THE DEVELOPMENT OF

A SCIENCE POLICY INFORMATION SYSTEM

AT THE MINISTRY OF STATE

FOR SCIENCE AND TECHNOLOGY

March, 1975

P. S. ROSS & PARTNERS
MANAGEMENT CONSULTANTS

THE DEVELOPMENT OF

A SCIENCE POLICY INFORMATION SYSTEM

AT THE MINISTRY OF STATE

FOR SCIENCE AND TECHNOLOGY

March, 1975

P. S. ROSS & PARTNERS
MANAGEMENT CONSULTANTS

THE DEVELOPMENT OF A SCIENCE POLICY INFORMATION SYSTEM AT THE MINISTRY OF STATE FOR SCIENCE AND TECHNOLOGY

March, 1975

P. S. ROSS & PARTNERS

P. S. ROSS & PARTNERS
MANAGEMENT CONSULTANTS 90 SPARKS STREET, OTTAWA KIP 5B4, CANADA, 236-9662

March 27, 1975

Dr. G.Y. Tremblay
Assistant Secretary
Resources and Information Branch
Ministry of State for Science & Technology
Room 1221 Martel Building
270 Albert Street
Ottawa, Ontario

Dear Dr. Tremblay:

We are pleased to submit our report on "The Development of a Science Policy Information System at the Ministry of State for Science and Technology". This report describes the requirements for policy information in the Ministry and outlines our suggested approach to providing a computerized document storage and retrieval capability. Additionally, we have developed a work program for the implementation of such a system at the Ministry.

In submitting this report we wish to thank the Ministry for the co-operation extended to us, and in particular to the assistance provided by Drs. Quadling and Guttormson.

We have appreciated this opportunity of working with you on the first phase of this program to develop a computer based documentation storage and retrieval system and look forward to being of continued assistance to you. If you wish to discuss any aspect of this report or the work it represents, we will be pleased to do so at your convenience.

Sincerely,

P.S. ROSS & PARTNERS

A.P. Bogie
Principal

TABLE OF CONTENTS

I. INTRODUCTION

The Ministry of State for Science and Technology was established in 1971 for the purpose of assessing and developing policies for the Canadian Government in the fields of science and technology. Since that time, the Ministry has evolved through a series of organization formats to the present structure which comprises the following principal functions:

* Policy Formulation
* Program Review and Assessment
* Resources and Information
* Personnel, Finance and Administration

The Resources and Information Branch includes, among its responsibilities, a role in the functions of policy making and review and assessment. As part of this responsibility, the Branch recently undertook a preliminary review of the requirement for computer-based systems to support information requirements within the Ministry. This review involved preliminary research of the types of information within the Ministry that were appropriate for some form of mechanized storage and retrieval, along with an exploration of the types of approach that could be considered in such a computerization process. This research resulted in a decision to undertake a more complete study of the feasibility and approach to a computer-based information storage and retrieval system, and accordingly, P.S. Ross & Partners was asked to undertake the work.

It is envisaged that such a system initially will meet internal Ministry needs and provide prompt access to a broad range of Ministry data related to the development, monitoring, and understanding of policy related topics in Canada and around the world. It is foreseen that the system subsequently will be made available for use by other federal government departments, providing them access to selected information and facilitating the creation of their own collections of policy related information.

Other parts of the information coordination process relate to the provision of information available from external sources including other library collections, statistical data, and data analysis services provided by agencies such as Statistics Canada.

Our work has involved an examination of the nature of the data flowing into the Ministry, and has identified a number of such data sources including federal government departments, provincial governments, industry organizations, universities, research organizations, and foreign governments. Further, we have conducted interviews with a significant number of senior professional staff in the Ministry to gain an appreciation of their functions and of their storage and retrieval requirements. Additionally, consideration has been given to the planned availability of the PDP-11/45 computer as a basis for supporting this system.

This report describes the scope of the work conducted, and the requirements of Ministry users. A review of systems that would reasonably satisfy these requirements is provided and followed by our recommendations.

Finally, the report outlines a suggested plan for establishing such a capability for the Ministry and provides some preliminary estimates of the related costs and timing.

II. OBJECTIVES

The objectives of this study are:

* To determine general systems requirements of the Ministry in regard to the filing and retrieval of science policy information;

* To identify the systems available to meet immediate needs;

* To identify costs of alternative systems; and,

* To outline the activities involved in the implementation of such systems.

III. POLICY INFORMATION REQUIREMENTS

A major part of our work program focused on the users of information in the Ministry. Accordingly, an extensive interview program was undertaken with some twenty senior managerial and professional staff. The interviews were designed to obtain a good understanding of the functions of each user and his perspectives, the kind of documentation and material available to him and the way it is used. During the course of each interview, we attempted to prompt the users with ideas and to elicit suggestions on alternative approaches to satisfy anticipated future requirements.

As a result of this work, we have been able to draw some conclusions on the nature of the major categories of information and documents required by the Ministry personnel. This section of our report documents our analysis of requirements, and provides an assessment of the resulting implications.

A. INFORMATION REQUIREMENTS

Our analysis of information and documents used by the Ministry determined that there are two broad classes of information available to the Ministry:

(i) Documents filed or retained at the Ministry; and

(ii) Documents available from external organizations.

The various categories of documents and materials are described briefly in the following sub-sections.

1. General Documents

This category embraces a range of documentary material processed by the Ministry including memoranda, letters, position papers, cabinet decisions, policy statements, and press releases. The materials are derived from the activities of various science and technology groups in the federal and provincial governments, universities, industry and private sector groups. The volume

of such materials in the Ministry is considerable and concerns were expressed by many individuals that it was becoming increasingly difficult to be fully aware of all such documents available in the Ministry.

2. Personal Data Collections

Many of the types of material in this broad class of information are available as part of personal files, data collections, or personal knowledge. In certain cases the systems are quite extensive and comprise a useful collection of selected documents.

3. Scientific Activities

This category of information relates to the knowledge of scientific activities in Canada (including such items as research programs), and the level of expenditures associated with them. These science and technology activities relate to federal and provincial government programs, private industry programs, and bilateral programs where other countries may be involved. The information is required to support the policy formulation and review and evaluation processes of the Ministry.

4. Policy Inventory

Knowledge of existing policies in federal and provincial governments emerged as an item of significance in policy formulation activities. The knowledge of such policies assists the Ministry in the avoidance of conflicts in policies under development and in the identification of areas where policies are indistinct or non-existent.

5. International Information

The Ministry policy formulation processes require that information on the policies of foreign governments and other organizations be available. It would also be beneficial to the Ministry to understand the framework within which foreign policies are developed. Additionally, the policy makers

and the Ministry must be aware of decisions taken by international organizations such as the UN and the OECD which may impact on Canadian policy, and must also be aware of the related view points of key groups within these organizations.

6. Who's Who

This category of data refers to key persons and key organizations involved in specific functional areas of science and technology who are involved in the policy formulation process. Availability of the names of these people and their areas of relevance is particularly useful to individuals new to the organization.

7. Quantitative Data

This requirement is for access to quantitative data collected by other departments and agencies such as Statistics Canada, both directly and on behalf of the Ministry. MOSST accumulates and analyzes financial data on federal departmental programs in the area of science and technology. Descriptions of the type of data and analyses that are available would be a useful component of a data base.

8. Published Documents

Published documents cover the broad category of both materials produced by universities, governments, research institutions and private sector organizations in the form of books, journals, conference proceedings and published research papers, etc. The Ministry makes extensive use of such materials as they relate to the areas of science and technology. The awareness or knowledge of the Ministry staff of specific materials in this category are brought about, in addition to the Ministry library staff function, by personal efforts and as a result of references from other contacts in the field. While this process provides reasonable access to a certain range of materials in this category, there are distinct limits on the extent of an individual's awareness of publications in the wide ranging application areas with which the Ministry is concerned.

Over recent years, several organizations have undertaken the development of extensive and comprehensive data bases of publications, references, and literature in certain fields of endeavor. Examples of such data bases are those developed through the efforts of the National Research Council and the CAN / SDI Program. The library and other Ministry staff are aware of the availability of certain of these data bases. It is likely that the availability of such data base information will expand considerably over the next few years. A requirement emerging from our interview program was for a better and more complete awareness of such systems and reference materials, and the provision for access to them by Ministry staff to support their generally expanding requirements.

9. Other

Other general categories of information requirements within the Ministry have been expressed as follows:

* Policy Review and Assessment - in this area certain information types could be developed to support the process, depending on the categories of information pertinent to any particular topic of science policy that is current at a specific point in time. As part of this process, it would be useful to consider the development of impact and effectiveness indicators.

* Internal Staff Skills - it was suggested that in view of the varying technological and science disciplines and experience required in the Ministry, combined with the factors of high skill level required and a relatively high rate of replacement as new areas of of interest emerge, that an inventory of staff skills and experience would be useful.

B. OPERATIONAL ENVIRONMENT

In accordance with Federal Government policy on bilingualism, due consideration was given to the needs in this area. As well, the sensitive nature of the material involved has been appropriately considered so that provision may be made for adequate security arrangements.

1. Bilingual Capability

An important factor that emerged relative to a data storage and retrieval system is the requirement to interrogate the system in either English or French. There are two requirements to be considered:

(a) English and French interactive capability for the operations staff using the system.

(b) Provision of automatic translation equivalents for selected keywords.

2. Security Requirements

The question of security arose in a number of areas, and it is clear that, in the long term, any storage and retrieval system must provide an adequate level of security. Greatest security can be assured by storing information and materials on the Ministry's premises and making use of the PDP 11/45. This approach necessitates the development of new computer software with implied higher initial costs.

3. It is foreseen that the computer-based system will be accessed by the user in two alternative ways:

(a) Directly - in this mode, the user personally will enter requests for data directly through computer terminals. The physical operation will not be difficult to learn, but it may take time for the user to access the material required in the most effective manner.

(b) Through the Information Exchange function - in this mode, the user will relay his requirements to the Information Exchange function who will, in turn, enter requests through the computer terminals, obtain the required results, and relay them back to the user.

C. GENERAL IMPLICATIONS

The analysis of requirements described in the first part of this section of the report confirms the need for:

* The creation of a computer-based information system to provide storage and retrieval capability for document abstracts and bibliographies for the categories of material and documents identified in this section of the report that are not otherwise available through other services;

* The development of adequate manpower resources to maintain the information data base and to assist the staff to access the information; and,

* The development of a system to provide a framework for the classification of science policy related documents. This will enable the retrieval system to function effectively in the storage and retrieval of data and documents.

IV. SYSTEMS FOR DOCUMENT STORAGE AND RETRIEVAL

Section III of this report identifies requirements for the storage and retrieval of nine broad categories of policy related documentation. These categories may be divided into two broad classes: documents held by the Ministry; and, documents or literature available from other organizations for which computer based lists of references have been created or can be expected in the future. This study explored the alternatives available to MOSST for the storage and retrieval of documents of the above categories using computer-based systems, and evaluated their applicability to MOSST.

A. COMPUTER STORAGE AND RETRIEVAL SYSTEMS

The creation of a computer-based capability for the storage and retrieval of references held or identified internally by the department requires that MOSST acquire a hardware and software capability. Such a capability must be able to store key information concerning a document that will enable it to be stored and retrieved easily through identification criteria and descriptions of the document.

A document storage and retrieval system of this kind can be supported by special purpose software operating on either an "in-house" computer or at an external service bureau, each with access to local terminals on MOSST's premises. In each case, however, there must be a sufficient guarantee of security to protect the information from unauthorized access.

This study included a review of the capabilities of external computer service bureaus to supply services of this type to the Ministry. In addition, the market was surveyed to determine the availability of software suitable for use (directly or with modifications) on the PDP-11/45 computer being installed by the Ministry.

Several organizations have been identified with the capabilities to provide the external services in flexible user format with minimal start-up costs.

Potentially, however, there is less data security associated with the service bureaus, as their systems are located remote from their clients and connected by communications lines. Such an environment provides less opportunity for control and security of data than provided by the "in-house" option. However, it should be pointed out that considerable effort is put into the provision of a well controlled and secure

environment for the processing, transmission and storage of data, by external service bureau organizations. Those who merit serious consideration for the processing of the Ministry's data take extensive security precautions which, in some cases, have been approved by the RCMP and the Department of Supply and Services.

Our survey of some twenty organizations who indicated a capability to provide the services found that the most suitable, satisfying all the identified requirements, is the VIP system provided by COMSHARE Limited. The requirements indicated were for a general purpose storage and retrieval system able to store document abstracts with related key words and bibliographic and classification information. The system would operate on an interactive environment with data searches occurring by use of combinations of specific references and key words. The system would be flexible and able to support the variable quantities and formats of bibliographic data for the different storage requirements represented by each different category of documents.

The VIP system would enable the Ministry to implement a selected portion of its overall documentation storage and retrieval system requirements in a relatively short period of time. While permitting the Ministry to move rapidly in this area, it would also provide experience in the use of such systems. Furthermore, this option would allow the conversion of the Ministry's present files and documents to a computerized form that could also be converted to operate on an "in-house" system.

Another option, the use of the PDP-11/45 computer, is more suitable to the Ministry in the longer term because reference and abstract information can be retained entirely on the Ministry's premises. This provides the most secure environment for sensitive material and documents which the Ministry has in its possession. However, despite an extensive survey in Canada and the U.S., our study was unable to locate any software available for use on the PDP-11/45 which satisfies the range of requirements defined above. However, a number of developed programs that satisfy some of the requirements were identified.

In addition, there are a number of organizations with experience in the development of software systems of the type described. Preliminary estimates indicate that the requisite software can be developed for application on the PDP-11/45 in 9 to 18 months for a cost of $45,000 to $150,000, depending on the ultimate degree of sophistication required. Some of the factors that would impact on the cost and timing are:

* the extent to which the software is parameterized for general application;

* the incorporation of a "key word in context" capability for the abstracts;

* the inclusion of a technical synonyms capability; and,

* the degree to which a bilingual capability is incorporated.

B. DOCUMENT RETRIEVAL TECHNIQUES

The retrieval of documents ideally is effected by the use of key words in combination with a policy classification system. The key word system searches the data files for documents which contain the words specified in the search criteria. The policy classification system enables the user to obtain all references to documents in a particular area of policy.

C. MANPOWER DEVELOPMENT

Whatever approach is selected, there is a need within the Ministry, to upgrade the capability of the information gathering and dissemination function to meet the increased demands inherent in the planned system development. Such a function would develop and maintain an awareness of external data bases and provide the capability to obtain data from the data base. Systems already available would include major data bases in use by the federal government such as the CANOLE and WATDOC bibliographic access systems.

Manpower will also be required to provide the staff of the Ministry with assistance in submitting data to the data base and performing data retrieval searches on their behalf. The facility also should exist for the staff to carry out these functions on their own behalf, although not all staff members will wish to operate in this manner.

As part of the overall data gathering function for the total system, primary responsibility for the content of portions of the data base may reside with selected individuals or groups. For example, foreign government intelligence information may be the responsibility of one group within a branch. The central function must support and coordinate these groups in their use of the system.

The approaches identified in this section incorporate the use of the latest computer hardware equipment in combination with state-of-the-art software development. The implementation of such a system is well within the capabilities of the services available to the Ministry.

V. RECOMMENDATIONS

The following are the recommendations of the study team with respect to the development of a science policy information system.

* A computer storage and retrieval capability should be developed for documents and information maintained by the Ministry.

Such a system will require computer software that will enable relevant information concerning documents to be stored in a computer and retrieved through a range of selection criteria. The system would provide support to the library function, staff members, and eventually services to other government departments.

* The Ministry should make use of an external service bureau to provide this capability, in the short term.

A review of several software systems meeting certain of the requirements of the Ministry shows that the most appropriate service is the VIP system available from COMSHARE Limited. Such a system would be accessible through the terminals on the PDP-11/45 computer in the Ministry. Some minor modifications may be required to the equipment for this specific purpose. Costs will vary depending on frequency of use and the size of the data bases, but a range of $15,000 to $25,000 per year with 10,000 documents on file and 6 retrieval sessions per day should be considered typical.

The above approach will enable the conversion of part of the Ministry's operation to a computerized retrieval system. The data collected at that time will be stored in a form readily transcribable to the in-house PDP-11/45 referred to in the following recommendation.

* The Ministry should develop the computer software necessary to implement the system in-house on the PDP-11/45.

This subsequent implementation is important in order that security requirements can be met. At the same time it will remove the operational costs of using an external computer service.

* The Ministry should define job specifications and responsibilities for an information exchange function with

responsibility for the ongoing operation of the information storage and retrieval capability.

The information exchange function will provide the services associated with abstracting, document storage, computer input and document searches. The function also will have responsibility for maintaining awareness of other documentation systems available to the Ministry from other sources.

* An indexing system of keywords should be developed.

This task should be undertaken early, since the availability of a suitable thesaurus will significantly affect the success of the storage and retrieval system. The development of the thesaurus must take account of the requirements for bilingual capability, and accordingly, synonyms should be developed for application in both languages.

The indexing system would make use of keyword and bibliographic data based on policy classification, for data storage and retrieval purposes.

VI. A PLAN FOR IMPLEMENTATION

This section of our report details an approach to implementation that will provide a sound basis for implementing our recommendations and a framework for the management of the process. Exhibit 1 overleaf, shows a phased work program with parallel activities on which the approach is based, identifies whether the activity is carried out by MOSST or externally, shows the man-time requirements, and identifies likely external costs and the time frames involved.

Inherent in the approach is the need to provide capability for document storage and retrieval at an early stage while at the same time making it part of the longer term process to establish a computer software system that will operate on the Digital Equipment Corporation, PDP-11/45, which is planned for installation at the Ministry. The approach has seven major activities which are outlined in more detail in the remaining part of this section.

1. Project Management

It is important in a project of this kind that an integrated approach to the overall management of the work program be established to ensure adequate control and supervision throughout all the phases. The project manager should be familiar with both the user environment and the technologies involved in view of the on-going requirement for liaison and supervision with MOSST officers, and with computer/technical personnel. Additionally, the project manager will participate in other activities inherent in the program, including the development of the information exchange organization.

2. User Coordination

In order to provide ongoing MOSST involvement and increasing responsibility, the primary user team interface should involve a responsible staff member from MOSST. This will also ensure that an adequate knowledge of the system and its development remains with MOSST throughout the implementation phase, as well as ensuring maximum co-operation and participation of MOSST branches in the development program.

IMPLEMENTATION ACTIVITIES

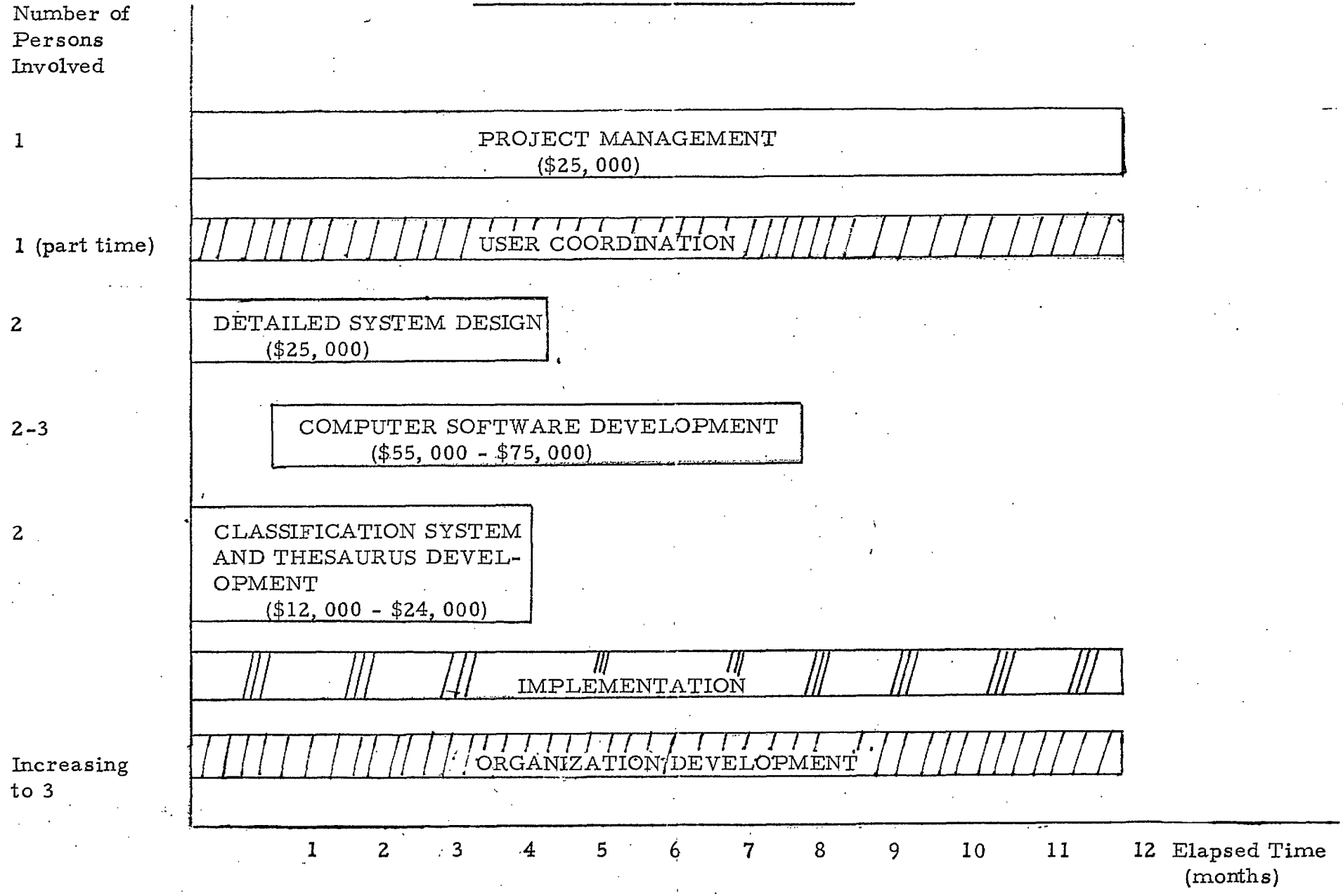

Exhibit 1

Legend:
- external resources:
- MOSST staff:
- mixed MOSST staff and external resources:

N.B. The dollar amounts on this chart are approximate, and represent expenditures for both the service bureau (V.I.P.) and mini-computer options.

3. Detailed Systems Design

During this activity, a conceptual design should be prepared after having examined and analyzed the details of existing systems. Following this, an assessment can be made of the implications on the computer hardware/software systems and on the user operations. Finally, detailed specifications can be developed for computer software development.

4. Computer Software Development

Using the specifications developed in activity 3, the computer programs should be developed and tested. Maximum advantage should be taken of utilizing existing techniques, methodologies and logic for storage and retrieval.

5. Classification System and Thesaurus Development

The classification system should be conceived and examined against the framework of general criteria developed to assess its viability. Following this, a detailed design should be developed and an acceptability test conducted. Instructions can then be prepared for using the system.

The thesaurus concept should be developed having regard to the maximum desirable size, and the approach adopted to French/English equivalents. A search should be made of other available thesauri which relate to the science and technology field, and also to the type of documents in use by MOSST.

6. Implementation

This process relates to a number of objectives, including the implementation of the system on an external computer service bureau using existing software, the training of staff, the preparation of procedure documentation required to implement, operate and use the systems, and the subsequent conversion of the system to the Ministry's PDP-11/45 computer. The process will be ongoing from the inception of the project, and will initially include data gathering for document analysis and for eventual system definition.

7. Organization Development

This activity is concerned with the ongoing planning for, and staffing of, the information exchange function that will be responsible for the operation of the system. It is envisaged that this group will be involved in maintaining and searching the document data bank of behalf of staff of the department and, where required, assisting the staff to carry out this function on their own. A more detailed description of the functions of this organization is contained in Section IV of this report.

DATE DE RETOUR