

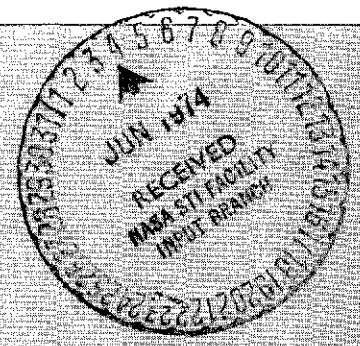
NASA TM-X-69903

NSSDC

74-09

AN OVERVIEW OF THE
NATIONAL SPACE SCIENCE DATA CENTER
STANDARD INFORMATION RETRIEVAL SYSTEM
(SIRS)

MAY 1974



NATIONAL SPACE SCIENCE DATA CENTER

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION - GODDARD SPACE FLIGHT CENTER, GREENBELT, MD.

(NASA-TM-X-69903) AN OVERVIEW OF THE NATIONAL SPACE SCIENCE DATA CENTER STANDARD INFORMATION RETRIEVAL SYSTEM (SIRS) (NASA) ⁴⁷ p HC \$5.50 CSCL 05B N74-22613 Unclas 76 G3/34 38553

AN OVERVIEW OF THE
NATIONAL SPACE SCIENCE DATA CENTER
STANDARD INFORMATION RETRIEVAL SYSTEM
(SIRS)

By

Arthur Shapiro
National Space Science Data Center

Sheldon Blecher
and
Ellen E. Verson
PMI Facilities Management Corporation

Technical Editor

Margaret L. King
PMI Facilities Management Corporation

May 1974

National Space Science Data Center
National Aeronautics and Space Administration
Goddard Space Flight Center
Greenbelt, Maryland 20771

CONTENTS

	<u>Page</u>
INTRODUCTION	1
Purpose	1
Background	1
Brief Description of SIRS	2
SIRS FILE STRUCTURE	5
General Structure	5
Types of Fields	7
Types of Data	8
SIRS DICTIONARY FORMAT	11
SIRS FUNCTIONS	13
Update	13
Search and Retrieval	15
Sort	16
Special Report Writers	18
General Report Writer	21
Extract and Merge	27
Derived Items	28
APPENDIX A. BRIEF DESCRIPTIONS OF NSSDC FILES MAINTAINED BY SIRS	A-1
APPENDIX B. A SAMPLE SIRS FILE	B-1

PRECEDING PAGE BLANK NOT FILMED

ILLUSTRATIONS

<u>Figure</u>		<u>Page</u>
1	The Hierarchical Structure of a SIRS File	6
2	Fields Allowed by SIRS	9
3	Data Allowed by SIRS	10
4	The Process of Creating a User's Data File	12
5	Updating a User's Data File	13
6	Example of Search and Retrieval	17
7	Sample of an Edit List Report	19
8	Sample of a Line List Report	20
9	Sample of a Histogram Report	22
10	Process for Generating a Report	23
11	Example of an Index Report Created by the General Report Writer	24
12	Example of a Data Catalog Created by the General Report Writer	25

INTRODUCTION

PURPOSE

This document presents a general overview of the National Space Science Data Center (NSSDC) Standard Information Retrieval System. It describes, in general terms, the information system that contains the data files and the software system that processes and manipulates the files maintained at the Data Center. Emphasis in this document is placed on providing users with an overview of the capabilities and uses of the NSSDC Standard Information Retrieval System (SIRS). Examples given in this document are taken from the files at the Data Center. Detailed information about NSSDC data files is documented in a set of File Users Guides, with one user's guide prepared for each file processed by SIRS. Detailed information about SIRS is presented in the SIRS Users Guide.

BACKGROUND

NSSDC was established as part of the Goddard Space Flight Center in 1965. The major functions of the Data Center are the collection, organization, storage, announcement, retrieval, dissemination, and exchange of space science data obtained from spacecraft experiments, sounding rocket probes, and high-altitude aeronautical and balloon investigations. To assist in performing these functions, a number of automated files have been developed. These files contain information pertaining to: (1) all spacecraft launchings, experiments carried aboard these spacecraft, and data sets emanating from these experiments (Automated Internal Management File); (2) all identified scientific sounding rocket launchings (Rocket File); (3) names and addresses of NSSDC data users and the data requested (Request Activity and Name Directory File); (4) documents of interest to NSSDC and its users (Technical Reference File); (5) the NSSDC data holdings (Data Set Inventory File); and (6) the nonsatellite data available at NSSDC (Nonsatellite Data File). More complete descriptions of the files just mentioned are given in Appendix A. To obtain a better understanding of the information system in use at the Data Center, the user should become familiar with these files. A future document, An Overview of the National Space Science Data Center Information System, will present more detailed descriptions and discussions of these files. Although the types of data contained in these data files and the specific structure of the data base vary considerably, all files are being processed using SIRS.

BRIEF DESCRIPTION OF SIRS

SIRS was designed specifically for use at NSSDC. It is a tape-oriented system implemented on the IBM 7094 computer. Many files may be maintained uniformly by SIRS because it is a "dictionary-driven" system. The term "dictionary-driven" means, simply, that all data contained in a user's data base is defined, or described, in the user's dictionary file (discussed in a later section) and is used to control operations performed on the data base. This allows changes to be made to the data base without the need for reprogramming.

SIRS provides NSSDC with a comprehensive and reliable information storage and retrieval system. It contains the basic capabilities required of any generalized information storage and retrieval system. SIRS provides a means by which a user can easily define and create a data base, update the data base, edit the data base, search and manipulate the data base, and produce formatted report displays of the data base content. Each of these basic capabilities is described briefly here and in detail later in this document.

To create a data base under SIRS, a user must design a dictionary that describes the data file to be created. The user must build the dictionary using SIRS and then create the data file using the dictionary and actual data. The user is able to update the data file by adding new information, replacing data already in the file, and deleting data from the file. Appendix B illustrates an example of a simple SIRS file extracted from the Automated Internal Management (AIM) File.

There are several editing features available to a user of SIRS. These features include checking minimum and maximum values for numeric data, checking for defined types of data, and checking to ensure that proper file structure is maintained. Also, the user may define tables of acceptable or unacceptable data values. These tables are contained in a SIRS file (Validation Table File) and are used as a means of checking data being entered into the system.

In addition to such editing features, SIRS offers many file manipulation capabilities. These include the ability to sort a data file, to extract selected data from a data file and to form a subfile containing these data, to merge two data files into one, and to search and retrieve records in the data file that meet certain user-specified criteria.

Several types of report capabilities are also incorporated into SIRS. These vary from a simple type of formatted listing showing the location of all data in the data file to a capability that allows the preparation of complex types of user-defined reports.

The following sections discuss the file structure of SIRS, its dictionary concept, and all of the major capabilities offered by SIRS. If more information is required than is given in this document, consult the SIRS Users Guide, an NSSDC publication that supplies information necessary to use SIRS.

SIRS FILE STRUCTURE

This section describes the general structure of a data file that is used with SIRS. There are certain restrictions SIRS imposes on the structure of a data file. These restrictions include maintaining a hierarchical structure of records within a file, grouping data into categories within records, and defining data in accordance with the acceptable SIRS formats. All these restrictions are discussed in detail in this section.

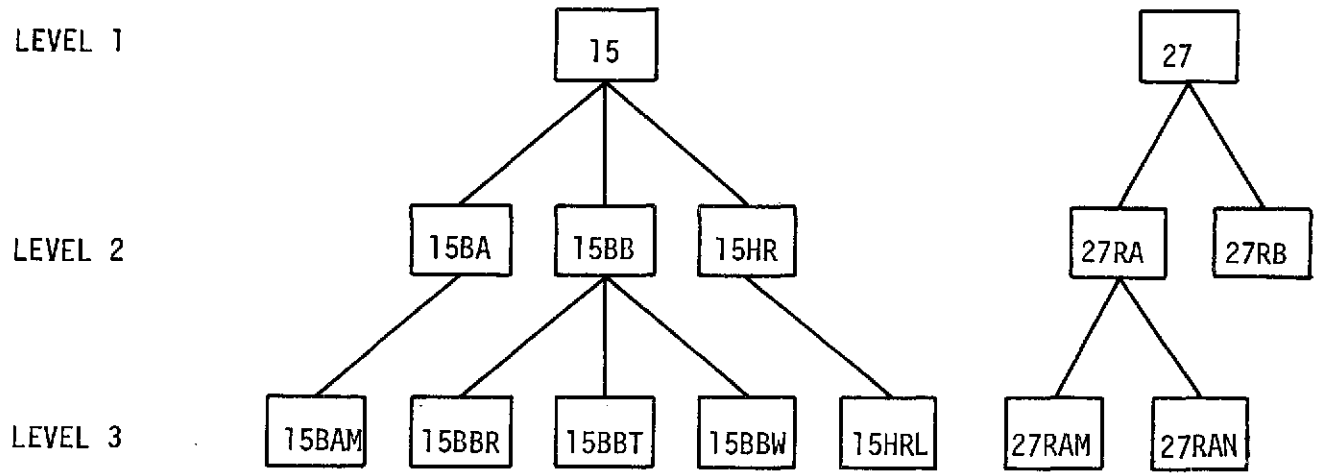
GENERAL STRUCTURE

The SIRS data files are stored on magnetic tape. At the beginning of each tape, there is a "tape label record" that uniquely identifies the tape and ensures that the proper tape has been mounted by the computer operations staff.

SIRS data files are stored in a hierarchical (or tree) structure, extending to as many as three "levels." A "level" within a SIRS data file may be defined as a group of records containing similar types of information. For example, the AIM File consists of three basic types of records: records that contain information about spacecraft, records that contain information about experiments carried aboard spacecraft, and records containing information describing the data emanating from these spacecraft experiments. In this instance these three types of records comprise the three levels of a file.

SIRS files consist of a series of records, each with a unique identification, called a record ID. The length of the record ID field for each of the possible three levels is user defined. This record ID may be a maximum of 30 characters in length. According to the SIRS hierarchical file structure, a level 2 record must have a longer record ID than a level 1 record, and a level 3 record must have a longer record ID than a level 2 record. The level 2 record ID is an extension of the level 1 record ID, and the level 3 record ID is an extension of the level 2 record ID. Thus, it is possible to trace a path from any level 2 or level 3 record to the appropriate level 1 record. (See Figure 1 for further clarification.)

PRECEDING PAGE BLANK NOT FILMED



Each box represents a record. The characters in the box are the record ID. In this example a level 1 record has a two-character ID; a level 2 record has a four-character ID; and a level 3 record has a five-character ID.

Figure 1. The Hierarchical Structure of a SIRS File

To further describe the SIRS file structure, certain terms should be defined. Fields in a SIRS file are grouped into categories, and groups of categories comprise a record. A field is information named or identified in any SIRS file. A logical grouping of fields is a category. Because a record containing many fields is easier to manipulate when divided into several smaller categories, the grouping of fields serves as a convenient method for a user to organize a data file. Categories are also a device to accommodate data that are variable in length and, with a minimum of programming, to optimize space used. Fifty categories are allowed per record in the user's dictionary, with the user specifying the categories needed and what will go into each category. Categories one and two are always reserved for system use and contain the record ID and dates a record is created and last modified.

TYPES OF FIELDS

There are five types of fields allowed by SIRS.

1. A Single-valued Field - This field allows only one data value in a field for any given record.
2. A Fixed-length Vector - This vector is a repeatable field that may contain more than one data value in a field for a given record. The maximum number of values allowed is user specified and is called the repetition count. Space is allocated in the record for the maximum number of values specified when the dictionary definition for this field is encountered.
3. A Variable-length Vector - This vector, like a fixed-length vector, is a repeatable field that may contain more than one data value in a field for any given record; however, only the number of occurrences containing data will appear in a record. Thus, the number of characters actually stored in a variable-length vector field will vary from no data values to the maximum number of data values allowed, depending on the input to the file.
4. A Matrix - This is a group of fields joined together by a common field name, called the "parent name." This feature allows the user to show direct relationships between fields containing more than one data value per record. A matrix may contain more than one row of data values. Each row is labeled with a two-character row identification. This identification is used as a subscript when the user wishes to access a specific row to update a file. The user specifies a repetition count, which is the maximum number of rows for a particular matrix, although only those rows put into the file are allocated space.

5. A Column Vector - This is any column in a matrix field. All column vectors must be given unique field names by the user. Specific values in a column vector may also be accessed by using the row identification as a subscript.

Figure 2 shows examples of each type of field.

TYPES OF DATA

There are six types of data allowed in a SIRS file.

1. Integer Data - These are whole numbers that do not contain a decimal or an exponent.
2. Real-number Data - These contain a decimal point. Exponents are also allowed.
3. Alphabetic Data - These consist of the characters from A to Z, plus blanks.
4. Alphanumeric Data - These consist of any printable character.
5. A Date Field - This field may be either four or six characters in length, referencing a specific date. A six-character date is entered in the form MMDDYY, where MM represents month, DD represents day, and YY represents year. A four-character date is entered in the form MMY Y.
6. A Free-text Field - This field contains alphanumeric data. This data type is considered special because it specifies a field of varying length in a SIRS file. The free-text field type allows the user to insert, into a record, variable-length descriptions in sentence and paragraph form.

Figure 3 lists examples of these six types of data.

In summary, a SIRS data file consists of records that consist of categories that, in turn, consist of fields. A single-valued field is the smallest accessible portion of a SIRS file. SIRS files are structured in a hierarchical manner, with up to three levels available to the user. All data fields allowed in a user's data file are pre-defined in what is known as a "user's dictionary." The contents and structure of this dictionary are discussed in the following section.

TYPES OF FIELDS

1. SINGLE-VALUED FIELDS
2. FIXED-LENGTH VECTORS
3. VARIABLE-LENGTH VECTORS
4. MATRIX
5. COLUMN VECTORS
(ANY COLUMN OF A MATRIX)

EXAMPLES

AMES RESEARCH CENTER
JOHN R. SMITH
37.856

[37.9, 18.6, 12.2, , ,]
[JONES, SMITH, BROWN, WHITE, , ,]

[37.9, 18.6, 12.2]
[JONES, SMITH, BROWN, WHITE]

01	CP	EL	4.0E4	1.E38	← ROWS OF A MATRIX
02	CP	EL	2.5E5	1.E38	
17	CP	PN	5.0E5	1.E38	
24	CP	DT	4.0E6	1.E38	

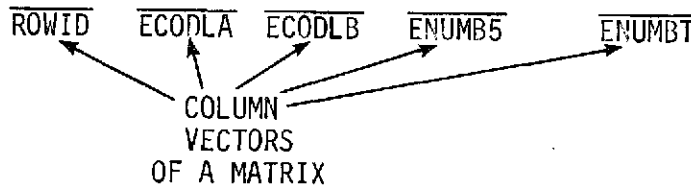


Figure 2. Fields Allowed by SIRS

TYPES OF DATA

EXAMPLES

1. INTEGER

4

2. REAL NUMBER

4.0

3. ALPHABETIC

VERSION A

4. ALPHANUMERIC

VERSION 1

5. DATES

031751, 0351

6. FREE TEXT

INJUN 3 WAS A MAGNETIC-FIELD ALIGNED
SPACECRAFT INSTRUMENTED FOR A STUDY
OF GEOPHYSICAL PHENOMENA.

Figure 3. Data Allowed by SIRS

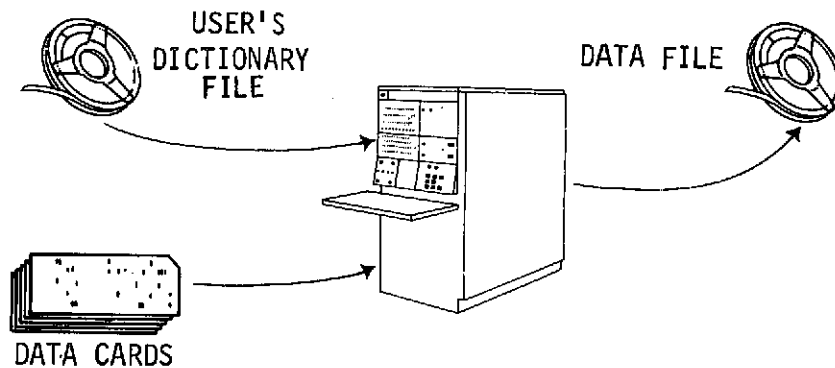
SIRS DICTIONARY FORMAT

This section describes the information a user must supply to create a dictionary and the methods used to create a data file. SIRS is a dictionary-driven system. This means that all data fields contained in a SIRS file have been precisely defined in a user's dictionary. Characteristics such as the field name, the field type, and the data type are part of a user's dictionary file.

The user's dictionary file is a SIRS file, and all rules concerning file structure apply to the dictionary file. Dictionary field entries will be referred to as descriptors in the remainder of this document. Fields will normally refer to the location of data values in a data file. Once the user's dictionary file has been created, it may be used as an input file to create a data file. Figure 4 illustrates the process of creating a user's data file. Data specification cards are used as input, along with the user's dictionary file, and the resulting output is the user's data file. By using this data file in conjunction with the user's dictionary file, any of the SIRS capabilities may be requested.

For each data field defined in the user's dictionary, certain characteristics must be given. These characteristics include the dictionary record identification, the field name, the character length of the field, the update permit flag ("YES" or "NO"), the level where the field is located, the category where the field is found, the starting character position of the field within the category, the data type, the field justification (right or left), the number of times the field repeats, and the field type.

Many other field descriptors may be entered into the user's dictionary file. These descriptors are optional, for the most part, and a complete discussion of both required and optional descriptors can be found in the SIRS Users Guide. It is also possible to subdivide a field and to define fields in the user's dictionary that overlap one another. One field may have more than one name or portions of that field may be given unique names. Contiguous parts of two contiguous fields may also be given another unique name. These features allow the user a great deal of flexibility in defining fields. For example, a six-character date field may be defined as an entity, and each two-character segment may also be given a name, so that the month, day, and year are individually accessible.



The user's dictionary file is read into the computer, along with data specification cards, and a user's data file is created.

Figure 4. The Process of Creating a User's Data File

The dictionary file is an integral part of SIRS. The dictionary is, in itself, a SIRS data file, and all SIRS functions are applicable to it. All fields contained in a user's data file are defined in the dictionary. The user has almost complete control over the structure of data files and, through foresight and thoughtful file construction, shares the responsibility for the efficient use of SIRS.

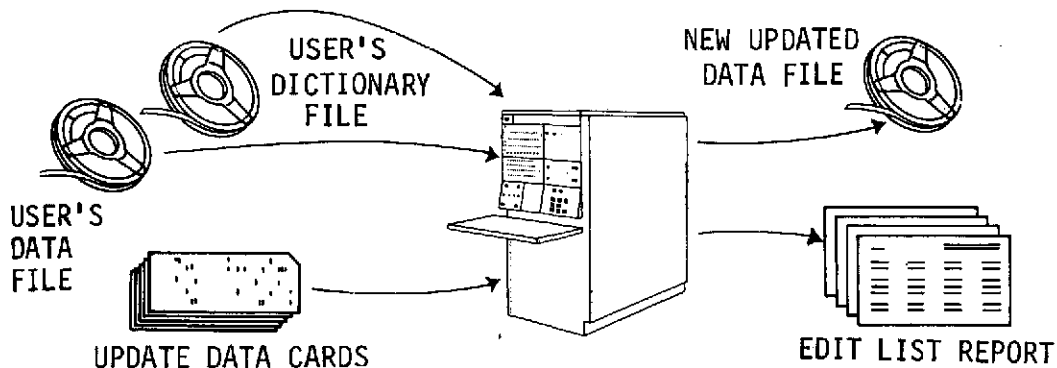
The final sections of this document contain a discussion of each of the capabilities offered by SIRS.

SIRS FUNCTIONS

This section describes the major user-oriented capabilities offered by SIRS. These are update, search and retrieval, sort, special report writers, general report writer, and extract and merge. Also included in this section is a discussion of special capabilities offered by SIRS and performed by derived item routines.

UPDATE

SIRS provides the user with a comprehensive update capability. The term "update" is used to mean any creation, modification, or deletion of data values in a field, category, or record in a SIRS data file. This section includes a discussion of the features of the update segment of SIRS. These features include the editing of input data, the automatic maintenance of the hierarchical structure of a SIRS file, and the ability to group input data for update operations. Figure 5 shows the process of updating a user's data file.



This is the process of updating a user's data file. The user's dictionary file, data file, and the changes required (update data cards) are read into the computer, producing an updated data file. A standard report, called an edit list, is automatically produced when updating a data file. See Figure 7 for an example of an edit list report.

Figure 5. Updating a User's Data File

Editing of input data may be done by several methods. Checks are performed automatically when minimum and maximum data value boundaries for numeric fields are specified in the user's dictionary, defining the data file being processed. Value checks are also performed automatically, according to the dictionary descriptor, identifying the type of data to be contained in a specific field. The user may exercise control over the contents of fields in the file by specifying, in the user's dictionary, the fact that a table containing either acceptable or unacceptable data values for a particular field or fields is to be checked for all data being entered into the field. This table is called a validation table and is, in itself, a SIRS data file stored on a magnetic tape. By specifying in the dictionary a table name and comparison criteria, only those allowable values specified by the user will be entered into the file.

When data are being entered into the file for a record that is not already in the file, a check is done automatically to ensure that the correct hierarchical structure for that file is maintained. Records that do not fit into the specified structure (see File Structure section) will not be entered into the file.

SIRS allows the grouping of replacement data values for two or more fields into one single data input group. The grouped replacement data values are split into individual update entries according to a predefined pattern created by the user. For example, to update six fields -- SAGENT (spacecraft acquisition agent ID code), SVHICLE (spacecraft launch vehicle), SITE (launch country/site code), LDATE (launch date), LTIME (launch time), and DECAY (spacecraft decay date) -- a user would supply the following six input data cards:

```
SAGENT=JJB
SVHICLE=THOR-AGENA
SITE=UA      VNBC
LDATE=121362
LTIME=0405
DECAY=082568
```

To avoid supplying six input data cards describing a particular spacecraft, the user may define a group field name that specifies the pattern to be used for separating the grouped data values. The following input data card will cause the six fields (SAGENT, SVHICLE, SITE, LDATE, LTIME, and DECAY) to be modified appropriately.

```
FIELDS=JJB|THOR-AGENA|UA      VNBC|121362|0405|082568
```

This capability not only minimizes the number of card transactions required to update a user's data file but also allows for the design of simple fixed-field input forms.

SEARCH AND RETRIEVAL

The capabilities to search a file and to retrieve records that meet specified search criteria are an important part of any information processing system. This section includes discussions of the search and retrieval capabilities of SIRS. Boolean, or logical, and relational operators are used extensively in the search and retrieval segments of SIRS and are described in this section.

The relational operators consist of equal to (.EQ.), less than or equal to (.LE.), greater than or equal to (.GE.), not equal to (.NE.), and the inclusive "between" (.BT.). Boolean operators are used with field names contained in a data file to join sets of relational search criteria.

Three Boolean operators are available to the user: AND, OR, and NOT. These operators are used in conjunction with the five relational operators just described, allowing more than one relational expression to be combined into a complex expression by using AND, OR, and NOT. A user may specify any one or combination of Boolean operators to be used with one or more of the relational operators to satisfy any search requirements.

Retrieval of records from the AIM File is performed routinely at the Data Center by searching for specific values in the Phenomenon Measured Keyword String matrix. This matrix is used to describe the phenomenon measured by a particular spacecraft experiment. The string consists of 26 fields; 18 are two-letter codes and 8 are floating-point numbers. The first six of the 18 two-letter codes identify what is measured, the next four specify where the measurements are taken, the following four codes state when the measurements are taken, and the last four codes identify the meanings of the four pairs of numerical values (eight floating-point numbers) that are also part of the keyword string. This entire string may repeat up to 10 times.

The first of the 18 two-letter codes is used to divide all space phenomena into five classes: (1) electromagnetic fields (EM), (2) charged particles (CP), (3) microscopic neutral particles (MN), (4) macroscopic bodies (MB), and (5) other experiments (OT). The field name given to this two-letter code is ECODLA. Thus, if a user wishes to retrieve all experiment records that measure charged particles in a file, the search criterion consists of 'ECODLA .EQ. CP'.

Occasionally, a user may wish to use search criteria that are more selective than the examples just mentioned. To perform this task, a "subset" of search criteria is defined. A subset consists of two or more search criteria that are given a name for future reference as one entity. When performing a search based on a matrix, such as the

Phenomenon Measured Keyword String, the user must specify that the field to be searched is a matrix field. This guarantees that a record will be retrieved only if all specified search criteria are met within any one row of the matrix.

For example, a user may wish to retrieve experiment records that measure charged particles. However, the user may be interested in only two types of charged particles, protons (PN) or hydrogen ions (DT). In this case a second field name, ECODLB, must be used as a part of the search criteria. ECODLB defines the type of charged particle that is measured. The user's search criteria would now consist of two parts, 'ECODLA .EQ. CP' and 'ECODLB .EQ. PN OR DT'. By giving this set of search criteria a name and by specifying that the fields are part of a matrix, SIRS will retrieve experiment records in which CP and PN or CP and DT occur in the same matrix row. Figure 6 is an example of this. The user may create many subsets of search criteria and, by joining them with one of the Boolean operators, may perform very complex searches of a data file.

Searching and retrieving may be performed on all field types used in SIRS, with the exception of free-text fields. Also, a particular occurrence of a repeatable field may be used as a search criterion, and multilevel searching may be performed.

SORT

This section contains descriptions of the sort capability of SIRS. This capability allows the user to duplicate a data file and sort this duplicate copy in an order different from the usual record ID order. The field types that may be used to order the copy of the data file are single-valued fields, fixed-length vectors, variable-length vectors, matrices, and column vectors. Repeatable fields, such as vectors and matrices, may be subscripted to select a particular occurrence that will be a part of the "sort key."

The data values of user-specified fields, for which the file is to be sorted, are stored in a special area within each record of a user's data file. This area is called a sort key. The fields that will comprise the sort key are user specified; the data values contained in the sort keys are used to sort, in the desired sequence, records of the input data file onto the output data file.

Two types of sort sequences are available to a user. The first type is called the master file alphabetic sort order. This sort order places special characters first, followed by numbers, then letters. The second type of sort is called the logic sort. In this case, characters are sorted according to the internal codes assigned to each

Below are three records containing the Phenomenon Measured Keyword String.
 The search criteria are 'ECODLA .EQ. CP' and 'ECODLB .EQ. PN OR DT':

CODED SEARCH CRITERIA	}	SUBA	MATRIX						
			ECODLA	.EQ.\$CP					
			ECODLB	.EQ.\$PN\$DT					
		START							
		SUBA							

62-067B-01 RECORD → RETRIEVED	[01	CP	EL	4.0E4	1.E38	62-067B-02	[01	CP	PS	4.0E4	6.0E4
		02	CP	EL	2.5E5	1.E38			02	CP	EL	8.0E4	1.1E5
		03	CP	PN	5.0E5	1.E38							
		04	CP	PS	4.0E6	1.E38							
		<u>ROWID</u>	<u>ECODLA</u>	<u>ECODLB</u>	<u>ENUMBS</u>	<u>ENUMBT</u>		<u>ROWID</u>	<u>ECODLA</u>	<u>ECODLB</u>	<u>ENUMBS</u>	<u>ENUMBT</u>	

62-067B-03 RECORD → RETRIEVED	[01	CP	AP	5.2E5	4.0E6	62-067B-03 RECORD → RETRIEVED	[01	CP	AP	5.2E5	4.0E6
		02	CP	EL	5.2E5	4.0E6			02	CP	EL	5.2E5	4.0E6
		03	CP	DT	9.0E5	1.8E6			03	CP	DT	9.0E5	1.8E6
		04	CP	PS	9.0E5	1.8E6			04	CP	PS	9.0E5	1.8E6
		<u>ROWID</u>	<u>ECODLA</u>	<u>ECODLB</u>	<u>ENUMBS</u>	<u>ENUMBT</u>		<u>ROWID</u>	<u>ECODLA</u>	<u>ECODLB</u>	<u>ENUMBS</u>	<u>ENUMBT</u>	

RECORD 62-067B-01 IS RETRIEVED BECAUSE
 'CP' AND 'PN' OCCUR IN THE SAME ROW.
 RECORD 62-067B-03 IS RETRIEVED BECAUSE
 'CP' AND 'DT' OCCUR IN THE SAME ROW.
 RECORD 62-067B-02 IS NOT RETRIEVED
 BECAUSE NEITHER 'PN' NOR 'DT' OCCURS
 IN THE SAME ROW AS 'CP.'

Figure 6. Example of Search and Retrieval

character by the computer. The logic sort capability exists in SIRS because there are some cases where numbers will not sort in an acceptable order when using the master file alphabetic sort.

When a repeatable field is specified by the user to define a part of the sort key, one record on the input data file may create more than one record on the output file. This occurs because there are as many different values of a sort key for one input record as there are elements in the specified repeatable field; i.e., there are as many distinct output records as there are elements in the specified repeatable field. Therefore, if a repeatable field is specified as part of the sort key definition and that field repeats five times in one particular input record, five duplicate records containing five different sort keys will be created on the output file as a result of the sort.

SPECIAL REPORT WRITERS

SIRS provides three types of special report writers for the user. This section describes each of these three types - the Edit List, the Line List, and the Histogram report. The term "special report writer" used in conjunction with SIRS means that the user has little control over the appearance of the report, but does maintain control over the contents of the report.

1. Edit List - The simplest type of report available is called an Edit List. An Edit List of a file provides a picture of the contents of the records. This report allows the user to view the contents of the data file as it appears on magnetic tape. Figure 7 is a sample Edit List report.
2. Line List - Another type of report that is available to the user is called the Line List. The user may specify that only certain fields in the file be printed. Field names from the dictionary are chosen to be printed, along with any report title the user may choose. Centering of the report title and columns of data is done automatically, as well as printing the date of the computer run, page numbers, and record identifications. Figure 8 is a sample Line List report.
3. Histogram Report - A third type of special report is the Histogram report. This shows a count of the number of occurrences of a data value in a given field versus the occurrences of a data value in any other field. Abscissa and ordinate field names are specified as input to the Histogram report writer, and the output is a tally of occurrences of data values. These field names may be single-valued fields or vectors. The

RECORD ID	CATEGORY NUMBER	
[65-068A-01]		*** 204 WORDS ***
CATEGORY 1	1	RECORD ID
CATEGORY 2	2	0 [65-068A-01] \$ ← CATEGORY DELIMITER
CATEGORY 3	3	0 063072082873
CATEGORY 4	4	0 MATRIX ROW ID
CATEGORY 5	5	0 [1]NSSDC ← ZODIACAL LIGHT PHOTOGRAPHY
CATEGORY 6	6	0 GEMINI 5, ZODIACAL LIGHT PHOTO NEY CCDWF 45\$
CATEGORY 7	7	0 2040473U OF MINNESOTA
CATEGORY 8	8	0 A HANDHELD CAMERA (F/1) EQUIPPED WITH AUTOMATIC TRIGGERING TOGRAPHS OF AIRGLOW, ZODIACAL LIGHT, THE MILKY WAY, AND STAR CONSTRUCTED FOR THE EXPERIMENT, HAD A 50-BY 130-DEG FIELD OF VIEW. THE CAMERA WAS PROGRAMMED TO TAKE FRAMES IN A SEQUENCE IN WHICH THE INTERVAL BETWEEN FRAMES WAS DOUBLED, AND ENDED AT 3 MIN. THE SHUTTER WAS CLOSED FOR 20 SECONDS BETWEEN EACH FRAME REORIENTATION. \$
		0 SEE TYPESCRIPT ENTRY IN NOTEBOOK. \$
		0 SEE TYPESCRIPT ENTRY IN NOTEBOOK. \$

1 2 3 4 5
 12345678901234567890123456789012345678901234567890123456789
 CHARACTER POSITION COUNTER

Figure 7. Sample of an Edit List Report

DATE AND PAGE NUMBER → DATE - 07/12/73 PAGE 1
 PROVIDED AUTOMATICALLY

TITLE (AUTOMATICALLY CENTERED) ← LINE TEST - TO BE USED TO CHECK HISTOG

RECCRD ID	SAGENT	DAGENT	EAGENT	AGENT	SPECIFIED FIELD NAMES (AUTOMATICALLY PRINTED)
62-067A					
62-067B	JJB			JJB	
62-067B-00A	JJB	JJB		JJB	
62-067B-00B	JJB	JJB		JJB	
62-067B-01	JJB		JJB	JJB	
62-067B-01A	JJB	CDW	JJB	CDW	
62-067B-01B	JJB	JJB	JJB	JJB	
62-067B-01C	JJB	JJB	JJB	JJB	
62-067B-02	JJB		EGS	EGS	
62-067B-02A	JJB	EGS	EGS	EGS	
62-067B-03	JJB		JJB	JJB	
62-067B-03A	JJB	JJB	JJB	JJB	
62-067B-03B	JJB	JJB	JJB	JJB	
62-067B-04	JJB		JJB	JJB	
62-067B-04A	JJB	JJB	JJB	JJB	
62-067B-05	JJB		JJB	JJB	
62-067B-05A	JJB	JJB	JJB	JJB	
62-067B-06	JJB		JJB	JJB	
62-067B-06A	JJB	JJB	JJB	JJB	
62-067B-07	JJB		JJB	JJB	
62-067B-07A	JJB	JJB	JJB	JJB	
62-067B-08	JJB		RNH	RNH	
62-067B-08A	JJB	RNH	RNH	RNH	
62-067B-09	JJB		LLD	LLD	
62-067B-09A	JJB	LLD	LLD	LLD	
62-067B-10	JJB		JJB	JJB	
62-067B-10A	JJB	JJB	JJB	JJB	
62-067B-11	JJB				
62-067B-11A	JJB				
62-067B-12	JJB		LLD	LLD	

RECORD ID (AUTOMATICALLY PRINTED)

DATA VALUES

20

Figure 8. Sample of a Line List Report

record-count option, which is part of the Histogram report writer, produces a tally of only the first occurrences of a data value within a field and within a record. This tally lists the number of records in which a specified data value occurs, without regard to the number of occurrences of the data value within a record. Figure 9 is an example of a Histogram report.

GENERAL REPORT WRITER

The general report writer provides the capability to create reports in formats other than those offered by the special report writers. This part of SIRS allows the user to specify not only the fields to be printed but also their position on a page. In addition, data values may be modified by the user, and spacing and pagination are user controlled. These and other features of the general report writer are discussed in this section.

Figure 10 shows the process by which a user generates a report. Report specification cards are used as input, along with the user's dictionary and data file. The general report writer is similar in function to a compiler, or a language translator. The report specification cards are translated by the general report writer, and the actions described by the report specification cards are performed. Thus, the user has a high-level language available for use in generating reports. Figures 11 and 12 are examples of reports that are generated at NSSDC through the use of the general report writer. These examples illustrate the versatility of the report writer. Both examples use the same data file with two different report specifications, resulting in different output formats.

Some features of the general report writer include the following:

1. Fields to be printed are specified by their field names.
2. Reports may contain report headers, page headers, report footers, and page footers.
3. Values contained in fields may be printed anywhere on a page, as specified by the user.
4. Totals and subtotals of either integer fields or real-number fields may be computed.
5. The number of occurrences of a particular value of a field may be counted.
6. Fields may be truncated in printing.

HISTOGRAM EXAMPLE

REQUEST CATEGORY

ABSCISSA

DATA VALUES

<u>COUNTRY</u>	<u>ROCKET</u>	<u>SATELLITE</u>	<u>BALLOON</u>	<u>OTHER</u>	<u>TOTAL</u>
DATA VALUES — USA	2	2	1	1	6
— USSR	1	1	1	0	3
— CHINA	1	0	2	2	5
— OTHER	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>
TOTAL	4	3	4	3	14

ORDINATE {

Figure 9. Sample of a Histogram Report

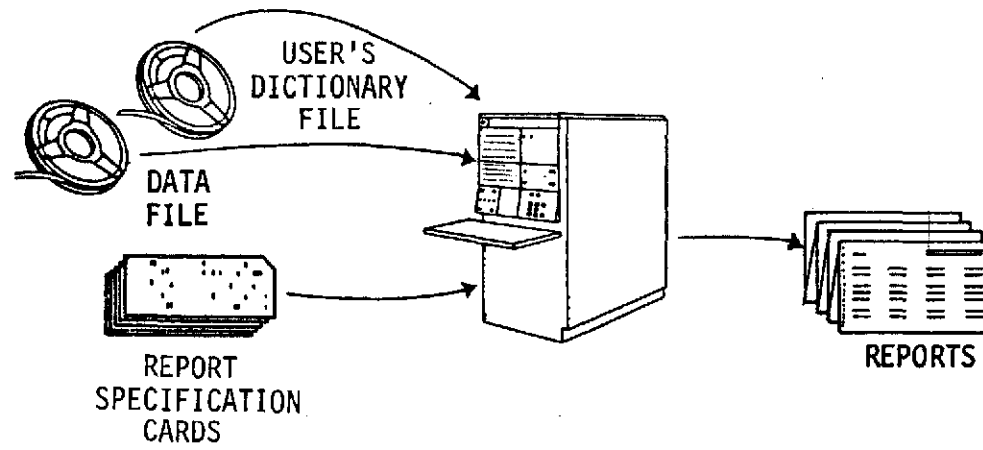


Figure 10. Process for Generating a Report

SPACECRAFT NAME INDEX

SPACECRAFT NAME -----	PAGE ----
IMP 4 (EXPLORER 34 - 67-051A).....	1163
IMP 5 (EXPLORER 41 - 69-053A).....	1289
IMP-A (EXPLORER 18 - 63-046A).....	1036
IMP-B (EXPLORER 21 - 64-060A).....	1046
IMP-C (EXPLORER 28 - 65-042A).....	1058
IMP-D (EXPLORER 33 - 66-058A).....	1112
IMP-E (EXPLORER 35 - 67-070A).....	1176
IMP-F (EXPLORER 34 - 67-051A).....	1163
IMP-G (EXPLORER 41 - 69-053A).....	1289
INJUN 1 (61-0158).....	1009
INJUN 2B (INJUN 3 - 62-067B).....	1030
INJUN 3 (62-067B).....	1030
INJUN 4 (EXPLORER 25 - 64-076B).....	1050
INJUN 5 (EXPLORER 40 - 68-066B).....	1230
INJUN IE-C (EXPLORER 40 - 68-066B).....	1230
INJUN-C (EXPLORER 40 - 68-066B).....	1230
INJUN-SR-3 (INJUN 1 - 61-0158).....	1009
ISIS 1 (69-009A).....	1254
ISIS 2 (71-024A).....	1350
ISIS-A (ISIS 1 - 69-009A).....	1254
ISIS-B (ISIS 2 - 71-024A).....	1350
ISIS-X (ALOUETTE 2 - 65-098A).....	1083
ISIS-X (EXPLORER 31 - 65-098B).....	1088
ITDS 1 (70-008A).....	1322
ITOS-A (NOAA 1 - 70-106A).....	1336
LEM 14 (APOLLO 14 LM - 71-008C).....	1343
LEM 15 (APOLLO 15 LM - 71-063C).....	1377
LEM 16 (APOLLO 16 LM - 72-031C).....	1395
LUNAR MODULE 14 (APOLLO 14 LM - 71-008C).....	1343
LUNAR MODULE 15 (APOLLO 15 LM - 71-063C).....	1377
LUNAR MODULE 16 (APOLLO 16 LM - 72-031C).....	1395
LUNAR ORBITER 1 (66-073A).....	1121
LUNAR ORBITER 2 (66-100A).....	1137
LUNAR ORBITER 3 (67-008A).....	1148
LUNAR ORBITER 4 (67-041A).....	1158
LUNAR ORBITER 5 (67-075A).....	1195
LUNAR ORBITER-A (LUNAR ORBITER 1 - 66-073A).....	1121
LUNAR ORBITER-B (LUNAR ORBITER 2 - 66-100A).....	1137
LUNAR ORBITER-C (LUNAR ORBITER 3 - 67-008A).....	1148
LUNAR ORBITER-D (LUNAR ORBITER 4 - 67-041A).....	1158
LUNAR ORBITER-E (LUNAR ORBITER 5 - 67-075A).....	1195
MARINER-1 (MARINER 9 - 71-051A).....	1354
MARINER 2 (62-041A).....	1020
MARINER 4 (64-077A).....	1051
MARINER 5 (67-060A).....	1172
MARINER 6 (69-014A).....	1258
MARINER 7 (69-030A).....	1266
MARINER 9 (71-051A).....	1354
MARINER MARS 69A (MARINER 6 - 69-014A).....	1258
MARINER MARS 69B (MARINER 7 - 69-030A).....	1266
MARINER MARS 71 (MARINER 9 - 71-051A).....	1354
MARINER R-2 (MARINER 2 - 62-041A).....	1020
MARINER VENUS 67 (MARINER 5 - 67-060A).....	1172

Figure 11. Example of an Index Report Created by the General Report Writer

NATIONAL SPACE SCIENCE DATA CENTER
 DATA CATALOG OF SATELLITE EXPERIMENTS
 SUPPLEMENT NO. 1 TO NSSDC 71-20

SPACECRAFT COMMON NAME- EXPLORER 6
 ALTERNATE NAMES- ABLE 3, 1959 DELTA 1 NSSDC ID 59-004A

LAUNCH DATE- 08/07/59 SPACECRAFT WEIGHT IN ORBIT- 64. KG

FUNDING AGENCY- DOD-USAF
 FUNDING AGENCY- NASA

DATE LAST USABLE SPACECRAFT DATA RECORDED-100659
 SPACECRAFT STATUS OF OPERATION- INOPERABLE

EPOCH DATE- 12/19/59 ORBIT TYPE- GEOCENTRIC ORBIT PERIOD- 754. MIN
 APODAPSIS- 41900.0 KM ALT PERIAPSIS- 237.000 KM ALT INCLINATION- 47.0 DEG

SPACECRAFT BRIEF DESCRIPTION

EXPLORER 6 WAS A SMALL, SPHEROIDAL SATELLITE DESIGNED TO STUDY TRAPPED RADIATION OF VARIOUS ENERGIES, GALACTIC COSMIC RAYS, GEOMAGNETISM, RADIO PROPAGATION IN THE UPPER ATMOSPHERE, AND THE FLUX OF MICROMETEORITES. IT ALSO TESTED A SCANNING DEVICE DESIGNED FOR PHOTOGRAPHING THE EARTH'S CLOUD COVER. THESE STUDIES WERE CARRIED OUT IN INTERPLANETARY SPACE AND WITHIN THE MAGNETOSPHERE. THE SATELLITE WAS LAUNCHED INTO A HIGHLY ELLIPTICAL ORBIT WITH AN INITIAL LOCAL TIME OF APOGEE OF 2100 HR. THE SATELLITE WAS SPIN STABILIZED AT 2.8 RPS, WITH THE DIRECTION OF THE SPIN AXIS HAVING A RIGHT ASCENSION OF 217 DEG AND A DECLINATION OF 23 DEG. FOUR SOLAR CELL PADDLES MOUNTED NEAR ITS EQUATOR RECHARGED THE STORAGE BATTERIES WHILE IN ORBIT. EACH EXPERIMENT EXCEPT THE TELEVISION SCANNER HAD TWO OUTPUTS, DIGITAL AND ANALOG. A UHF TRANSMITTER WAS USED FOR THE DIGITAL TELEMETRY AND THE TV SIGNAL. TWO VHF TRANSMITTERS WERE USED TO TRANSMIT THE ANALOG SIGNAL. THE VHF TRANSMITTERS WERE OPERATED CONTINUOUSLY. THE UHF TRANSMITTER WAS OPERATED FOR ONLY A FEW HOURS EACH DAY. ONLY THREE OF THE SOLAR CELL PADDLES FULLY ERECTED, AND THIS OCCURRED DURING SPIN UP RATHER THAN PRIOR TO SPIN UP AS PLANNED. CONSEQUENTLY, INITIAL OPERATION OF THE PAYLOAD POWER SUPPLY WAS 63 PERCENT NOMINAL, AND THIS DECREASED WITH TIME. THE DECREASED POWER CAUSED A LOWER SIGNAL TO NOISE RATIO AFFECTING MOST OF THE DATA, ESPECIALLY NEAR APOGEE. ONE VHF TRANSMITTER FAILED ON SEPTEMBER 11, 1959, AND THE LAST CONTACT WITH THE PAYLOAD WAS MADE ON OCTOBER 6, 1959, AT WHICH TIME THE SOLAR CELL CHARGING CURRENT HAD FALLEN BELOW THAT REQUIRED TO MAINTAIN THE SATELLITE EQUIPMENT. A TOTAL OF 827 HR OF ANALOG AND 23 HR OF DIGITAL DATA WAS OBTAINED.

EXPERIMENT NAME- SCINTILLATION COUNTER NSSDC ID 59-004A-02

ORIGINAL EXPERIMENT INSTITUTION- SPACE TECHNOLOGY LAB

EXPERIMENT PERSONNEL (PI=PRINCIPAL INVESTIGATOR, OI=OTHER INVESTIGATOR)
 PI - C.P. SONETT NASA-ARC MOFFETT FIELD, CA
 OI - A. ROSEN TRW SYSTEMS GROUP REDONDO BEACH, CA
 OI - T.A. FARLEY U OF CALIFORNIA, LA LOS ANGELES, CA

DATE LAST USABLE EXPERIMENT DATA RECORDED-091059
 EXPERIMENT STATUS OF OPERATION- INOPERABLE

1005

Figure 12. Example of a Data Catalog Created by the General Report Writer

7. Fields may be right or left justified in printing.
8. Free-text data may be printed, with the user specifying the margin size and the maximum number of characters per line.
9. The user may selectively print elements of a vector or rows in a matrix by specifying that only those elements or rows that match certain specified values be printed.
10. Values in a field may be substituted with more descriptive phrases by means of a table-lookup option.
11. Optional printing of fields (depending on data values) may be controlled by using a logical "IF-THEN" procedure.
12. Binary numbers may be converted to binary-coded-decimal (BCD) format for printing.

Specifications for the general report writer are similar to various FORTRAN format specifications; the "X" specification is used to insert blanks into a print line, the Hollerith (H) specification is used to designate literal fields, slashes (/) are used to skip lines, and FORTRAN carriage control characters are used to control line spacing.

To create a report, the user specifies the field names desired and the file level in which the field name is located. Along with field names, the user specifies any title desired and spacing between data values. The report writer provides the user with the capabilities to print the page number, the line number, and the date. The user may access these capabilities by specifying them as field names in the print line.

A capability exists within the report writer that permits continual output or data modification at print time to provide for most of the features previously described. This is accomplished by means of report writer functions. These report writer functions are user specified and modify a data value associated with the field name. This modified data value replaces the original data value. If a second report writer function is specified, the modified data value is used as input to the function. Thus, many report writer functions may be used in conjunction with only one field name. Processing proceeds from left to right and uses the previous modified data value as input to the function currently being processed.

For example, the field CSNAME (spacecraft common name) is defined to be 25 characters long and right adjusted. However, for reporting purposes, the field is left adjusted and truncated to a length of 14 characters. This means that the last 11 characters will be deleted.

An entry for the field CSNAME is as follows:

INJUN 3, 1962 BETA TAU

If truncation is performed before using an adjustment function, CSNAME-TRUNC will produce the following result:

INJUN 3, 19

This result is incomplete because some of the significant data are missing. However, by using the adjustment function first, then truncating that result (CSNAME-ADJUST-TRUNC), the following result is obtained:

INJUN 3, 1962

and the significant data are retained.

The logical "IF-THEN" procedure is one of the most powerful of the report writer functions. The user may specify that a field be printed only if certain conditions are met. These conditions include less than, less than or equal to, equal to, greater than, greater than or equal to, and not equal. The "IF" portion of the function specifies the condition, and the "THEN" portion specifies the action to be taken if the condition is satisfied. Many "IF-THEN" pairs may be associated with a single-field name. Thus, logical relationships may be specified to control the printing of a field name or names.

By using one or a combination of the general report writer features, the user is able to generate the kind of report required. Management reports, data indexes, statistical summaries, and catalogs are types of reports that are frequently created at NSSDC.

EXTRACT AND MERGE

This section describes the extract and merge segments of SIRS. The extract process allows the user to create a subfile from an original data file by selecting records or portions of records from the original file and producing a new file. The merge process is a method of combining two data files into one.

1. Extract - The user may create a new data file from an existing data file by extracting the desired data, category by category, within a record, to build a subfile. This capability allows the user to reduce the number of categories within records of a data file, reorder those categories that remain, reorder all the categories within the records of a data file, combine categories from records in various levels of a multilevel file into level 1 records of a single-level file, and move categories from records in level 1 of a file to another level in another file.

2. Merge - Two data files may be combined to form one data file by using the merge capability of SIRS. There are several options available when merging two tape files. The choice of options depends upon the desired format of the resulting merged data file. Two data files containing no identical record IDs may be merged, and the resulting data file will contain all records from each input data file. These records will be sorted by record ID.

If the two input data files contain some matching record IDs, three options are available to the user. First, only selected portions of those records with identical record IDs will appear on the resulting data file. The user specifies the category numbers desired from each original data file and the category numbers to be put on the resulting data file. Second, the user may request that portions of records on either one of the original data files be transferred to the resulting data file. As with the first option, the user must specify both category numbers for the original data file and category numbers for the output data file. Third, portions of all records on both original data files may be transferred to the output data file if the user so specifies. Again, both original and new output category numbers must be specified.

DERIVED ITEMS

Occasionally, a SIRS user needs special capabilities that are not routinely provided by SIRS. In such cases, SIRS allows the installation of special routines, called derived item routines. As the term "derived item" denotes, the user may perform an operation on several fields within a record to derive a temporary field, which may be used for sorting, selecting, or reporting. This section discusses the two options available to the user: preprogrammed derived item routines and a generalized derive capability.

There are five preprogrammed derived item routines. Each routine performs a specific task, with the user specifying only the fields to be used with a particular routine. This is done by entering the specific derived item routine name desired, along with any other required information, into the dictionary descriptors that are reserved for this purpose. When using the generalized derive capability, the user specifies not only the field names to be used in a derivation but also the type of operation desired. The options available under generalized derive capability are discussed later in this section.

1. Preprogrammed Derived Items - Five preprogrammed derived item routines are currently available to a user of SIRS. Each routine allows changes to data values at the time of a SIRS computer

run. One of these routines is used to search and retrieve the affiliation name portion of an address field, depending on the content of another data field. Another routine allows the inversion of a name field stored in a data file in the order "first name, last name," to the order "last name, first name." This routine is useful when a user wishes to sort a file based on last names.

A third derived item routine performs several functions. Vectors may be searched for minimum or maximum data values. If the vector being searched is a column vector of a matrix, this routine may be used to obtain data values of other column vector fields that are in the same matrix row as the data value retrieved by the minimum or maximum option. Also, data values within a vector may be totaled using this routine. Another option allows the user to retrieve one specified data field value in a row of a column vector (within a matrix) whenever another specified data field value in the same row of another column vector (of the same matrix) matches a given value provided by the user.

A fourth SIRS derived item routine is used after a data file is sorted based on the data values in a column vector of a matrix. After the sort is performed, the user may print data values contained in another column vector within the same matrix. The fifth derived item routine available under SIRS allows the user to convert data values within a column vector to a common unit, hertz. This modification is done to perform a sort based on the values in the column vector. Another option within this routine allows the user to convert data values expressed in hertz to different types of units.

2. Generalized Derive Capability - The SIRS user has the capability of using a generalized derive item routine in addition to the specialized preprogrammed routines described previously. Several functions may be performed by using the generalized derive item routine. Two fields may be added, subtracted, divided, multiplied, or concatenated (linked together) to form a temporary third field that may be used for sorting, selecting, reporting, or as an intermediate value to be used in another step in the same derived item computation. Exponentiation is also an available option when using the generalized derive item routine.

A conditional operator may be used in conjunction with the functional operators described previously. This conditional operator allows the user to specify search criteria that a

data value must satisfy before a specified derivation is performed. For example, a column vector within a matrix may contain many different data values, all of which represent photon energies. However, these data values may be expressed in one of three units: hertz, electron volts, or centimeters. For comparison purposes the user may wish to express all data values in hertz. The conditional operator allows the user to specify the following conditions: If units are already hertz, make no change to the original value; if units are electron volts, divide the data value by 4.136×10^{-15} , giving a new data value in hertz; if the units are centimeters, divide 3×10^{10} by the data value, giving a new data value in hertz. All data values in the vector will now be expressed in a common unit, and a meaningful comparison may be made. In addition the relational expressions, less than or equal to (.LE.), greater than or equal to (.GE.), equal to (.EQ.), and not equal to (.NE.), may be used as conditions for the computations. All input to the generalized derive item routine is user specified.

APPENDIX A
BRIEF DESCRIPTIONS OF
NSSDC FILES MAINTAINED BY SIRS

The following section contains a brief description of each of the major files currently maintained by SIRS.

AUTOMATED INTERNAL MANAGEMENT (AIM) FILE

The AIM File provides descriptive information to characterize each spacecraft, experiment, and data set that NSSDC has identified. The file is primarily designed to support the acquisition, processing, and announcement of data, and the searches required for the selective retrieval of information concerning space science experiments and data.

The information in the AIM File is contained in discrete records that are organized into a three-level hierarchical structure. The first-level records contain information that relate to a specified spacecraft planned to be launched or launched. Fields contained in level 1 include spacecraft name(s), names and addresses of spacecraft project officials, spacecraft launch date and time, spacecraft launch vehicle and launch site, operational spacecraft status and performance information, and descriptions of the spacecraft mission objectives.

The second-level records contain information that describes experiments carried aboard these spacecraft. Any number of second-level records can be associated with each spacecraft record. Fields contained in level 2 include experiment name(s), names and addresses of experiment personnel, the experiment operational status and performance information, brief descriptions of the experiment objectives, and encoded data to categorize the phenomenon observed by the experiment and the related scientific disciplines.

The third level of the AIM File contains records presenting information related to a set of data originating from a particular experiment on a particular spacecraft. Any number of level 3 records can be associated with an experiment. Fields contained in level 3 include data set name(s), names and addresses of data set personnel, a brief description of the data set, the form and quantity of data in the data set, data set availability and processing status information, and the overall time period of coverage for the complete data set.

Some AIM fields, such as NSSDC acquisition agent ID codes, acquisition group codes, and priority codes are contained in all three levels of the file.

The AIM File is used to produce a variety of useful management reports pertinent to NSSDC data acquisition activity, including actual and predicted acquisition rates, acquisition workload, and acquisition status information. The AIM File is also used to produce indexes and to perform logical searches that aid in responding to specific user requests for data and information. AIM is also used to produce camera-ready copy for several publications, including the Data Catalog of Satellite Experiments, the Report on Active and Planned Spacecraft and Experiments, and the OSS and OA quarterly report.

TECHNICAL REFERENCE FILE (TRF)

The Technical Reference File contains information that accounts for and characterizes all technical documents of a reference nature required to support NSSDC and the World Data Center A (WDC-A) for Rockets and Satellites. The main functions of the file are to serve as a locator/retrieval aid for all relevant technical reference material accessible at or through NSSDC, and to serve as a bibliographic tool in space science and related disciplines for NSSDC users.

The information fields in the TRF are contained in discrete records that are organized into a two-level hierarchical structure. Level 1 records contain information fields such as author(s), title, and key-words for each unique article or document entered in the file. Level 2 records for each document contain information fields pertaining to the bibliographic description(s), its form, location, availability, and related items. There may be one or several level 2 records for each level 1 record.

Reports produced from the TRF data base are used to provide input to the OSS and OA quarterly report, to supply NSSDC data users with bibliographies pertinent to the experiment data stored at NSSDC, and to furnish accounting and retrieval information concerning the physical location and availability of a particular document.

REQUEST ACTIVITY and NAME DIRECTORY (RAND) FILE

The RAND File consists of two major types of data. The first type, stored in level 1 of the file, contains the names and addresses of all rocket and spacecraft experimenters, of people who are receiving specific types of documents from NSSDC and WDC-A on a regular distribution basis, and of people who have requested materials from the Data Center. The second type of data, stored in level 2, contains bookkeeping information on all requests for information or data. Information fields contained in level 2 include the requester's current organizational affiliation, date the request is received, encoded data to identify the person and group(s) at NSSDC responsible for satisfying the request, request processing status, cost of fulfilling the request and amount charged to requester, identification of products supplied to the requester, processing services performed, and remarks identifying problems associated with request processing.

Reports produced from RAND are used to generate mailing labels for document distributions, provide current and complete addresses of NSSDC users, provide up-to-date information on the status of requests being processed by NSSDC, provide data on costs associated with the request activity at NSSDC, generate action reminders for request agents when

estimated completion dates are past, aid in the construction of a flexible model of NSSDC by supplying information describing the NSSDC user community, the types of requests received, the responses required, and the data sets most likely to be requested.

ROCKET FILE

The Rocket File is a one-level file containing information that names and characterizes all scientific sounding rocket launchings identified by NSSDC.

Information fields represented in the file include rocket identification number, rocket type, launch date and time, launch site, rocket and experiment performance data, name and current organizational affiliation of project and experiment officials, and codes describing the scientific disciplines and experiment instrumentation pertinent to a particular launch. The main use of the file is to produce the monthly Sounding Rocket Launching Report and the rocket portion of the World Data Center A for Rockets and Satellites Catalogue of Data.

DATA SET INVENTORY FILE (DSIF)

The DSIF is actually a number of SIRS subfiles used to account for and locate each physical unit of data, such as a reel of microfilm, a reel of magnetic tape, or a photograph, all stored at NSSDC. The DSIF contains basic information to identify and retrieve the physical data unit. Files that are presently, or will soon be, a part of the overall DSIF include the ISIS and Alouette Ionogram Inventory Files, the Magnetic Tape Inventory File, the Nonsatellite Data Inventory File, and the Mariner, Apollo, and Lunar Orbiter Photographic Inventory Files. Standard fields contained in each of the DSIF subfiles include an NSSDC data set ID code, form and location of the data, date the data are received, start and completion dates of data coverage for the data unit, and additional comments and remarks. The DSIF also contains other supportive information useful to data requesters in selecting and utilizing any data available at NSSDC.

NONSATELLITE DATA FILE (NSDF)

This file provides basic information describing the nonsatellite data available at NSSDC. These data consist of program decks, model environment data, rocket data, aircraft data, balloon data, and ground-based correlative data. The NSDF is a three-level file; the first level contains fields to describe the type of data and the method of accumulating the data, the second level contains fields to describe the overall content of the data, and the third level contains fields describing the specific data sets available at NSSDC.

APPENDIX B

A SAMPLE SIRS FILE

The following section illustrates a simple SIRS file that was extracted from the AIM File. Note the structure of the file: level 1 records describe spacecraft, level 2 records describe experiments on these spacecraft, and level 3 records describe the data sets related to the experiments.

Following is a list of the fields extracted from the AIM File for use in this sample:

<u>Level 1 Field Name</u>	<u>Description</u>
ID1AIM	Spacecraft Record ID Code
SINTUP	Date the Spacecraft Record is Entered
SLSTUP	Date the Spacecraft Record is Last Updated
SPNAME	Matrix Containing the Spacecraft Name Code and Common or Alternate Names of the Spacecraft
SCODEN	Spacecraft Name Code - Column Vector Within SPNAME
SNAME	Spacecraft Name - Column Vector Within SPNAME
CSNAME	Spacecraft Common Name
LNDAT	Spacecraft Launch Date
SAGENT	Spacecraft Acquisition Agent ID Code
SBRIEF	Brief Description of the Spacecraft

<u>Level 2 Field Name</u>	<u>Description</u>
ID2AIM	Experiment Record ID Code
EINTUP	Date the Experiment Record is Entered
ELSTUP	Date the Experiment Record is Last Updated
EXNAME	Matrix Containing the Source of the Name Given to an Experiment and the Actual Names Given to the Experiment
ECODEN	Source of the Experiment Name - Column Vector Within EXNAME
ENAME1	Experiment Name - Column Vector Within EXNAME
ENAME2	Experiment Name - Column Vector Within EXNAME
ECONTC	Name of the Experimenter
EAGENT	Experiment Acquisition Agent ID Code
EBRIEF	Brief Description of the Experiment Conducted On Board the Spacecraft

Level 3 Field NameDescription

ID3AIM	Data Set Record ID Code
DINTUP	Date the Data Set Record is Entered
DLSTUP	Date the Data Set Record is Last Updated
DSNAME	Matrix Containing the Source of the Data Set Name and the Actual Names Given to the Data Set
DCODEN	Source of the Data Set Name - Column Vector Within DSNAME
DNAME1	Data Set Name - Column Vector Within DSNAME
DNAME2	Data Set Name - Column Vector Within DSNAME
DAGENT	Data Set Acquisition Agent ID Code
DBRIEF	Brief Description of the Data Set

The initial step is to define fields of the file, thereby creating a dictionary.

The following represents a sample of level 1 dictionary entries taken from the AIM File.

<u>FIELD NAME</u>	<u>FIELD TYPE</u>	<u>DATA TYPE</u>	<u>FIELD SIZE</u>	<u>REPEATABLE</u>	<u>CATEGORY</u>	<u>LEVEL</u>
IDIAIM	SINGLE VALUE	ALPHANUMERIC	7	1	1	1
SINTUP	SINGLE VALUE	DATE	6	1	2	1
SLSTUP	SINGLE VALUE	DATE	6	1	2	1
SPNAME	MATRIX	ALPHANUMERIC	30	6	3	1
SCODEN	COLUMN VECTOR	ALPHABETIC	2	6	3	1
SNAME	COLUMN VECTOR	ALPHANUMERIC	24	6	3	1
CSNAME	SINGLE VALUE	ALPHANUMERIC	25	1	4	1
LNDAT	SINGLE VALUE	ALPHANUMERIC	8	1	4	1
SAGENT	SINGLE VALUE	ALPHABETIC	3	1	5	1
SBRIEF	SINGLE VALUE	FREE TEXT	-	1	6	1

- . Fields are combined to form categories.
- . Categories are combined to form a record.
- . Category 1 always contains the record ID.
- . Category 2 always contains the date the record is first written into the file and the date the record is last updated.

The following examples, taken from the AIM File, represent two level 1 records, 62-067B and 64-054A.

CAT1	62-067B	ID1AIM	64-054A	CAT1
CAT2	052472	SINTUP	052472	CAT2
	062172	SLSTUP	090172	
CAT3	CN INJUN 3	SPNAME	CN OGO 1	CAT3
	AN 1962 BETA TAU 2	SCODEN, SNAME	AN EOGO 1	
CAT4	INJUN 3 1962 BETA TAU 2	CSNAME	OGO 1, EOGO 1, S-49, OGO A	CAT4
	12/13/62	LNDAT	09/05/64	
CAT5	JJB	SAGENT	JJB	CAT5
CAT6	INJUN 3 WAS A MAGNETIC-FIELD ALIGNED SPACECRAFT INSTRUMENTED FOR A STUDY OF GEOPHYSICAL PHENOMENA ...	SBRIEF	THE PURPOSE OF THE OGO 1 SPACECRAFT WAS TO CONDUCT MANY DIVERSIFIED GEOPHYSICAL EXPERIMENTS ...	CAT6

The following represents a sample of level 2 dictionary entries taken from the AIM File.

<u>FIELD NAME</u>	<u>FIELD TYPE</u>	<u>DATA TYPE</u>	<u>FIELD SIZE</u>	<u>REPEATABLE</u>	<u>CATEGORY</u>	<u>LEVEL</u>
ID2AIM	SINGLE VALUE	ALPHANUMERIC	10	1	1	2
EINTUP	SINGLE VALUE	DATE	6	1	2	2
ELSTUP	SINGLE VALUE	DATE	6	1	2	2
EXNAME	MATRIX	ALPHANUMERIC	90	5	3	2
ECODEN	COLUMN VECTOR	ALPHABETIC	7	5	3	2
ENAME1	COLUMN VECTOR	ALPHANUMERIC	40	5	3	2
ENAME2	COLUMN VECTOR	ALPHANUMERIC	41	5	3	2
ECONTC	SINGLE VALUE	ALPHABETIC	9	1	4	2
EAGENT	SINGLE VALUE	ALPHABETIC	3	1	5	2
EBRIEF	SINGLE VALUE	FREE TEXT	-	1	6	2

The following examples, taken from the AIM File, represent two level 2 records associated with the level 1 record, 62-067B.

CAT1	62-067B-03	ID2AIM	62-067B-04	CAT1
CAT2	052472	EINTUP	052472	CAT2
	060172	ELSTUP	053073	
CAT3	NSSDC	EXNAME	NSSDC	CAT3
	MAGNETIC DIFFERENTIAL ELECTRON SPECTROMETER	ECODEN	INTEGRAL MAGNETIC ELECTRON SPECTROMETER	
CAT4	LAUGHLIN	ENAME1		CAT4
		ENAME2		
CAT5	JJB	ECONTC	LAUGHLIN	CAT5
		EAGENT	JJB	
CAT6	A MAGNETIC DIFFERENTIAL SPEC- TROMETER WAS USED TO DETECT LOCALLY MIRRORING ELECTRONS IN THE ENERGY RANGES 40-60 KEV AND 80-110 KEV ...	EBRIEF	AN INTEGRAL MAGNETIC SPEC- TROMETER WAS TO BE USED TO STUDY MIRRORING HIGH-ENERGY FISSION ELECTRONS (E .GT. 1.5 MEV) ...	CAT6

The following represents a sample of level 3 dictionary entries taken from the AIM File.

<u>FIELD NAME</u>	<u>FIELD TYPE</u>	<u>DATA TYPE</u>	<u>FIELD SIZE</u>	<u>REPEATABLE</u>	<u>CATEGORY</u>	<u>LEVEL</u>
ID3AIM	SINGLE VALUE	ALPHANUMERIC	11	1	1	3
DINTUP	SINGLE VALUE	DATE	6	1	2	3
DLSTUP	SINGLE VALUE	DATE	6	1	2	3
DSNAME	MATRIX	ALPHANUMERIC	90	4	3	3
DCODEN	COLUMN VECTOR	ALPHABETIC	5	4	3	3
DNAME1	COLUMN VECTOR	ALPHANUMERIC	41	4	3	3
DNAME2	COLUMN VECTOR	ALPHANUMERIC	42	4	3	3
DAGENT	SINGLE VALUE	ALPHABETIC	3	1	5	3
DBRIEF	SINGLE VALUE	FREE TEXT	-	1	6	3

The following examples, taken from the AIM File, represent two level 3 records associated with the level 2 record, 62-067B-03, and with the level 1 record, 62-067B.

CAT1	62-067B-03A	ID3AIM	62-067B-03B	CAT1
CAT2	052472	DINTUP	052472	CAT2
	070672	DLSTUP	100673	
CAT3	NSSDC MASTER FILE ON MAGNETIC TAPE, ELECTRON SPECTROMETER COUNTS	DSNAME DCODEN DNAME1 DNAME2	NSSDC ANALYZED MAGNETIC DIFFER- ENTIAL ELECTRON ELECTROMETER FLUX PLOTS ON MICROFILM	CAT3
CAT5	JJB	DAGENT	JJB	CAT5
CAT6	THE DATA SET CONSISTS OF A TIME-ORDERED MASTER FILE FOR INJUN 3 OF REDUCED DATA ON FIVE 7-TRACK, IBM 7094, BCD MAGNETIC TAPES ...	DBRIEF	THIS ANALYZED DATA SET CONSISTS OF MACHINE- GENERATED PARTICLE FLUX PLOTS ON ONE 16-MM REEL OF MICROFILM ...	CAT6