

SCIENCE INFORMATION AND INFORMATION SCIENCE IN
THE FEDERAL BUDGETARY PROCESS

Text of a speech to the
Canadian Association for Information Science
May 17, 1977

by
Morley S. Lipsett
Ministry of State for Science and Technology

Ottawa, Canada K1A 1A1

Introduction

In the spring of 1977 the federal budgetary process culminated once again in the tabling in the House of Commons of the "Estimates" for the forthcoming year (1977-1978).¹ But this time a snapshot had been taken of the Government's intentions with respect to scientific activities. A data base was created containing the historical, current, and planned allocations for science of all government Departments and Agencies. Science, in Canada, had become a more conspicuous part of the federal budgetary process.

This enabled a new publication by the Ministry of State for Science and Technology, entitled "Federal Science Programs 1977-1978"², to appear shortly after Estimates 1977-1978 was published. In the foreword, the Minister of State for Science and Technology, the Honourable J. Hugh. Faulkner, explained that:

"This publication reports on the scientific and technological activities that will be undertaken in the 1977/1978 fiscal year with federal government funds. It is the first time such a comprehensive account of future science programs has been published by the federal government.....".

Aimed at a wide readership, "Federal Science Programs 1977-78" concentrates on aggregate budgetary figures for science and narrative descriptions of major scientific and technological programs.

Details of the underlying data base were presented concurrently in a companion MOSST report entitled "Federal Science Expenditures 1975/1976 - 1977/1978"³. A more specialized publication, it presents numerous statistical tables and charts, without commentary, and is directed to a technical audience.

Having cited these publications, my purpose in this paper is not to discuss them further nor to elaborate on their contents. It is to describe the system of data collection and display which was used in the preparation of these publications. I wish to demonstrate the workings of an information system consisting of people, departments, surveys, reports, computers, terminals and software. This system enables users to look into the future, in step with the process used by the government to lay its budgetary proposals before Parliament, prior to the start of each new fiscal year.

The Budgetary Process

The federal budgetary process has been documented on an annual basis in the federal government publication entitled "How Your Tax Dollar is Spent". A particularly clear explanation of this highly complex process is given in the 1975 edition.⁴ The main events are illustrated in Figure 1. It is important to note that the process is a continuing cycle of proposals and approvals of budgetary estimates for Government Programs.

To be considered in the same budgetary context, science must march along with the same dynamics, and be measured accordingly. This was not always the case.

Two main events in the budgetary process concern us here. These are:

- (a) Program Forecasts: This is the Treasury Board call for a springtime report on budgetary proposals by each Department or Agency for the next year, in particular, and for future fiscal years. This part of the process is carried out in accordance with general guidelines issued to Departments by the government but, in general, is an internal management matter between Departments and the Treasury Board. It contains an aspect of continuity for existing departmental programs and an aspect of wishful thinking for new or expanded ones.

(b) Main Estimates: Another call by Treasury Board, this is a request to Departments and Agencies for a report, in the fall, on budgetary estimates derived from the earlier spring-time Program Forecasts. This time, however, the program forecasts have been modified to conform to the government approvals following a lengthy and analytical process of consultation and negotiation with Treasury Board. These reports are consolidated by the Treasury Board Secretariat to become the new Estimates for the forthcoming year. At this point, when the Estimates are presented to Parliament, this part of the process becomes public.

This year's Estimates (1977-78)¹ is the presentation to Parliament of the government's spending plans totalling 45.1 billion dollars (of which 41.1 billion is budgetary) for the programs of some 93 Departments and Agencies. I mentioned earlier that we had detailed information concerning the science content of these programs within the same context as Estimates 1977-78. The information had been requested via a Science Addendum when the Treasury Board Secretariat issued its traditional call for Main Estimates and this request had been preceded by an addendum to the earlier call for Program Forecasts.

The Infrastructure

Departments and Agencies willingly responded to the request for information. They did so because they had previously agreed to participate collectively with the Ministry of State for Science and Technology, Statistics Canada, and the Treasury Board Secretariat in this process, and because they had agreed that this was an appropriate way to unify science statistics with budgetary figures. Sounds simple enough, but we are dealing with huge institutions where science and technology as such are not necessarily high on the list of their preoccupations, and where questionnaires, however well-intentioned, impose substantial burdens on the respondents.

Therefore let me explain the infrastructure that led to this process being accepted and which made it successful. This necessitated much goodwill on the part of many people; in particular by official representatives of each Department and Agency having any interest whatsoever in science and technology. There exists a standing committee of such officials, the so-called Interdepartmental Committee on Science Expenditures (ICSE). Established many years ago, initially by the predecessor

to MOSST, the Science Secretariat of the Privy Council Office, ICSE is the general interdepartmental mechanism by which consultation and coordination are achieved with respect to content, timing, and purpose of federal science surveys. Statistics Canada is the traditional operational agency for such surveys. MOSST is the major policy-oriented client and provides the secretariat for ICSE. Departments and Agencies are the major suppliers of the data and frequently find themselves to be users as well.

Before federal science surveys became integrated with the budgetary process, they occurred as annual, retrospective events without involvement in the budgetary machinery. The resulting data was historically accurate, but insufficiently timely to indicate, at the outset of a new fiscal year, which directions the scientific activities of departments and agencies were going to take. It was therefore agreed by all of the various participants to integrate these surveys into the budgetary process. The separate surveys by Statistics Canada became replaced by the aforementioned addenda to the Program Forecasts and Main Estimates, and the data obtained in this way had to be suitable to be merged with Statistics Canada's historical series.

The deputy heads of Statistics Canada and MOSST agreed to a very interesting experiment in government organization which greatly simplified the interactions necessitated by this process. They

arranged for the establishment of a satellite of Statistics Canada at MOSST. The former Science Statistics Section of Statistics Canada moved in toto adjacent to MOSST facilities to become the new Science Statistics Centre, and its work program was integrated with that of MOSST. The objectives of the Centre are shown in Appendix A.

However, complete organizational integration has not taken place nor was it intended. Statistics Canada operates under an act that prohibits unrestricted government access to many of its data bases. Similarly, MOSST, because of its role as scientific advisor to other departments as well as to Treasury Board and the Cabinet, has access to planning-type data that is transitory and lacks official sanction and which, therefore, would not be appropriate for inclusion in Statistics Canada's data bases.

The integration can be rather extensive, nonetheless, in joint activities such as the Program Forecast and Main Estimates science addenda exercise. In this instance, a joint project team was assembled by both Statistics Canada and MOSST which worked closely with departmental respondents in carrying out the various steps leading to the creation, updating and accession of the data base on federal scientific activities.

With several months of experience behind us we can now say the experiment in organization appears to be working very effectively. This is partly due to a common purpose; partly because of effective working relationships; and some innovative EDP facilities at MOSST which are shared by the various participants.

EDP Aspects

Such a heading usually signals an introduction to a highly technical description of hardware, software, data-base organization, and other esoterica. However, I am not going to talk about machine architecture, bottom-down programming, or random access files. The closest I'll be getting to "technical" is in Table 1, and Figure 2., which summarize the salient aspects of the EDP facilities and relevant computer programs, respectively. My aim is to concentrate here on how the MOSST EDP facilities appear to the users. For this purpose let's consider some of the major steps in the process.

1. Asking departments for information

Questionnaires such as the Main Estimates Science Addenda are sent to departments and typically ask for data on the past year (historical), current year (forecast), and new year (estimates). At the same time as the questionnaires are issued, each department receives a computer printout in similar, if not identical, format containing any existing data that had previously been provided by the respondent department. Thus only updates, or corrections to the existing data base, need be provided. The burden on the respondent is simplified and the EDP facilities are used, in this instance, to reformat or re-present existing information in such a way as to assist in filling out what has now become a rolling report.

The general rule we apply is as follows: where we are asking for an update to existing information, a handy copy of that information is supplied.

2. Dealing with the returns

The project team receives the forms from the respondents, duly filled in, but with the inevitable errors, i.e., transpositions, arithmetic errors, etc. In answer to prompting-type questions appearing on a CRT screen, this information is fed into

an interactive terminal. While this is going on some reasonableness checks are carried out internally and gross errors are flagged. This step is of course preceded by an examination of the forms to see whether they look right to the experienced eye. Input in this way, the returns, one for each of the human and natural science components of each departmental program become part of a potential on-line science expenditure data base. The design philosophy here is that the terminal should help the statistician, clerk or other user feed the information into the computer and help ensure that the data is accurate. It should not make him feel stupid.

3. Making sure

I've mentioned that the data is only potentially part of the science expenditure data base at this stage. This is because the next step is to make sure that the data is self-consistent and correct. For this purpose a number of programs are available to the project team to perform a wide variety of checks and cross checks. The programs assist in the process of recognizing and correcting errors, but the corrections are made by experts in the meaning of science statistics on a department by department, program by program, scientific activity by scientific activity

basis. If these experts know anything at all about computers, however, it is fortuitous or irrelevant. The emphasis here is on the user-machine interaction and assumes pedestrian users, i.e., those without EDP expertise.

4. Presenting the data

To me this is the most exciting aspect of what you all would recognize as a specialized information management and reporting system. All too often technical virtuosity is used to create an elegant data base, only to have it remain stashed away on a disc pack because the potential user is turned off by problems of data inaccessibility or the other extreme, too much data. I think that most of us have known the humiliating experience of trying to understand a query language only to find out the answers we were looking for were buried somewhere in a two-inch thick pile of computer print-out. Let me tell you about how we have tried to avoid these problems. Our design philosophy has been as follows:

1. The users of the output of the system don't have time to digest voluminous reports. They prefer concise tables, and they want to be able to change their minds about the format or contents of the tables at their convenience, not at the convenience of a programmer.

2. Tables are fine, but they should be self-contained and include footnotes, if necessary, to establish definitions or to cite references.
3. Graphs and charts are easier to understand than tables and are the easiest way for the user to visualize the significance of the figures being dealt with. Needless to say, items 1 and 2 above apply equally well to graphics.
4. Hard copy is generally useless unless it represents exactly what the user wants to see and use. Thus hard copy should be discouraged at intermediate stages and encouraged only at the end.
5. The user is not only entitled to change his mind about what he wants, he should be encouraged to do so! I know this is redundant with 1, but the point here is that our design philosophy is not simply to tolerate human nature, it is to recognize that EDP should expand the user's horizons and not contract them to someone else's programming constraints.

Back then to presenting the data. Two families of programs have been developed for accessing our data bases. One is for general purposes, and assumes only that the user wants to prepare tables, graphs, charts or combinations of all three from existing or new data. For example, the tables and graphs in the "Trends" section of reference 2 were prepared in this way. The other family, which obviously derives from the first, is for the routine accession of specific data bases and for the automated production of updated reports. An example, here is the preparation of all of the tables and charts for ref. 3 . As a matter of fact once the data base on science expenditures for 1977-78 had been created, it only took about two days for Federal Science Expenditures 1975/76-1977/78 to be made ready for printing in both languages. This was facilitated for example by program options which automatically change routine headings to correspond to another year's data added to the data base.

Looking at things differently

In my opinion one of the most important purposes served by an information management and reporting system is to enable its clients to look at data differently as a result of unanticipated questions. Let me give you two examples, one of them occasioned

by a new set of guidelines on contracting out recently issued by
the Treasury Board Secretariat,⁵ the other occasioned by this
meeting and your particular interest in Scientific and Technical
Information.

In the case of the Contracting-out (Make or Buy) Policy, those
scientific activities which are included in the policy are
specified in the guidelines and are defined in accordance with
standard conventions. Corresponding to international usage,
these definitions are given official status in Statistics Canada
and MOSST publications, and reflect considerable consultation
with members of ICSE. As mentioned earlier, each of these
scientific activities exists as a separate entry for each program
of each department in the data base, and much of this detail is
published in ref. 3. Thus there exists current information which
can be projected to show possible trends in scientific activities
under the influence of the contracting-out policy. It will be no
surprise therefore that programming facilities are available in
MOSST to access this data base and to model the consequences of
various policy options.

This was a valuable aid at the policy-development stage -
especially because of the graphics-orientation, i.e., the visual-
display aspects.

My second, and concluding example concerns your field of interest, information. Let us be somewhat legalistic at this point and confine ourselves to Scientific and Technical Information (STI). There is no official definition for STI as such, but we can build up a definition by reference to those scientific activities which would logically fit into the broader term STI. These are the so-called Scientific Data Collection and Scientific Information activities in the Natural Sciences, and the General Purpose Data Collection and Information Services activities in the Human Sciences. Their official definitions can be found in refs. 3, 5, 6 and 7.

Let us suppose you would like to find out how much money is budgeted for STI by the federal government for 1977/78; how it compares with total science expenditures? And how it might have changed over the past few years? I asked these and similar questions and the answers, including numerous revisions, were generated on-line in a few hours. The results are shown in figs. 3 through 5. Incidentally much of this same information could have been derived from the data given in ref. 3. but more laboriously.

Concluding Remarks

There were a couple of messages I hoped to convey in this talk.

The first is that "information science" can make dealing with science information exciting. This occurs when the information base we are using is relevant; when it tells us something about what is happening right now; and about what will be happening in the future. The particular information base I was discussing has this sort of relevance. I find special significance in the fact that we have progressed from an after-the-fact survey of how federal resources for scientific activities were divided up in the past, to a forward-looking window into Federal intentions for science in the context of the federal budgetary process.

The second message is possibly the more important one to his audience; namely that information science has not very much to do with information and even less with science. It is an art.

The end user of information only knows that he has a question that needs answering. It is our job to provide the answers using every resource at our disposal. When this resource happens to be Electronic Data Processing, remember that it's unnatural to converse with a machine and it's difficult to read computer printouts. It is much more natural to visualize the significance

of something, and it's mainly to assist in this process that we should be emphasizing EDP. Behind the scenes we should use computers to eliminate drudgery, to make it easier for respondents to respond, clerks to do clerical things, analysts to analyze, planners to plan, statisticians to create statistics and users to use.

The job of an information scientist, your job, can be expressed very simply. It is to help people find things out.

Acknowledgements

I wish to acknowledge that in preparing this paper I drew principally upon the contributions of others. It would be specious to mention everyone by name and subject area, but I do wish to thank the members of ICSE for their patience and constructive interaction.

Errors of fact or interpretation are however my responsibility, alone, and should not reflect adversely on any of my colleagues.

References

1. Estimates 1977-78, Government of Canada,
(DSS catalogue No. BT 31-2/1978)
2. Federal Science Programs 1977-78, Ministry of State
for Science and Technology, Canada
(Cat. No. ST 21-3/1978)
3. Federal Science Expenditures 1975-76 - 1977-78, Report
No. 100-3, Ministry of State for Science and Technology,
Canada.
(ISSN 0701-8142)
4. How Your Tax Dollar is Spent, Treasury Board, Canada
(DSS catalogue no. BT 47-1/1976) p.60)
5. Policy and Guidelines on Contracting-Out the
Government's Requirements in Science and Technology,
Treasury Board, Canada
(Cat. No. BT 32-10/1977) April 1977.
6. Federal Government Activities in the Natural Sciences,
Statistics Canada.
(Cat. 13-202 Annual)
7. Federal Government Activities in the Human Sciences,
Statistics Canada
(Cat. 13-205 Annual)

TABLE 1. EDP FACILITIES

HARDWARE

DEC PDP 11/45
256 kb core
floating point
4 - 2.5 Mb cartridge disc drives
1 - 88 Mb disc pack drive
1 - 800 bpi. 9 track tape drive
600 lpm, 132 char electrostatic printer
16 line, programmable asynchronous multiplexor
varous modems for local and remote terminals

TERMINALS

5 - CRT, alphanumeric terminals (DECVT05B)
various hard copy, low-speed alphanumeric terminals
2 - storage CRT, alphanumeric, graphic terminals
(Tektronix 4015)
Electrostatic printer/plotter (Versatec)
Sharable between storage CRT terminals for
hard copy.

SOFTWARE Operating System - RSTS/E Version 6B

Principal language - BASIC plus

Special software - Graphics handler to drive
the 4015's in graphics mode.

Sub-objectives associated with the development of statistical indicators of the use of S&T are:

- a) To examine and evaluate the indicators used by other countries such as the U.S.A. and members of the Council for Mutual Economic Assistance.
- b) To investigate the relationships between S&T activities and economic indicators.
- c) To provide new statistical series relevant to the use of S&T in Canada, and to the formulation and implementation of specific science policies.

APPENDIX A: OBJECTIVES OF THE SCIENCE STATISTICS CENTRE

The Centre has three principal objectives:

1. To carry out a comprehensive program of statistical data collection concerning the resources devoted to science and technology in Canada.
2. To provide statistical support to departments, in general, and to the Ministry of State for Science and Technology, in particular, for studies relating to the resources utilized on science and technology.
3. In cooperation with other interested departments, to develop a system of statistical indicators appropriate for the analysis of the use and effect of science and technology in Canada.

Sub-objectives associated with the measurement of resources committed to S&T are:

- a) To develop and maintain a national statistical framework for R&D consistent with the needs of the users of such data in Canada, and, insofar as possible, compatible with the international models of the OECD and UNESCO.
- b) To conduct surveys of the industrial and government sectors.
- c) To develop satisfactory surveys or estimating techniques for the university sector.
- d) To develop statistical series for the measurements of non-R&D aspects of S&T.

Sub-objectives associated with the statistical support of projects of other departments are:

- a) To advise other departments on the relevant available statistics and their characteristics.
- b) To carry out special surveys or tabulations for such projects.
- c) To ensure that any statistical work done by the Centre for such projects meets the requirements of the project and is of maximum general utility.

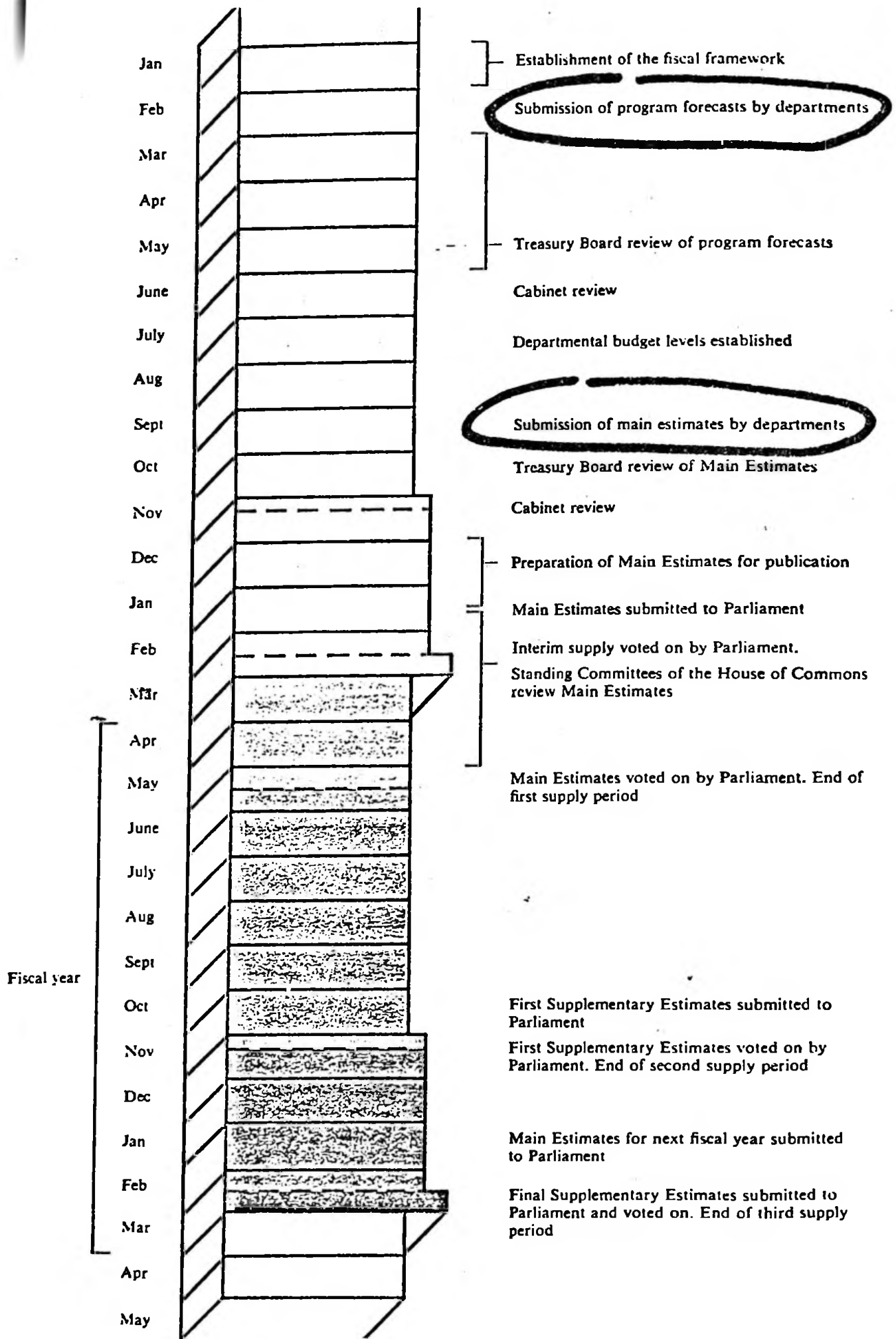


Figure 1. The Budgetary Process

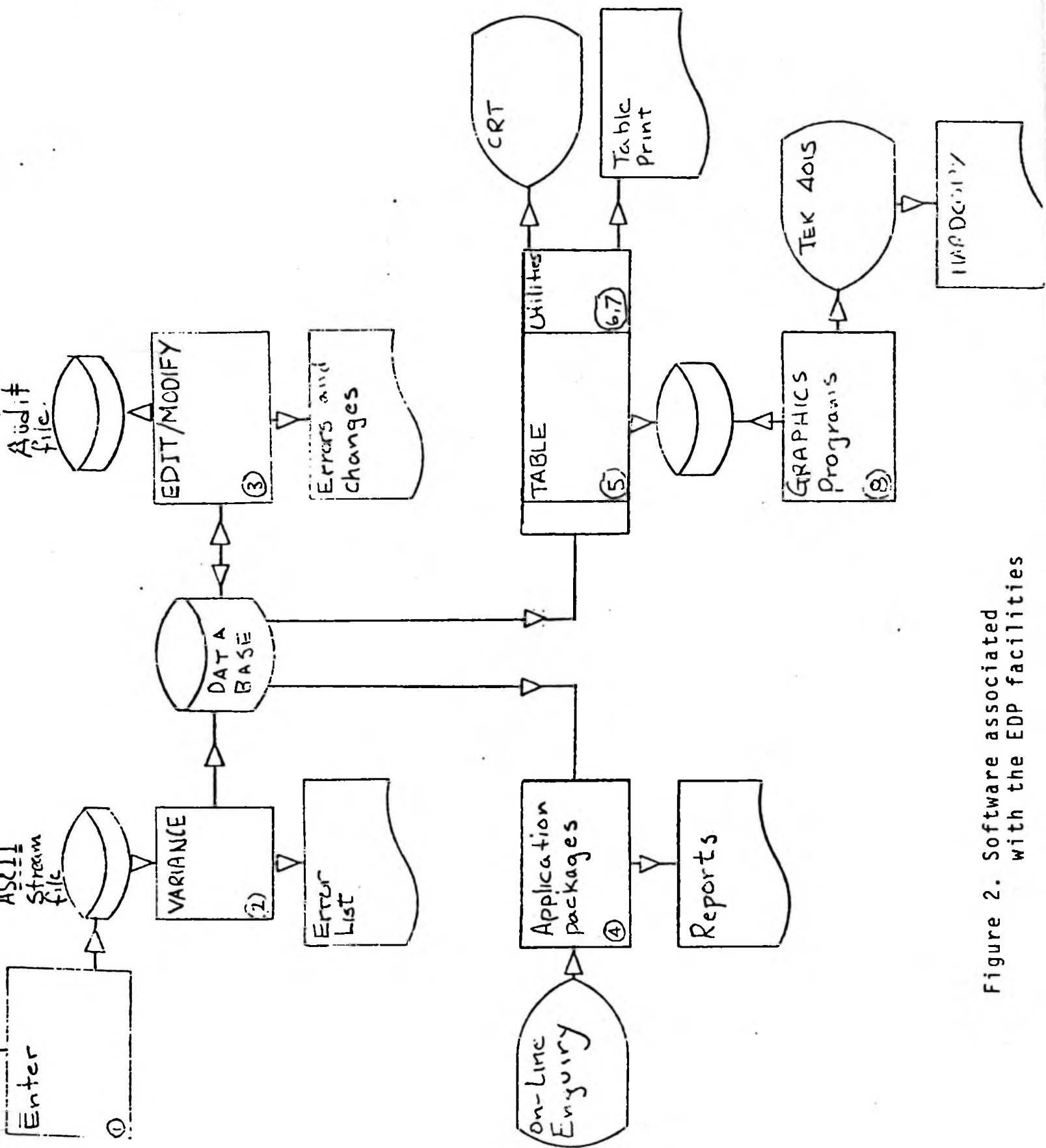


Figure 2. Software associated with the EDP facilities

FEDERAL EXPENDITURES ON SCIENTIFIC & TECHNICAL INFORMATION (STI)
IN THE HUMAN AND NATURAL SCIENCES

NATURAL AND HUMAN SCIENCES

(\$ MILLIONS)

	1972-73	1973-74	1974-75	1975-76	1976-77	1977-78
RESEARCH AND DEVELOPMENT	666.5	749.9	808.6	890.0	999.5	1,078.1
SCIENTIFIC AND TECHNICAL INFORMATION(a)	253.9	279.7	329.5	339.5	413.7	412.1
TOTAL SCIENCE EXPENDITURES(b)	1,042.1	1,136.6	1,275.4	1,367.8	1,577.4	1,664.1
STI AS A PERCENT OF TOTAL SCIENCE	24.4%	24.6%	25.8%	24.8%	26.2%	24.7%

SOURCE: STATISTICS CANADA AND MOSST

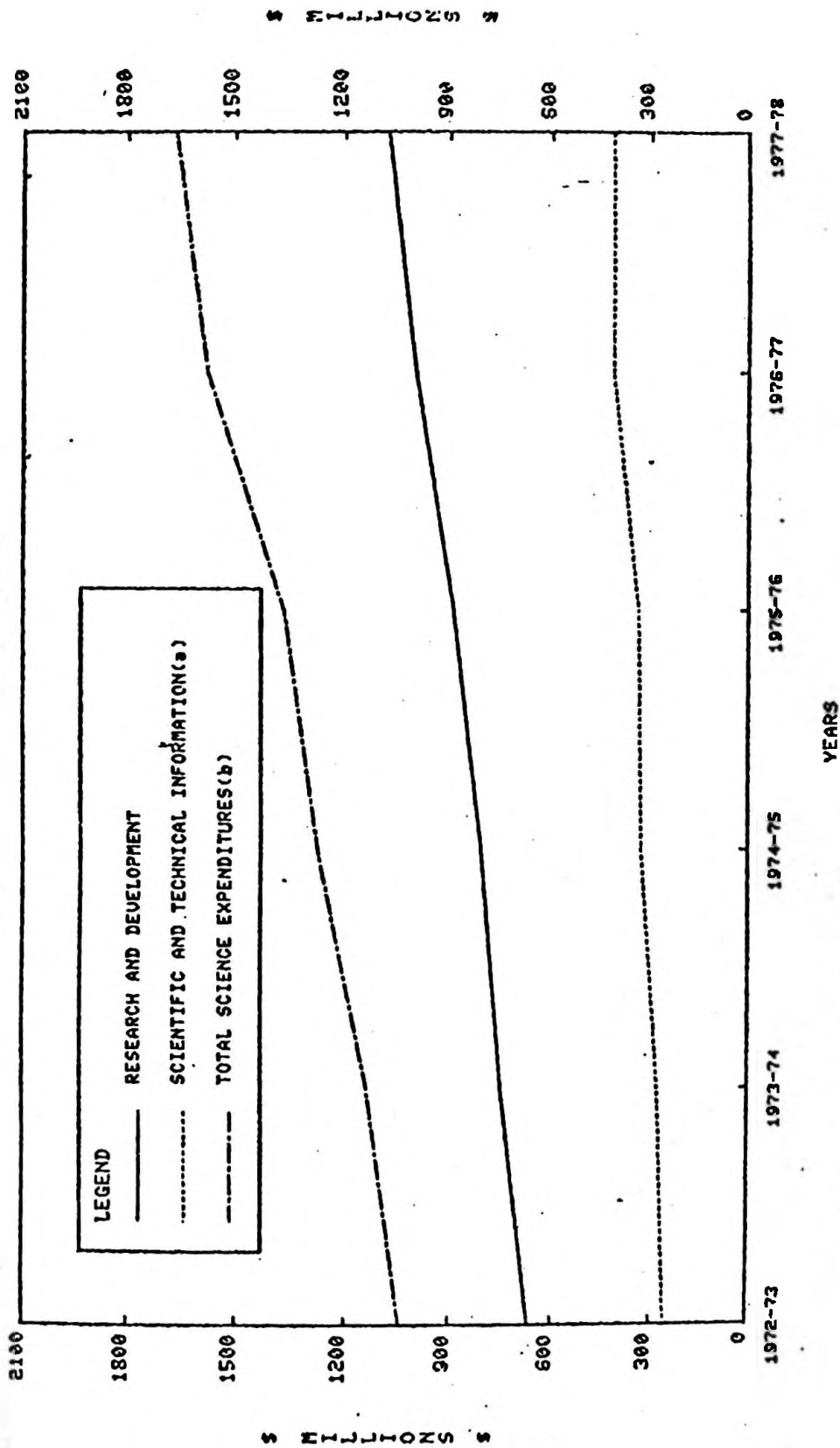
(a) COMPRISED OF SCIENTIFIC DATA COLLECTION AND SCIENTIFIC INFORMATION (STATISTICS CANADA AND MOSST DEFINITIONS).

(b) INCLUDES RELATED SCIENTIFIC ACTIVITIES OTHER THAN STI

ED ON 04-May-77

Figure 3. Computer generated table on Federal Expenditures for STI.

FEDERAL EXPENDITURES ON SCIENTIFIC & TECHNICAL INFORMATION (STI) IN THE HUMAN AND NATURAL SCIENCES

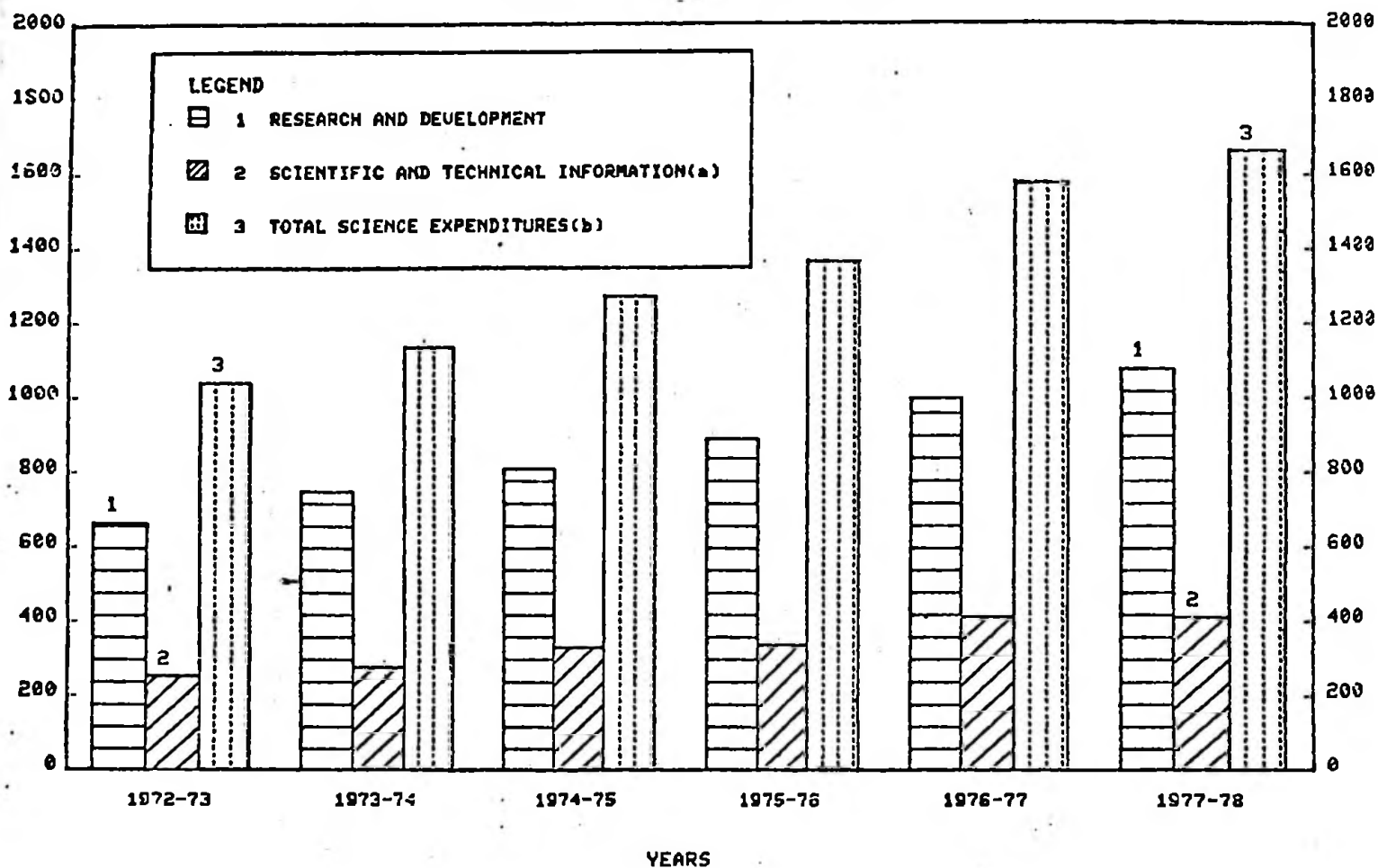


SOURCE : STATISTICS CANADA AND MOSST

- a - COMPRISED OF SCIENTIFIC DATA COLLECTION AND SCIENTIFIC INFORMATION (STATISTICS CANADA AND MOSST DEFINITIONS).
- b - INCLUDES RELATED SCIENTIFIC ACTIVITIES OTHER THAN STI

Figure 4. Computer generated graphs of Federal Expenditures on STI.

**FEDERAL EXPENDITURES ON SCIENTIFIC & TECHNICAL INFORMATION (STI)
IN THE HUMAN AND NATURAL SCIENCES**



SOURCE : STATISTICS CANADA AND MOSST

a . COMPRISED OF SCIENTIFIC DATA COLLECTION AND SCIENTIFIC INFORMATION (STATISTICS CANADA AND MOSST DEFINITIONS).

b . INCLUDES RELATED SCIENTIFIC ACTIVITIES OTHER THAN STI

Figure 5. Computer generated histograms of Federal Expenditures on STI.