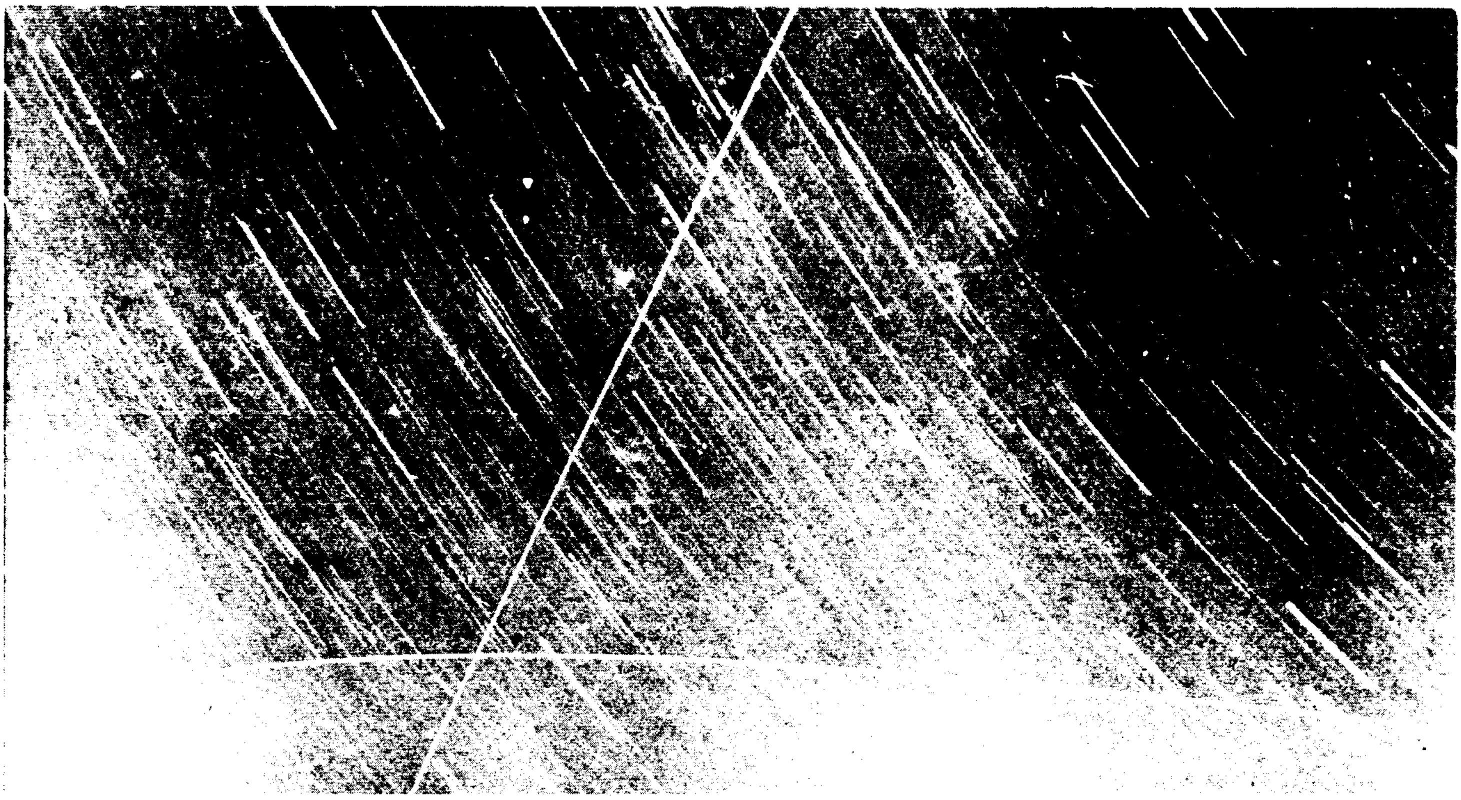

FLOW AND USE OF INFORMATION AT THE NATIONAL SPACE SCIENCE DATA CENTER

NSSDC 69-02

JANUARY 1969

NATIONAL SPACE SCIENCE DATA CENTER

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

FLOW AND USE OF INFORMATION AT THE NATIONAL SPACE SCIENCE DATA CENTER

NSSDC 69-02

Nick Karlow
James I. Vette

JANUARY 1969

PRECEDING PAGE BLANK NOT FILMED.

FLOW AND USE OF INFORMATION AT THE NATIONAL SPACE SCIENCE DATA CENTER

Nick Karlow
James I. Vette
National Space Science Data Center

ABSTRACT

An integrative, overall view of the flow and use of information at the National Space Science Data Center (NSSDC) is presented by examining the tasks facing the NSSDC, present resources, and the flow of information. The information system used to handle this flow is then discussed in terms of its four main subsystems: the Automated Internal Management file, the Machine-Oriented Data System, the Technical Reference File, and the Request Accounting Status and History file. Also examined are five special purpose files: Computer Program Status Report, Magnetic Tape Unit Record, NSSDC Computer Program File, Rocket File, and Distribution File. This is followed by a look into the future of information systems development at NSSDC.

PRECEDING PAGE BLANK NOT FILMED

CONTENTS

LIST OF FIGURES

Figure		Page

FLOW AND USE OF INFORMATION AT THE NATIONAL SPACE SCIENCE DATA CENTER

INTRODUCTION

It is apparent that satellite measurements produce a tremendous amount of data to be processed and analyzed. For example, during 1967 there were over one trillion bits of data transmitted to the earth. This huge volume of data is further complicated by the wide diversity of experiments. To illustrate, scientific experiments measure those quantities which will lead to a general understanding of the natural phenomena; engineering measurements advance techniques and procedures which lead to more sophisticated and reliable spacecraft systems; applications satellites are concerned with communications, navigation, earth resources and weather; biomedical experiments assist the manned flight program.

Many of the categories described above could well fit into already existing data activities or centers. However, there are some differences. The useful lifetime of engineering data is generally much shorter than that for scientific data, since rapid strides in technology soon produce new devices with different and more desirable properties. Likewise, operational data is used in near-real time for most purposes. Consequently, the dissemination of this information must occur rapidly. The nature of the scientific measurements, on the other hand, requires that a much longer active life be provided for the data. The processes being observed are not completely understood, and they interact with each other in a vast complexity of ways. Consequently, the results vary significantly with time and location. Large volumes of data are necessary to obtain patterns, interactions, and relations. As new ideas develop in understanding the phenomena, scientists will want to analyze further much of the existing data. This, in preference to repeating a measurement, is necessitated in space science because it now takes approximately 5 years between experiment conception and receipt of data for analysis.

For some of the earth and environmental sciences, satellite experiments represented only a new technique of measurement, not the real beginning of new scientific disciplines. Thus, most of the data concerning geodesy, meteorology, weather, oceanography, and the ionosphere could be handled by the existing data centers for these disciplines. However, those disciplines which comprise space science were really developed as a result of satellite and rocket experiments. It was clear a new data center needed to be established to handle this function. As a result, the National Space Science Data Center (NSSDC) was founded in 1965 with the primary function of providing the means for the dissemination and analysis of space science data beyond that provided by the original experimenter. As such, the Data

Center is responsible for the active collection, organization, storage, announcement, retrieval, dissemination, and exchange of space science data.

NSSDC has grown considerably in the last 2 years. There is a well-established flow of data and information into and out of the Center. In order to effect this transfer of knowledge and keep track of the large number of transactions which occur, an information system has been developed. It is the purpose of this paper to present an overall view of the Data Center as a system, trace the flow of information, and to discuss the changes envisioned for the future. Hopefully, it will prove useful to Data Center users, to people interested in information transfer, and to management personnel who require an understanding of the detailed processes involved in setting up, operating, and developing an effective data center. In addition, the various Data Center subsystems are tied together for the people working at NSSDC to show them how they contribute to the overall system.

THE TASK OF NSSDC

The Data Center is concerned with serving the following user groups:

- Space scientists
- Scientists in related fields
- Engineers and system planners
- Management
- Educational activities

In order to provide effective service to these different segments of society, a number of capabilities must exist at the Data Center.

To insure that the experimental data can be analyzed independently in the future by other scientists, it is necessary to collect and preserve this data in the proper form. Raw data transmitted directly from the satellite is not the appropriate form because it is impossible to acquire all the supporting information that is necessary for independent use. Reduced data records seem to be the basic ones which NSSDC should acquire. A discussion of the processes included in going from raw to analyzed data is given in reference 1. However, a definition of reduced data will be given at this time to insure an understanding of terms.

> *Reduced Data Records* – Data records prepared from raw data records by a compacting, editing, correcting, and merging operation performed under the supervision of the principal investigator. Data in this form contain all the basic usable information obtained from the experiment and generally include the instrument responses measured as functions of time along with appropriate position, attitude, and equipment performance information necessary to

analyze the data in an independent fashion. The engineering corrections such as temperature, voltage, dead time, gain changes, and other similar corrections to the instrument response will have been made. Unusable noisy data and periods of questionable instrument performance will have been removed as well as duplicate portions of information. Time averaging and the conversion of the instrument response to physical units will not have been accomplished in most cases. Visual data, such as photographs derived from data processing techniques, may also be considered as reduced records.

It should be apparent that the acquisition of reduced data and the supporting information requires people who are familiar with the instruments, calibration techniques, data processing, and the interpretations of the data. This supporting information includes descriptions of the instrument, specific mechanical and electrical details of the experiment, calibration procedures and results, and unusual operating conditions. Thus, professional people who have specialized in the various space science disciplines are necessary to the successful operation of the Data Center.

For some scientific uses of the data it is not necessary to re-analyze or re-evaluate. In most cases the interpretations given by the original investigator are the most valid and respected ones. These analyzed results are often incorporated with those of other experiments in order to gain new understanding of the various phenomena. For this reason, the Data Center must also collect analyzed data records which are defined as follows.

Analyzed Data Records – Data records prepared from reduced data by the principal investigator, his co-workers, and other space scientists which display the scientific results of the experiment. In general, the physical quantities derived from the sensor responses are displayed in various appropriate coordinate systems and correlated with other geophysical measurements. The results may be time averaged over meaningful intervals, displayed in the form of parameters of specific physical models or theories or as best-fit parameters of empirical descriptions. This form may include charts, graphs, photographs, and tables which are the results of data processing and analysis techniques employed by the analyzing scientist. Examples of these appear in his published works, but the total number are usually too large to be published in their entirety.

To provide for the engineer or system planner, a new product, the space environment, must be developed from the raw materials which flow into the Data Center. Most knowledge of the space environment comes directly from the scientific measurements. However, only a small segment comes from each individual experiment. It is necessary to study the analyzed and/or reduced data from many experiments in order to obtain a fairly comprehensive description of the space environment. This translation or synthesis into useful data summaries, compilations, or environments is a natural professional activity for those space scientists associated with the Data Center. It represents a type of analysis not done very often by the original investigators and contributes a useful product for dissemination to all user groups, including space scientists. At the present time, only the charged particle trapped radiation environment is being generated by NSSDC. However, this activity will be expanded to include the other portions of the space environment as the professional staff grows.

Finally, it is necessary to include in the information system those facts about space research which pertain to funding, on-orbit performance, publication of research, instrument characteristics, satellite orbits, and transmitter frequencies. Retrieval of this information in a variety of ways serves the management community as well as the previously mentioned users. When all of the capabilities discussed above have been incorporated, then the general public, and the educational activities in particular, can also be adequately served by NSSDC. That this Data Center can play an important supporting role in the university graduate program is pointed out by Dessler (ref. 2). He indicates that the cost of a space-hardware program could perhaps be reduced from about $400,000 per Ph.D. to about $100,000 by carrying out one or two space experiments on campus while making use of data stored at NSSDC. To be effective in this role, NSSDC must strive to get the data deposited as early as possible and yet not compromise the rights of the original investigator.

To satisfy the needs of the various user groups, data must be obtained, processed, analyzed, entered into a system, stored, retrieved, and disseminated. Supporting information including reports and publications must also be entered. The whole task requires specialists in the various space science disciplines, systems analysis, computer programming, data processing, technical writing, publication, and reproduction.

The complexity of the job to be done has required the adoption of a total systems approach and the automation of the Data Center. A second reason for automation is the large volume of data which requires the handling of numerous magnetic tapes, cards, pictures, microfilm, and copies of written, graphical, and tabular materials. A few examples may be appropriate at this time to understand the problem more fully.

Within the next few years, the Data Center expects to receive annually about 5-10,000 magnetic tapes, 150,000 linear feet of roll charts, and 2,000 100-ft rolls of microfilm. For comparison purposes, the data base of a year ago and the present one are shown in figure 1. More indicative of the anticipated volume of data, however, is the steady growth of incoming magnetic tapes, as shown in figure 2. In this context, because tape damage or loss of data can occur during usage, it is necessary to maintain a spare copy. Translating this into reality, a digital data base of 30,000 tapes would imply storage facilities for 60,000 tapes, not to mention the time and effort required in the duplicating process. Figure 3 emphasizes the anticipated diversity of incoming data in still another way by showing the number of experiments for which some data is present in NSSDC and the total number of space science experiments successfully flown. This is even more significant if one considers the fact that one experiment may produce several sets of data organized in different ways. It is felt that one acquisition agent can handle between 20-25 experiments of a similar nature and can process completely into the system about 5-10 per year. Finally, the NSSDC request history, as portrayed in figure 4, shows the trend of use of the Data Center by all user groups.

VOLUME OF DATA AT NSSDC

MEDIUM	AUGUST 1967	SEPTEMBER 1968
SHEETS AND BOUND VOLUMES, SHEETS	175,000	215,000
DIGITAL MAGNETIC TAPES, ½" × 2400'	291	2,000
MICROFILM,100-FT ROLLS	7,800	10,621
PHOTOGRAPHIC FILMS:		
9½" WIDTH, LINEAR FEET	14,000	18,000
70-mm WIDTH, LINEAR FEET	33,200	143,500
35-mm WIDTH, LINEAR FEET	0	81,000
4×5 INCH, EACH	2,100	2,400
8×10 INCH, EACH	0	400
16×20 INCH, EACH	93	93
20×24 INCH, EACH	2,200	5,100
PHOTOGRAPHIC PRINTS:		
9½" WIDTH, LINEAR FEET	0	9,000
70-mm WIDTH, LINEAR FEET	0	6,000
8×10 INCH	600	2,500
11×14 INCH	200	500
16×20 INCH	93	93
20×24 INCH	2,200	3,700

Figure 1—Growth of the Data Base at NSSDC

PRESENT RESOURCES

Before examining the system, it may be enlightening to look at the present staffing and facilities used to fulfill the NSSDC mission. A list of equipment will be given in the appendix.

The Data Center is operated by 60 contractor personnel and seven Civil Service employees. There also are two Air Force officers from the Environmental Technical Applications Center (ETAC) assigned to NSSDC as acquisition agents. To gain a realistic insight into the scope and type of effort involved, figure 5 presents the main functions, the type of individuals involved, and the number of actual people used in each function. This staff is only adequate to process into the NSSDC system about 60-70 experiments per year. Consequently, present plans call for a buildup to approximately 110 people within the next two years. Furthermore, if NSSDC is to keep up with the vast amount of data that will be coming in and the advances of modern

NSSDC TAPE ACCUMULATION
(DATA + DUPLICATE)
SEPT. 30, 1968

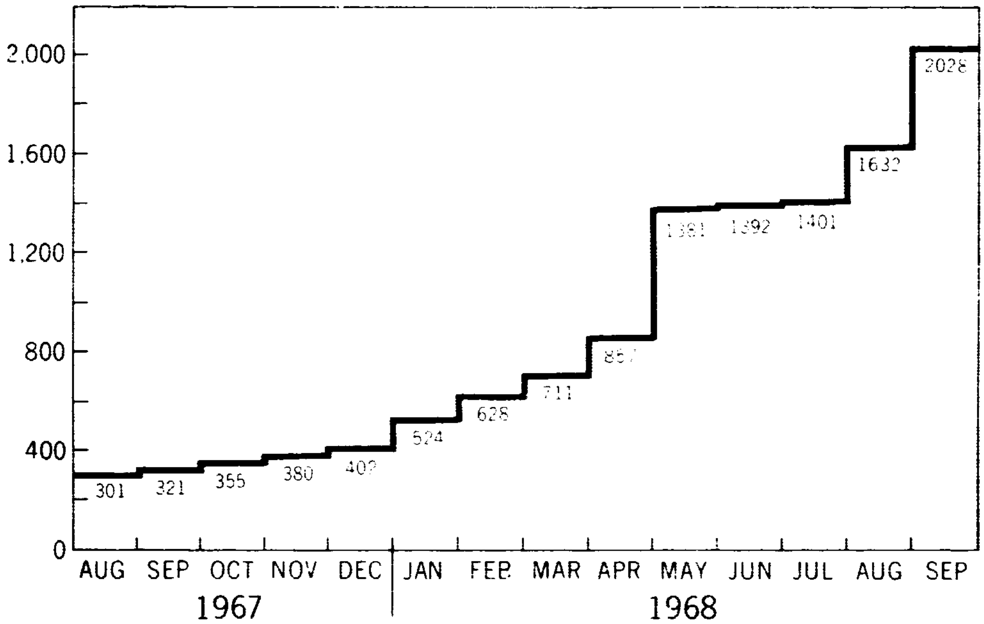

Figure 2—Growth of Incoming Tapes

SPACE SCIENCE EXPERIMENTS

DISCIPLINES	SUCCESSFULLY FLOWN EXPERIMENTS *	SOME DATA AT NSSDC**
IONOSPHERES & RADIO PHYSICS	86	12
PLANETARY ATMOSPHERES	96	20
PARTICLES & FIELDS	396	54
SOLAR PHYSICS	28	11
ASTRONOMY	35	3
PLANETOLOGY (INC. SELENOLOGY)	127	45
TOTAL	768	145

* AS OF MAY 17, 1968
** AS OF SEPTEMBER 17, 1968

Figure 3—Data on Hand vs. Successful Experiments

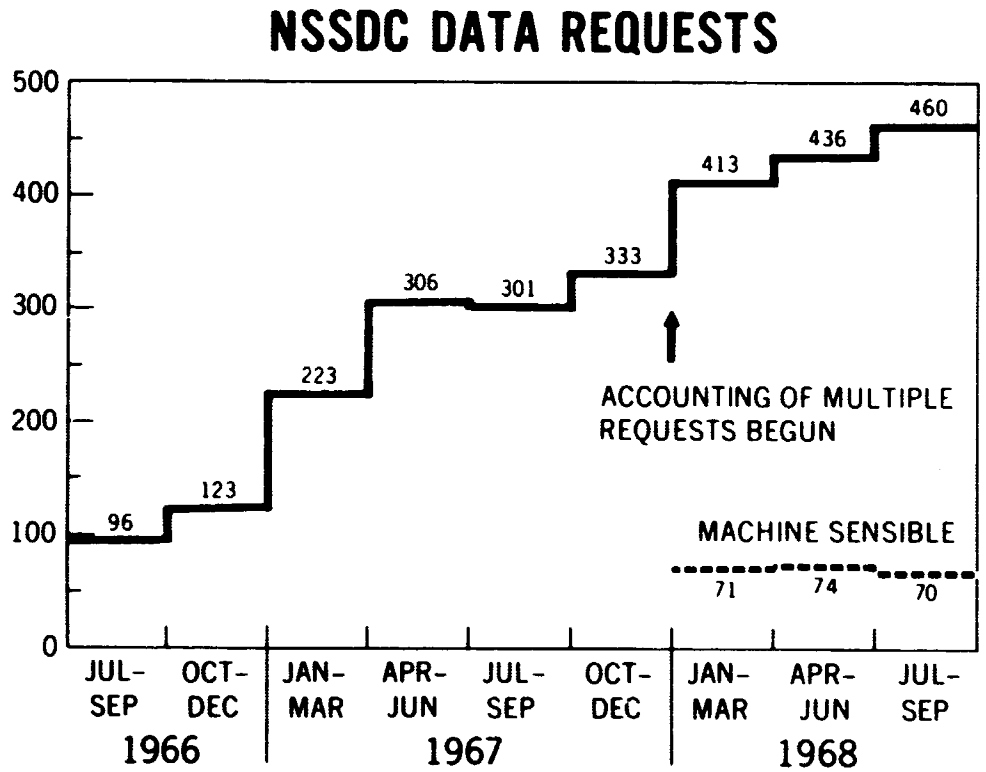

Figure 4—Growth of Requests. As of January 1968, requests requiring machine processing were identified apart from other requests, and single requests requiring different forms of data were treated as multiple requests.

systems technology, it has to employ the third generation of computer hardware and software effectively. If growth plans cannot be fulfilled, the most reasonable alternative is to reduce the amount of incoming data by being highly selective in the acquisition of experimental results. Thus, only the most significant data would be preserved for further use. Committees would then have to be established to determine which would be preserved by NSSDC.

As far as facilities are concerned, the Data Center is housed in a modern 3-story building (No. 26) at GSFC. In addition to housing personnel, space has been allocated for equipment, including photo laboratory, for visiting scientists who wish to perform research using stored data, for classrooms and lecture rooms, and for storage of microfilm and data tapes. The usable space can be roughly broken down as follows:

Office Space		9,650 sq ft
Work Areas		10,555 sq ft
Special-Purpose Areas		7,515 sq ft
	Total	27,720 sq ft .

NSSDC SCOPE OF EFFORT

Function	Types of People	No. of People
Management	R&D Administration	3
Systems	Planning & Development Analyst Operations Maintenance Technician	2 4 2 1 1
Data Acquisition	Space Scientist	9
Programming	Programmer Keypunch	8 1
Computer Services	Data Technician	7
Publications	Technical Information Writer/Editor Technician MTST* Operator	1 4 1 1
Storage & Reproduction	Technician (requests) Technician (data) Technician (photo) Technician (microfilm)	2 5 5 1
Research & Analysis	Space Scientist Technician	3 1
Administration	Administrator Clerical	2 5
	TOTAL	69
*Magnetic Tape Selectric Typewriter		

Figure 5—Type and Number of Personnel

The Data Center is presently occupying about two-thirds of the available office space (other GSFC units are using the remaining offices). The work areas include such items as data reference room, indexing and classification room, mailroom, duplication room, and computer area (3,450 sq ft). Some of the special-purpose areas are:

Photo Lab	1,250 sq ft
Tape Storage	3,000 sq ft

Microfilm Processing	1,300 sq ft
Document Storage	320 sq ft
Auditorium & Classrooms	2,660 sq ft

The volume of space allocated for tapes will hold about 25-35,000 reels. This represents the maximum size to which the active data base can grow with the present facilities.

As the staff expands, the space being used by others will have to become available. With this in mind, the building is considered adequate for the next several years.

FLOW OF INFORMATION INTO THE SYSTEM

It was within this framework that the Data Center planned and developed its current integrated information system (see figure 6). To oversimplify the mission of NSSDC, one must first arrange for obtaining the space science data and understand the form/format of incoming data. Once the data begin to arrive, there must be a central source of information concerning these data. This need is satisfied by the subsystem called Automated Internal Management (AIM). Upon arrival, one must process the machine-sensible data, prepare it for retrieval, and be able to handle special types of data in different forms and formats – this is done through the Machine-Oriented Data System (MODS). (The steps for processing non-machine-sensible data are generally analogous.) In their work, the professionals at NSSDC must have ready access to the documentation relating to appropriate satellites, experiments, and data – the Technical Reference File (TRF) serves this purpose. Finally, statistics must be kept on the processing and use of data, and management must have a variety of reports in this area – this is greatly facilitated by the use of the Request Accounting Status and History (RASH) file. These subsystems are supported by five additional special-purpose files: Computer Program Status Report, Magnetic Tape Unit Record, Computer Program File, Rocket File, and Distribution File.

To obtain a better understanding of the NSSDC information system and to get a broad picture of what happens during this process, it will be helpful to follow the path of information flow from the experimenter to the system. This information flow is summarized in figure 7. First, a space data scientist, referred to previously as an acquisition agent, is assigned to each satellite/experiment/data set, as appropriate. He then obtains advance prelaunch information from such sources as the project office, news releases, bulletins, reports, and personal contacts. This and subsequent information is entered into a working acquisition file, and, at this time, an AIM entry is generated. The agent establishes contact with the experimenter and his data processing personnel to arrange for bringing in data and related documentation. It should be pointed out that long periods of time are normally involved between this first stage and getting the actual data into the NSSDC system. This usually takes two or more years after launch.

Once the preliminaries are over, the acquisition agent remains in constant contact, through visits or phone, with the experimenter and his data processing staff to help solve problems relating to the submission of data and documentation. Thereafter, the data and

NATIONAL SPACE SCIENCE DATA CENTER (NSSDC) INFORMATION SYSTEM

AIM — AUTOMATED INTERNAL MANAGEMENT
FOR INFORMATION RETRIEVAL AND REPORT GENERATION

MODS — MACHINE - ORIENTED DATA SYSTEM
FOR PROCESSING AND REFORMATTING MACHINE - SENSIBLE DATA

TRF — TECHNICAL REFERENCE FILE
FOR INTERNAL OPERATIONS/USER TOOL

RASH — REQUEST ACCOUNTING & STATUS HISTORY
FOR REQUEST BOOKKEEPING

PLUS

- COMPUTER PROGRAM STATUS REPORT
- MAGNETIC TAPE UNIT RECORD
- COMPUTER PROGRAM FILE
- ROCKET FILE
- DISTRIBUTION FILE

Figure 6—Subsystems Comprising Information System

information come in almost on an automatic basis, except where special problems arise. The first items that should arrive at the Data Center are usually calibration curves, unpublished information, instrument descriptions, and data processing documentation. These are analyzed and selectively entered into the acquisition file and TRF, and notices are placed into the AIM file for subsequent use in processing incoming data and preparing publications.

Next, the reduced data, consisting mainly of magnetic tapes, arrive. At this time, the acquisition agent, together with a programmer, as required, verify and analyze this data, prepare duplicates of the tapes, and prepare data set catalogs (indexing) using the MODS subsystem to accomplish these tasks. At this point, tape reformatting often has to be accomplished. The agent then feeds appropriate information into special-purpose files such as the tape and program status files. The AIM entry is brought up to date. (These subsystems and special-purpose files will be discussed in the next section.)

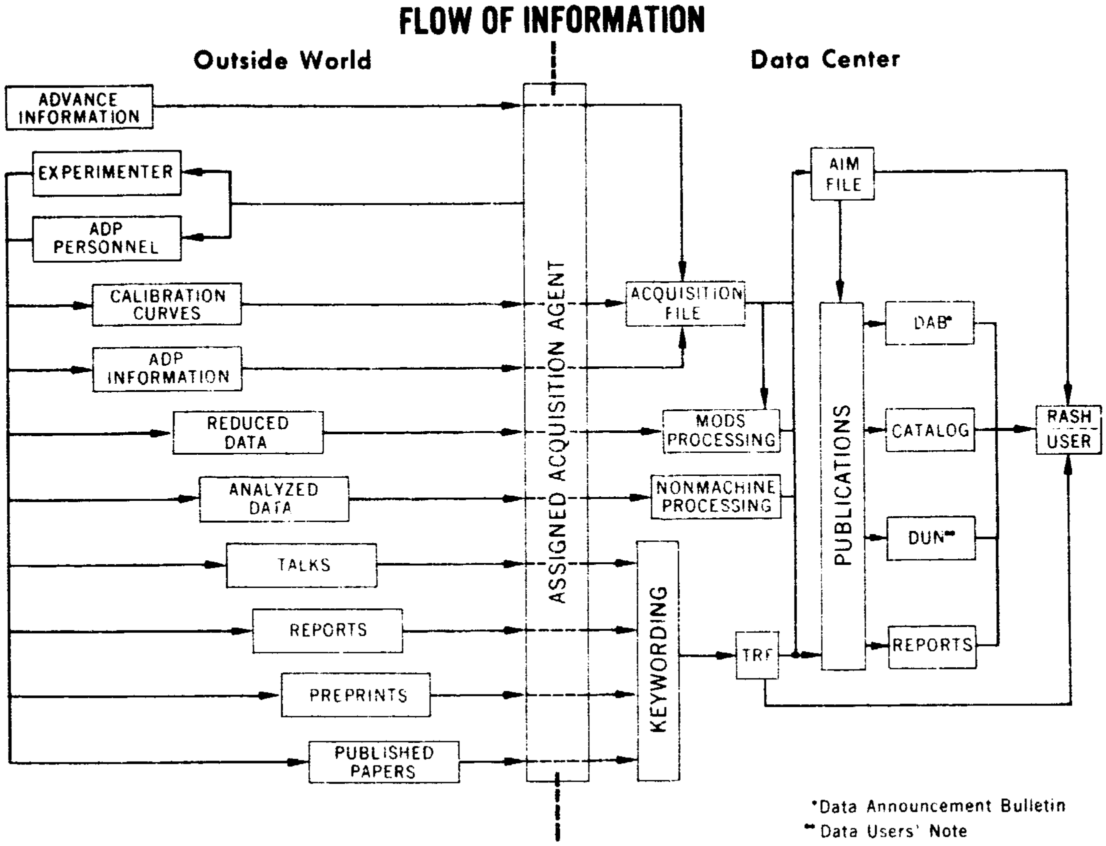

Figure 7—Flow of Information from the Experimenter into the NSSDC System

The analyzed data, normally made up of plots, graphs, and tables, arrive quite a bit later. Of course, for older experiments that are not yet in NSSDC, analyzed and reduced data may arrive in any order. The acquisition agent again goes through a similar processing cycle as in the case of machine-sensible data. Data are verified, analyzed for information content, logged, indexed, copied or microfilmed, and the information entered into the AIM information subsystem.

The working relationship with the experimenter is beneficial for the information transfer in other ways. Through this association and contacts at professional meetings, the acquisition agent receives copies of appropriate talks, reports, preprints, and published papers, as well as gaining a deeper understanding of the experiment and the implications of the data. (These items are supplemented by a thorough screening of the current literature.) Each of these documents is carefully analyzed, keyworded (by the acquisition agent), and entered into the TRF. Appropriate information is extracted for entry into the AIM.

It should be clear by now that the acquisition agent spends a great amount of his time studying and working with the data to put all the necessary information together so that it may be useful to others. Using the information from the AIM and TRF subsystems and special-purpose files, the acquisition agent and publications staff prepare, as necessary, *Data Announcement Bulletins* and entries for the *Catalog of Satellite and Rocket Experiments*. This does not necessarily mean that all data from a particular satellite experiment have arrived or have been completely processed. Many other contacts and correspondence with the experimenter may still be necessary. The preparation of a *Data Users' Note* concerning a particular experiment normally occurs after the final stage of data acquisition and processing. This document shows where the supporting information is, in what forms the data are available, what literature of previous work relating to the experiment is available, and offers a key to the use of the data.

The information flow is not complete without a mention of the RASH subsystem. The acquisition and request agents work through the RASH subsystem in satisfying users' requests on a daily basis. These requests may involve copies of data or publications, logical searches of the information files, or may even require further detailed data analysis on the part of the acquisition agent to help solve a particular problem.

THE NSSDC INFORMATION SYSTEM

The NSSDC information system used to handle this flow of information is comprised of four main subsystems and five special-purposes files, each of which will be examined in detail. It is integrated at the operating level through the capability of logical searches of the AIM, RASH, and TRF subsystems. For example, AIM search items include: satellite, launch date, silent date, experiment group, rank (priority), and acquisition agent. Standard satellite names, numbers, and discipline keywords are used. Such discipline keywords are listed in figure 8.

Automated Internal Management (AIM)

AIM, as the centralized source of information, is the hub around which the other subsystems revolve. It is built upon detailed descriptions of the data, experiment, and spacecraft, along with the status of acquisition activity. The purposes served by the AIM subsystem are detailed in figure 9. In addition to supplying this type of information, AIM has the capability to perform information retrieval on a more general scale.

The contents of the AIM file are organized into a hierarchical structure. The most significant level is the spacecraft. Information which relates to the spacecraft and is implicitly true of all experiments on that spacecraft is included here. The second level relates to the experiment. All experiment identification, detector descriptions, and commentary about a single experiment are contained in this section. The third level deals with a single data set.[1]

[1]Defined as a body of data reduced by one group of investigators in one specific way in a form, format, or organization which uniquely describes it. It can be a unit of machine-sensible or nonmachine-sensible data which can contain one to several hundred magnetic tapes, rolls of film, etc. The same information appearing on different storage media is generally given a separate data set number.

DISCIPLINE KEYWORDS

• ASTRONOMY	• SOLAR PHYSICS
UV Astronomy	Solar X Rays
IR Astronomy	Solar Gamma Rays
Visible Astronomy	Solar UV
X-Ray Astronomy	Solar Lyman-Alpha
Gamma-Ray Astronomy	Solar EUV
Radio Astronomy	Solar Magnetic Fields
Gravitational Astronomy	
Neutrino Astronomy	

Figure 8a

• PARTICLES & FIELDS	• IONOSPHERES & RADIO PHYSICS
Terrestrial Trapped Rad.	Planetary Ionosphere
Planetary Trapped Rad.	Ionosphere Events
Galactic Cosmic Rays	Ionosphere Structure
Solar Particle Rad.	Ionosphere Mechanisms
Solar Wind	Ionosphere Observing
Energetic Plasma	Ionosphere Wave Propagation
Interplanetary Magnetic Fld.	
Geomagnetic Fields	
Electric Fields	
Magnetosphere Boundary Proc.	
Energetic Auroral Part.	

Figure 8b

• PLANETARY ATMOSPHERES	• PLANETOLOGY
Comets & Meteors	Lunar Internal Prop.
Micrometeorites (Interplanetary Dust)	Lunar Mineralogy
Atmosphere Composition	Planetary Internal Prop.
Atmosphere Structure	Planetary Mineralogy
Atmosphere Dynamics	Selenography
Aurora	Planetography
Terrestrial Radiations	
Atmospheric Radiations	
Martian Atmosphere	
Venusian Atmosphere	
Jovian Atmosphere	

Figure 8c

• APPLICATIONS	• BIOSCIENCE
Geodesy	Planetary Bio.
Communications	Environmental Space Bio.
Navigation	Behavioral Space Bio.
Meteorology	Physical Space Bio.
Oceanography	
Earth Resources Geology	

Figure 8d

Figure 8—Discipline Keywords at NSSDC

AIM AUTOMATED INTERNAL MANAGEMENT

- LOGICAL SEARCHES TO ANSWER QUERIES
- WORKLOAD/VOLUME OF EXPECTED DATA
- ACTION REMINDERS

- ACQUISITION MANAGEMENT REPORT
  - SPACECRAFT/EXPERIMENTS/DATA SETS
  - RESPONSIBLE AGENT
  - PRIORITY
  - STAGE OF ACQUISITION
  - HOURS EXPENDED
  - CURRENT ACTIVITY
  - NEXT CONTACT

- FILE INDEX
  - LISTING OF SPACECRAFT/EXPERIMENTS/DATA SETS

Figure 9—Uses of AIM

These levels are generally tied together in the following manner, depending on identification of experiment and availability of data: the satellite-level entry will be followed by all experiment-level entries which pertain to that spacecraft; similarly, the data set-level entries are associated with the experiment. However, the satellite-level entry need not have any experiments associated with it, as in the case of satellites with no scientific payloads; an experiment-level entry need not have data sets associated with it, as in the case of recently launched spacecraft where no data is yet available for deposit. This concept can be perhaps better visualized by examining the typical AIM File Index entry shown in figure 10. Point 1 shows the satellite entry, point 2 the experiment level, and point 3 the data set entry, with pertinent information following across the remainder of the line.

Within each of these levels in a full AIM entry, there are specified categories of information. However, certain categories apply to all levels. These are:

Title or Name
TRF Accession Number
Fixed Format Information
Action Reminders
Remarks
Acquisition Management Information
Date of Last Entry

TYPICAL AIM FILE INDEX ENTRY

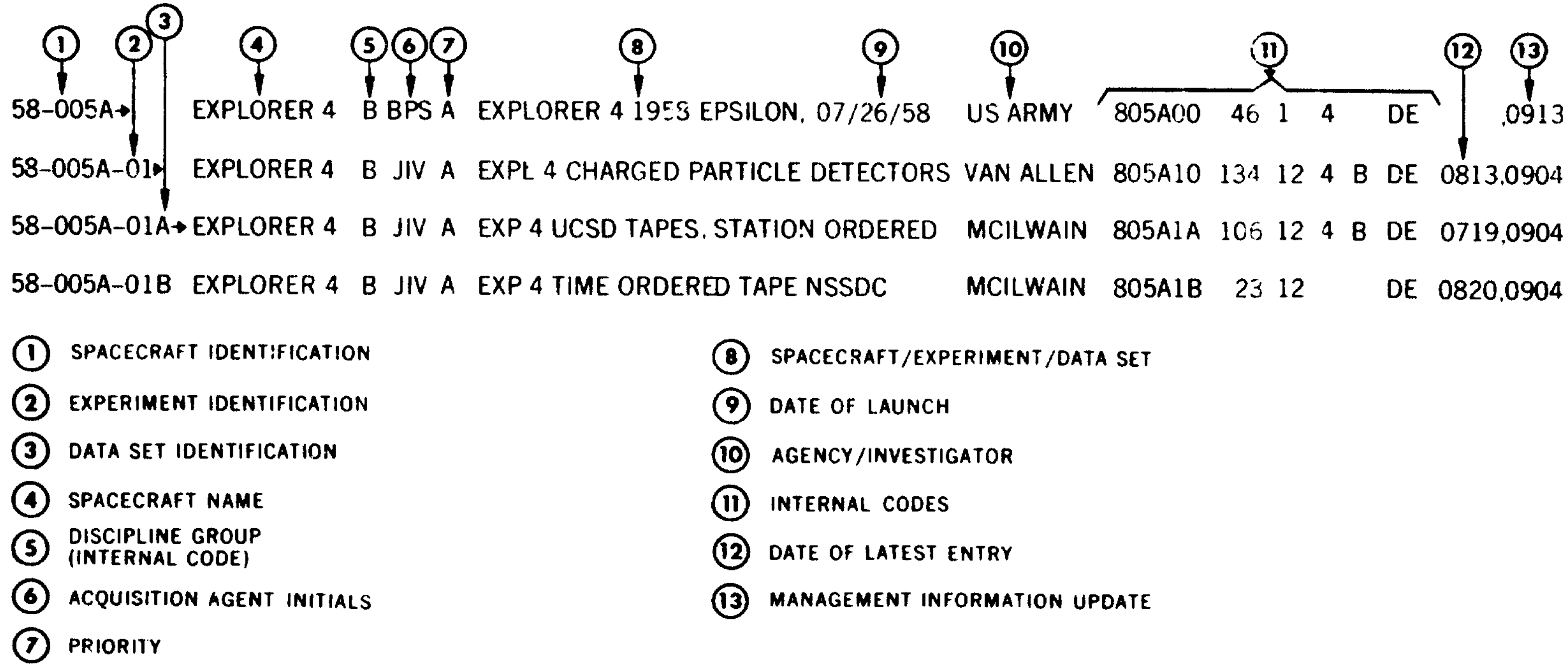

Figure 10—Typical AIM File Index Entry

Additional categories for each level are as follows:

Spacecraft	Experiment	Data Set
Personnel	Personnel	
Objectives	Instrument Description	Accession Units
Telemetry	Detector Characteristics	Description
Instrumentation	Experiment Performance	Data Set Contents and Catalog
Orientation	Acquisition Information	Programs Available
	Expected Date Set Information	
	Calibration Information	

Again, it is easier to translate this concept into reality by examining the sample full listing shown in figure 11.

As noted earlier, AIM is also used for providing management information. Based on the same levels just discussed, the following information is provided concerning spacecraft, experiment, and data set.

NSSDC Identification

Descriptive Information (satellite, experiment, data set)

Experiment Group (internal assignment)

Acquisition Agent

NSSDC Rank (priority assigned to data set)[2]

Total Acquisition Man-hours Used

Man-hours Used Last Month

Last Visit

Last Contact

[2]Of necessity, a significant number of desirable data sets are assigned low priorities simply because the total time required for top priority experiments cannot exceed the time an acquisition agent has available.

Next Scheduled Contact

Acquisition Stage

Task Wait (processing stopped/reason)

Data Set Form/Size (type/units)

Date of Data Arrival

Date of Completion

TYPICAL AIM FULL LISTING

65-093A-01 EXP 30. SOLAR RADIATION KREPLIN

SOLAR RADIATION

EXPERIMENT INSTRUMENT DESCRIPTION

EIGHT (8) ION CHAMBERS AND TWO (2) GEIGER COUNTERS COVERING THE SPECTRAL REGIONS FROM 0.5A TO 60A AND 1080A TO 1350A WERE MOUNTED PERPENDICULAR TO THE SPIN AXIS OF THE SPACECRAFT. THE ANALOG DETECTOR OUTPUTS WERE TELEMETERED CONTINUOUSLY ON SIX (6) IRIG CHANNELS. A COMPLETE DESCRIPTION OF THE INSTRUMENT IS FOUND IN TRF A-01329.

EXPERIMENT PERFORMANCE

THE EXPERIMENT PROVIDED GOOD DATA FOR ALL DETECTORS FROM NOV. 27, 1965, TO AUG. 24, 1967, WITH THE FOLLOWING EXCEPTIONS - (1) THE LYMAN-ALPHA DETECTOR AND THE U.V. DETECTORS WERE SATURATED FOR NORMAL ASPECT ANGLES, (2) THE CORE MEMORY FAILED AT LAUNCH SO THE DATA WAS COLLECTED IN REAL TIME TELEMETRY ONLY, (3) A GRADUAL DECREASE IN SPIN RATE CAUSED ASPECT TO DRIFT AWAY FROM NORMAL IN THE SECOND YEAR OF OPERATION. EXPERIMENT PERFORMANCE IS DISCUSSED IN TRF A-01329.

CALIBRATION INFORMATION

DETECTOR CALIBRATION CONSISTS OF TWO ELEMENTS, (1) EFFICIENCY AS A FUNCTION OF WAVELENGTH, (2) EFFICIENCY AS A FUNCTION OF ASPECT ANGLE. THESE CURVES ARE FOUND FOR ALL DETECTORS IN TRF A-01329. CONVERSION FACTOR USED TO PRODUCE FLUX VALUES IS FOUND IN TABLE 11 OF A-01329.

ENTRY INFORMATION

DATE OF LAST ENTRY 07/08/68

DATE OF LAST ACQUISITION ENTRY 09/10/68

ACQUISITION MANAGEMENT INFORMATION

GROUP C	AGENT CLM		RANK A
TOTAL ACQ. HOURS SPENT	30	EXPERIMENT DOC.	9
ACQ. HOURS SPENT LAST MO.	0	CALIBRATION DOC.	A
MONTHS SINCE LAST VISIT	4	DATA SET DOC.	N
MONTHS SINCE LAST CONTACT	4	DATA RECEIVED	N
MONTHS TO NEXT CONTACT	3	PROGRAMMING	N
DATE DATA FIRST REC'D	07/67	PROCESSING	N
DATE ACQUISITION COMPLETE	06/68	PUBLICATIONS	N
DATA		TASK WAIT	

Figure 11(a)—AIM Full Listing Showing Three Levels

TYPICAL AIM FULL LISTING
(CONT'D)

65-093A-01A	EXP 30. SOLAR RADIATION	KREPLIN
	1 MIN FLUX AVG FOR 12 DETECTORS	MCCLINTON

1 MINUTE AVERAGES OF X-RAY FLUX FROM 12 DETECTORS IN ERGS PER SQ CM SEC WITH DATA FLAGS

ACCESSION UNITS
D00207, D00204, D00148, D00149, D00100, D00176, D00118, D00084-D00087, D00711-D00717

DATA SET INFORMATION

DATA FIRST RECEIVED	03/01/67	DUN BEGUN	07/24/67
ALL DATA RECEIVED	05/01/68	DUN COMPLETED	
DSC BEGUN	01/01/68	DUN NUMBER	------
DSC COMPLETED	/ /	ANNOUNCED	
DAB BEGUN	/ /	REFORMATTED	/ /
DAB COMPLETED	/ /	FILE REVIEW	
DAB NUMBER	------	PROCESSING TERMINATED	
NUMBER OF REQUESTS			

DATA SET DESCRIPTION
THIS DATA SET CONSISTS OF 18 BCD CARD IMAGE TAPES WHICH CONSTITUTE THE FINAL TOTAL OF REDUCED DATA OF 65-093A-01. THE FOLLOWING ERRORS EXIST IN THIS DATA SET, (1) FOR CERTAIN PASSES F1 AND L1 DATA WERE INTRODUCED AS VALID, (2) FOR CERTAIN PASSES TWO FIELDS WERE INTERCHANGED, (3) FOR CERTAIN PASSES ONE FIELD WAS MULTIPLIED BY 100, (4) FOR ASPECT ANGLES. GT. 14 ONE DETECTOR HAD AN IMPROPER CONVERSION FACTOR.

ENTRY INFORMATION
DATE OF LAST ENTRY 07/19/68
DATE OF LAST ACQUISITION ENTRY 09/10/68

ACQUISITION MANAGEMENT INFORMATION

GROUP C	AGENT CLM		RANK A
TOTAL ACQ. HOURS SPENT	100	EXPERIMENT DOC.	A
ACQ. HOURS SPENT LAST MO.	0	CALIBRATION DOC.	A
MONTHS SINCE LAST VISIT	4	DATA SET DOC.	A
MONTHS SINCE LAST CONTACT	4	DATA RECEIVED	A
MONTHS TO NEXT CONTACT	3	PROGRAMMING	A
DATE DATA FIRST REC'D	07/67	PROCESSING	A
DATE ACQUISITION COMPLETE	06/68	PUBLICATIONS	9
DATA DIGITAL MAG. TAPE	18 TAPES	TASK WAIT	

REMARKS:
PROGRAMMING IN PROGRESS TO PRODUCE A COMPACT AND CLEANED UP DATA SET.
DUN IS IN PROGRESS.

Figure 11(b)—AIM Full Listing Showing Three Levels

To sum up the AIM subsystem in terms of output, it is used to produce the following reports, most of which have been discussed, with the remainder being self-explanatory by virtue of their titles.

- Acquisition Management Report
- Acquisition Management Update

- Acquisition Notebook
- AIM File Index
- Action Reminders
- Satellite List
- Edit List
- AIM General Printout

Machine-Oriented Data System (MODS)

In order to be responsive to the users who request data in digital form, as well as to those who provide the original data, NSSDC must have the flexibility to accept the data in any format and to provide it in any format. Since both the giver and taker will have restrictions imposed by their existing computer hardware and software, the Data Center facility must provide the "impedance matching." The tools to do this job, as summarized in figure 12, are found in the MODS subsystem. MODS is used for processing the data into the NSSDC computerized data base, for data set analysis, generation of data set catalogs, tape reformatting (when the interchange of information is inhibited by the diversity of hardware), and production of allied reports. Perhaps the best way to examine the composition of this subsystem is to follow incoming machine-sensible data sets through their processing cycle and then look at the tape reformatting process.

MODS

MACHINE-ORIENTED DATA SYSTEM

- DATA SET ANALYSIS
- DATA SET CATALOGS
- TAPE REFORMATTING

Figure 12—MODS Results

Processing Incoming Data Sets[3] – All magnetic tapes received by NSSDC are first entered into the storage records by filling out the proper forms and assigning a unique accession number. At this point, an acquisition agent, to be assisted by a programmer as

[3]Since the steps involved in processing nonmachine-sensible data are analogous, they will not be discussed here.

necessary, is assigned to the data set for preliminary analysis. The following steps, as appropriate, are taken by the acquisition agent.

1. Data and documentation reviewed.

2. Processing determined.

3. AIM entries made.

4. Computer Program Status Report (to be discussed) entry generated.

5. Receipt acknowledged (letter) and discrepancies noted.

The joint objectives of the acquisition agent and programmer in the preliminary processing are: (1) ability to read the entire tape in its logical format; (2) ability to list out any function or special record; (3) ability to detect errors (logical and physical); and (4) verification of the acquisition agent's understanding of the contents.

During this preliminary processing, the programmer writes all the necessary routines to manipulate the data and reformat it, if necessary. These programs are entered into the Computer Program File (to be discussed). The preliminary analysis stage is completed when NSSDC has the ability to use and interpret all data in the data set. This may require additional contacts with the experimenters.

At this time, the acquisition agent and the programmer define the format of the Data Set Catalog, the functions of which are to:

- Provide an index to the contents of the data set
- Provide a series of error checks
- Calculate bounds or distributions of functions
- Provide a useful tool to the request agent for identifying data
- Provide a coarse description of the information in the data set

Some items included in the catalog are: ranges of important items, tape-oriented descriptions of major functions, tables of gaps and overlaps, etc.

After the requirements are defined, the programmer writes a program to produce the Data Set Catalog. This routine should also produce a copy of the original tape or a reformatted version. After checking the program and turning it over to the computer people, the rest of the tapes in the data set are processed.

The Data Set Catalog is considered available for distribution only after the following steps have been taken.

1. Table of contents is prepared, material is bound, and catalog is stored.

2. Program is entered into the Computer Program File.

3. Information for the Magnetic Tape Unit Record (to be discussed) is completed.

4. All entries for the AIM file are completed.

5. The *Satellite and Rocket Experiment Catalog* entry is completed.

Tape Reformatting – The processing of normal machine-sensible data is well taken care of by using the procedures just outlined. However, experience has shown that people will not use data if it takes a lot of time and effort to convert it to a format which allows for direct entry into their own computer. Consequently, one of the major problems confronting the Data Center is the processing of magnetic tapes produced by different computers and operating systems. This presents two main difficulties: the physical problem of reading tapes which cannot be used with the normal hardware available to NSSDC and the logical problem of interpreting data where word size, word format, data arrangement, etc., cannot be easily handled with standard software. The approach used is to achieve the desired flexibility by producing software which is bit-oriented rather than character- or word-oriented. The Data Center currently has routines available to read tapes generated by a number of operating systems,[4] as well as BCD (binary coded decimal) tapes with arbitrary and variable record sizes, physically formatted binary tapes, and FORTRAN generated tapes. For achieving compatibility with systems using 9-track tape, NSSDC uses other computers at GSFC. The hub of the MODS subsystem will be a package called PIFT (Package for Information Formatting and Transformation) which will accomplish the functions just outlined and at the same time will produce densely packed machine- and media-independent data sets that may be accessed in the man-machine mode. A basic building block of this subsystem is a bit manipulator program which has recently been completed.

Technical Reference File (TRF)

The Data Center professionals must have internal documentation support and a tool for satisfying the bibliographic needs of space science data users. This is why the TRF comes into the system. It provides access to documents used for evaluating and verifying acquired data and for publishing catalogs, *Data Users' Notes,* and bulletins. It provides a useful record of the documentation available at NSSDC, as well as that which exists in the published literature. The references include published and unpublished documents relating to the spacecraft, experiment, or data set which are or will be preserved at the Data Center. No attempt is made to physically maintain references of a general scientific nature or in any way duplicate the services of the NASA library facilities. However, the classification of the documents

[4]BESYS, APLOS, IBM-DCS operating systems.

and the descriptors are usually different because the depth of indexing is tailored specifically to the space science community and to NSSDC retrieval use.

Figure 13 outlines the TRF subsystem. As can be seen, the computer can display pertinent information in a variety of ways. Open (subjective) and controlled keywords are used to cover standard satellite/rocket identification, type and contents of publication, and discipline-oriented keywords (see figure 8). The methodology for describing the type and content of publication can be seen in figure 14. A typical TRF entry is presented in figure 15.

To overcome the common gap between indexer-selected terms and user-selected terms, the acquisition agents themselves, who are space data scientists and the prime users of this subsystem, review the literature, select the entries, and keyword the inputs.[5] Thus, each member of the acquisitions staff devotes a portion of his time to building up the TRF base and verifying the output. In this manner, up to 120 items per week are entered into the file, which now contains well over 3,000 entries.

TRF TECHNICAL REFERENCE FILE

- WORKING SPACE SCIENCE DOCUMENTS
- ACCESS BY
  - ACCESSION NUMBER
  - TITLE
  - AUTHOR
  - KEYWORDS
    SPACECRAFT/EXPERIMENTS/DISCIPLINE
- TIED TO AIM PROGRAM
- CONTACT WITH AUTHORS
- SPACE SCIENCE BIBLIOGRAPHIES

Figure 13—TRF Summary

[5]Document storage and retrieval is based upon randomly assigned accession numbers.

TRF

PUBLICATION CLASS		CONTENT CODE	
A	JOURNAL ARTICLES	**0**	BIBLIOGRAPHIES
B	BOOKS	**1**	THEORETICAL PAPERS
C	GOVERNMENT PUBLICATIONS	**2**	SCIENTIFIC PAPERS – EXPERIMENTAL RESULTS
D	UNIVERSITY REPORTS	**3**	INSTRUMENT DESCRIPTION PAPERS
E	INDUSTRY REPORTS	**4**	DISCIPLINARY REVIEW PAPERS
F	MAGAZINES, PRESS RELEASES AND NEWSPAPER ARTICLES	**5**	SATELLITE & MISSION DESCRIPTION
G	PROCEEDINGS, SYMPOSIA & OTHER COLLECTIONS	**6**	NEWS RELEASES
H	UNPUBLISHED	**7**	DATA PROCESSING PAPERS
		8	WORKING PAPERS, MINUTES, ETC.
		9	DATA TABULATION

Figure 14—Codes to Identify Publications. Code letters and numerals are combined to identify both type and content of publication by the use of two or more characters, e.g., BG 23.

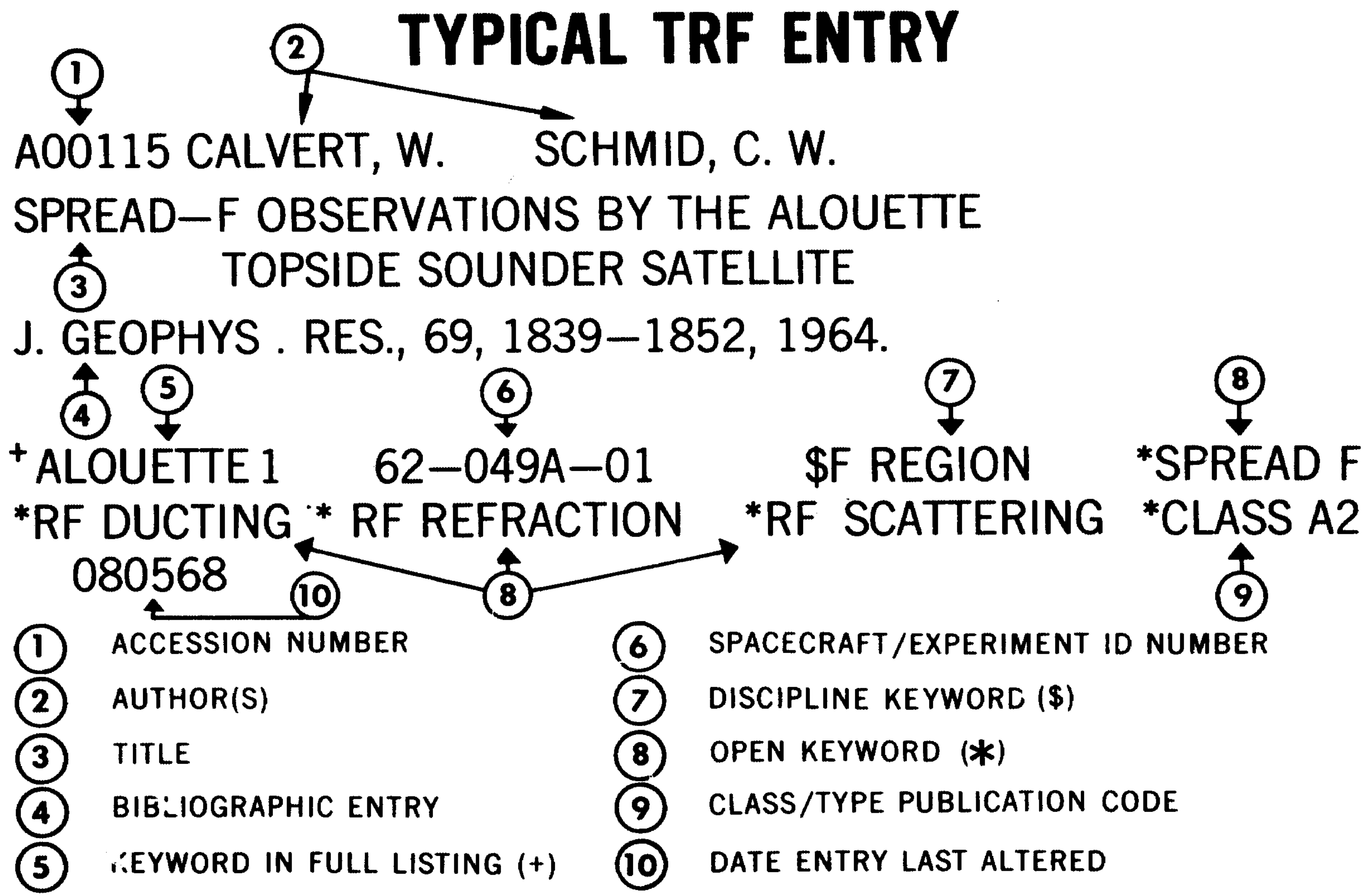

Figure 15—View of TRF Entry

As concerns the external uses of the TRF, considerable effort is presently being devoted to the production of a notebook-sized TRF output. Once this is fully implemented, NSSDC will have the capability of producing space science bibliographies ordered by author, discipline, experiment, or spacecraft. To produce special bibliographies upon request, a logical routine is in the process of being integrated into the TRF program which will allow for the usual Boolean logic searches of AND, OR, and NOT among the entries. Present and additional keywords are also being studied to eventually derive a meaningful thesaurus.

For the purposes of accuracy and completeness, the Data Center will eventually send each principal experimenter (and his colleagues) the NSSDC listing of his publications and request that he make necessary corrections and additions. This will not be done until it is felt that the file is nearly complete. Copies of current preprints and publications are presently being requested to ensure that the TRF will be maintained in an updated condition.

To sum up the present TRF subsystem, there are three basic outputs, with variations.

1. Complete entry in fixed format. It may be sorted by:

   a. Accession number
   b. Author
   c. Acquisition agent

2. One line per entry listing. These are identical to listings by:

   a. Author
   b. Title
   c. Keyword

3. Notebook listing (external distribution). It may be ordered by:

   a. Accession number
   b. Discipline
   c. NSSDC identification

Request Accounting Status and History (RASH)

At this point, the data from the space science experiments have been obtained, entered into the system, and prepared for retrieval. The next step is to facilitate the acquisition and request agents' contacts with the users of these data – this need is satisfied through the RASH subsystem. (Upon occasion, the Data Center will obtain data not at NSSDC for a user.) As shown in figure 16, this subsystem provides much valuable information. It is used to aid in keeping track of the progress of requests received by the Data Center and providing management with bookkeeping information. Specifically, RASH is designed to display up-to-date information relating to the number of requests, their status, estimated and actual costs, processing time, and necessary action reminders.[6] This variety of information can be

[6]This subsystem also can aid in the construction of a model of the NSSDC by supplying viable information describing the user community, types of requests and responses, and data sets most likely to be used.

RASH — REQUEST ACCOUNTING STATUS & HISTORY

- NUMBER OF REQUESTS
  - STATUS OF REQUESTS
    - COSTS
      - PROCESSING TIME
        - ACTION REMINDERS
          - WORK QUEUES
            - REQUESTER HISTORY
              - DATA SET USE
                - REPORTS

Figure 16--Some Answers Provided by RASH

retrieved by data set, requester, affiliation, date of request, date filled, request agent, status of request, and so forth. A typical RASH entry is shown in figure 17.

The six stages of request processing should be touched upon briefly for a better understanding of this subsystem.

1. Generation – letter, phone, visit

2. Validation

   Necessary contacts/correspondence
   Identification of requester
   Charges (if applicable)
   Selection of request agent
   Anticipated delays

3. Definition

   Request form filled out

      Units required
      Work to be performed
      Coordination
      Estimated completion date
      Processing start

   Entry into RASH

   Acknowledgement of request

4. Processing

   Reproduction/assembly
   Actions taken (remarks)
   Programming required (if applicable)

5. Completion

   Reproduction/assembly completed
   Documentation completed
   Letter of transmittal
   Receipt acknowledgement form
   Shipment
   RASH update

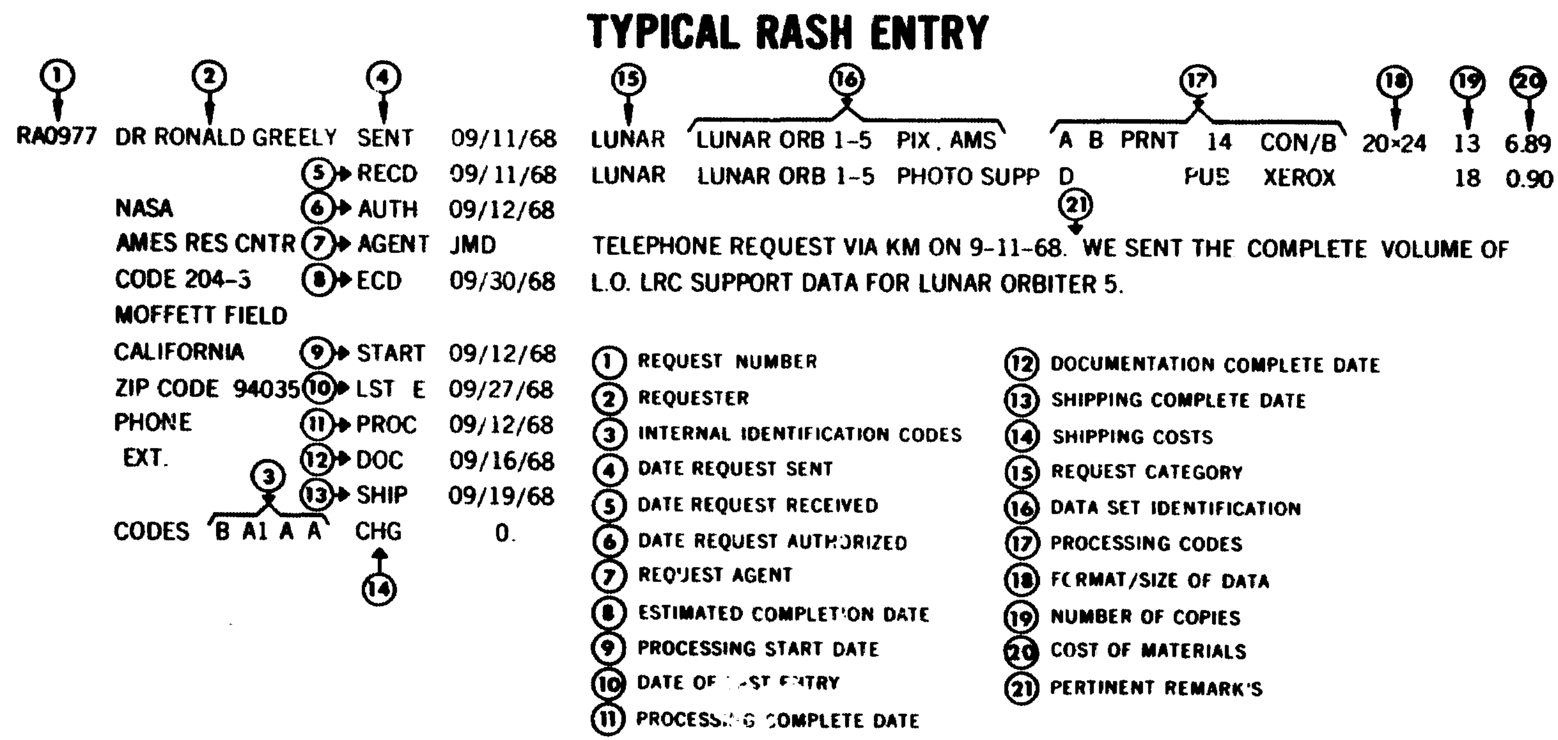

Figure 17—View of RASH Entry

6. Follow-up and Customer Satisfaction

Condition of data
Usefulness
Other user responses
Distribution file update
Entry into RASH

On the output side, RASH is used to produce five weekly reports (first five listed) and five reports on an as-needed basis, the purposes of which are explained by the titles.

1. RASH File Listing
2. Action Reminders
3. Special Request Classification (ordered by requester, affiliation, and data set)
4. Request Accounting Status History Report
5. Summary of Data Request Services
6. Data Request Master Log
7. Costs for Request Services
8. Payments to other Agencies
9. Machine-sensible Data Reproduction
10. Nonmachine-sensible Data Reproduction

Computer Program Status Report

This file is the first of five special-purpose files set up to help accomplish the Data Center mission. The intent of this listing is to provide a convenient and up-to-date report of the programming jobs requested and their status. At the same time, it serves as a tool for measuring and establishing programming work-level estimates at the Data Center. This report gives the following categories of information:

- Program Identification
- Date Requested
- Program Sponsor (requester)
- Program Category[7]
- Status[8]
- Date Program Assigned
- Programmer

[7]Data Set Catalog, request, systems, analysis, or exchange program.
[8]Unassigned, active, suspense, or completed.

- Estimated Completion Date
- Percent Defined

  Compiled

  Checked

  Documented
- Date Completed

Magnetic Tape Unit Record

Magnetic tapes are identified at NSSDC by serially assigned accession numbers, each of which is associated with a rack position in storage. Because of the large number of tapes involved, tape identification indexing records have been converted to a machine-sensible form. As such, they can be sorted to produce listings of tapes ordered by accession number, data set, or in other desired ways.

NSSDC Computer Program File

This file contains listings of complete programs written at the Data Center as well as any general-purpose routines which may be of use to NSSDC personnel. Each index contains a one-line, alphabetically ordered catalog of the contents and a sequentially ordered index with the abstract.

Rocket File

The rocket file is a card-oriented, machine-sensible file designed to simplify the announcement, acquisition, and request processing of rocket experiments; to automatically generate the rocket sections of the *Data Catalog of Satellite and Rocket Experiments;* to provide up-to-date records of all rockets fired and reports published; and to provide a variety of reports for NASA management. The general policy is that rocket experiments will be identified and listed, but there is no planned activity to systematically acquire the data obtained from these experiments. On the other hand, if rocket data are obtained in the course of other acquisition efforts, these are archived and cataloged. The reason for this policy is that rocket and balloon data are very often contained in the published literature. Consequently, there is no need to actively collect such data.

As in the case of TRF entries, an acquisition specialist (in rocket data) is charged with screening rocket reports and assigning field and discipline codes.

Figure 18 shows the depth of breakout of the rocket discipline codes, and figure 19 depicts a typical, self-explanatory entry found in the *Data Catalog of Satellite and Rocket Experiments.*

ROCKET DISCIPLINE CODES

1. Aurora and Airglow
   - A. Gegenschein
   - B. Auroral emissions
   - C. Airglow emissions
   - D. Airglow composition
   - E. Atmospheric radiations

2. Atmospheric Physics
   - A. Meteorology
   - B. Winds
   - C. Pressure
   - D. Temperature
   - E. Albedo
   - F. Planetary radiations (IR)
   - G. Neutral density
   - H. Neutral composition
   - I. Electromagnetic waves
   - J. Acoustics

3. Ionosphere
   - A. Wave propagation
   - B. Electric currents (mag. fields)
   - C. Ion/electron density
   - D. Ion/electron composition
   - E. Ion/electron temperature

4. Energetic Particles
   - A. Galactic cosmic rays
   - B. Solar particle radiation
   - C. Terrestrial trapped radiation
   - D. Particle precipitation

5. Magnetic Fields
   - A. Geomagnetic fields
   - B. Electric fields

6. Solar Physics
   - A. Radio (1–1000mm)
   - B. Infrared (.8–1000μ)
   - C. Visible (3000–8000A)
   - D. Ultraviolet (2000–3000A)
   - E. Extreme UV (100–2000A)
   - F. X Rays (.001-100A)

7. Astronomy
   - A. Radio (1–1000 mm)
   - B. Infrared (.8–1000μ)
   - C. Visible (3000–8000A)
   - D. Ultraviolet (2000–3000A)
   - E. Extreme UV (100–2000A)
   - F. X Rays (.001–100A)
   - G. Gamma Rays (<.001A)

8. Planetology
   - A. Micrometeorites
   - B. Zodiacal light
   - C. Gravity
   - D. Terrain photographs

9. Biology

0. Test & Other
   - A. Rocket performance
   - B. Communication systems
   - C. Satellite experiment test

X. Not Determined

Figure 18—Breakout of Rocket Discipline Codes

Test and Support	Rocket Number	Launch Date	Launch Time (UT)	Launch Site	Peak Altitude (KM)	Experimenter	Organization	
Albedo	4.60 GT	08/08/62	1655	WI	150	A. Boggess	GSFC	P
Cosmic Rays	4.48 GT	05/25/62	1243	WI	200	O. E. Berg	GSFC	P
Electron Density/Temp	4.48 GT	05/25/62	1243	WI	200	O. E. Berg	GSFC	P
Galactic Noise	8.01 GT	12/22/59	0256	WI	900	M. W. Oleson	NRL	
Ionospheric Comp/Propagt	17.02 GT	08/17/66	1950	WI	358	D. G. Cartwright	MINN	*
Micrometeorites	4.12 GT	03/25/60	1840	WI	248	E. F. Sorgnit	GSFC	P
	4.48 GT	05/25/62	1243	WI	200	O. E. Berg	GSFC	
	14.343GT	08/05/67	0159	WSMR	150	C. L. Hemenway	DUDL	P

*Data at NSSDC. **Data at NSSDC, description in following section. P indicates partial success.

Figure 19—View of Rocket Experiment Entry

Distribution File

The other special-purpose file is the consolidated distribution list, which for obvious reasons, is tied to request reporting of the RASH subsystem. It is used for file maintenance and document distribution. In addition to the production of self-adhesive labels, this subsystem will provide listings by recipient, affiliation, publication, etc. It also will have a built-in capability for SDI (Selective Dissemination of Information), should this tool be used in the future.

As concerns the users of NSSDC data and documents, the distribution file contains the names and addresses of recipients, what they receive, and general organizational classifications in terms of recipient categories. Typical distribution file entries are shown in figure 20. Looking at this figure from left to right, the numbers under "Entry" are unique numbers assigned to each entry to facilitate updating. This is followed by the recipient category code, which is explained in figure 21. The numbers *following* "Distribution" also are identified in figure 21; the numbers *below* represent the number of copies per item for each special distribution list. Note that there is a special distribution list for each of the documents that the Data Center produces or for which it has responsibility.

An interesting application of the recipient category code is its use to identify those sections of the potential user community which receive inadequate coverage of the available space science experiment data and documents.

The last point to consider in the Distribution file is the method of updating and purging. This is accomplished through the following mechanisms:

NSSDC DISTRIBUTION FILE

Entry	Address	Distribution–1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8
20	Dr. Philip Abelson, Consultant Carnegie Institution of Washington 2801 Upton Street, N.W. Washington, D.C. 20008				1														
40	Carolyn D. Abeny, Librarian Geophysics Institute University of Alaska College, Alaska 99735	1																	
60	H. K. Abraham Philadelphia Bulletin Philadelphia, Pa. 19105	1																	
80	Acquisitions National Science Library National Research Council Susses Drive Ottawa 2, Ontario, Canada	1																	
100	Advanced Plans Staff GSFC, Code 110	1	1																

Figure 20—View of Distribution File Entries

1. Dissemination of copies of the distribution list to all concerned activities at the Data Center
2. Continual hand-updates by each section
3. Use of RASH
4. Verification (annual) by recipients
5. Consolidation, machine update, and redistribution of updated copies

A LOOK TO THE FUTURE

It is apparent that the use of the Data Center will continue to grow at such a rate as to double by 1971. As more staff are added, the rate at which data can enter the system will increase. It is hoped that 200 data sets per year can be processed within the next few years. As the capability of NSSDC grows, the services that it can render will multiply, thereby

DISTRIBUTION FORM CODES

KEY CODE OF ASSIGNED DISTRIBUTION LISTS

1. *Data Catalog of Satellite and Rocket Experiments*
2. *Catalog of Correlative Data*
3. Lunar Orbiter
4. Nimbus & Tiros
5. "Models of the Trapped Radiation Environment," List 1
6. OGO Program
7. *GSFC Computer Newsletter*
8. "Models of the Trapped Radiation Environment," (NASA SP-3024 only)
9. Surveyor

LIST OF RECIPIENT CATEGORIES

A NASA, GSFC Only
B NASA, Other Than GSFC
C DOD
D ESSA (Environmental Science Services Administration)
E Other Government Agency
F Contractor (Industry)
G Museum, Planetarium, Observatory
H Academic Facility, Including Institutions
I Government Affiliated Contractor
J Foreign Government Agency
K Foreign Academic Facility, Including Institutions
L Foreign Museum, Planetarium, Observatory
M Foreign, Other
N Miscellaneous

Figure 21—Codes Used in Distribution List

encouraging further use. However, it is realized that only certain resources will be available for this data activity. Present estimates indicate that if 1–2% of the budget available for the research areas which produce the original data handled by NSSDC is given to the Data Center, then an adequate facility can be developed to handle the complete data needs.

As the Data Center grows, so must its information system. It must be able to handle large varieties and amounts of data and p[illegible]e them for a multitude of different uses. It must rely on new and better software and hardware to effectively perform these tasks, bearing in mind the guiding principle of furthering the effective use of data from space science experiments. The present system software must be upgraded with respect to processing incoming tapes for verification of inputs and quality control – two goals are *immediate* detection of errors or omissions and *standard* maintenance and systems quality control programs. Effective purging of the active data base will have to be accomplished. Consideration must be given to a good long-term archival medium as the life time of magnetic tape cannot compare with photographic or printed matter, although recent tests are encouraging. Time-phased data compression will be another vital area of concern. Considerations include higher density storage techniques and the actual compression of data. This data compression can occur in various steps. The first step would involve the retirement of alternate forms of the data in which the most useful form would be retained. Then, even this most useful form of data could be subjected to the removal of derived variables, which are computed from basic positional and attitude information. This would still permit recalculation of these variables at a later date, should this prove necessary. At this point, no reduction in the basic information content has occurred. However, if one wishes to use this data, more time and resources will have to be utilized than previously. One is balancing this cost against the maintenance cost of keeping all the derived bits in the active data base. There is a break-even point depending on data usage. As one starts the irreversible process of destroying information content, a sensible approach would be to separate the background information (ambient, quiet time) from the event information (disturbed time). This will permit time averaging the background information, say over hours or days, for subsequent use in determining long-term changes. A sizable reduction in the number of bits for a given data set will occur in this process, and yet the information most likely to be used in future studies will still be available. Clearly, data of historical significance would be preserved as long as possible.

In other words, certain points should be considered when planning for the retirement of data.

1. Large volume plus cost of maintenance plus fixed resources dictate the *orderly* retirement of data.

2. *Early* in its life, various forms of the data are useful, e.g., time ordered, space ordered, etc.

3. Data can be reduced without losing information content:

   - By eliminating certain forms of data
   - By removing derived variables
   - By keeping only the significant number of bits, not the full computer words

4. Information content of data can be reduced:

   - By breaking out event data from background
   - By averaging the background over suitable time intervals
   - By preserving only special data for historical purposes
   - By preserving only outstanding geophysical event data

- By compressing data into analyzed forms so that general understanding of phenomena is retained

In short, data can shrink in size and in information content, but knowledge of it never disappears from the scene.

Some thought is already being given to the next generation of the NSSDC information system. One consideration is to provide the Data Center with much greater flexibility and capability by developing varied analysis programs which can be readily applied to the data. Although complete requirements have not yet been defined, it is envisioned that scientists, experimenters, and acquisition agents should be able to interact, on-line, through a computer, to data bases and data sets held at NSSDC. It is also anticipated that in this way the resulting dialogue between two or more scientists can be used to synthesize new information in the process.

These concepts are not too far from reality. With the progression of time, the central processing facilities are performing more work on the raw data before it is sent to the experimenters for analysis. In the beginning, the raw data was sent directly. Now, tapes are digitized and edited, noise flags are inserted, time overlaps are removed, and decommutation is performed. There is interest at present in having the orbit and attitude information merged with the data before it is sent to the experimenter. As high-speed data links become available across the country, there will be no need to send the data to the experimenter. Instead, standard processing will be performed up to the point where detailed analysis can begin. The data in this form could be sent directly into the Data Center where it could be reached via the high-speed terminals and manipulated on large computers by the principal investigators using many standardized analysis programs. Special-purpose analysis programs would be constructed on-line by the individual users as the needs arise. At that point, the processing facility and the Data Center will have blended into one operation. There exists today an on-line retrieval system with a 10^{12} bit capacity (ref. 3). This is capable of handling a years worth of space science data at the present rate of generation. It is quite clear that this new type of facility could be a reality in 5–10 years.

REFERENCES

1. "The Operation of the National Space Science Data Center," NASA/NSSDC 67-41, October 1967.

2. Dessler, A. J., "The Role of Space Science in Graduate Education," *Transactions, American Geophysical Union,* vol. 49, no. 3, pp. 549-554, September 1968.

3. Kuehler, J. D., and Kerby, H. R., "A Photo-Digital Mass Storage System," *Proceedings-Fall Joint Computer Conference, 1966,* pp. 735-742.

APPENDIX

EQUIPMENT AT THE NATIONAL SPACE SCIENCE DATA CENTER

The following computer and periphery units currently are available at the NSSDC in Building 26, GSFC.

Machine	Feature/Model	Description	Quantity
711	2	Card Reader	1
716	1	Printer	1
721	1	Card Punch	1
7109	1	Arithmetic Sequence Unit	1
7111	1	Instruction Processing Unit	1
7151	2	Console Control Unit	1
7302	3	Core Storage	1
7606	2	Multiplexer	1
7607	4	Data Channel	1
7607	3	Data Channel	1
7608	1	Power Converter	1
7617	1	Data Channel Console	2
7618	1	Power Control	1
729	6	Magnetic Tape Units	10
729	4	Magnetic Tape Units	6
7094	II	Computer	1
2030	D	Processing Unit	1
	F5960	First Selector Channel	
	F6961	Second Selector Channel	
	F7915	1051 Attachment	
	F3237	Decimal Arithmetic	
1051	N1	Control Unit	1
	F3130	CPU Attachment	
	F4410	Punch Attachment	
	F4411	Reader Attachment	
1052	6	Printer Keyboard	1
2540	1	Card Read Punch	1
1403	N1	Printer	1
	F8640	Universal Character Set	
1416	1	Train Cartridge	1
2821	1	Control Unit	1
	F1990	Column Binary	
	F3615	1100 LPM Adapter	
	F8637	Universal Character Set	

Machine	Feature/Model	Description	Quantity
2841	1	Storage Control	1
	F4385	File Scan	
2311	1	Disk Drive	1
2403	6	Tape Unit and Control	1
	F3228	Data Conversion	
	F3471	Dual Density	
2403	2	Tape Unit and Control	1
	F3228	Data Conversion	
	F7125	7-Track Compatibility	
2402	6	Tape Units (2)	1
	F7161	Simultaneous Read-Write	
	F3472	Dual Density	
2401	3	Tape Units	2
	F5121	Mode Compatibility	
	F7160	Simultaneous Read-Write	
407		Accounting Machine	1
557	1	Alphabetic Interpreter	1
514		Reproducing Punch	1
26	1	Printing Card Punch	4
29	A22	Card Punch	1
56	1	Verifier	1

In addition, the Data Center has access to the family of computers on Goddard Space Flight Center, up to and including the IBM 360/95.

Other capital equipment, with the exception of typewriters and small calculators, includes:

Description	Mfg./Model	Quantity
Calculator	FRI EC132	1
Calculator	RER DM991205	1
Camera, 4 × 5 view	CLU CC400	1
Camera, Rotary, 16mm	RD-3	1
Camera Planetary, 35mm	F1088	1
Cleaner, Air	H/I T38A	3
Copier, Xerox	914	1
Cutter, Paper	PAK 212	1
Dehumidifier	H/I AFAH152	2
Densitometer	MAC T0140	1
Dryer	OME 26M	1
Dryer, Film		1
Dryer, Pako	26W	1
Duplicator, Film	OZA FP2	1

Description	Mfg./Model	Quantity
Enlarger	OME D3V	2
Frame, Vacuum	EKC	1
Hydromixer	PAK 3655	2
Light Table	NUA ULT23	1
Light Table	RCC GFL940	1
Memorex Typewriter	IBM 785	1
Microfilm, Rotoline		1
Microscope	OLY HSB	1
Porta Mix	EKC 10	1
Printer, Contact	EKC A	1
Printer, Contact	REC 12	1
Printer, Strip	LOG SP1070	1
Processor, Film	EKC DVR	1
Processor, Pako	1-3-1	1
Reader, Microcard	MDP FR5	1
Reader/Printer	MMM 100T11	1
Reader/Printer	MMM 300	1
Reader/Printer	MMM 400M	3
Sensitometer	EGG MARK V1	1
Splicer	M/H 1635	1
Stereoscope	WIL 3X	1
Straightner, Print	EKC G	1
Tank, Photo Process	PAK 178	1
Transport, Paper	PAK SIMATIC	1
Viewer, Photo	RCC GSL09188	1