variables. The Alpha score was .80055, indicating that the items form a homogeneous scale.

A new variable, USCALE, representing user awareness of the EDP activity, was computed from a composite of Questions 19(a), 19(b), 20(a), 21(a), 22(a), 22(b), 23(a) and 24(a). First, answers to these 7 questions were converted to Z scores. Standardization is required so that each question is based on the same scale and therefore direct addition can be performed

⁴⁰ A description of the reliability test can be found in the SPSS updates to Nie et al. (1975) Alpha is a single index of the strength of the relationship between variables. If a number of variables purport to measure a particular construct (in this case 'existence of system development standards'), then they should be highly intercorrelated.

without one question weighting the sum disproportionately.⁴¹ Next, a reliability test was performed on the sum of the standardized variables. The Alpha score was .80247, indicating that the items form a homogeneous scale.

No attempt was made to construct from Question 11 responses a single measure of the criteria used by Senior Management in evaluating the performance of the EDP activity because the objectives can conflict with each other. For example, user satisfaction may require that the budget be exceeded.

The linear regression equation then becomes:

$$\begin{aligned}
 DA = & A0 + A1(\text{Question 2}) + A2(\text{Question 13}) + A3(\text{Question 14}) \\
 & + A4(\text{Question 16}) + A5(\text{Question 17(a)}) \\
 & + A6(\text{Question 11(a)}) + A7(\text{Question 11(b)}) \\
 & + A8(\text{Question 11(c)}) + A9(\text{Question 11(d)}) \\
 & + A10(\text{LEVEL}) + A11(\text{CHARGE}) + A12(\text{USCALE}) \\
 & + A13(\text{SSCALE}).
 \end{aligned}$$

Question 15 was not included in the regression equation because it correlated highly with Question 2 (over .8).

The regression analysis indicated that R Squared (proportion of the variance accounted for by the linear regression) was only .12890. The result indicates that the linear combination of 'maturity' variables cannot be used to predict the existence of a Data Administrator. The correlation between each independent variable and DA is shown in Table XIV.

⁴¹ Nie et al, ibid, page 187.

There is no reason to believe, based on these data, that there is a relationship between any of the so called 'maturity' variables and the existence of a Data Administrator.

In view of these results, a reliability test was performed to see if the "maturity" variables were, in fact, intercorrelated, a necessary but not sufficient condition for the variables to measure a common construct. The Alpha was .53896 indicating that these variables are not aspects of a common "maturity" concept.

TABLE XIV

Correlations of DA With Maturity Variables

<u>Maturity Variable</u>	<u>Correlation</u>
Number of EDP employees	.1890 (n=251, p=.001)
Strategic Planning Committee	.1373 (n=252, p=.015)
Data Flow Model	-.0793 (n=251, p=.105)
Experience with computers	.2020 (n=252, p=.001)
Fraction of EDP budget spent on operational level systems	.0919 (n=239, p=.078)
User awareness	.1530 (n=251, p=.008)
Systems development standards	.1427 (n=247, p=.012)
EDP reporting level	.1157 (n=253, p=.033)
Chargeback of costs	.1059 (n=253, p=.046)
Performance-cost savings	-.0762 (n=253, p=.113)
Performance-user satisfaction	-.0324 (n=253, p=.304)
Performance-meeting budgets	.0602 (n=253, p=.170)
Performance-organizational goals	.0614 (n=252, p=.166)

Since larger EDP activities were more likely to have a Data Administrator than smaller ones, as shown above, the researcher considered that perhaps "maturity" applied only to the very large EDP activities. The same linear regression was performed on a selected subset of the sample - Question 2 "greater than 28 employees". The resulting R Squared was only

.12244.

Based on this analysis, the researcher must conclude that there is little if any relationship between the existence of a Data Administrator and scores on several 'maturity' variables. Hypothesis H1 must be rejected.

Data Administration Specialization

Analysis was performed on 69 cases comprised of 53 organizations who answered 'yes' to Question 7 on the EDPPQ and 16 organizations who answered 'no' but were recoded to 'yes' answers. Three organizations, although they responded positively to Question 7 on the EDPPQ, did not return a DAPQ.

Organizational Structure

Only 5 organizations indicated that the Data Administrator did not report within the EDP activity (Question 1 answer of 'no'). Table XV gives a frequency distribution to Question 3, the reporting level of the Data Administrator in the organization. The number of employees in the Data Administration function ranged between 1 and 21. Over 80% of the Data Administration functions consisted of 5 or fewer individuals (Table XVI).

Experience

Experience of the Data Administration function (Question 6) ranged from 0 to 8 years with almost 70% having 3 years or less experience (Table XVII). A response of 0 indicated that the function was in the process of being established when the survey package was received.

TABLE XV

Reporting Level of DA

Number of Levels Between DA and Head of Organization	Frequency	Relative Frequency (%)
1	4	5.8
2	18	26.1
3	28	40.6
4	16	23.2
5	2	2.9
6	1	1.4
	---	---
	69	100.0

TABLE XVI

Size of DA-Number of Employees

Number of DA Employees	Frequency	Relative Frequency (%)
1	15	21.7
2	15	21.7
3	16	23.2
4	3	4.3
5	6	8.7
6	4	5.8
7	3	4.3
> 7	5	7.0
	---	---
	67	100.0

TABLE XVII

Experience of DA

Years of Experience	Frequency	Relative Frequency (%)
0.0	3	4.3
0.5	5	7.2
1.0	15	21.7
1.2	1	1.4
1.5	3	4.3
2.0	14	20.3
2.5	2	2.9
3.0	5	7.2
3.5	1	1.4
4.0	9	13.0
5.0	6	8.7
6.0	3	4.3
7.0	1	1.4
8.0	1	1.4
	69	100.0

Scope

Responses to Question 12(a) indicated that few Data Administrators are responsible for non machine readable data (53 or 77% responded 0%). Responses to Question 12(b) indicated that Data Administrators are not even responsible for all the machine readable data (Table XVIII). The responses to Question 13 ranged between 15 and 110,000 millions of characters of data (Table XIX). Little of this data is shared between applications but there is some expectation that more could be shared in the absence of technical and other constraints (Table XX).

TABLE XVIII

Scope of DA
Fraction of Machine Readable Data
Under DA Responsibility

Fraction of Machine Readable Data	Frequency	Relative Frequency (%)
25 % or less	20	29
26 - 50 %	10	14
51 - 75 %	6	9
76 % or more	30	44
missing	3	4
	69	100

TABLE XIX

Size of Machine Readable Data Resource

Millions of Characters	Frequency	Relative Frequency (%)
< 500	12	17
500 - 1000	12	17
1001 - 1500	6	9
1501 - 3000	7	10
3001 - 7500	10	15
> 7500	13	19
missing	9	13
	69	100

Table XX
Sharing of Data

Range (%)	Number of Respondents	
	Actually Shared	Potentially Sharable
0 - 25	44	16
26 - 50	16	16
51 - 75	3	18
76 -100	3	12
missing answer	3	7
	69	69

Tools

Questions 20 (Data Dictionary), 25 (Online Generalized Query Facility), 26 (Batch Generalized Query Facility), 27 (Online Parametric Query Facility), 28 (Batch Generalized Query Facility), and 30 (Information Requirements Methodology) were transformed to dichotomous variables with a value of 1 indicating the use of the tool and 0 indicating the non-use of the tool. Table XXI gives a frequency distribution of the usage of the various tools. The Data Base Management System is not used extensively for the management of data (Table XXII). Hypothesis H2, was tested using Pearson correlation analysis. As shown in Table XXIII, Data Administration function experience is not highly correlated with the use of any of the tools. Hypothesis H2 must therefore be rejected.

TABLE XXI

<u>Use of DA Tools</u>		
Tool	With	Without
Data Base Management System	64	5
Data Dictionary Directory	44	25
Online generalized query facility	31	38
Online parametric query facility	31	37
Batch generalized query facility	45	22
Batch parametric query facility	24	44
Information requirements methodology	28	40

Table XXII

Percentage of Computer Applications Using DBMS

Percentage	Number	Frequency (%)
< 20	28	40.6
21 - 40	13	18.8
41 - 60	10	14.5
61 - 80	10	14.5
81 -100	3	7.2
no DBMS	5	7.2
	69	100.0

TABLE XXIII

DA Experience vs Use of Tools

Tool	Pearson Correlation
Data Base Management Systems	.2179 (n=69, p=.036)
Data Dictionary Directory	.0623 (n=69, p=.306)
Online Generalized Query Facility	.0798 (n=69, p=.257)
Batch Generalized Query Facility	.0303 (n=69, p=.404)
Online Parametric Query Facility	.2369 (n=68, p=.026)
Batch Parametric Query Facility	.2270 (n=68, p=.031)
Information Requirements Methodology	.0617 (n=68, p=.308)

Data Administrator Background and Experience

Table XXIV suggests that Data Administrators have strong computer technology backgrounds and experience but they do not have strong managerial backgrounds or experience.

TABLE XXIV

Background of Data Administrator

Item	No. (%)
Computer technology background	61 (91)
1 year's exp. in a functional area	12 (18)
>2 years' exp. in a functional area	24 (35)
Bachelor degree-Computer Science	20 (30)
Bachelor degree-Business Admin.	11 (16)
Bachelor degree-Arts	5 (8)
Bachelor degree-Science	13 (19)
Technical school diploma	6 (9)
Post grad degree-Business Admin.	3 (5)
Post grad degree-Computer Science	8 (12)
Post grad degree-other	2 (3)
>3 years' supervisory experience	51 (76)
DBMS experience	41 (61)
Systems development experience	62 (93)

Missing cases= 2

Data Administration Responsibilities

In order to test hypotheses H3 and H4, Questions 38 to 72 first had to be classified as either "administrative" or "technical" responsibilities. Second, a single measure of the time spent on each set of responsibilities was required. The researcher performed the classification using Secrest's dichotomy (see Chapter II). Questions 38 to 49, 51 to 55, 57, 66, and 68 to 72 were classified as "administrative" responsibilities while all others were classified as "technical". In order to determine if the classification was done correctly, reliability tests were performed on the "administrative" and "technical" responsibilities (responses to "A Data Administration Responsibility?"). Alpha for "administrative" responsibilities was .81700. Alpha for "technical" responsibilities was .86117. These results suggest that the questions had been classified correctly. All responses to 'PERCENT OF TIME SPENT ON ACTIVITY' were first adjusted to a fraction of 100%, if the total percentage for all questions exceeded 100%. Where the total was less than 100%, no adjustment was made because these respondents were assumed to be spending time on other activities not queried in this questionnaire. Before adjusting the percentages, fractional responses (eg. .5%) were rounded up to the next whole number (eg. 1%). Second, all responses to the adjusted 'PERCENT OF TIME SPENT ON ACTIVITY' were transformed to Z scores (mean of 0, variance of 1). Third, APSCALE was computed as the sum of the adjusted and standardized responses to the 'administrative'

questions 'PERCENT OF TIME SPENT ON ACTIVITY'. For 'technical' questions, a similar sum, TPSCALE, was computed.

Table XXV gives the correlation between the composite scales and Data Administration experience (Question 6). The low correlations suggest that there is little relationship between the years of experience of the Data Administration function and the time spent on administrative or technical activities. The time spent on technical activities would, however, appear to be greater for more experienced functions than less experienced ones. Further testing of the experience versus responsibility hypotheses (H3 and H4) was done using one-tailed t-tests.

TABLE XXIV

DA Experience and Time Spent on
Technical and Administrative Responsibilities
Pearson Correlation

<u>Composite Variable</u>	<u>Correlation with Experience</u>
APSCALE	-.0716 (n=63, p=.288)
TPSCALE	.2589 (n=64, p=.019)

A one-tailed t-test was performed on partitions of Data Administration experience (Question 6) against TPSCALE and APSCALE. Table XXVI and XXVII show the results of a partition of Question 6 into 2 groups- those with 2 years or less of experience and those with more than 2 years experience. There is a significant difference in means for TPSCALE but not for

APSCALE. The results suggest that more experienced Data Administration functions spend a greater percentage of their time on technical responsibilities than less experienced ones. There is little difference in the times spent by experienced and inexperienced functions on administrative responsibilities.

The analysis indicates that the two hypotheses on experience and responsibilities (H3 and H4) must be rejected.

TABLE XXVI

DA Experience and Time Spent on
Technical Responsibilities
One-tailed T-test

Item	< 2 Years Experience	>= 2 Years Experience
Number of cases	35	27
Mean time spent (Standardized)	-1.5684	2.0432
Standard deviation	4.486	4.490
F Value		1.000
2 tailed probability		.983
T Value		-3.14
Degrees of freedom		60
1-tailed probability		.0015

TABLE XXVII

DA Experience and Time Spent on
Administrative Responsibilities
One-tailed T-test

Item	< 2 Years Experience	>= 2 Years Experience
Number of cases	34	27
Mean time spent (Standardized)	-.0090	.7117
Standard deviation	6.888	5.544
F Value		1.54
2 tailed probability		.258
T Value		-.44
Degrees of freedom	59	
1-tailed probability		.165

The sum of the mean times spent on administrative responsibilities was 63.971% whereas the sum of the mean times spent on technical responsibilities was 19.257%. Overall then Data Administration functions spend more time on administrative responsibilities than technical ones. The hypothesis (H5) must be accepted. The mean times spent on the activities are evenly distributed and no one activity stands out above the others (Table XXVIII).

Most respondents indicated that they would be spending more time 2 years from now on nearly all the responsibilities. Questions 42, 45, 49, 50 and 55 can be viewed as start up tasks and therefore one might expect less time to be spent on them 2 years from now.

TABLE XXVIII

Responsibilities

Question	Resp. (%)	Mean Time Spent (%)	Standard Deviation	2 Years Hence (%)	
				LESS	MORE
35	66.7	2.277	3.720	2.9	58.0
36	82.4	5.263	5.572	0.0	68.1
37	79.7	3.832	4.406	2.9	66.7
38	79.7	2.011	2.508	2.9	63.8
39	69.6	1.194	2.157	2.9	65.2
40	97.1	3.901	3.928	0.0	71.0
41	76.8	3.076	4.822	1.4	68.1
42	92.8	4.389	4.688	5.8	60.9
43	65.2	1.059	1.706	1.4	47.8
44	84.1	1.567	1.945	0.0	53.6
45	82.6	2.687	3.425	5.8	44.9
46	49.3	.999	1.695	0.0	46.4
47	91.3	3.043	3.367	2.9	53.6
48	53.6	1.035	2.097	2.9	50.7
49	82.6	2.900	3.585	7.2	34.8
50	76.8	2.833	3.840	11.6	36.2
51	68.1	2.487	6.567	2.9	60.9
52	17.4	.194	.721	1.4	17.4
53	76.8	5.672	7.603	1.4	50.7
54	68.1	1.609	3.063	1.4	46.4
55	78.3	4.586	5.117	8.7	37.7
56	75.4	3.929	4.444	4.3	36.2
57	46.4	.599	1.675	1.4	46.4
58	82.6	1.659	3.303	1.4	59.4
59	76.8	1.318	1.852	2.9	55.1
60	63.8	.902	1.619	1.4	49.3
61	82.6	1.920	2.916	1.4	42.0
62	72.5	1.321	2.003	0.0	36.2
63	58.0	.958	1.577	1.4	33.3
64	63.8	1.896	3.137	1.4	37.7
65	79.7	1.067	1.600	1.4	37.7
66	31.9	.340	1.004	0.0	36.2
67	55.1	1.433	2.437	2.9	33.3
68	31.9	.333	.969	0.0	33.3
69	43.5	.527	1.641	1.4	36.2
70	81.2	5.228	10.203	2.9	50.0
71	50.7	1.179	3.337	1.4	46.4
72	66.7	1.163	1.485	0.0	50.7

CHAPTER V

SUMMARY, DISCUSSION AND CONCLUSIONS

Survey Package Response

The mail survey achieved a response rate of 47.12% (Table I). As shown in tables II and III, the responses were evenly distributed over the private and public sectors.

The need to recode several responses to Question 7 on the EDPPQ (dealing with the existence of a Data Administration function) leads to speculation that the reliability of this question is low. A reliability index could have been constructed from a random sample of the respondents; unfortunately time and funds were not available. Nevertheless, the reliability of this question is probably high enough to warrant use of the data for projections over the entire population. The measurement "error" in this question can result only if respondents answer 'yes' when the "true" answer is 'no', or respondents answer 'no' when the "true" answer is 'yes'. In the first situation, the researcher can easily verify the reliability of the answers by examining the returned DAPQs or the explanations given for failing to return a completed DAPQ. Only 3 organizations answered 'yes' and did not return a DAPQ. One of these organizations explained that their Data Administration function had only recently been established and, as a consequence, they did not want to spend the "great deal of time" that might be required to research the

answers to the DAPQ. Only 4 organizations with 'yes' answers that returned a DAPQ were recoded to 'no'. Their answers to the DAPQ indicated that they did not have a separate Data Administration function (ss Chapter IV).

In the second situation, the researcher could verify the response if a DAPQ was returned. The researcher recoded 16 cases where the respondent had indicated the non-existence of a Data Administration function but returned a completed DAPQ (see Chapter IV). In many of these cases the person completing the DAPQ had a title of "Data Base Administrator". Some recent literature⁴² differentiates between a Data Administrator and a Data Base Administrator. The latter reports to the former and is responsible for the more technical activities associated with managing data. Since the results of this study suggest that there may be several "different" types of Data Administrators,⁴³ it is highly probable that these recoded organizations had in mind a "corporate" Data Administrator when responding to Question 7 of the EDPPQ. If this question was similarly interpreted by other respondents who did not return a DAPQ, the results of this study, clearly, would not be materially affected. Since considerable interest in the

⁴² GUIDE, Establishing the Data Administration Function, *ibid*, page 25.

⁴³ This conclusion is discussed later in this chapter.

results was shown by the participants,** it is not likely that responding organizations with a "corporate" Data Administrator purposefully answered 'no'.

The Data Administration Function in Large Canadian Organizations

The results of this study suggest that:

1. Few large Canadian organizations have a Data Administrator.
2. Organizations with very large EDP activities are more likely to have a Data Administrator than smaller ones.
3. Organizations with more than 10 years of experience with computers are more likely to have a Data Administrator than those with less experience.
4. Organizational types that could be considered to have discretionary funds available are more likely to have a Data Administrator than those with more limited funds.

The correlation between the existence of a Data Administrator and the size of the EDP activity, and the correlation between the existence of a Data Administrator and the experience of the organization with computers are both very low, suggesting that size and experience are probably necessary

** A comment received from a large integrated mining company is typical of many responses received.

"I feel that the subject of Data Administration - as perceived by your study - is a long overlooked essential component needed to successfully bring computer based Management Information Systems to play their expected role in large organizations." Only two organizations bothered to return questionnaires with negative commentary (a large steel producer and a large trust company).

but not sufficient conditions for the presence of a Data Administrator.

According to GUIDE,⁴⁵ considerable funds might be required to establish a Data Administration function. GUIDE also suggests that the benefits attributable to the implementation of a Data Administration function are achievable only through the implementation and enforcement of certain data management policies. Some of these policies cannot be enforced without a "rewrite" or redesign of existing application systems. As a consequence, the expenditure of considerable funds, beyond merely staffing the Data Administration function and purchasing a Data Base Management System, would be required to achieve such benefits as improved data availability, reduced data redundancy, improved data security, and a higher degree of data integrity. In short, the costs are very visible in the short run, but the benefits are more intangible and probably not realizable for many years after the establishment of the Data Administration function. Therefore, organizations with discretionary funds available would be more likely to embark on the path towards achieving the long term benefits of Data Administration. Some other reasons for the low occurrence of Data Administrators in large Canadian organizations are discussed later in this chapter.

⁴⁵ GUIDE, ibid.

Stage Model of EDP Growth

The "maturity" of the organization's EDP activity was not found to be a good predictor of the existence of a Data Administrator. Furthermore, the validity of the "maturity" construct used must be questioned since the "maturity" variables were not found to be highly positively intercorrelated. The researcher could not, however, validate the "maturity" measurement criteria used in this study against Nolan's studies because Nolan has never published an explicit formulation for his measurement of "maturity", nor any psychometric evidence of the construct validity of this concept or its measure.

Nolan's writings also fail to mention the size of the companies (EDP employees or hardware expenditures) he studied. United States organizations are, as a whole, larger than similar Canadian organizations.⁴⁶ Therefore, it is possible that Canadian organizations are just not 'large enough' for the problem of data management to have become serious. Several other reasons were given by the respondents to explain why they didn't have a Data Administrator:

- a) a lack of top management support and therefore the necessary budget,
- b) organizational decentralization (functional and/or

⁴⁶ In "The Fortune Directory of the 500 Largest US Industrial Corporations", the 500th organization had sales of \$379,948,000 (FORTUNE, May 7, 1979). This would have placed this organization at about the 80th position in the top 400 Canadian Industrial organizations (Canadian Business, July 1978).

geographical) is not compatible with the Data Administration concept,

c) a high degree of user involvement in the systems development process is not compatible with the establishment of a Data Administration function,

d) user departments are responsible for managing their data resource,

e) a Data Base Management System is necessary before there is a need for a Data Administrator.

The comments about decentralization and user involvement suggest that the Data Administration concept may not be well understood by the sampled organizations. Clearly, for example, decentralization does not hinder the establishment of a Data Administrator as Data Administration is a policy setting and coordinating function.

Nolan does not seem to have included the possibility of such individual misperception in his model. Closely related to the misperception issue is corporate "politics", which also does not seem to be accounted for by Nolan. These items will be discussed later in this chapter.

Data Administration Specialization

Organizational Structure

The results of this study suggest that:

1. Very few Data Administrators report outside the EDP

activity.

2. The Data Administrators report fairly low in the organizational hierarchy.
3. Few people work in the Data Administration function.

Experience

The sample Data Administration functions were found to be quite inexperienced as most had existed for less than 3 years. The first published literature on the Data Administration function only began to appear in 1971, or about 8 years ago. Perhaps a longer time period is required for organizations to become aware of and evaluate the concept.

Scope

This study found that:

1. Except for a few cases, Data Administrators were not responsible for any of the non-machine-readable data resource.
2. Only one-third of the Data Administrators were responsible for the entire machine-readable data resource.
3. The size of the machine-readable data resource varied considerably among the sampled organizations.
4. Little of the machine-readable data resource was shared among application systems.
5. There was little agreement among the sampled Data Administrators as to how much of the data resource was potentially shareable.

Tools

This study found that:

1. Nearly all the respondents used a Data Base Management System (DBMS) and there were 14 organizations that used more than one DBMS.
2. The Data Administration function was established after the acquisition of a DBMS in 67% of the cases.
3. The DBMSs were not used extensively by the sampled Data Administration functions.
4. Many of the respondents did not have a Data Dictionary/Directory. Only 7 respondents ventured a guess as to the number of data elements that exist in their organization.
5. About 50% of the Data Administrators had online query facilities but their use was limited to a small fraction of the machine-readable data resource.
6. In 59% of the cases the approach to be used for determining the information needs of managers was left to the discretion of the systems analyst.
7. The use of certain tools was not related to the level of experience of the Data Administration function.

Data Administrator Background and Experience

Most of the Data Administrators had a strong computer technology background and a strong systems development background with several years of supervisory experience. Few had any experience in the functional areas of management or a university degree in business administration.

Data Administration Responsibilities

This study found that:

1. All but 5 Data Administrators were responsible for the technical activities associated with managing the data resource (Question 7 of the DAPQ).
2. Few Data Administrators had implemented standards for the capture, maintenance or dissemination of data, but most had developed standards for the definition and storage of data.
3. All Data Administrators spent considerably more time on administrative responsibilities than technical ones.
4. More experienced Data Administrators spent relatively more time on technical activities than less experienced ones.
5. The relative amount of time spent by Data Administrators on administrative activities was not a function of their experience.
6. The major administrative tasks performed by the Data Administrators were associated with providing a liason function between the "data bases" and system analysts and users. Few of the policy setting activities were performed by the Data Administrators.
7. About 20% of the time spent by Data Administrators was unaccounted for by the responsibilities queried in this survey.
8. The sampled Data Administration functions exhibited a wide dispersion in both the activities performed and the amount of time spent on each.

9. Very few Data Administrators indicated that they would be spending less time 2 years from now on any of the activities. Depending on the responsibility, anywhere from one-third to two-thirds of the respondents indicated that they would be spending more time on the activity 2 years hence.

It is interesting to note, however, that although all the responsibilities queried were described in the literature as constituting the major responsibilities of a Data Administrator, there was no consensus among the respondents that these activities were in fact their responsibilities. Many activities that are strongly managerially oriented had fewer than 75% of the respondents answering 'yes'.

The above results would suggest that there may be "different" types of Data Administration functions.⁴⁷ The lack of a trend away from "technical" and towards "administrative" responsibilities may also be explained in this way. Organizations can choose one of several roles for their Data Administrator. Therefore a function with little experience could exhibit the same characteristics as one which has been in existence for several years.

Figure 1 depicts the different roles of the Data Adminis-

⁴⁷ In a study of 25 Data Administration functions in the New York area, Weldon obtained similar results (J. L. Weldon, Data Base Administration: Organization and Tasks, Working Paper #78-143, November 1978, New York University, Graduate School of Business Administration)

trator. In this simplified view, the scope of the data under the control of the Data Administrator can either be application specific or corporate wide. Similarly, the tasks performed can be categorized as either technical or administrative. An organization can choose either an application or corporate perspective and then can assign technical and/or administrative responsibilities to the Data Administrator. If the Data Administrator is to play a support or advisory role within the EDP activity, then the benefits attributable to a corporate Data Administration function are not likely to be achieved: policies can not be created at this level.⁴⁸

FIGURE 1

Roles of the Data Administrator

		PERSPECTIVE	
		Application	Corporate
TASKS	Technical	1	2
	Administrative	3	4

⁴⁸ A large insurance company replied that they had made two false starts at getting the Data Administration concept established in the EDP activity. They are now convinced that the function, if it is to manage the corporate data resource, must be set up outside the EDP activity. This company is now establishing the function in its corporate planning department.

Role of Data Administrator

The above results suggest that the Data Administration function in large Canadian organizations is not at an advanced stage of evolution. The high incidence of DBMSs suggests that the function was probably formed around the acquisition of a DBMS. Note that the sampled functions were staffed by persons competent in the use and application of a DBMS. The extent to which this tool is used, which seems to be small, may indicate the role that the Data Administrator is expected to play in the organization. That is, if the DBMS is used to manage a small fraction of the machine-readable data resource, then the Data Administrator likely has a minor role in the EDP activity. This minor role will be discussed more later in the chapter.

The lack of a positive and high correlation between the use of Query Facilities and the experience of the Data Administration function may be explained by an "application" perspective to the design of systems. Some applications, such as banking, may require the use of Online Query Facilities for routine transaction processing so that the existence of this tool does not necessarily indicate an "advanced" Data Administration function. The low incidence of use of Data Administration tools, other than the DBMS, plus the fact that the tools tend to be applied to a small fraction of the data resource, suggest that the Data Administrators are not taking a coordinated approach to the management of the machine-readable data resource.

In order to explain the finding that the Data

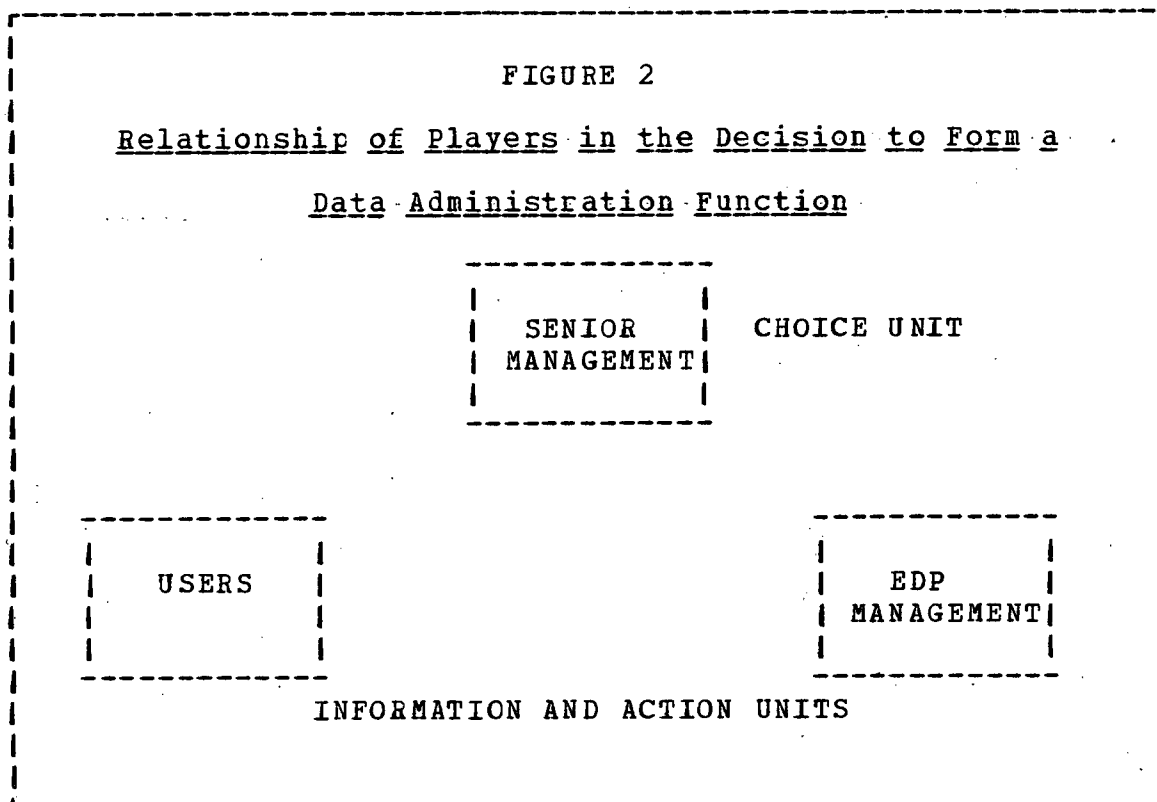
Administration function usually occupies a relatively minor role within the EDP activity, we consider the following hypothetical decision:

Should our organization form a separate Data Administration function? If so, what role should it perform in managing the data resource?

Since very few small EDP activities (i.e. < 28 persons) had separate Data Administration functions, those organizations have either not considered this decision or concluded that the coordinating role of a Data Administrator is unlikely to be required when so few people are involved in the management of the data resource. The researcher was therefore primarily interested in the larger EDP activities in the sample. If the Data Administration function is to perform a minor role in the EDP group, then EDP Management can probably implement the function without consultation with Users or Senior Management. Some EDP Managers might have done just this, anticipating that the role would later expand in scope.

The literature, however, stresses that the Data Administrator should perform a coordinating and policy setting role. Therefore Senior Management should be the "choice unit" - the one to make the decision to have a Data Administrator. The Users and EDP Management are considered to be "information units"-suppliers of information to the choice unit, and "action units"-implementors of the choice unit's decision (see Figure 2). Having little exposure to the Data Administration concept and perhaps being relatively uninterested in the problems associated with managing the data resource, Senior Management

is unlikely to initiate the implementation of a corporate Data Administrator (see Chapter I).



Misperceptions of the Data Administrator role by Users and EDP Management or conflicts between them may also account for the low level of Data Administration activity in large Canadian organizations. Three possible scenarios illustrate this conclusion.

1. EDP Management, because of its involvement with all functional areas of management and exposure to the Data Administration concept, perceives a need for a separate Data Administration function. This need is corporate wide and therefore requires the support of Users. As with any other project, the Users and EDP Management develop a

cost/benefit statement. Benefits are found to be entirely intangible in the short run yet costs are extremely visible. As a consequence, the Data Administration project is placed at the end of a long list of projects with measurable benefits. When this backlog has been reduced, Data Administration will be implemented.

2. When approached by EDP Management about a Data Administration project, Users suggest that they are already performing the function, and very well at that.⁴⁹ Without the support of Users, EDP Management would be unlikely to make a presentation to Senior Management, for fear of weakening the User-EDP activity relationship so necessary for successful implementation of computer systems.
3. EDP Management examines the Data Administration literature and concludes that the concept is not applicable to their organization (eg. the decentralized environment of their company is considered to be incompatible with a corporate Data Administration concept).

Conclusions

The results of this study suggest that Data Administration has not emerged to any large extent as a separate functional speciality in large Canadian organizations. Organizational

⁴⁹ The accounting and finance departments have traditionally played a major role in supplying information to management for decision making purposes. These groups likely have instituted over many years several policies and procedures for the control of the non-machine readable data resource.

conflicts and a general misunderstanding of the function by EDP Management have likely held back the development of the function beyond one involved primarily with the support of DBMS application systems.

Future research should be directed at understanding these conflicts and misperceptions through an analysis of the decision process involved in establishing the Data Administration function; in particular, how a cost/benefit analysis of the Data Administration is or might be performed. An attempt should be made to more fully understand the data resource and how it might differ among organizational types. In-depth interviews could be conducted within the very large Canadian EDP functions using their responses to this survey as a point of departure.

Before further use is made of Nolan's Stage Growth model, serious thought should be given to determining in more precise terms what the EDP growth process variables are and how they might best be measured. Without such a base, empirical testing of the model is not possible. It might also be useful to conjecture what other variables, such as corporate "politics", exist that affect the growth of an EDP activity.

REFERENCES

- [1] Ackoff, Russell L., "Management Misinformation Systems", Management Science (14:4), 1967 December, pages B147 156.
- [2] Anthony, Robert N., Planning and Control Systems: A Framework for Analysis, Harvard Graduate School of Business Administration, Boston, 1965.
- [3] Berdie, Douglas R. and John F. Anderson, Questionnaires: Design and Use, Scarecrow Press Inc, N. J., 1974.
- [4] Canning, Richard G., "The Data Administrator Function", EDP Analyzer, (10:1), November 1972, 14 pages.
- [5] Canning, Richard G., "Toward the Better Management of Data", EDP Analyzer, (14:12), December 1976, 13 pages.
- [6] Canning, Richard G., "Installing a Data Dictionary", EDP Analyzer, (16:1), January 1978, 13 pages.
- [7] Charach, L., An Exploratory Investigation Toward The Development of a Research Design For a Study of Youth Work Transition, unpublished master's thesis, University of British Columbia, 1977.
- [8] Davis, Gordon B., Management Information Systems- Conceptual Foundations, Structure, and Development, McGraw-Hill Book Company, 1974.
- [9] Deardon, John, "MIS is a Mirage", Harvard Business Review, (50:1), January-February 1972, pages 90-99.
- [10] Everest, Gordon C., "Database Management Systems Tutorial", Fifth Annual Midwest AIDS Conference Proceedings, Vol. 1, Minneapolis, Minnesota, 1974 May 10-11, edited by Norman L. Chervany, Minneapolis: University of Minneapolis, College of Business Administration, 1974, pages A1-A12.
- [11] Gorry, G. A., and M. S. Scott Morton, "A Framework for Management Information Systems", Sloan Management Review, Fall 1971.
- [12] GUIDE International Corporation, The Data Base Administrator, November 1972.
- [13] GUIDE International Corporation, Managerial Considerations of Data Base Systems (GP 44), 1977.
- [14] GUIDE International Corporation, Establishing the Data Administration Function, 1977.

- [15] GUIDE International Corporation, Data Administration Methodology Draft Paper, May 1978.
- [16] GUIDE International Corporation, The Data Administrator: A Catalyst of Corporate Change, October 1975.
- [17] Kimberly, John R., "Organizational Size and the Structuralist Perspective: A Review, Critique, and Proposal", Administrative Science Quarterly, (21:4), December 1976, pages 571-597.
- [18] Lyon, John K., The Data Base Administrator, John Wiley & Sons, New York 1976.
- [19] Lyon, John K., "The Role of The Data Base Administrator", Database, Winter 1971, pages 11-12.
- [20] Mantha, Robert W., An Analysis of Skill Requirements In Data Processing Environments, unpublished master's thesis, University of British Columbia, 1978.
- [21] Martin, James, Principles of Data Base Managemet, Prentice Hall, Englewood Cliffs, N. J. 1976.
- [22] Nie, Norman H. et al, Statistical Package for the Social Sciences, Second Edition, McGraw Hill Book Company 1975.
- [23] Nolan, Richard L., "Computer Data Bases: The Future Is Now", Harvard Business Review, September-October 1973, pages 98-114.
- [24] Nolan, Richard L., Managing the Data Resource Function, West Publishing Co. 1974.
- [25] Nolan, Richard L., "Thoughts About the Fifth Stage", Database, (7:2), 1975, pages 4-10.
- [26] Nolan, Richard L., "Organizational Response and Information Technology", Proceedings of the National Computer Conference 1978, pages 517-524.
- [27] Nolan, Richard L., "Managing the Crises in Data Processing", Harvard Business Review, (57:2), March-April 1979.
- [28] Oppenheimer, A. N., Questionnaire Design and Attitude Measurement, Basic Books, Inc. New York, 1966.
- [29] Secrest, Richard D., "The Data Administrator", in McFarlan, F. W. and Richard L. Nolan, The Information Systems Handbook, Dow Jones-Irwin, Inc. Homewood, Illinois, 1975.
- [30] Seltiz, C. et al, Research Methods in Social Relations,

Holt, Rinehart and Winston, Inc, Revised Edition, 1967.

- [31] Sprowls, R. Clay, Management Data Bases, Wiley Hamilton, Santa Barbara, 1976.
- [32] Tillinghast, J., "Establishing a Framework for Data Base Planning", Auerbach Publishers Inc. 1976.
- [33] Weldon, J.L., Data Base Administration, Organization and Tasks, Working Paper #78-143, November 1978, New York University, Graduate School of Business Administration.
- [34] Wiorkowski, G. K. and J. J. Wiorkowski, "Does a Data Base Management System Pay Off?", Datamation, April 1978, pages 109-114.
- [35] Yasaki, E. K., "The Many Faces of the DBA", Datamation, May 1977, pages 75-79.

APPENDIX A
EDP PROFILE QUESTIONNAIRE

DATA RESOURCE MANAGEMENT SURVEY

EDP PROFILE QUESTIONNAIRE

(TO BE COMPLETED BY THE MANAGER OF THE EDP ACTIVITY)

Name of organization _____

Please indicate your job position/title _____

This questionnaire is designed to assess the stage of development of the Electronic Data Processing (EDP) activity in large Canadian organizations. For the sake of this study, the EDP activity is considered to be the development, operation and maintenance of computer based information or data processing systems. This questionnaire will be used to generate a profile of the EDP activity in large Canadian organizations.

Throughout the questionnaire the phrase "your organization" will mean your entire company, agency or institution. If we want information about only the EDP activity, we will indicate that explicitly.

Please answer all of the questions asked. You may be assured that all information obtained will be used specifically for research purposes and under no circumstances will individual responses be disclosed.

The questionnaire will not take more than 30 minutes of your time.

Once completed, please mail the questionnaire in the attached prepaid self addressed envelope. We would be grateful if you could return the completed questionnaire to us by 6 February, 1979.

If you wish to receive a copy of the research results please attach your business card to the returned questionnaire.

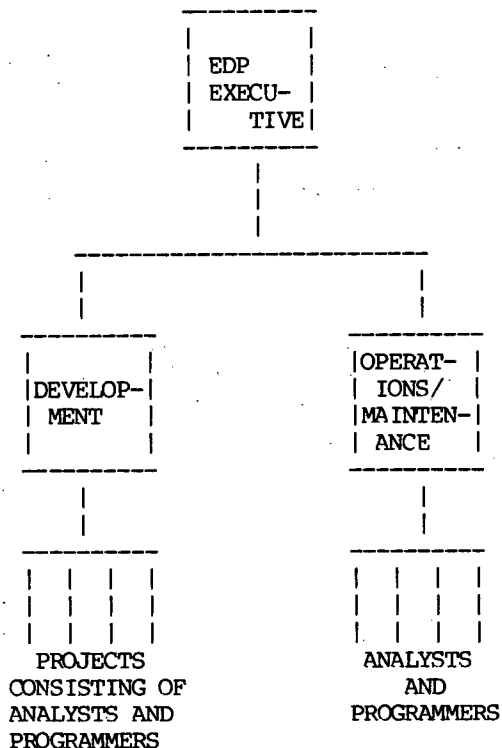
We would like to take this opportunity to thank you in advance for taking part in this study and to emphasize that we do need your cooperation.

Dr. R. C. Goldstein
Faculty of Commerce and Business Administration
The University of British Columbia

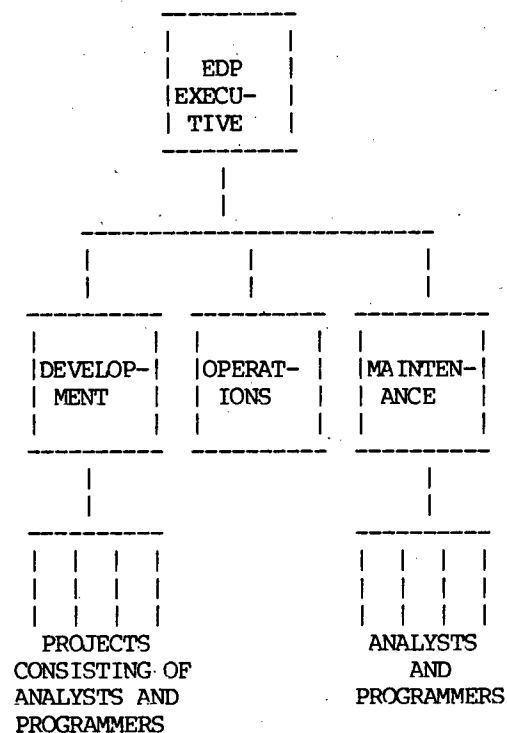
SOME QUESTIONS ON ORGANIZATIONAL STRUCTURE

Which of the following best describes the structure of the EDP activity in your organization? Please circle a letter.

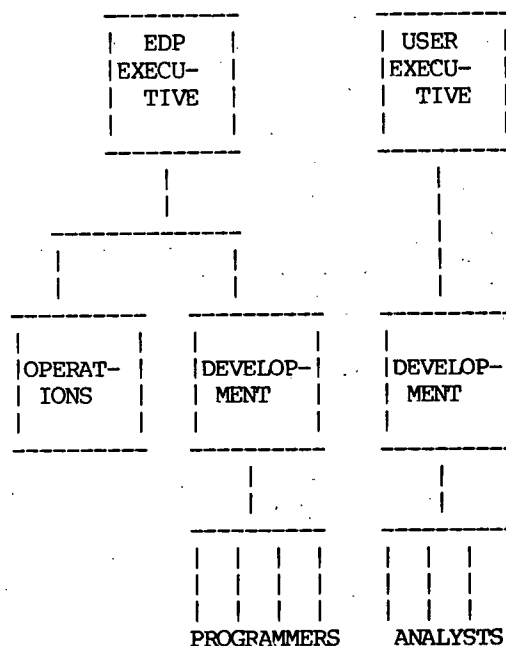
A.



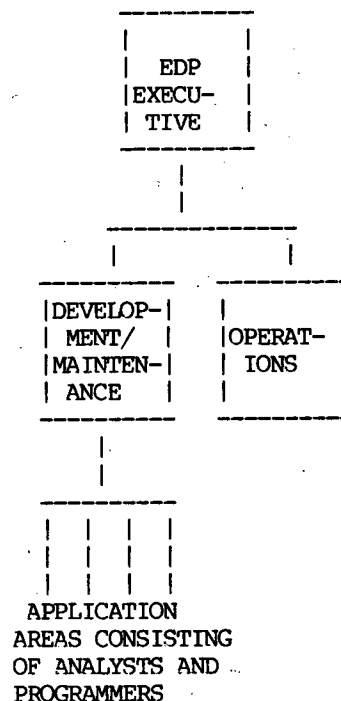
B.



C.



D.



E. None of the above describes the structure of your EDP activity. Please describe your structure on the back page of the questionnaire. Alternatively, please attach a copy of the organization chart for the EDP activity in your organization.

OVER

What is the total staff complement reporting to the Manager of the EDP activity in your organization?

_____ individuals

3. Please indicate, with reference to the following structure, the organizational location of the EDP activity in your organization. Please circle a number.

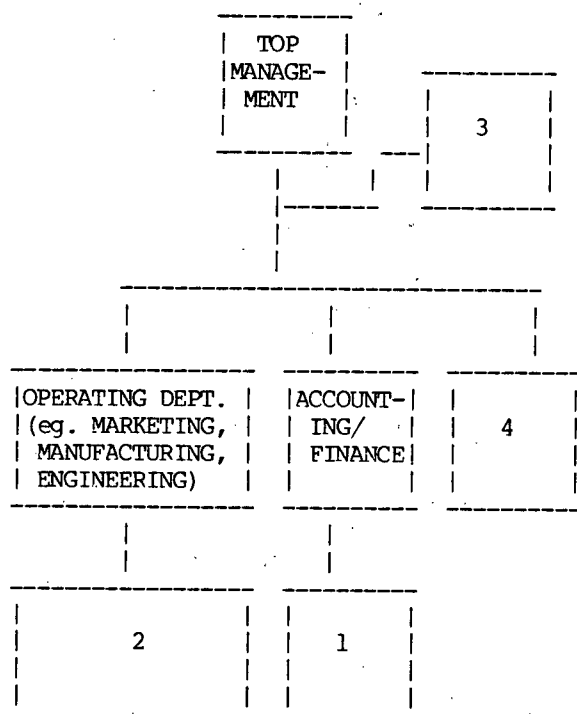
1. In Accounting/Finance department.

2. In Operating department such as Marketing, Manufacturing, or Engineering.

3. Service bureau or equivalent.

4. Independent operating department reporting to Top Management.

5. None of the above. Please describe the organizational location in the space following or on the back page of the questionnaire. Alternatively, please attach a copy of the organizational chart for your organization depicting the location of the EDP activity.



4. What is the total number of employees in your organization?

_____ employees

5. Is your organization a subsidiary of a larger one?

_____ No.

_____ Yes, parent organization is _____

6. What is the approximate value of your organization's assets and annual sales?

Assets _____ Not applicable.

_____ dollars

Annual sales _____ Not applicable

_____ dollars

7. In the past few years, many organizations have recognized that data are one of their most valuable corporate resources and have taken steps to better manage and control this resource area. For some organizations, these actions have resulted in the establishment of a separate group to manage the corporate data resource. This group is often known as the Data Administrator.

Has your organization formally established such an organizational unit?

_____ Yes, the name of the group is _____

_____ No, but we are thinking of establishing such a function.

_____ No, we used to have such a function but we decided to discontinue the function because _____

_____ No, because _____

_____ No.

Which of the following best describes the business environment of your organization? Please circle a number.

SOME QUESTIONS ON PLANNING AND CONTROL

9. Which of the following best describes the accounting for EDP activity expenditures in your organization? Please circle a number.

1. Wood, pulp and paper
 2. Primary metals and fabrication
 3. Petroleum and coal products
 4. Chemical and chemical products
 5. Food, beverage and tobacco
 6. Accomodation and food services
 7. Electrical products
 8. Automotive products
 9. Printing and publishing
 10. Miscellaneous manufacturing
 11. Industrial equipment
 12. Construction
 13. Transportation
 14. Telecommunications, communications
 15. Department, supermarket, variety stores
 16. Household and personal products
 17. Property development, management
 18. Information processing equipment
 19. Electric, gas or water utility
 20. Bank, deposit or credit agency
 21. Life insurance
 22. Government department or agency
 23. University
 24. Other. Please specify
-

1. EDP costs are treated as corporate overhead and are not allocated proportionately to users.
2. EDP costs are treated as corporate overhead and are allocated proportionately to users.
3. Users are charged (in real dollars) for some of the EDP resources used and some costs are treated as corporate overhead.
4. Users are charged (in real dollars) for the amount of EDP resources used.
5. None of the above describes accounting for EDP expenditures in your organization. Please specify in the space below or on the back page of the questionnaire.

10. For each of the following functions in the EDP activity, please indicate whether formal standards, which are enforced, have been established in your organization. Circle either Yes or No.

1. Yes No Project management
2. Yes No Systems definition
3. Yes No Systems design
4. Yes No Programming
5. Yes No Systems operation
6. Yes No Systems maintenance
7. Yes No Systems documentation

OVER

11. Please indicate how important the following criteria are in senior management's evaluation of the performance of the EDP activity in your organization. Circle a number from 1 to 5 where 1 indicates UNIMPORTANT and 5 indicates VERY IMPORTANT.

A. Cost savings due to clerical staff reductions or increased efficiency of operation support systems.

UNIMPORTANT 1 2 3 4 5 VERY IMPORTANT

B. User satisfaction.

UNIMPORTANT 1 2 3 4 5 VERY IMPORTANT

C. Meeting budgets.

UNIMPORTANT 1 2 3 4 5 VERY IMPORTANT

D. Contribution to organizational goals as stated in the overall organizational plan.

UNIMPORTANT 1 2 3 4 5 VERY IMPORTANT

12. Has your organization established Project Review Boards to review the progress of new systems development projects?

_____ Yes _____ No

13. Has your organization established a Long Range or Strategic Planning Committee (or equivalent) to develop the EDP activity strategy for your organization?

_____ Yes _____ No

14. Has your organization developed a model of the organization that shows data flows across functional areas and hierarchical management levels?

_____ Yes _____ No

SOME MISCELLANEOUS QUESTIONS

15. Please indicate the approximate average monthly rental cost of computing hardware, including communication expenditures, in your organization for the past 12 months (use rental equivalent if leased or purchased).

_____ dollars per month

16. Indicate the approximate number of years your organization has been using computers.

_____ years

17. What approximate percentage of the EDP budget is spent on the following three categories of systems (includes development, operation and maintenance)?

1. _____ % Operational Support Systems.
Systems which perform the routine transaction level activity required in the daily operation of the organization and report on the operational status of the firm so that management is aware of day-to-day activities. (includes order entry systems, invoicing, payroll, etc.)

2. _____ % Management Control Systems.
Systems which provide control information required by managers of departments, profit centres, etc to measure performance, track the efficiency and effectiveness of operations, decide on control actions, formulate new decision rules to be applied by operational personnel, allocate resources and provide for coordination between several departments. (includes manufacturing cost control systems, sales analysis systems, etc.)

3. _____ % Planning Systems.
Systems which provide information for strategic level management (top management). This information will permit these managers to carry out their planning activities, such as formulating and revising company objectives, determining long-term goals (over 3 years) and establishing company policies. (includes financial planning systems, corporate models, etc.)

_____ % TOTAL

18. What approximate percentage of the EDP activity budget is spent producing ad hoc reports for users for which special or "one-shot" programs are developed?

1. Less than 5%
2. 6% to 15%
3. 16% to 25%
4. More than 25%

Questions 19 to 25 are concerned with the degree that Users, Senior Management and the EDP activity are involved in various new system development activities. Please circle a number from 1 to 5 to indicate the degree that the group is involved in the activity, where '1' indicates NOT INVOLVED and '5' indicates VERY INVOLVED.

19. Identifying new systems development projects.

USER	NOT INVOLVED	1	2	3	4	5	VERY INVOLVED
SENIOR MANAGEMENT	NOT INVOLVED	1	2	3	4	5	VERY INVOLVED
THE EDP ACTIVITY	NOT INVOLVED	1	2	3	4	5	VERY INVOLVED

20. Determining costs of new systems development projects.

USER	NOT INVOLVED	1	2	3	4	5	VERY INVOLVED
SENIOR MANAGEMENT	NOT INVOLVED	1	2	3	4	5	VERY INVOLVED
THE EDP ACTIVITY	NOT INVOLVED	1	2	3	4	5	VERY INVOLVED

21. Determining benefits of new systems development projects.

USER	NOT INVOLVED	1	2	3	4	5	VERY INVOLVED
SENIOR MANAGEMENT	NOT INVOLVED	1	2	3	4	5	VERY INVOLVED
THE EDP ACTIVITY	NOT INVOLVED	1	2	3	4	5	VERY INVOLVED

22. Ranking or prioritizing new systems development projects.

USER	NOT INVOLVED	1	2	3	4	5	VERY INVOLVED
SENIOR MANAGEMENT	NOT INVOLVED	1	2	3	4	5	VERY INVOLVED
THE EDP ACTIVITY	NOT INVOLVED	1	2	3	4	5	VERY INVOLVED

23. Developing a project schedule for new systems development projects.

USER	NOT INVOLVED	1	2	3	4	5	VERY INVOLVED
SENIOR MANAGEMENT	NOT INVOLVED	1	2	3	4	5	VERY INVOLVED
THE EDP ACTIVITY	NOT INVOLVED	1	2	3	4	5	VERY INVOLVED

24. Managing a new systems development project.

USER	NOT INVOLVED	1	2	3	4	5	VERY INVOLVED
SENIOR MANAGEMENT	NOT INVOLVED	1	2	3	4	5	VERY INVOLVED
THE EDP ACTIVITY	NOT INVOLVED	1	2	3	4	5	VERY INVOLVED

25. Designing a new computer based information or data processing system.

USER	NOT INVOLVED	1	2	3	4	5	VERY INVOLVED
SENIOR MANAGEMENT	NOT INVOLVED	1	2	3	4	5	VERY INVOLVED
THE EDP ACTIVITY	NOT INVOLVED	1	2	3	4	5	VERY INVOLVED

APPENDIX B**DATA ADMINISTRATION PROFILE QUESTIONNAIRE**

DATA RESOURCE MANAGEMENT SURVEY

DATA ADMINISTRATION PROFILE QUESTIONNAIRE

(TO BE COMPLETED BY THE DATA ADMINISTRATOR OR SIMILAR POSITION)

Name of organization _____

Please indicate your job position/title _____

In the past few years many organizations have recognized that data are a valuable corporate resource and have taken steps to better manage and control this resource. This survey is designed to assess the extent to which managing the data resource in large Canadian organizations has been recognized as a separate functional speciality, and to determine how this is being accomplished.

Throughout this questionnaire the phrase "your organization" means your entire company, agency or institution. For the sake of this study, the EDP activity is considered to be the development, operation and maintenance of computer based information or data processing systems.

The questionnaire is divided into two (2) parts. Part 1 is concerned with various characteristics of the Data Administration function such as organizational structure, tools, planning and controlling. Part 2 is concerned with the activities of the Data Administration function in your organization.

Please answer all of the questions asked. You may be assured that all information obtained will be used specifically for research purposes and under no circumstances will individual responses be disclosed.

This questionnaire will not take more than 60 minutes of your time. You may not be able to answer some of the questions precisely without considerable research. For these questions please give an approximate answer.

Once completed, please mail the questionnaire in the attached prepaid self addressed envelope. We would be grateful if you could return the completed questionnaire to us by 6 February, 1979.

If you wish to receive a copy of the research results please attach your business card to the returned questionnaire.

We would like to take this opportunity to thank you in advance for taking part in this study and to emphasize that we do need your cooperation.

Dr. R. C. Goldstein
Faculty of Commerce and Business Administration
The University of British Columbia

PART 1

SOME QUESTIONS ON ORGANIZATIONAL STRUCTURE

1. Does the Data Administration function in your organization report within the EDP activity? Please check one.

_____ Yes CONTINUE WITH QUESTION 2.

_____ No SKIP TO QUESTION 3.

2. How many levels of supervision are there between you (Data Administrator) and the Manager of the EDP activity? For example, if your supervisor reports directly to the Manager of the EDP activity there is one (1) level of supervision between you and the Manager of the EDP activity.

_____ levels

3. How many levels of supervision are there between you (Data Administrator) and the head of your organization (eg. President, Deputy Minister)? For example, if your supervisor reports to the Vice-President of Administration, who in turn reports to the President, there are two (2) levels of supervision between you and the head of your organization.

_____ levels

4. What is the job title of your immediate supervisor?

5. How many full-time clerical, technical, professional and management individuals are employed within the Data Administration function in your organization?

_____ people

6. How long ago was the Data Administration function established in your organization?

_____ years

7. Is the Data Administration function responsible for the more technical activities associated with managing the data resource(eg. Data base design, selection of storage and access methods)? Please circle a number.

1. Yes PLEASE SKIP TO QUESTION 12

2. No PLEASE CONTINUE WITH QUESTION 8

8. Does your organization have a separate organizational unit, not reporting to the Data Administration function, that is responsible for the more technical activities associated with managing the data resource? Please circle a number.

1. Yes CONTINUE WITH QUESTION 9

2. No SKIP TO QUESTION 12

9. How many levels of supervision are there between this group and the Manager of the EDP activity?

_____ levels

10. What is the name of this organizational unit?

11. How many full-time clerical, technical, professional, and management individuals are employed within this organizational unit?

_____ people

SOME QUESTIONS ON THE TYPE OF DATA MANAGED BY THE DATA ADMINISTRATOR

Data, for the sake of this questionnaire, are groups of non-random symbols which represent quantities, actions, things, etc. We refer to data that has regular routines associated with its definition, capture, storage, maintenance and dissemination as Formal data. Formal data can be further categorized as being Machine Readable or Non Machine Readable.

12. What percentage of your organization's Formal data falls under the responsibility of the Data Administrator.

1. _____ % of Non Machine Readable Data.

2. _____ % of Machine Readable.

OVER

13. What is the approximate size of the Machine Readable data resource in your organization? To answer this question, you may want to consider the number of disk packs of storage your organization has on its computers.

_____ millions of characters.

14. What approximate percentage of the Machine Readable data is actually shared among application systems in your organization? What percentage of this data, in terms of characters, could in the absence of technical or other constraints, usefully be shared among application systems?

1. _____ % is actually shared among applications.
2. _____ % could usefully be shared among applications (in the absence of technical or other constraints).

SOME QUESTIONS ON TOOLS USED BY THE DATA ADMINISTRATOR

Data Base Management System (DBMS) is the software that handles all access to the data base. A DBMS can be characterized as a generalized software tool that provides a single, flexible facility for accomodating different data files and operations, while demanding less programming effort than convential programming languages. A DBMS facilitates operations on data (definition, capture, storage etc.); it facilitates reference by name rather than by physical location; and it provides an environment that is not tied to a particular set of application programs or files.

15. Does your organization use a DBMS? Please circle a number.

1. Yes CONTINUE WITH QUESTION 16.
2. No SKIP TO QUESTION 20.

16. Does your organization use more than one DBMS?

_____ Yes we use _____ DBMS's.
 _____ No we use only one DBMS.

17. For each DBMS your organization uses, please indicate the name of the package, whether the package was obtained commercially or developed inhouse and the number of years since the package was first installed in your organization.

NAME	COMMERCIAL	INHOUSE	YEARS SINCE FIRST INSTALLED
	(Please tick one)		
_____	_____	_____	_____
_____	_____	_____	_____

18. Was the Data Administration function established before or after the aquisition or development of a DBMS? Please circle a number.

1. Before.
2. After.

19. What percentage of your organization's computer applications use the DBMS's. Compute an approximate percentage by estimating what portion of the EDP activity budget is spent on the development, operation and maintenance of computer applications that use the DBMS's. Circle a number.

1. Less than 20%
2. 21% to 40%
3. 41% to 60%
4. 61% to 80%
5. 81% to 100%

Data Dictionary / Directory (DD/D) is a software tool used to list, describe, and locate each data element in a data base. It provides a centralized repository of information about each data element in order to facilitate management of, control of, and access to the data base. The DD/D may also contain information about files, transactions, source documents, reports, systems, programs, users, departments, projects, standards, security levels and personnel.

20. Does your organization use a DD/D to record the characteristics of the data resource? Please circle a number.

1. Yes, a commercial package.
2. Yes, an inhouse developed package.
3. Yes, a manual record keeping system.
4. No . SKIP TO QUESTION 25.

21. What is the name of your DD/D?

22. How long ago was the DD/D acquired or developed in your organization?

_____ years and _____ months

23. How many data elements are there in your organization?

1. _____ don't know
2. _____ data elements

24. How many data elements are recorded in your DD/D?

_____ data elements

Query facilities can be considered as batch or online and parametric or generalized. Periodic processing of queries is generally referred to as 'batch processing', as contrasted with immediate or 'online processing'. Parametric means that the user interfaces with the system in an anticipated and highly structured way by invoking transactions or procedures which have been predefined to the system. Generalized means that the user interacts with the system in an unstructured, unanticipated and ad hoc manner.

25. Does your organization use an online generalized query facility software package? If so please indicate the name of the package, the years since it was first installed at your organization and the percentage of the Machine Readable sector that is accessible by the package.

1. ^Yes a commercial package

2. ^Yes an inhouse developed package

3. No

NAME	YEARS	%

26. Does your organization use an batch generalized query facility software package? If so please indicate the name of the package, the years since it was first installed at your organization and the percentage of the Machine Readable sector that is accessible by the package.

1. ^Yes a commercial package

2. ^Yes an inhouse developed package

3. No

NAME	YEARS	%

27. Does your organization use an online parametric query facility software package? If so please indicate the name of the package, the years since it was first installed at your organization and the percentage of the Machine Readable sector that is accessible by the package.

1. ^Yes a commercial package

2. ^Yes an inhouse developed package

3. No

NAME	YEARS	%

28. Does your organization use a batch parametric query facility software package? If so please indicate the name of the package, the years since it was first installed at your organization and the percentage of the Machine Readable sector that is accessible by the package.

1. ^Yes a commercial package

2. ^Yes an inhouse developed package

3. No

NAME	YEARS	%

OVER

Data Definition Language (DDL) is a software tool that provides the means for writing and describing the storage structure and access strategy for the data base.

29. Does your organization use a Data Definition Language (DDL) for defining storage structures and access strategies? Is the DDL a commercial package, an inhouse developed package or built into the DBMS? How long ago was the package acquired or developed? What percentage of applications, by EDP budget, use the DDL?

	NAME	YEARS SINCE ACQUIRED DEVELOPED	%
1. ^Yes a commercial package			
2. ^Yes an inhouse developed package			
3. ^Yes built into the DBMS			
4. No			

Information Requirements Analysis (IRA) is the process of determining the information needs of managers. Examples of IRA techniques are interviews, questionnaires, flowcharts, decision tables, simulation models and data base tracking.

30. Does your organization have a standard methodology or approach for determining the information needs of managers? Please indicate if the methodology was developed inhouse or acquired as a commercial package. Circle a number.
1. Yes an inhouse developed methodology.
 2. Yes an acquired package methodology.
 3. No the approach is left to the discretion of the analyst.
 4. Other please describe in the space below or on the back page of the questionnaire.

SOME QUESTIONS ON THE BACKGROUND OF THE DATA ADMINISTRATOR

31. For each of the following statements please circle YES if the statement applies to you and NO if the statement does not apply to you.
1. Yes No - a technical background in computer technology
 2. Yes No - at least one year's experience in each of the major functional areas of management (Personnel, Accounting, Marketing, Logistics, Manufacturing, etc.)
 3. Yes No - more than two year's experience in at least one of the functional areas of management.
 4. Yes No - a college bachelor's degree or equivalent in computer science.
 5. Yes No - a college bachelor's degree or equivalent in business administration.
 6. Yes No - a college bachelor's degree or equivalent in the Arts.
 7. Yes No - a college bachelor's degree or equivalent in the Sciences other than computer science.
 8. Yes No - a technical school diploma in computer science or business administration.
 9. Yes No - a post graduate degree in business administration.
 10. Yes No - a post graduate degree in computer science.
 11. Yes No - a post graduate degree in a discipline other than computer science or business administration.
 12. Yes No - more than three years supervisory experience.
 13. Yes No - a background in DBMS software technology.
 14. Yes No - a background in the definition, design, and implementation of computer based information systems.

32. Is the Data Administrator a member of the EDP Long Range or Strategic Planning Committee or similar group?

_____ Yes _____ No

33. Does the Data Administrator maintain a formal set of standards or procedures governing the following activities? Circle either Yes or No.

1. Yes No the definition of data in the organization.
2. Yes No the capture of data in the organization.
3. Yes No the storage of data in the organization.
4. Yes No the maintenance of data in the organization.
5. Yes No the dissemination of data in the organization.

34. Which of the following best describes how your organization processes requests from the User community to develop sets of data? Please circle a number.

1. The application systems analyst, working with Users, decides what data should be collected and how the data should be stored.
2. The Project Manager reviews the analysis of the application systems analyst and users, and then decides what data should be collected and how the data should be stored.
3. The Data Administration function reviews the analysis of the application systems analyst and users, and then decides what data should be collected and how the data should be stored.
4. None of the above. Please describe in the space below or on the back page of the questionnaire.

THIS IS THE END OF PART 1 OF THE QUESTIONNAIRE. IF YOU HAVE ANYTHING YOU WISH TO ADD, PLEASE USE THE SPACE BELOW. OTHERWISE, PLEASE PROCEED TO PART 2 ON THE NEXT PAGE.

In this part you will be asked to indicate whether an activity is the responsibility of the Data Administration function in your organization. For each activity, you will also be asked to indicate the relative amount of time that the Data Administration function spends on the activity and whether you think that the Data Administration function will spend more, less or the same amount of time on the activity two years from now.

***** EXAMPLE *****

Document and communicate what data is available to information system users.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

A. If this activity is a responsibility of the Data Administration function in your organization then you would circle YES. On the other hand, if this activity is not a responsibility of the Data Administration function in your organization then you would circle NO.

B. If no time is currently spent by the Data Administration function on this activity then you would enter '0'. In all other cases, enter the approximate percentage of the Data Administration function's time spent on this activity.

C. If you feel that 2 years from now the Data Administration function will spend less, the same or more time on the activity than you have indicated for the present, then you would circle LESS, SAME, MORE respectively.

35. Know how the corporation conducts business from general and overall points of view. Understand the general thrust, policies, goals and objectives of the corporation.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

36. Understand information requirements of the users of the information system.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

37. Identify effective corporate data base applications (through understanding of present applications and planned direction of future growth).

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

38. Mediate conflicting data needs and requirements of information system users.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

39. Establish criteria for determining the ownership of data in the corporate data base.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

40. Integrate users' data requirements for common information to ensure data non-redundancy as well as security and privacy.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

41. Develop tactical and strategic plans regarding the evolution of the corporate data base and information system.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

42. Establish, communicate, and control the policies and procedures governing operations upon the corporate data base and related dictionaries, profiles, glossaries, libraries and logs.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

43. Control the definitions and rules which provide for the privacy of information system users.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

44. Control the definitions and rules which provide for the security of data within the corporate data base.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

45. Control the definitions and rules which provide for the integrity of data within the corporate data base.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

46. Define and control the auditing and monitoring requirements of the information system with respect to the corporate data base.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

47. Insure standardization of terminology and codes associated with the corporate data base.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

48. Control the content of the corporate data base by reviewing and authorizing all related maintenance specifications.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

49. Define backup, recovery, restart and reorganization requirements associated with the corporate data base.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

50. Develop backup, recovery, restart and reorganization processes associated with the corporate data base.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

51. Document and communicate what data is available to information systems users.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

52. Define and control the priorities and scheduling of computer processes against the corporate data base.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

53. Provide consultation and help to information system users.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

54. Define information system user educational requirements which will contribute to proper maintenance and utilization of the corporate data base.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

55. Design the logical data structures in the corporate data base.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

56. Design the optimal physical storage schemes and access strategies for supporting computer processes against the corporate data base.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

57. Develop and maintain a model of the organization that shows data flows, including sources and uses, across functional areas and hierarchical management levels.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

58. Formulate synchronized maintenance procedures which will ensure the integrity of the corporate data base and related dictionaries, profiles, glossaries, libraries and logs.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

59. Define parameters to be measured and used in predicting, monitoring, and optimizing the performance of the corporate data base.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

60. Establish facilities for conducting meaningful simulation and performance analysis activities as each relates to the corporate data base.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

51. Evaluate and recommend the disposition of software technologies interfacing with the corporate data base, including DBMS, DD/D and query packages.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

52. Participate in corporate configuration planning (hardware and software) to ensure that the corporate data base needs are adequately met.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

63. Define modifications and enhancements which will improve the software technologies interfacing with the corporate data base.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

64. Establish a flexible, reliable, and efficient corporate data base operating environment.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

65. Provide for efficient migrations from one technological environment to another as these movements relate to the corporate data base.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

66. Determine the cost of acquisition of various types of data and relate this cost to the usage of that data.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

67. Provide a testing facility for computer processes interacting with the corporate data base.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

68. Establish, communicate and control the policies and procedures governing the capture of data for the application systems.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

69. Establish, communicate and control the policies and procedures governing the archiving of data.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

70. Participate, as a consultant, in the conduct of feasibility studies and Information Requirements Analyses for the development of application systems.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

71. Establish, communicate and control the policies and procedures relating to the accuracy and currency of data.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

72. Establish, communicate and control the policies and procedures governing access to the corporate data base.

A. A Data Administration responsibility?

YES NO

B. PERCENT OF TIME SPENT ON ACTIVITY

_____ %

C. TIME 2 YEARS LESS SAME MORE
FROM NOW

Thank you for your time. If you have anything to add which is not covered by this questionnaire, please use the back cover.

APPENDIX C
COVERING LETTER

APPENDIX D
FOLLOW UP CARD