NASA Conference Publication 2343
Part 1

Planetary Data Workshop

Proceedings of a workshop
held at Goddard Space Flight Center
Greenbelt, Maryland
November 29-December 1, 1983

NASA

NASA Conference Publication 2343
Part 1

Planetary Data Workshop

Hugh H. Kieffer, *Chairman*
NASA Office of Space Science and Applications
Washington, D.C.

Proceedings of a workshop
held at Goddard Space Flight Center
Greenbelt, Maryland
November 29–December 1, 1983

NASA
National Aeronautics
and Space Administration

**Scientific and Technical
Information Branch**

1984

Contributing Authors

Hugh Kieffer	U.S. Geological Survey
Raymond E. Arvidson	Washington University
William A. Baum	Lowell Observatory
Larry Bolef	Washington University
Larry H. Brace	Goddard Space Flight Center
Roger N. Clark	University of Colorado
Randal Davis	University of Colorado
Richard Elphic	University of California - Los Angeles
John Pearl	Goddard Space Flight Center
Chris Russell	University of California - Los Angeles
Stephen R. Saunders	Jet Propulsion Laboratory
Richard A. Simpson	Stanford University
William Smythe	Jet Propulsion Laboratory
Laurence A. Soderblom	California Institute of Technology
I. A. Stewart	University of Colorado
R. J. Walker	University of California - Los Angeles
Charles Acton	Jet Propulsion Laboratory
Isadore Adler	University of Maryland
Joseph K. Alexander	Goddard Space Flight Center
Donald E. Anderson	Naval Research Laboratory
Donald L. Anderson	Arizona State University
John Anderson	Jet Propulsion Laboratory
Daniel Baker	Los Alamos National Laboratory
Edwin S. Barker	McDonald Observatory
Reta F. Beebe	New Mexico State University
Jay T. Bergstralh	Jet Propulsion Laboratory
Michael L. Bielefield	Space Telescope Science Institute
A. Lyle Broadfoot	University of Arizona
James Brown	
Donald B. Campbell	Arecibo Observatory
G. R. Carignan	University of Michigan
Ara Chutjian	Jet Propulsion Laboratory
Pamela E. Clark	Jet Propulsion Laboratory
Neal Cline	University of California - Los Angeles
Thomas Cravens	University of Michigan
G. Edward Danielson	California Institute of Technology
Merton E. Davies	Rand Corporation
Imke DePater	Lunar and Planetary Laboratory
John Dickel	University of Illinois
Thomas Duxbury	Jet Propulsion Laboratory
E. Eliason	U. S. Geological Survey
Nancy Evans	Science Applications Incorporated
Elizabeth Feinler	SRI International
Paul D. Feldman	Johns Hopkins University
Peter Ford	Massachusetts Institute of Technology

Dave Fulker	National Center for Atmospheric Research
James Garvin	Brown University
Lawrence P. Giver	NASA, Ames Research Center,
Rob Gold	Applied Physics Laboratory, Johns Hopkins University
Jonathan Gradie	Cornell University
James Green	NASA, Marshall Space Flight Center
Samuel Gulkis	Jet Propulsion Laboratory
Carole Hamilton	Jet Propulsion Laboratory
John K. Harmon	Arecibo Observatory
Christopher Hartsough	Jet Propulsion Laboratory
James W. Head, III	Brown University
Kent H. Hills	National Space Science Data Center
Garry Hunt	Imperial College, London, England
Garrett Jernigan	University of California - Berkeley
Ken Johnson	EROS Data Center
Torrance Johnson	Jet Propulsion Laboratory
Ray Jurgens	Jet Propulsion Laboratory
Ralph Kahn	Cornell University
Michael Kaiser	NASA, Goddard Space Flight Center
Joseph H. King	National Space Science Data Center
Arvydas Kliore	Jet Propulsion Laboratory
Charles Klose	Jet Propulsion Laboratory
Robert Kraus	University of Wisconsin
William S. Kurth	University of Iowa
Arthur L. Lane	Jet Propulsion Laboratory
Alan J. Lazarus	Massachusetts Institute of Technology
Dale Lumb	NASA/Ames Research Center
William C. Maguire	NASA, Goddard Space Flight Center
Beth Martin	Jet Propulsion Laboratory
Michael D. Martin	Jet Propulsion Laboratory
Terry Martin	Jet Propulsion Laboratory
Dennis L. Matson	Jet Propulsion Laboratory
Mark S. McDonagh	NASA Headquarters
G. E. McGill	University of Massachusetts
Joel A. Mosher	Jet Propulsion Laboratory
Peter Mouginis-Mark	University of Hawaii
Duane Muhleman	California Institute of Technology
Andrew Nagy	University of Michigan
Robert Nelson	Jet Propulsion Laboratory
Glenn S. Orton	Jet Propulsion Laboratory
Steven Ostro	Cornell University
Albert J. Pallman	St. Louis University
Frank Don Palluconi	Jet Propulsion Laboratory
Stanton J. Peale	University of California - Santa Barbara
Alan R. Peterfreund	Brown University
Gordon Pettengill	Massachusetts Institute of Technology

Donna Peuquet	University of California - Santa Barbara
Rick Pomphrey	Jet Propulsion Laboratory
Robert Reasenberg	Massachusetts Institute of Technology
Edward Rothe	NASA/Goddard Space Flight Center
Ryan Sanders	University of Colorado
Don Sawyer	National Space Science Data Center
Gerald G. Schaber	U.S. Geological Survey
A. W. Schardt	NASA, Goddard Space Flight Center
F. Peter Schloerb	University of Massachusetts
Peter H. Schultz	Lunar and Planetary Institute
Peter Shames	Space Telescope Science Institute
Robert Singer	University of Hawaii
Bill Sjogren	Jet Propulsion Laboratory
J. Slavin	Jet Propulsion Laboratory
Brad Smith	University of Arizona
Sean Solomon	Massachusetts Institute of Technology
Michael Spanner	
Nelson Spencer	NASA, Goddard Space Flight Center
Tom Stevenson	Smithsonian Astronomical Observatory
Joseph Stoltzfus	Jet Propulsion Laboratory
Donald Sweetnam	Jet Propulsion Laboratory
S. P. Synott	Jet Propulsion Laboratory
Gary Thomas	University of Colorado
T. W. Thompson	Jet Propulsion Laboratory
Jim Tillman	University of Washington
Alan T. Tokunaga	University of Hawaii
Larry D. Travis	NASA, Goddard Space Flight Center
G. Leonard Tyler	Stanford University
Prasad Varanasi	State University of New York, Stony Brook
Joseph Ververka	Cornell University
Anthony Villasenor	NASA Headquarters
Ray J. Wall	Jet Propulsion Laboratory
Jim Williett	Jet Propulsion Laboratory
A. T. Young	University of Wisconsin

TABLE OF CONTENTS

1. INTRODUCTION

1.1 Introductory Statement

The community of planetary scientists believes that two general problems concerning planetary data must be addressed.

(1) Important datasets are being permanently lost.

(2) Utilization of planetary data is constrained by difficulties in identifying, locating and accessing data sets; finding supporting documentation; and finally manipulating these data to address specific scientific questions.

An increasing number of planetary scientists feel that the present system is inadequate and that changes are necessary to insure retention, access, and use of this data now and in the future.

Past successes of the planetary programs supported by NASA have provided numerous data which can now be brought to bear in exploring answers to several complex planetary data analysis problems. Specific among these are those problems which require extensive cross-disciplinary data and expertise in their solution; and those for which solutions lie in the study of detailed phenomena over long periods of time.

Several limited attempts have been made to consolidate related data sets and make them readily available to support investigations such as these. Scientific results from those efforts have emphasized the significance of the scientific problems, and have stressed the importance of easy access to appropriate data for effective data analysis. These facts support the need for a Planetary Data System (PDS). Such a system should embody working archives of past and future planetary mission data, and should include those functions which support access, identification, use, analysis, and distribution of data in such a way as to constitute a direct analysis tool for planetary scientists.

It is noted with increasing alarm by many in the science community that valuable data sets are disappearing. Some become lost because of deterioration of the media upon which they are stored. Some sets are effectively lost because the documentation was not retained or the software required to read and interpret the data no longer compiles on current computer systems. In a few cases, the most knowledgeable individuals (of the data set) have left the field through career changes, retirement, or death. Loss of data or data knowledge can be expected to accelerate in the next few years unless appropriate action is taken.

Data archived at the National Space Science Data Center (NSSDC) are not immune from this processes. Additionally, data were often submitted to NSSDC for archival without documentation and poorly organized for use by others. Mission Instrument Principal Investigators and Team Leaders are contractually obligated to submit reduced data records to be archived by the NSSDC. The funding to support data preparation for archive is included in the Mission data analysis budget. This arrangement has suffered from two major shortcomings: (1) Priorities are difficult to set when budgets are tight; and scientists often chose to analyze the data rather than to improve the state of documentation and organization of that data prior to submission to NSSDC; and (2) The reduced data product designated by missions and investigation teams as the archival record, is (in retrospect) not the most appropriate for use with current data analysis methods and technologies.

Computer technology has experienced tremendous growth over the last decade. The growth has been expressed not only in the size, speed, and capability of central processors; but also in the proliferation of the workstation concept, the capacity of both on-line and off-line storage devices, and the sophistication of software to support data management and data analysis procedures. Advances in communications technology have provided networks which join equipment and users nationwide. At the same time, and for many reasons, infiltration of these technology advances into the planetary community has been slow. The business of planetary data analysis continues to be performed in much the same way as it was a decade ago, though perhaps with the aid of a faster computer.

There is a clear perception on the part of many in the community that advantage should be taken of these technology advancements. Implementations which include some of these new technologies should permit analysis of a greater fraction of the returned data; results from science analyses should be available more rapidly; and planetary data should be made available to significantly more members of the community than currently obtain data - without a great investment of additional resources. In order for this to come about through the implementation of a Planetary Data System (PDS), increased cooperation is required among scientists in specific disciplines, between disciplines, and within and between NASA missions. This involves some reduction of independence among individuals and groups and the application of a broader perspective when examining system-wide costs of alternative approaches.

It is possible to develop a Planetary Data System, which takes advantage of the present and future advances in technology, avoids past mistakes in the archiving and curating of data, and restores essential 'old' datasets for continued and productive use by present and future generations of scientists. The purpose of this document is to outline the present status of planetary science data, to assess the community's needs for data and data analysis support, to evaluate the state of technology today and extrapolate a short way into the future, and to suggest ways in which a Planetary Data System could be implemented.

There is an implicit assumption that the improved access to and increased usability of planetary data (afforded by a PDS) would be accompanied by appropriate funding of scientists to use that data. Past under-utilization of archived planetary data by scientists (both inside and outside of flight projects) is felt to have occurred not only because data quality was poor and the data components were difficult to access, but also because the scientific activity has been dollar limited. The Planetary Data System, as a whole, should enhance all three essential components of a healthy planetary science program: good data, improved access to that data, and sufficient funding for scientists to interpret the data and publish their results.

1.2 The Current State of Planetary Science Analysis

NASA is a mission oriented agency. As such, the analysis of data is centered around flight projects - both the analysis of past mission data sets to support new mission planning and the analysis of current mission data by science investigation teams. Flight projects are usually composed of several investigations carried on simultaneously from a single spacecraft. The instruments which support the investigations are built under the supervision of Principal Investigators (PI) and/or Team Leaders. Major differences between these two types of individuals fall into two areas: (1) The Team Leader generally has less control over the fabrication of the facility-class instrument (usually the largest on the spacecraft) than does a PI over the instrument which he designs and builds. (2) PI investigation proposals also include a team of Co-Investigators (CoI), while the Team Leader and members of facility-instrument teams are selected from individual proposals to participate and the team is put together by NASA HQ. These various groups of scientists, led by their respective Team leaders or Principal Investigators, are afforded the first opportunity to study the data as it is returned during the mission.

A recent addition to the flight analysis effort is the selection of Interdisciplinary Scientists (IDS) who have no direct hardware responsibility. The IDS's participate throughout the planning and execution of the mission and have the right of early access to calibrated data from instruments of their choice. Another group of scientists, the Guest Investigators, are selected to join the analysis effort for one year terms. All of these scientists are supported out of project funds, and are, in return, expected to perform data analysis using the facilities at their home institutions.

The Planetary Data Analysis Programs (locally referred to as VDAP for Venus Data Analysis Program, MDAP for Mars, JDAP for Jupiter, and OPDAP for Outer Planets), has increased the data analysis effort by funding scientists directly from NASA Headquarters. These programs are a relatively recent development and have been instituted during the late mission phases when sufficient calibrated data have been deposited in the National Space Science Data Center(NSSDC) to support the investigations. There is no restriction as to who may propose for PDAP funds. Current mission Principal Investigators may be funded as well as the scientist who has no mission experience.

While the NSSDC is the designated facility for the archiving of planetary mission data, not all data presently reside there. Many important data sets are resident only at centers or other investigator institutions. The digital imaging data reside at JPL and at the USGS at Flagstaff, while the designated archive product - hard-copy images and negatives - are maintained at the NSSDC. In fact, receipt of all high-resolution data from all NASA science missions would certainly exceed the archive capacity of the NSSDC. At the present time, NSSDC maintains catalogs of its holdings and fills requests for copies of data in its archive. A catalog of data outside its holdings is not maintained; however, personnel at NSSDC will seek to direct requestors to data held by investigator institutions when requests for such data are received.

Planetary science is multi-disciplinary. Hence, there are not only a variety of types of data to be considered but also a multitude of methods by which they are to be manipulated and analyzed. The primary data return from some experiments is a scalar or vector versus time information. Other instruments return spectra which relate to a specific look direction, while still others return a two dimensional matrix of measurements. Radio science data uses the original tracking station tapes and synopses of those data created at the tracking station. Radio science and other disciplines often employ data of relevance to planets which are not obtained in space. Earth-based and Laboratory spectra are often crucial to the interpretation of spacecraft data. Planetary data handling is not coordinated at the present time. Not all data are readily available or usable in the archived format, and some data may be lost. Information transmission is slow. Even the identification of what data are available is a lengthy procedure. Similar software is often reinvented at multiple sites and the lack of protocols for data exchange require substantial new programming efforts for most exchanges of data. It is felt that improvements in this system could be made that would readily lead to a highly cost conservative, net increase in the scientific output of the planetary community. For example, there is wide, general acceptance of a need for standards in the encoding and handling of data. While the design and specification of standards require a significant effort, they should ultimately have widespread application and should reduce the total effort and cost for supplying the community at large with similar services.

Many of the problems discussed in this report are not confined to the planetary science community, but are experienced by all users of digital data. A study of data problems related to space science in general was carried out by the Space Science Board: Committee on Data Management and Computation (CODMAC). Findings of the early portion of their study were published as, Issues and Recommendations", commonly known as the CODMAC Report. A second volume is in preparation, as a result of their 1983 Summer Study. The concepts for a Planetary Data System, developed during the Planetary Data Workshop, are in concert with the broader recommendations of the CODMAC report.

In the period, during which final preparations were being made for the publication of these proceedings, the National Research Council released a report entitled "Solar-Terrestrial Data Access, Distribution, and Archiving". This report is the result of concerns about data and data management expressed by the solar-terrestrial science community. The conclusions drawn by two communities, working independently, are strikingly similar.

1.3. A Search for Solutions

Throughout the latter half of 1982, the Solar System Exploration Division was exploring ways to respond to the recommendations of the CODMAC report, primarily: Facilitate data access, availability, and analysis within the planetary science community and involve the scientists in the solution. The steps preceeding the workshop activity are described below in roughly the order in which they were instituted.

In mid-1983 the conclusion was drawn from the various exploratory activities that changes in the current method of handling planetary data were required. The Planetary Science Data Steering Group was appointed by the Director to guide the evolution of a system for data handling which met the needs of the planetary community. That group has oversight over both the Pilot Planetary Data System (PPDS) and the Planetary Data System (PDS) efforts.

1.3.1 The Pilot Planetary Data System

Creation of a Pilot Project to explore the technology needs of the planetary community is the prologue to the Planetary Data System. Any substantial change in a method of doing business should be tested. The PDS is no exception, especially since some of the potential components of such a system are undergoing rapid evolution. A Pilot Planetary Data System (PPDS) is under development at JPL and in representative institutions. It is structured around a set of coordinated scientific tasks designed to test certain key elements in a planetary data system such as networking, optical-disk storage technology, navigational support techniques, database management, and catalog construction.

Scientific areas selected for the pilot demonstrations are: Planetary Imaging, Mars Surface Properties, Mars Aeronomy, the Jovian Magnetosphere, and Planetary Rings. Selection was based not only on appropriateness to current and future planetary exploration goals, but also because the particular studies lent themselves to the study of solutions to two important data management and analysis problems. The first problem involves the coordination of data handling and analysis techniques to support studies which require data from many instruments upon the same spacecraft, each data set having different encoding format as well as different spacial and temporal resolution. The second supports studies requiring data obtained from many different spacecraft of the same target using the same or similar instruments, again with each data set having widely varying comparative parameters. The experience gained in the Pilot effort will be inherited by the designers and implementers of the Planetary Data System.

1.3.2 The Planetary Data Survey

In order to identify the location of existing (digital) planetary data sets and to determine the nature of computer hardware and computing processes in planetary investigators' home institutions, a survey was initiated in early 1983. The data portion of the survey met with limited success, primarily because those investigators with poor inventory control on their data bases were the least likely to answer the questionnaire adequately. A follow-on effort is required which will concentrate upon locating the most complete data and documentation for each investigation flown on past NASA missions. The second half of the survey was more revealing. It indicated that a Planetary Data System, which was to utilize equipment presently in place, would have to accommodate a large variety of dissimilar computers and operating systems; i.e., the system would need to be largely device independent.

1.3.3 The Planetary Data Workshop

The next logical step in this search for solutions was to bring the affected community together to define the failures of the current system in meeting needs and to recommend appropriate changes to be incorporated into a Planetary Data System. With guidance from the Planetary Science Data Steering Group (PSDSG - which had been appointed to oversee appropriate evolution of the system), a workshop of concerned users and experts was convened from November 29 through December 1 of 1983.

The intense planning and coordinating activities which took place prior to the Planetary Data Workshop were responsible, to a large degree, for its success. Splinter chairmen were selected and, in turn, splinter groups were formed to address specific topics. These included six (6) groups which defined instrument data requirements, six (6) which addressed technical issues, and two (2) others which addressed requirements from the point-of-view of both the user and the system. Groups and Chairmen were as follow:

Workshop Chairman

HUGH H. KIEFFER
U. S. Geological Survey
Flagstaff, Arizona

Splinter Group Chairmen

Data Definition Groups

Imaging
STEPHEN R. SAUNDERS
Jet Propulsion Laboratory
Pasadena, California

Non-Imaging Remote Sensing
JOHN PEARL
Goddard Space Flight Center
Greenbelt, Maryland

In-Situ
LARRY H. BRACE
Goddard Space Flight Center
Greenbelt, Maryland

Laboratory Data
ROGER H. CLARKE
U. S. Geological Survey
Denver, Colorado

Radio/Radar Science
RICHARD A. SIMPSON
Stanford University
Stanford, California

Earth Based Observations
WILLIAM A. BAUM
Lowell Observatory
Flagstaff, Arizona

Hardware Groups

Imaging Workstations
LARRY BOLEF
Washington University
St. Louis, Missouri

Non-Imaging Workstations
RAYMOND J. WALKER
University of California
Los Angeles, California

Networking
WILLIAM D. SMYTHE
Jet Propulsion Laboratory
Pasadena, California

Software/Data Analysis Groups

Imaging
LAURENCE A. SODERBLOM
U. S. Geological Survey
Flagstaff, Arizona

Non-Imaging
RICHARD ELPHIC
University of California
Los Angeles, California

Data Base Management
RANDAL DAVIS
University of Colorado
Boulder, Colorado

System Description

Facility Configuration
RAYMOND E. ARVIDSON
Washington University
St. Louis, Missouri

Functional Requirements
A. IAN STEWART
University of Colorado
Boulder, Colorado

Splinter group chairmen met first on September 30, 1983 to plan the workshop architecture and operations and to discuss possible membership of the splinter groups. Subsequent to this meeting, splinter groups formed, met, and developed position papers which represented the issues for their topic. These papers formed the nucleus for discussions during the workshop in late November. At that time, the positions were strengthened through interaction with members at large who brought broader or different views from a wider segment of the planetary community or during interaction with other groups studying similar or related topics. In several instances, two or more groups joined together to develop continuity between related areas of study. A compilation of the basic material for these proceedings resulted from the workshop activities. A draft version was completed at a meeting of splinter group chairmen on March 6 - 7, 1984.

1.3.4 A Planetary Data System (PDS)

Solutions to problems encountered in optimizing analysis and curation of planetary science data transcend project, mission, and discipline boundaries. It is felt that all involved in the business of planetary data can benefit from changes in the system; whether their function be in obtaining data from planetary spacecraft, handling and/or documenting that data, reducing raw data or applying algorithms to calibrated data to yield scientific results, synthesizing data from many missions, or obtaining results to stretch the minds of the next generation of scientists.

A new system should be more cost effective, reduce the time interval from the conception of an idea to its fruition, extend the longevity of data sets, and broaden the dissemination of data. In short, at a fixed level of funding, the system should increase scientific productivity while preserving the data for future generations.

A well documented archive of existing and (future) planetary data, with the documentation and ancillary information necessary to make the science data useful, is considered to be of basic importance to the system. Although, the archive may consist of several levels, i. e., master or vault archive through active, working archives at the data nodes, it is neither implied nor intended that these should exist in a central facility. This system of archives should be supported by an efficient method for access, inspection, acquisition, analysis, and finally distribution of the various data holdings. A central catalog with pointers to the various data sets has been suggested as the most appropriate tool for initial access.

Users of the Planetary Data System are as varied as the data holdings. This group is anticipated to include all past and current investigation team members, Interdisciplinary Scientists, and Guest Investigators; scientists in planetary and other disciplines who analyze planetary data or use it to support other science findings (including those who receive no NASA funding); graduate students, sophisticated undergraduates, and post-doctoral appointees who use planetary data in their research; and Mission Planning and Advanced Planning personnel who require planetary data to make decisions. This document attempts to define the requirements of these various users and to suggest means by which their needs can be met.

2. PLANETARY DATA DEFINITION

2.1 Introduction

Planetary data include all of those data which have resulted from measurements made by the instruments carried aboard planetary exploration spacecraft, and (for our purposes) exclude observations of Moon and Earth. The working, planetary data base is envisioned ton contain not only these data, but also a wide range of supporting measurements such as calibration files, navigation parameters, spacecraft engineering states, and the various earth-based and laboratory measurements which provide the planetary research scientist with historical and comparative data. These various elements of a planetary data base, and the rationale for their inclusion, are discussed in detail in the Data Definition Volume of this document.

No convention exists across the disciplines of the planetary community for defining or naming the various levels through which data pass in the progression from a sensed impulse at the spacecraft to a reduced, calibrated, and/or analyzed element in a planetary data set. Terms such as EDR (experiment data record), RDR (reduced data record), and SEDR (supplementary experiment data record) imply different meanings depending on the data set under consideration. Therefore, any discussion of data levels, processing states, or data products between scientists (with misunderstanding. It is clear that the development of standard terminology for the general levels of planetary data is necessary.

2.2 Downlink Data Flow

The General Downlink Data Flow, Figure 2.1, presents diagramatically the relationships among the ten (10) data levels which were defined to be common to all planetary data. Descriptions of these levels and intervening processing steps follow.

1. SPACECRAFT DATA BUS: The format of instrument and spacecraft system data at level 1 (digital vs. analog) is determined by instrument and spacecraft conversion. The data which pass through the spacecraft data bus to be stored in the on-board tape recorder are formatted according to recorder specifications. All data passing from the Spacecraft Data Bus or the tape recorder, during transmission to Earth, undergo coding procedures. Examples of the coding conventions in wide use for transmission of planetary data are: Reed-Solomon, Golay, or Block coding; data compression, when in use, is also applied here. Spacecraft clock time tags are associated with all data and remain with the data until level 6 where they are converted to UT (Universal Time, previously GMT).

2. DOWNLINK TELEMETRY: Coded data which have been transmitted from the spacecraft prior to receipt by antennae on Earth make up the downlink telemetry stream. Time delay and frequency changes resulting from the distance and relative velocity of the spacecraft can provide valuable information on gravity fields and on media along the radio path.

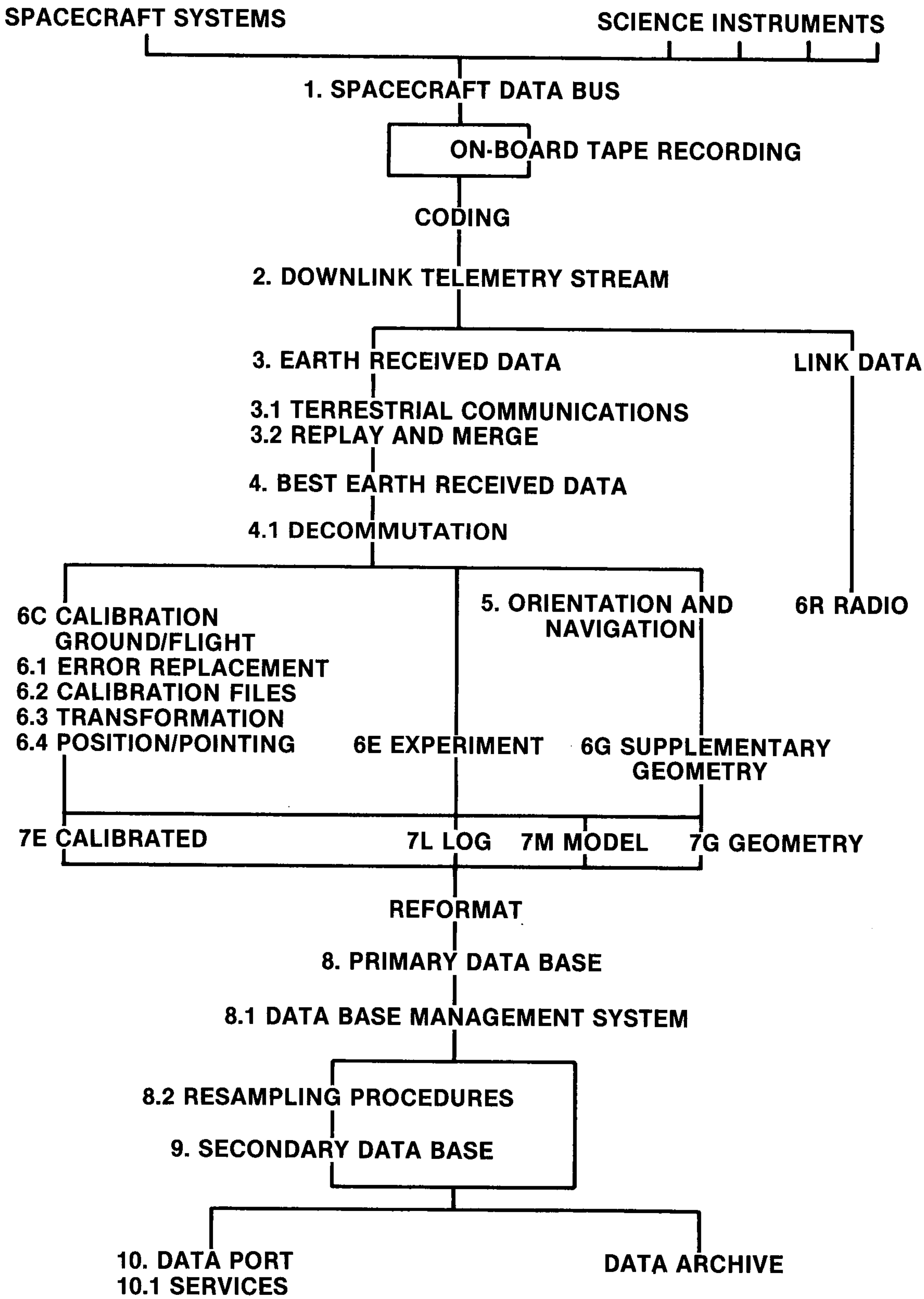

FIGURE 2.1 GENERAL DOWNLINK DATA FLOW

3. EARTH RECEIVED DATA: Coded data transmitted from the spacecraft are considered Earth-received when they have been acquired by the antenna and recorded onto magnetic tape at the ground station. LINK DATA, which provides the condition of the link (time, doppler, strength) for all received times, is included in level 3. Some portion of the link data is considered to be a primary data product for Radio Science.

Two influence factors are encountered at this data level:

3.1 Changes induced by terrestrial communications links during transmission from the station may alter the data quality.

3.2 Spacecraft replay and ground station merge activities improve the quality of poor data transmissions. A better quality data product is produced than that which usually results from any single transmission during bad transmitting or receiving conditions.

4. BEST EARTH RECEIVED DATA: Data which have been transmitted from the ground station to the mission operations data center. This is considered to be the most complete and best quality data available.

4.1 Decommutation involves the removal of coding characteristics, data decompression (if appropriate), and the stripping (separation) of the various data sets and the engineering data from the merged data stream transmitted from the spacecraft.

5. ORIENTATION/NAVIGATION: Spacecraft position, orientation and pointing information is extracted through the application of orientation and navigation processing subroutines to the engineering and link data. Relationships between the spacecraft clock and UT (light-corrected Universal Time) are established here. The experiment data and the radio data are not included in the level five (5) data level, as illustrated in figure 2.1.

5.1 Geometry Processing is the use of basic position and orientation information (level 5) to compute derived geometric relationships, such as the latitude and longitude of the center and four corners of an image.

6E EXPERIMENT: The raw data associated with an instrument. Apart from the errors and omissions associated with the data transmission, the measurements should be the same as those put onto the Spacecraft Data Bus (level 1) by the instrument.

6R RADIO: A subset (possibly complete) of link data, along with the best estimate of the spacecraft path. Data may have been averaged and may not contain phase information.

6G SUPPLEMENTARY GEOMETRY: Supplementary geometry contains geometric information about the target as well as about the spacecraft. These data have been traditionally called the SEDR. Time relationships between spacecraft clock and UT are included.

6C CALIBRATION GROUND/FLIGHT: Measurements taken to establish instrument performance parameters, whether made pre-launch to establish initial characteristics or in-flight to monitor changes in performance, are calibration data.

6.1 Error Replacement is the detection, flagging, and recording of unrealistic data spikes which are judged to be telemetry errors. A record of the original data should be kept if there is a significant chance that the spike is not an error.

6.2 Calibration files (6C) are applied to the experiment data to remove instrument generated characteristics from the measurements. Offset between spacecraft time and instant of measurement is corrected and converted to UT (UT must accompany data through levels 7 and 8).

6.3 Transformation processing is necessary to convert the levels detected by the instrument to physical units. This is especially true for interferometers and some particle detecting instruments.

6.4 Position and pointing processing involves application of Supplementary Geometry files (6G) to the data to establish the relation of the field of view to object space or to a target.

6.n This is no intended as a complete list of procedures which are applied to planetary data at this level.

7E CALIBRATED EXPERIMENT: Calibrated data in physical units represent the experimenters' best estimate of the actual values incident upon the instrument.

7R CALIBRATED RADIO: Measurements which have been converted to standard format and physical units.

7L Processing Log: A record of the software versions, calibration files, and command files (with parameters which affect calibration processing) used in production of level 7 data. This may not be practical for imaging data.

7G Geometry: The instrument - target geometry of the observation. This includes field of view location (latitude/longitude), scattering geometry (incidence, emission, and phase angles), location of possible, obstructing spacecraft structures, spatial resolutions, etc.

7M Models: Standard reference models (such as surface temperature or gravity models) are maintained in the archive in the level 7 format for application to level 7 data in performance of data analysis procedures.

7.1 Reformat procedures are applied to the level 7 data to convert it from the convention of the investigation team to the convention of the PDS. The cataloging of these data is implicit in the procedure. As the PDS matures, these conventions may become more closely aligned.

8. PRIMARY DATA BASE / Calibrated and Associated: The primary data base consists of level 7 data (calibrated experiment, geometric, and associated model data) which are cataloged and formatted to PDS conventions. The full measurement resolution of the instrument is maintained and the order is chronologic. The Primary Data Base may also contain derived values based upon the measurement data, e.g., ozone abundance based upon a reflectance spectrum.

8.1 The data base management system (DBMS) includes the data base functions by which the data are stored and the catalog functions by which they may be accessed and examined.

8.2 Resampled data are those which have been altered (geocoded, rectified, filtered, smoothed and/or averaged by further reduction or by processing so that a one-to-one relationship with instrument measurements no longer exists.

9. SECONDARY DATA BASE: Data which have been resampled, mapped, or significantly rearranged in collections (e.g., consortia) are contained in the Secondary Data Base. Catalog and format requirements will be similar to those for the Primary Data Base; however, the arrangement of the data is not necessarily chronologic.

10. DATAPORT/ARCHIVEPORT: The gateway to the Planetary Data System through which its contents are shared with the Planetary Science Community. Access to catalog information and data contained in the Primary Data Base (8) or the Secondary Data Base (9) is through the DATAPORT. The ARCHIVEPORT provides access to levels 6 and 7 data in the PDS archive. Access through the latter, however, implies a high level of system and processing proficiency on the part of the user. Connection to either port will be by terminal (through appropriate communications channels) from the user's home institution or from within one the the Discipline Centers. Data transmission by this mode will not alter the format of the data record nor will access alter the master files in the PDS.

10.1 Services include the software and/or software applications which enable access, manipulation and display of the data and help for system users.

10.2 Through n. Analysis Tools are software applications available at the DATAPORT for use with on-line computational devices or home institution computer.

NOTE: Through this report, when data terms are used without qualification, the following is implied: Raw data, level 6; Calibrated data, level 7 (available in the (8) Primary Data Base); and Resampled or mapped data, level 9, Secondary Data Base).

2.3 Data Content Requirements for the PDS

2.3.1 Levels of Processing

Data, at the following processing levels, should be available through the PDS for each planetary spacecraft, and for each instrument included on the spacecraft payload. Radio, laboratory and Earth-based observation data may vary in requirement according to the data set and the intended use.

Raw data: 6E, 6R, and 6G data are maintained in the archive and may be accessed or ordered along with the documentation and software necessary for their use. It is anticipated that there will be limited requests for these data sets; however, their value is high as a record of planetary conditions at a specific time and place and as a basis for comparison with later measurements.

Calibrated data: 7E data represent the best estimate of the actual flux incident upon an instrument (7R data are of a similar level). These measurements are represented in conventional physical units and are ordered chronologically. Associated data sets, 7G (geometry) And 7L (processing log), provide the orientation and reduction history of the measurements.

Resampled data: Level 9 data have been resampled in time, space, and/or a spectral dimension; are mapped; or have been organized into collections (consortia) and, as such are more readily compared with data from other instruments. Various procedures such as rectification, geocoding, filtering, smoothing, and averaging alter the data set so that a 1:1 correspondence between the instrument measurements is no longer reflected. Data, so altered, are not necessarily in chronologic order.

2.3.2 Completeness

The level 6 data sets should include all times when the instrument was operating and data were returned. Level 7 data sets should include all non-redundant data which contain scientifically useful information. The guidelines as to what information and measurements are useful, and are to be included in the level 7 processing and maintained in the level 8 or 9 databases should be determined by each PI. These recommendations should then be reviewed by a peer group such as the mission science steering group and additional experts as appropriate. Decisions as to what data will or will not be included will have a permanent effect on the utility of the PDS.

Additional data sets of direct relevance to planetary missions or to spacecraft experiments should accompany the planetary data in the data bases and archives of the Planetary Data System. Suggested for inclusion are the results of ground-based, suborbital, and earth-orbital observations, and laboratory measurements of materials likely to constitute the surface, atmosphere, or charged particle environments of planets and satellites.

2.3.3 Data Compression and Reliability

As the processing level (of data increases), data of uncertain reliability or limited usefulness should be progressively excluded from a given data set. Considerable judgement is required at each stage of data compression so that confidence and utility of a data set increase as the volume decreases in the progression between level 6 and level 9. Large volumes of data which are not useful for the purpose implied, such as satellite or ring searches which contain no information, calibration or cruise sequences taken of deep space, or mapping images of Mars taken during the dust storms which occluded the surface, need not be processed to level 7 status. However, the raw data and all information necessary to accomplish such processing should be submitted to the archive.

Prior to level 8, data of poor quality should be removed. These might include poorly exposed images or remote sensed data acquired at high emission angles through the atmosphere. It is of particular importance that level 9 (resampled) data be selected carefully to include data of uniform quality.

A consistent philosophy should be developed by the PDS as to the treatment of errors or spikes in the data. At present, data judged to be in error are eliminated, flagged (set to 0), or replaced by a "benign" valued based upon adjacent measurements. A log should be kept of all replacements, as well as the original value if there is a finite possibility that the data could represent a real, albeit unexpected, observation. To aid in this judgement, the log should note the probable cause of the erroneous measurement, i.e., telemetry spike, the instrument is known to have been drifting badly, the spacecraft or instrument were in an uncertain state or were unusually noisy.

There is also an increasing need for information on the reliability of data values as persons with less expertise in a particular discipline access data with which they are unfamiliar. General science users may not be aware of many kinds of data uncertainties, nor of their signatures which would alert a familiar user. In order that all users can proceed with confidence, data validation should be done as appropriate for each data level. Questionable data, so located, might be tagged with a level of uncertainty in data sets of level 7 or less and omitted from the levels 8 and 9 databases.

2.3.4 Longevity and Accessibility

All data at level 6 (raw data, calibration files, and geometry) should be archived indefinitely, under ideal conditions. Due to the fact that the content of these data is largely redundant with higher level of expertise for access. The Primary Data Base containing level 7 data and the Secondary Data Base of resampled level, 9 data should be readily accessible through the PDS dataport, and should remain so As long as they remain actively in use. As data are used less, or not at all, they should be moved to the archive and then to increasingly deeper levels of that archive to allow space for more active data sets. When compromises of significant impact concerning with an inter-disciplinary peer group. Once entered into the Planetary Data System, a data set will remain there, and will remain accessible (at some level) indefinitely.

2.3.5 Update and Change

The detailed knowledge of an investigation and instrument operation makes it unlikely that others, with exception of the original investigation team, would be able to produce an instrument calibration procedure or calibration files which were significantly better than those used during the active mission. It is common, however, for geometric knowledge to change with time through development of improved ephemerides for the spacecraft and improved pointing derived from the measurements of scientific instruments themselves (e.g., limb-crossings or feature identification using images). Geometric recalibration, therefore, can continue over a period of many years after data are acquired. Pointing improvements often apply to many instruments, especially when they are located upon a common scan platform. Major changes which could affect vast quantities of data in the PDS include the redefinition of the cartographic system of a target body due to revision of the orientation of the rotation axis, or a change in the radius or shape of the body; or the forthcoming update of the ephemeris from 1950 to J2000 coordinates.

Provision must be made to incorporate change into the PDS data bases. Where it is appropriate, altered data or conversions to alter the data upon access should be entered into the system, otherwise notation should be tagged to the erroneous data values. Complete documentation should be included to explain the phenomena.

2.3.6 Documentation

Several levels of documentation should accompany the submission of data into the Planetary Data System. Documentation at the mission level, the instrument or investigation level, and the data set level are important and should be required. Older data, grandfathered into the system at inception, should be documented to the extent possible.

A brief historical summary for each mission should accompany data submissions from that mission. This should include the spacecraft definition, mission description and trajectories, science rationale, and listings of science team members.

The investigation should be similarly documented, with the focus being placed on the instrument characteristics, measurement capabilities, modes of operation, and sequence rationale. A narrative discussion of instrument calibration and any know instrument peculiarities should be included. Publications resulting from the analysis of observations should be cross referenced and listed into a mission/instrument/investigator bibliography.

At the data level, the system should allow on-line access to information concerning parameter definition, derivation algorithms, and the types of calibration data used. The archive system should contain (and have available upon request) the actual computer programs used in reduction and processing, the calibration files, and processing logs used in the conversions of data from level 6 to level 9.

2.3.7 External References

Pointers to detailed discussions of both investigations and analysis results in the open literature should be included in the PDS catalog. It was also suggested that the inclusion of lists of current, known analysis activities and associated science contacts might avoid duplication of effort and/or unnecessary processing. As analysis were completed, these descriptions could be moved to a bibliography. Directly related measurements, such as those obtained from Earth-based observations or other missions and having similar or overlapping measurement capabilities, should be noted and cataloged for reference.

3. USER REQUIREMENTS

3.1 Introduction

Summarized in this chapter are the capabilities, enumerated during the Planetary Data Workshop, which the ideal Planetary Data System (PDS) should provide to its users. On occasion, elements of these required capabilites reach "wish list" proportions. No attempt was made during the workshop nor in subsequent activities to discourage this approach. System design and implementation activities must take into account that all requirements do not carry the same "weight". While some system requirements are essential and must be met early, other requirements are of a less immediate nature and can be implemented during later stages of development. Some requirements may be determined to be not essential to the operation of a PDS or impractical, and then not considered for implementation. Issues of practical implementation and compromise are deferred for discussion until Chapter 4.

It is assumed, for purposes of this document, that the PDS implementation will be modular in architecture and accomplishment. The modular plan is supported by many potential benefits in that it: simplifies the design and planning activities; improves the ability to respond to the changing demands of the user community; permits incorporation of emerging technologies; and allows the implementation to remain flexible in the face of fluctuating resources.

3.2 Uses and Users

The initial objective of the Planetary Data System is to support the active research scientist in the performance of various activities necessary to identify, access, and use planetary data. Therefore, all levels of familiarity with the data must be anticipated and accommodated, as well as all degrees of expertise in using the system and its services. The unsophisticated user has enormous impact upon the initial access into the system in terms of the help files necessary to educated him in system use and upon the extra levels of catalog functions to support his searches. The competent scientist with limited institutional computing capabilities places heavy computational requirements on the system; while heavy distribution requirements for data to be analyzed at the home institution stress the system in an entirely different way.

Users can be categorized in terms of their traditional relationships to NASA in general and with certain mission-specific and discipline-specific activities. This information is useful in considering implementation sequences, computational requirements, and management structures; and are discussed in this context in Chapter 4.

In terms of impact upon the functional design of the system, the identification of four kinds of users according to their familiarity with the specific data set (with which each will be primarily associated) is a more meaningful categorization. These four user groups, arranged in order of decreasing familiarity are as follow:

1. SUPPLIERS are scientists or teams who contribute basic data sets to the system. They will have in-depth knowledge of a particular data set and will require the least help in using those data. If, in the role of supplier, the individual becomes an Investigation Node, significant requirements may be placed upon computational and catalogue services to support this function. Members of this category may require assistance in preparing data and documentation for inclusion in the PDS.

2. OTHER EXPERTS are scientists having extensive knowledge in the field addressed by a specific data set - or in the case of the interdisciplinary scientist, in the object of investigation by many data sets - but will not function as suppliers of that data. Members of this group may require both raw and calibrated data, and may wish to apply calibration procedures to the raw data. Thorough documentation of instrument characteristics, calibration procedures, and mission rationale will be required by this group of users. Dependent upon their location, these scientists could place heavy demands upon the computational services of the system.

3. The GENERAL RESEARCH category includes scientists who are addressing related science problems, surveying the results of others in the field, or graduate students who are applying detailed results of an investigation to their own theses. NASA Mission Planners might also be included in the group as their data requirements are similar to the general research scientist. Access will be required to mapped, derived, or otherwise consolidated data, and small portions of many of these will be used. Thorough documentation of the uncertainties and limits of interpretation associated with the data sets will be of great importance to this group of investigators. It is anticipated that this group will rely heavily upon the "catalog services" and distributions of small quantities of data to support their studies.

4. EDUCATIONAL USERS are teachers seeking knowledge of the results of planetary investigations and students performing research projects. The potential range of user skills is very large including persons with only a modest knowledge of computers and planetary missions to advanced undergraduates with considerable sophistication and skills. The major requirement will be for mapped data, as well as non-digital products which are not considered here. For this reason, and because the less knowledgeable user places significant impact upon system requirements in terms of the introductory help required, the educational needs are considered to be ancillary to the needs of the research scientist. Accommodation of these users may be developed later in the implementation plan. It is felt, however, that support for incorporation of features designed specifically to handle these user demands should ideally be provided by the Public Information and Educational programs.

3.3 Data Requirements

The Planetary Data System (PDS) should function both as a tool for active research and as a permanent archive. It should contain all data and resulting data products from NASA's planetary exploration programs, as well as the documentation necessary to make those data useful. The data, data products, and documentation are discussed in the previous chapter and in the detailed data descriptions which follow in later chapters. Guides to the use of the data, use of the data system, and to locations of related data in other archive systems (not planetary in origin) would benefit research and researchers. The inclusion of an index of prior and current analyses of data sets and developments in data analysis and data handling procedures, would reduce duplication of effort in data analysis and software development.

Data should be available at several levels of processing and/or compression. Levels 6 through 10, as described in Chapter 2, were felt to encompass the data necessary for inclusion in the archive to meet the varied needs of the defined users. All data sets should be thoroughly explained and the derivation documented. As the level of processing or compression increases, data should be more tightly screened for reliability and accuracy upon submission for inclusion in the system. Provision should be included to update data bases using improved information or to include associated variables and additional parameters; these updates would be documented and include statements of impact and rationale.

3.4 Data Access Requirements

A user must be able to determine, easily, whether data useful for his purposes exist and retrieve them for analysis. Interactive access to the actual data sets was deemed less important than the sophisticated procedures which would lead the scientist to the location of data which fit within his research parameters, allow him to examine these in increasing detail, and finally select for delivery (on-line or otherwise) those data which are most appropriate to meet current analysis needs. Approached from this point of view, three groups of functions support data access by the planetary scientist and should be included in a PDS: 1) catalog, 2) examination, and 3) acquire.

3.4.1 Catalog

The catalog will be the initial entry point for all levels of users into the data system. Requirements placed upon the catalog functions were that they be on-line and interactive, that entry levels into the catalog will vary from system novice through system expert, that the level of access, security, and use of services will be controlled through "smart" sign on procedures into the catalog, and that all data and services available in the system will be identifiable through catalog access.

The three functions supported in the catalog follow. As these services are key to the operation of a Planetary Data System, high priority is placed upon their development early in the PDS implementation.

1. The ENCYCLOPEDIA is a general guide to the Planetary Data System. A Hierarchical structure is indicated which includes: a) Solar system entities such as planets, minor bodies, solar wind, etc., b) Missions (when, to where, how long, science objectives) and spacecraft (class, trajectory/orbit/lander site, payload); c) Instruments (objectives, descriptions, measured quantities); d) PDS Services and Facilities (where located, how accessed, kinds of data included in data bases)

   Two encyclopedia access modes are desirable: 1) a hierarchical, menu driven mode, and 2) a key-word driven, subject oriented mode in which the user is guided to experiments, data sets, or data products which relate to a specific topic. Both modes should permit direct return to the desired point at a later access.

2. The DATA INDEX will be the primary entry point for the active research scientist. Data and information contained in the archive are listed and described to a level sufficient to permit identification of data sets which might be appropriate to meet the researcher's objectives. The description for data from each investigation should include an instrument overview, temporal and spatial coverage, data levels of products available with data quantities and qualities indicated, and resource contacts.

3. The DATA INVENTORY expands the descriptions in the Data Index to allow an investigator to pinpoint the observations appropriate for analysis. Elements such as cartographic coordinates, surface resolutions, instrument states, and illumination angles were suggested for inclusion. The range of each possible search parameter for a reasonably continuous set of observations, i.e., an imaging mosaic, a radar ground track, or a mission phase, should also reside in the inventory.

3.4.2 Examination

The next set of services support the examination of data located through the catalog search to assure that the observations are, indeed, appropriate to satisfy the specified conditions set in the research parameters and that all available data have been presented for evaluation. The three levels which make up the examination services are arranged in order of increasing depth. These extend from a cursory examination of specific observations and related materials through the capability to obtain samples of the data for testing data analysis procedures.

1. BROWSE permits the examination of data at a level of detail equivalent to leafing through hard-copy. It should allow the user to skip over portions of the data and then look at detailed characteristics of some subset of that data. The user should be able to designate a step-function, halt the progress to observe some feature in detail, and then resume the function. The capability to "mark" observations for order or further study should be supported.

Access should be possible both directly and from within the catalog functions. The result of browse should be a description and/or a display of the data at the direction of the user.

2. SEARCH allows the user to conduct a detailed search for data contained in the Planetary Data System which satisfy user supplied conditions (temporal, spatial, or target specific.) Application of this function assures that all information regarding a certain subject has been located. Used in conjunction with the browse function, this should provide a powerful tool to the scientist performing interdisciplinary research in the identification of overlapping observations from many instruments. It was recommended that the first response to a search request should be a summary of the number and size of records found so that unexpectedly large, or small, data returns may be aborted.

3. SAMPLE provides the user with the capability to acquire a representative, small sample of a data set which may either satisfy his specific data requirements or may be used as realistic test material for development of analysis procedures. Additionally, some researchers find it difficult to judge the appropriateness of a data set prior to making a "hands-on" examination. Both the search and browse functions should provide sufficient data identification to formulate a request for sample data. When practical, the sample data should be available on-line to remote users.

3.4.3 Acquire

The Acquire Function enables the user to request data to be transmitted either electronically or shipped after encoding on a transfer media such as magnetic tape, magnetic or optical disk, etc. The procedure should be straightforward, well defined, and enabled while on-line to the system. The first response to a request should be an estimate of cost (if any), the quantity of data involved, and the approximate time required to fill the order. Where alternate modes of delivery are available, these should be given (including the time required for preparation and delivery) with the final choice of transport being selected by the requestor.

There should be a less structured means than from an on-line function, by which requests for data, documentation, or HELP can be placed into the system. It is thought that many of the requests for deliveries of large quantities of data or documentation would be more appropriately handled in this way. This would also reduce the load on the central functions during periods of heavy use. A mailbox served by an on-line communications service was suggested as a possible alternative.

3.5 Processing Requirements

Several kinds of basic data processing support are required and are described below. Software applications to perform many of these basic tasks are available within the community and in use at the current time. Software to permit many of the larger comparative and manipulative tasks suggested would require large software development efforts. Such developments would foster broader use of the data and could consolidate many individual software development efforts in the community. Prior to inclusion all analysis tools must be thoroughly documented in terms of the algorithms and limits of their validity. Decisions upon the inclusion and validation of analysis software in the PDS should be a function of the peer review group.

Products from these operations will be stored in a temporary, assigned storage space and will not influence the contents of the controlled data bases.

3.5.1 Data Manipulation

These procedures are generally defined as simple algebraic or trigonometric operations performed upon one or several data parameters to generate a new parameter. Manipulation in its simplest form should permit (by subtraction, division, etc.) reference of one set of measurement values to those from related measurement or to values obtained from standard models. More complex manipulation procedures would support a series of operations which combine values of various parameters to generate a dummy parameter, which in turn could be manipulated in the same manner as any other parameter in the data base.

3.5.2 Statistics

Statistical procedures should support such calculations as the determination of mean and standard deviation of an individual variable for a portion of the data set or of a subset resulting from a search of a data set or data base. Calculation of cross correlation and frequency of occurrence should be supported in at least two dimensions.

3.5.3 Registration/Resampling

Capability should be provided to average data in time, space, or energy domains to support analysis of low signal-to-noise data; and to resample data from one experiment to uniform increments of time or to the same time increments as data from another experiment. The system should also support resampling of non-uniformly spaced data onto a uniform grid (such as latitude and longitude), the linear scaling of image data, and the registration and comparison of that scaled image with another which has been similarly scaled.

3.5.4 Graphic and Image Display

Many kinds of data are studied most effectively in graphic form. It is, therefore, necessary that the system provide plotting capabilities. Examples of this capability include plotting of: points or lines which indicate the relationships between any two variables in a data set; latitude versus longitude in a coverage diagram; blue versus red intensities to display different material abundances; and draped mesh displays of relationships among three variables. Default to automatic scaling and axis generation is a desirable attribute of any plotting routine. Display of imaging data is imperative. Currently, monochrome display is adequate, but color display should be considered as a requirement for the future. It must be possible to view the entire image (even at the loss of resolution), and to view selected portions at full resolution. First-order quantitative information should be made available, such as: cursor position in both line and sample of the image file and latitude and longitude in surface coordinates; image intensity; and distance (Km) between two cursor locations.

3.5.5 Analysis Tools

Interpretation of observations commonly depends upon comparison of data to models (atmospheric radiative-convective or surface diurnal temperature models for example), analysis of time or energy spectral properties (position of a molecular absorption band, periodicity of an electric field, or gravitational spherical harmonics), or reduction of the data to a physical property of the object (Minnaert albedo or cloud mean particle size). These computations vary greatly in sophistication of the algorithms and in the computational power required for their performance. Potential analysis software (to perform these computations) ranges from a relatively straightforward generic code, such as determination of spherical harmonics, to model-dependent codes which make assumptions about the composition or physical state of the environment. Development of software routines to support these operations would allow broader use of the data in the PDS; however, the development of software is often a lengthy and expensive procedure. Decisions as to the development, inclusion, and validation of analysis software should be a function of a peer review group. All analysis tools, available for use on the PDS should be thoroughly documented as to the terms of the algorithms and the limits of validity.

3.5.6 Recalibration

Recalibration of data may be desirable when there is improved knowledge of instrument behavior. This generally occurs during the analysis of later inflight mission data and is particularly likely to affect data sets shortly after their submission to the archive. In these cases, the revised, basic calibration data and all routines necessary to apply them should be made available to PDS users and should be executable on the host PDS hardward. The explanations and supporting documentation can assume a high level of user expertise.

3.6 User Environment

The benefits afforded by a system such as the Planetary Data System depend significantly upon the amount of effort required to use it. To be effective, procedures must accommodate users with a broad range of expertise; and a high degree of commonality must exist in the use of various data sets. That is to say that access to each kind of data must not require a re-education process. The following requirements were placed on the system in terms of the face the system should present to the user.

3.6.1 Uniform Appearance

The key to system efficiency was thought to be that the system appear similar during the performance of any data service, disregarding the mode of access. Procedures for establishing communication (remote or from a discipline center) and means for accessing the catalog functions, data bases, and computational services should be similar for all institutions and for all data sets. Initial access, help, and requests for standard services should be standard and be supported by on-screen prompting.

3.6.2 On-line Help

Services supported for remote users should have on-line, interactive help available, preferably in a hierarchical form. The users to be supported are expected to change with time; and tutorials may be required for new and less-sophisticated users. Means should be provided whereby experienced users can inhibit the prompting procedures.

3.6.3 Off-line Help

The implementation of a Planetary Data System is, by no means, considered a replacement for human interactions in the analysis of planetary data. The concept of the Discipline Center and the Investigation Nodes is developed around the need for help and the nurturing of scientist-scientist activities. The on-line help system should point to off-line help sources when addressed by remote users; and a mailbox where users can place inquiries and requests which require human attention, was suggested as appropriate. In addition, hard-copy documentation should be available either through remote printout or by mail.

3.6.4 Resource Requirements Forecast

Requestors of system services should be informed in advance of the implications of a request; such as the amount of data product to be delivered, the cost (if any) to support the activity, the time involved in execution of a program or in the preparation of a distribution. Possible means for alerting the user were suggested: 1) The on-line help system can include approximate guides to resource requirements; 2) Individual routines can forecast the magnitude of processing based upon the parameters of a specific request; 3) The system could provide notification of the rate of progress on a specific request, the volume of the product, or the expended resource units such as dollars or CPU seconds (the ability to inhibit this response should also be available).

3.6.5 Display and Hardcopy

The use of graphed or mapped data, displayed on the terminal CRT is employed frequently in planetary data analysis in the search for data and in the assessment of the appropriateness of data to the research problem. Volatile displays of data are then used to check progress and results. Hard copy is usually requested to support protracted study or the dissemination of results. Both types of output should be available upon request; and each should identify the data displayed and the scale of the display.

3.6.6 Large Data Requests

A user should be able to obtain digital data and supporting information for analysis at his home institution. Upon request, the data should be transmitted to him through the most appropriate channel - determined by the size of the request and the immediacy of his need. The means of delivery available for transmission will grow as technology is perfected to support them. Suggested were: magnetic tape, disk media (optical, video, and audio) to be sent by mail or package delivery; or electronic transmissions such as those supported by packet telemetry networks. A mailbox supported by order-type prompts was suggested as a convenient mechanism for placing these requests.

3.6.7 Software Transport

Software resident within the PDS (with rare exception of that considered proprietary by the developer) should be available to all users in machine readable form. Transportability was stressed. It is desirable therefore, in the development of PDS software, to minimize the use of machine-specific capabilities (vendor specific enhancements to standard languages) and to insure that all hardware dependent code (machine language or peripheral device drivers) be isolated in modular routines which are documented adequately. Customization of these modules can be readily accomplished upon receipt of the software.

4. CONFIGURATION AND IMPLEMENTATION

4.1 Introduction

The main objectives of a Planetary Data System (PDS) are to curate planetary data--that is to store and maintain complete planetary data sets and the necessary supporting documentation to make them useful--and to facilitate scientific study of these data through improved organization and access. A more specific goal is to organize and document these data in such a way that in-depth knowledge of a data set is not required in order to use it. Use of data by interdisciplinary scientists or scientists performing interdisciplinary studies, for whom familiarity with all data sets required is impractical, will be facilitated in this way.

A hardware system, no matter how sophisticated and supportive, cannot replace the live colleague interaction needed for good science. System users with only a general knowledge of a data file or data set will need guidance in determining whether that data is most appropriate to support a theory. The PDS is meant to allow easy access to data, to provide useful presentations of that data, and to facilitate data analysis; however it is not designed to supercede the critical examination and interactions that data analysis requires.

The Planetary Data System will be a powerful research tool which allows the entire planetary science community access to a complete data base and will facilitate the archiving of past, current, and future mission data sets. The PDS should be designed to accommodate changing needs and evolving technologies in the solution of data organization, storage, access, and distribution problems.

4.2 System Structure Considerations

The need for effective communications across the planetary science community, the need to develop standards for all aspects of data handling, and requirements for long term continuity and permanent archival of a national treasure, all indicate the need for a dedicated facility. On the other hand, the active research community must be provided with working data sets presented in a flexible environment conducive to the performance of scientific research activities. Meeting the latter objective is felt to indicate that data should be left in the hands of the scientists who are actively performing research upon that data, and that existing facilities and equipment should be incorporated into the PDS in a networked, distributed environment of universities and research centers. The system proposed here combines these elements and attempts to resolve the opposing requirements for a secure data set and a flexible environment conducive to research.

A several tiered structure has been proposed for the Planetary Data System. A central catalog with several layers of search and identification capabilites would direct the user to the appropriate data, discipline center, and/or services necessary to meet his research needs. Figure 4.2 * demonstrates the tiered structure and the proposed networking scheme used to access and coordinate the functions of the PDS.

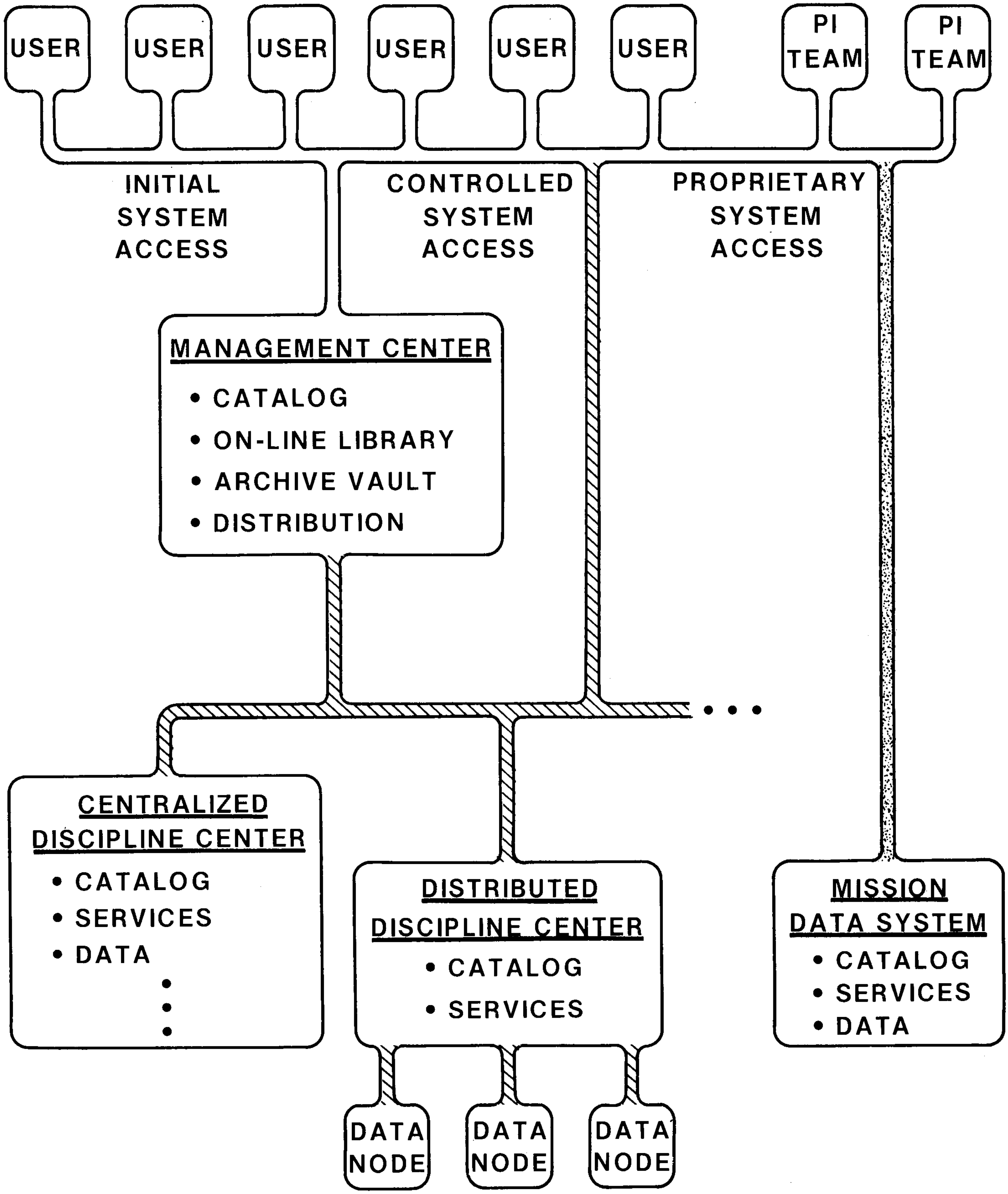

FIGURE 4.1 PRELIMINARY PDS FUNCTIONAL CONFIGURATION

4.2.1 Central Management Center

Many functions of the PDS are general in nature and should be centralized. These functions should be brought together and conducted from the proposed PDS central site, which should have as its primary purpose the long-term archiving of planetary data. The specific responsibilities mentioned for the Management Center include:

Management (funding) and coordination of active data base sites.

Format control of data, documentation, user interface etc.

Database administration

Communications network control

Central catalog, access to the system and all other catalogs

Flight project interface

Maintenence of historical archives;
- Primary for raw data and inactive processed data
- Primary for documentation
- Primary for navigation and other supporting data
- Redundant for active working data bases

Processing of large data requests (such as mission data set distributions)

Flight project interface

Community liaison

4.2.2 Discipline Centers

Planetary science has traditionally been divided into a number of active disciplines made up of scientists who work on closely related problems. Interactions between disciplines are commonly formal, while intradisciplinary interactions are detailed, working-level relationships. The concept of Discipline Centers has been proposed to coordinate data handling and analysis activities for a related group of science users. As such, Discipline Centers would serve two purposes: 1) to coordinate the activities relative to a number of data sets; 2) to provide nuclei for research activities.

There was disagreement and much discussion as to the boundaries which would contain a Discipline Center. Divisions which are instrument-based cross science discipline boundaries, while those based upon science disciplines fragment instrument boundaries. There will be no ideal division of planetary science into Discipline Center jurisdictions, nor will there be an ideal grouping of data sets. The desired nature of the PDS is that the location of any given data set will be invisible to the user--that is its access will not depend upon its physical location. This is fairly easy to accomplish, given the sophistication of network communications technology. The greater problem is a personal one which relates to: Who has responsibility for what data in what center and who provides the liaison between the user and the center?

The design of the PDS access, help, and documentation system should allow the data to reside in locations independent of any specific grouping chosen. One proposed division for PDS Discipline Centers is as follows:

Science Discipline Center	Pertinent Instrument Data
Aeronomy	UV Spectroscopy, Photopolarimetry, Laboratory reaction rates
Atmospheric Composition	High resolution IR Spectroscopy, Mass Spectroscopy, Laboratory Gas Spectroscopy, Cloud Measurements
Atmospheric Dynamics	Venus and Outer Planets Images; Lander or Probe Meterology, Entry Dynamics, and Radio Occultation
Planetary Geology	Solid-body Images by TV Camera, Imaging Spectrometer, and Radar; Lander Images
Surface Properties	Thermal Radiometry, Radar, Low-resolution Reflection Spectroscopy, Gamma- and X-ray Mapping, Landed Physical Properties Seismology, and Laboratory Solid Spectroscopy
Magnetospheres	Magnetometer, Electric Fields, Charged Particle, and Neutral Particle Instrument Measurements
Interplanetary Medium	Radio, Navigation, Meteroid Detector, Gravity Fields Data
Earth Observations	Ground, ballon, rocket and aircraft borne imaging and spectroscopy

It is strongly recommended that each Discipline Center be associated with an active research group. These centers may adopt either a centralized or distributed configuration.

A centralized Discipline Center would be self contained. It would contain all pertinent data sets and supporting information, would provide catalogue and search capabilities, aid in the interpretation and use of the data in residence, and have available for use those facilities which are appropriate to perform research in the specific discipline area.

The distributed Discipline Center implies that some or all of the data sets would reside at locations other than in the center, but that all of the services of the Discipline Center would be available within the networked system. Physically, the center would consist of a central location and Data Nodes (of two types) networked in such a manner that on-line access to the data would appear as if the data resided in the central file. Operationally, the difference would be significant.

It is felt by scientists that, for many of the more active or changing data sets, it may be most effective (for the scientist who generated the data and the scientist who wishes to use the data) for the data to be maintained where it is understood. Scientists who retain released data sets in a Data Node would be responsible for the management of the data set and for providing the various "help" and service activities necessary to promote effective use of the data by others. Data Nodes of two types have been defined.

Investigation Node - Data sets associated with active missions may reside in an Investigation Node for a variable period of time following release from proprietary status. The prime responsibility for the data set, documentation, and all supplimentary information necessary to use and analyze the data would remain with the Principle Investigator or Team Leader. It may be desirable for the data to be entered directly into the Investigation Node during mission acquisition, with access limited to Mission Scientists during the proprietary period. This scenario provides two majcr benefits: 1) services of the PDS would be available to support Mission Scientists and to transmit data between investigators and to make large data distributions; 2) the data would have already been placed in the desired format for distribution, properly catalogued and documented, so that further preparation for submission to the archive would be unnecessary.

Analysis Node - Data sets which result from extensive processing or from combining data from many sources may reside in an Analysis Node. Consortia and digital imaging mosaics are examples of these data sets. Scientists who have generated these are the most knowledgeable in their use, although they may have had no formal relationships with the missions from which the data resulted, and should have prime responsibility for the data sets. The obligations for documentation, catalog entry, help, etc. are the same as for the Investigation Node.

4.2.3 Communication/Distribution Network

The fact that communications technology will provide the method by which data are transported from center to center or center to user is given; the scale and sophistication of the communications network will depend upon the state of the technology available at the time of implementation and the existing communications links available for use by the PDS. Both are expected to evolve significantly on a decade time scale.

The method of distribution of large volumes of data will always be determined by financial considerations. Mail and package delivery of high-density encoded data will remain the mode of choice until electronic transmission becomes significantly less expensive. Advancing optical, digital disk technology will serve to make package delivery more attractive. It is less

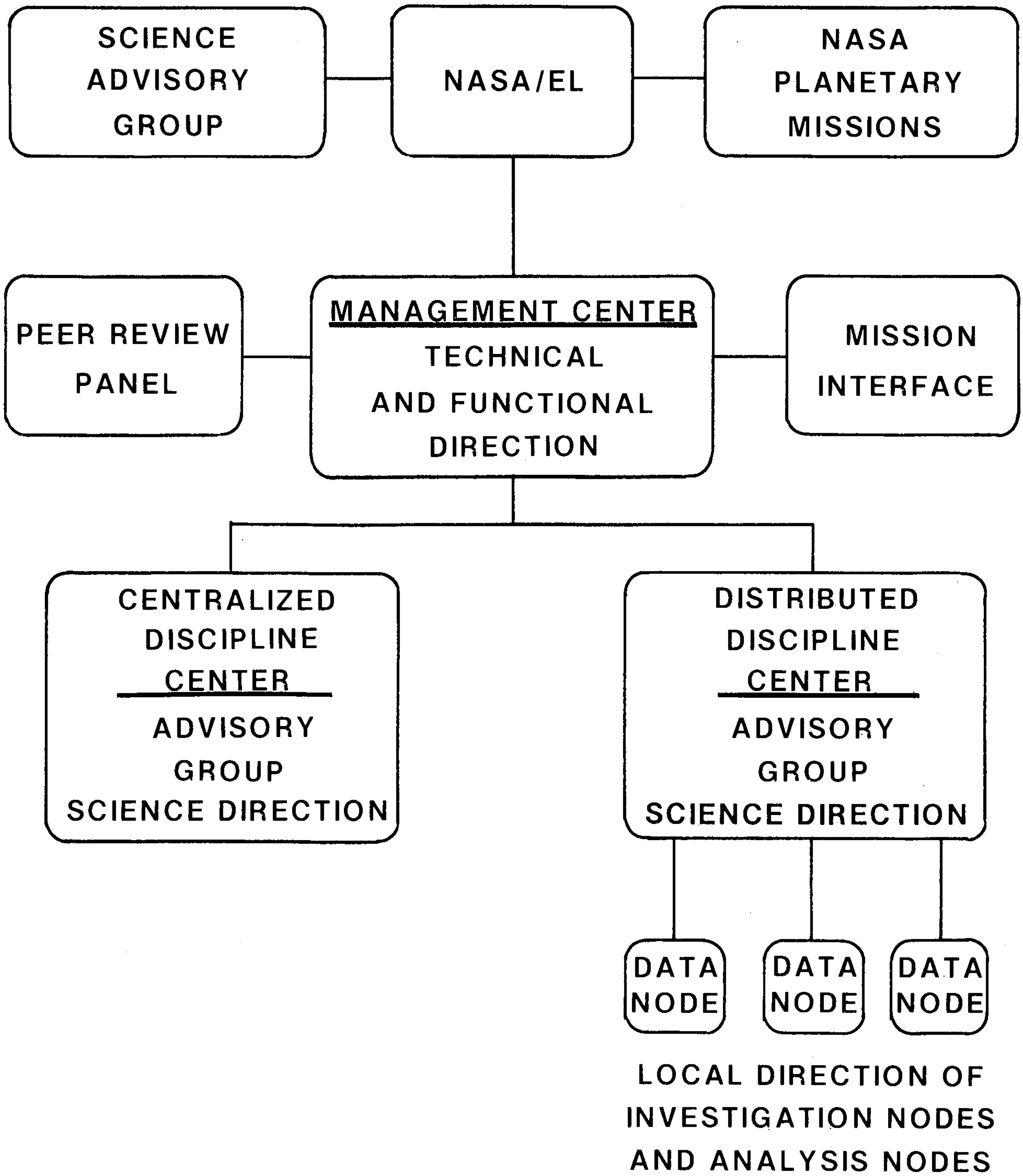

FIGURE 4.2 PRELIMINARY PDS MANAGEMENT CONFIGURATION

important which methods are used than that some affordable system exist. The PDS should prudently use the economy of scale to provide rapid data transport that is most cost effective in accomplishing the science tasks.

4.3 Management Consideration

Concerns were expressed during the workshop as to the management and control of the Planetary Data System. The issues discussed ranged from concerns about equality and abuse, through ensuring continued scientist direction, to system operations responsibilities and decisions. Figure 4.1*illustrates the tiered management structure and demonstrates the provision for science interaction at the level of each tier.

4.3.1 Science Advisory Committee

A committee of planetary scientists should be formed to advise NASA on issues of policy and progress in meeting goals for the PDS. This activity could initially be performed by the existing Planetary Science Data Steering Group (PSDSG); but as the oversight becomes more specific, a formal advisory group was thought to be necessary. This committee should advise and guide the Solar System Exploration Division in developing working relationships between the Planetary Data System and other NASA entities, particularly the Mission Project Offices, Mission Science Teams, Research and Analysis Programs, and Data Analysis Programs. In addition, the committee should monitor the function of the entire Planetary Data System to assure that the needs of the community are being served.

4.3.2 Management Center

The hierarchical arrangement of management centers, discipline centers, and data nodes is unfortunate and does not imply rigid reporting structures and funding constraints. A ringed structure, graphically more difficult to understand, would perhaps have been more appropriate. The management center performs those tasks which are more easily performed centrally than distributedly. The Management Center Advisory Group is envisioned as made up of Discipline Center Representatives so that the function becomes more coordinational than managerial. The data nodes are extensions of the discipline center of which they are a part, rather than separate entities under the control of a central organization.

A major requirement to be placed upon the system is that the management center, though it must remain highly visible, should be constrained so that it represents a very small part of the PDS operational system and budget. It was felt that appropriate administrative procedures must be used to insure that the management center does not grow to "facility" proportions with overhead burdens which impact the fiscal flexibility of the system or drain funds designated to support the major goals of the project.

4.3.3 Discipline Centers

Each Discipline Center will have responsibility for coordinating the cataloging, organizing, archiving, updating and otherwise maintaining of all science data, supporting data, and documentation which are deemed to be pertinent to the analysis activities conducted by scientists in a specific discipline. Responsibility for providing scientists access to these data and distributing data upon request also rests with the Discipline Center. The facility should provide those various functions, such as computational power and human contacts for advice and council, appropriate to the data analysis task at hand. Although the responsibility remains with the Discipline Center to meet these stipulations, a greater responsibility lies in ensuring that the facilities and functions employed in meeting the requirements are in accordance with the greater Planetary Data System. Discipline Center science advisory groups and system science advisory groups will work together to ensure that standards and consistency of interface be maintained across the various boundaries of the Planetary Data System.

4.3.4 Periodic Review

The PDS must be constantly aware of and respond to the changing needs of the planetary science community. It was felt that major reviews should be held at regular intervals of no less than four years. These reviews should address technology advancements appropriate for incorporation into the system, needs of the community which may be changing or growing, and current and future plans for NASA projects and data analysis programs. Topical reviews of a smaller scale, relating to developments in NASA of industry of immediate concern to the system, should be held as appropriate.

4.3.5 Resource Allocation

The PDS should be accessible to all scientists and other persons with a need for planetary data. Yet, there is the potential for any resource that appears "free" to be misused. There must be instituted policies and procedures for achieving fair and equitable access to the PDS system which will also prevent abuse of the system and it's services. The following means were recommended:

- o Inquiries, detailed access to the catalog, help, and other central services would be open to anyone with a password. Obtaining a password would require only the knowledge that the system existed and the telephone number of the Management Center.

- o Access to the deeper functions of the system and its services would require a proposal. The degree to which review of the proposal would be conducted would be determined by the extent of the use proposed. Major use of facilities and computing resources would be peer reviewed much as are current NASA data analysis proposals.

- o Quantitative accounting could be accomplished using conventional computer accounts with the user resources established as a result of the proposal review process. User justification requirements should increase in proportion to the consumption of resources.

- o Small data distribution requests are filled free of charge.

- o The cost of large data distributions would be partially borne by the requestor, and a queue and priority system should be developed to moderate the effect of large computational requests.

The peer review panels, PDS management, and oversight committee should periodically review resource control to maintain effective utilization of the system.

4.3.6 Mission Interface

The PDS should serve as a formal interface between future planetary missions and the users of planetary data. Future projects are invited to use the PDS as a primary archive and distribution center for raw and processed instrument and engineering data, a working data base to assist in the analysis of mission returns, and as a source for historical data and data appropriate for use in comparison with the new observations.

Future flight missions may substantially reduce their data distribution costs by using the PDS services. Instrument data, the supporting data, and even mission specific information such as sequences and commands could be on-line or transferred through the PDS system. As there will be standards governing the storage and exchange of information within the system, it would be advantageous that flight projects develop similar standards or participate with PDS in the development of common network communications protocol and conventions for labeling, formatting, and encoding data. It was suggested that a standing committee be formed to provide the necessary coordination between flight projects and the Planetary Data System.

4.4 Implementation Plan

Two issues pervade development of an implementation plan: the need to salvage and upgrade appropriate existing data sets while this is still possible; and the requirement to assure that appropriate funding is allocated to implement the PDS and provide scientists with sufficient resources to use it. In this chapter an attempt has been made to merge ideal goals with pragmatic considerations of costs, organizational structures, and existing facilities. The conceptual Planetary Data System Management Structure and Functional Structure, as presented in figures 4.1 and 4.2, meet most of the requirements outlined in this document. The proposed implementation plan provides for a modular, phased approach which supports the need to accommodate changing user needs and rapidly evolving technologies.

4.4.1 Prioritization of Activities

A major consideration in implementation of the Planetary Data System (PDS)is that many seemingly unrelated tasks must be conducted at the same time and in the most cost conservative manner. This calls for implementation by priority, but with a high level of coordination to assure continuity. It is in response to this need that the tasks associated with implementing a PDS have been divided into three broad, but overlapping, categories: critical, immediate, and developmental. Tasks in these categories will be summarized and some of the details of implementation procedures described. Discussion of various functional elements of these tasks may be found in the data definition and technology chapters.

Critical activities were considered to be: 1) rescue of previous mission data and documentation which is in danger of being permanently lost; and 2) influencing the ongoing planetary missions' AO preparations in regard to agreements regarding the submission of mission data products to the PDS.

Immediate implementation activities include: 1) initiation of planning activities, feasibility studies, and cost analyses necessary to proceed toward implementation of the PDS; 2) communication to the Pilot Planetary Data System (PPDS) of items needing long-lead development or demonstration; and 3) appointment of an on-going group to direct the design and implementation of the PDS.

Developmental activities include all other tasks leading to the implementation of an operational Planetary Data System. Specific among these are: 1) definition of critical elements such as data standards, management structure, and major software and hardware components; 2) evaluation of new and affordable technologies for incorporation into the system and upgrade of existing components; 3) definition and development of interfaces and contractual arrangements with ongoing and future flight missions; 4) selection of central management and discipline centers through an appropriate process; 5) development of the functional system including the data catalog, archive data bases, communications network, and supporting resources; and 6) incorporation of existing facilities and data sets into the system.

Major portions of the developmental activities may be conducted as a whole or in part by the PPDS. This necessitates strong communication bonds between the participants in both activities.

4.4.2 Scientist Involvement

Scientists who actively use and understand the data should be directly involved in the data system. The involvement should include scientific review of system performance and use, guidance in data preparation, participation in specification of system capability, and consultation upon details of the individual data sets. The planning and implementation structures, as well as the management and operation schemes must take these requirements for continuing science involvement into consideration.

Several review activities are included in the Preliminary Management and Functional Structures, figures 4.1 and 4.2. Specific among these are reviews of the entire system, peer review of proposed major uses of system resources, review of validity and completeness of data submissions, and internal reviews of system operations.

4.4.3 Recovery of Existing Data Sets by Proposal

The current planetary data complement consists of experimental results from early exploratory missions which are no longer returning data and from current missions, such as Pioneer 10, 11, and 12 and Voyager 1 and 2, where data is returned routinely from the spacecraft.

Because of the low priority placed upon data analysis during the past 20 years of planetary exploration, the processed state of data from individual instruments is in various stages of completion. Some data exist only in the form of raw data on magnetic tapes. Many valuable data are felt to exist upon these tapes which have never been examined in detail. The funding climate at the time of acquisition limited detailed science examination to data which were believed to be "the most interesting" or most appropriately applied to research problems of the greatest interest. These processed data have been reduced to images, physical units of the planetary environment, and various products (both analog and digital) which represent the science results of the various disciplines.

Some part of this data complement has been submitted to the National Space Science Distribution Center in response to meeting mission requirements for data distribution. Many data, not appropriate for distribution (or not required for submission), reside with the PIs who were responsible for the instrument or in various centers, institutions, and federal warehouses. In the case of many of these data sets, the magnetic substrate upon which the data are written is undergoing destructive decay processes.

The recovery involves locating all of the pertinent data and making complete data sets available through the PDS. Funding of each appropriate PI group to perform those tasks necessary to produce a viable data set -- which is accompanied by the documentation and supplementary information to make it usable -- will be required. PDS planners must determine the magnitude of the effort required to suitably update the data base. It was suggested that the most effective approach might be to issue an AO to existing PIs. Selection would be based upon the priority of the data as determined by its relative peril of loss, as well as the scientific benefits to be reaped by its reclamation.

It should be made clear that proposers will be required to reprocess their data into a format which is specified by and compatible with the overall PDS design. This may require standard labelling and formatting procedures to be performed.

These proposals will provide the information necessary to scope and cost the recovery activity, as well as to identify the PI groups and individuals who would be willing to commit themselves and their institutions to the activity. Prioritization of the elements of the recovery task can then be accomplished and recovery can proceed. In addition, the information can be used to develop a cogent scientific rationale for the initial configuration of the PDS and to scope its eventual configuration.

4.4.4 Revised Flight AO Process

The relationships between scientists and flight projects should be modified to account for the PDS requirements in regard to data submissions, and to invite use of PDS services in the conduct of mission data analysis and distribution tasks. Announcements of Opportunity (AOs) for flight investigation should include a requirement that investigators deliver data and documentation to the PDS; and that these requirements will be considerably more detailed that the current non-specific requirements for data distributions to the NSSDC. The AO should also encourage proposers to include use of the PDS in their data management plan, and anticipate that flight projects will use some aspects of the PDS in delivering data to PIs.

4.4.5 Competitive Proposals for Participation in PDS

Proposals for establishing and operating the Management Center and the various Discipline Centers will be evaluated upon a competitive basis. The Discipline Centers selection process should include peer review by planetary scientists. Selection of the Management Center should include assessment by scientists familiar with the proposing institutions. The longevity and control nature required of these centers, to promote continuity and system stability, necessitates that a clearly defined, long-term commitment from the proposing institution be included in the proposal.

4.4.6 Incorporation of Existing Facilities

The planetary scientific community is primarily distributed throughout a relatively small number of research groups located across the country. Characteristic of these groups is their expertise, their experience, and their long-term commitment to the planetary program.

These groups are supported by significant data analysis and data storage facilities which would be expensive to duplicate, and impactive to science if altered by drastic changes. Active data bases are used and maintained at many dispersed locations, and the experience gained in the development of prototype systems (i.e. the Pioneer-Venus data distribution system) should be considered in the design of the PDS.

Software and procedures have been developed in the community to support various data analysis activities. It is highly probable that incorporation of existing procedures could have widespread effects on the early operational status of the PDS. For example, the Mars Consortium is an active data base maintained by the USGS at Flagstaff, Arizona. This data base is widely used by members of the science community because: a) data are fairly well documented and up-to-date, b)personnel are available to answer questions about the data; c) data are centrally located; and d) data are in standard format. Although the format and organization of the Consortium data might not be appropriate for all data in a Planetary Data System, the experience gained in its development and use will be highly useful in developing a system to perform these activities on a wider scale.

4.4.7 Use of New Technology

New technology implementations should be incorporated into the PDS as they become available and affordable. Major items with immediate utility are being explored as part of the Pilot Planetary Data System (PPDS) activities, and include: long-life, high-density data storage equipment (optical digital disks); workstations; data base management systems, both hardware and software based; and network communications. The system design of the PDS must accommodate the incorporation of emerging technologies, and must do so without seriously impacting other portions of the system, the users, or existing standards. Major innovations which disrupt system operation must be reviewed at all system levels and benefits weighed against detriments from many points of view.

Optical-digital disks could significantly reduce the substantial effort necessary to maintain the integrity of a digital data base throughout the period of its use and existence. This medium holds the promise of being sufficiently economical to permit widespread distribution of large volume data sets. This option is especially attractive if the industry holds to standards, like those for magnetic tapes, which will permit wide use of the platters.

Workstations, essentially single-user computers, can handle many analysis and display tasks and reduce the load on host computers. As these are being produced in increasingly large numbers, the cost to acquire a workstation is becoming affordable to the research scientist.

Electronic communication between nodes on a network is required for optimum function of a PDS. The current state of communications technology is such that the demands can be easily met. Although the cost of communications networks is being reduced each year, it is still not an affordable option to create ones' own network to support the PDS. NASA is currently implementing a new communication system which should incorporate active PDS sites, to the degree possible.

4.4.8 Standards and Common Software

Increased coordination among scientists within disciplines, between disciplines, and across mission boundaries involves some reduction of independence and an increased perspective into system-wide issues and the cost benefits afforded by coordination of efforts. Recognition has been hastened by the rapid advances in computer technology and the complexity of digital analysis, as well as a realization that the basic procedures of data reduction and data analysis are quite similar. Individual research groups and missions can no longer afford to develop unique software and hardware systems to meet the needs for each application.

There exists a need for standards, coordination of efforts, and continuity throughout the implementation of the Planetary Data System. Early efforts must be in the specification and design of standards for the encoding of data and in the procedures by which data is accessed, examined, and processed. Standards developed during this phase will have wide-reaching effect upon system design and function. There are significant efforts of this nature underway in related fields such as the NASA Pilot Data Systems. The Space Telescope Institute and in the astronomical community. Developments in these areas should be examined for incorporation into the PDS.

4.5 Funding Issues

The principle reason for a Planetary Data System (PDS) is to improve access to and use of documented planetary data. A fundamental tenet in the consideration of such a system is that it must not siphon significant funds from the existing planetary R&A and data analysis activities.

Means whereby NASA should obtain funding to support the implementation and operation of a PDS are outside the purview of scientists and not discussed. However, means for distribution of these funds was discussed in detail. A synopsis of these ideas is presented in the following paragraph.

It was recommended that the most straight forward means of funding the PDS on a year-to-year basis was through blanket grants to the Management Center, or to the Management and Discipline Centers for operation. Users of significant PDS services could submit budgets for PDS support within their R&A or data analysis proposals. Major use by other than NASA supported scientists could be supported on a cost reimbursable basis, as is now done by NSSDC. Costs for enhancing the PDS to meet specific mission requirements should be borne by the missions, most especially if the nature of the enhancement was such that it was in effect only during mission operations, i.e., high-speed data links which would be disabled or removed at the end of data acquisition.

4.6 Implementation Schedule

The plan proposed for implementation is based upon a phased and modular approach. Determination of the approach was influenced by many factors: 1) flexibility in accommodating fluctuations in funding; 2) allowance for the orderly incorporation of existing facilities and data bases; 3) advantageous in utilizing results from the PPDS studies; and 4) ability to assimilate the rapidly advancing technologies into the system as they are proven and become affordable.

The PDS Implementation Schedule, figure 4.3, provides a framework from which plans for a Planetary Data System can begin. The layout of the schedule illustrates the phased approach with key milestones serving to indicate activities. These will vary as actual costs and the complexity of the system are better understood. The modular aspect of the approach is not obvious at this level of planning; nor are the many decision making activities which must be used to establish implementation priorities.

ACTIVITY FY	84	85	86	87	88
CRITICAL					
DATA SET RECLAMATION	* * * * * *	* * * * * *	* * * * * *	* * * * * *	
INFLUENCE AO PROCESS	* * *				
IMMEDIATE					
FUNCTIONAL RQMT. DOC.	* * *				
DETAILED SYST. DEF. & COST ANALYSIS	* * *	* * *			
MISSION INTERFACE DEF.		* * * * * *			
SELECT CENTERS (MANAGEMENT AND DISCIPLINE)		* * * * * *			
ONGOING ACTIVITIES					
PPDS INTERFACE	* * * * * *	* * * * * *	* * * * * *	* * * * * *	
PPDS RESPONSIBILITY TRANSFER					*
IMPLEMENTATION					
MANAGEMENT CENTER			* * * * * *	• • •	
DISCIPLINE CENTERS			* * *	* * * * * *	• • •
INCORPORATION OF DATA INTO DATA BASE	* * *	* * *	* * *	* * *	* * *
		(PPDS DATA BASE)		(PDS DATA BASE)	

FIGURE 4.3 PDS IMPLEMENTATION SCHEDULE

5. IMAGING DATA

5.1 Introduction

The primary, planetary imaging data base includes all images obtained in visible wavelengths by imaging camera systems aboard planetary spacecraft. (Excluded, due to the volume and complexity of the data, are images taken from Lunar or Earth orbital spacecraft). These estimated 145,000 individual images exist on 12,400 magnetic tapes in a digital format primarily as Experiment Data Records (EDRs); and most are accompanied by the Supplementary Experiment Data Records (SEDRs), which render the EDRs useful. Hardcopy pictures and negatives, the designated archive products for imaging experiments, along with color products and mosaics are maintained at the National Space Science Data Center (NSSDC). Also included in the discussion of imaging data are two-dimensional data produced by spacecraft instruments other than optical sensors which may be considered images in a broad sense. Examples of these instruments are imaging radar systems, scanning infrared radiometers, and photopolarimeters.

It has become evident that, with the advent of affordable image processing workstations, the planetary imaging community requires access to digital imaging data and the ancillary information necessary to make the instrument data useful. Two major road-blocks prevent ready access to the imaging data: 1) the volume and dispersed locations of the EDR tapes and ancillary information; and 2) the lack of commonality in format or organization of the data.

This chapter, therefore, contains an in-depth description of the elements of the data base (both digital and hard-copy), discusses the current location of data and suggests appropriate means for archive/preservation of these data, and establishes requirements for the Planetary Data System (PDS) in support of the data handling and analysis tasks unique to the planetary imaging community.

5.2 Elements of a Dataset

5.2.1 Digital Image Data

A vast majority of the digital images are stored in the form of Experiment Data Records (EDRs) on standard 12-inch reels of magnetic tape. EDRs (since Mariner 6/7) are generated in the Mission Test Imaging System (MTIS) and are delivered by the project to the imaging team. Backup copies are usually provided to the Image Processing Laboratory at JPL and to USGS, Flagstaff. Copies of EDR data for some missions (Viking Lander and Voyager Satellite images) have been provided to wider distributions. In general, EDR data consist of unprocessed (raw) telemetry records in mission specific formats. For some missions the EDR simply represents a string of imaging telemetry from the Master Data Record. However, variations such as the substantial reconstruction necessary to produce Viking Orbiter EDRs and the generation of Viking Lander EDRs by the Image Processing Laboratory (IPL) of JPL did occur and are discussed under the appropriate headings.

Other mission-specific digital data sets are the Team Data Records (TDR) of the Viking Lander and the Reduced Data Records (RDR) of Mariner 9. Both of these differ from standard EDRs as they are written in VICAR format. This differentiation is not nearly so significant now as before the conversion of all EDRs to the condensed versions which are stored in VICAR format.

Numerous specially enhanced second order data sets exist which are not considered formal data records but represent an enormous investment in computer processing and should be formally documented for retention in the image archive. In addition, numerous calibration data sets required for certain types of image processing and analysis have been produced and maintained by IPL analysts. These files (as well as those which may be generated in the future) and procedures for their use should be documented and submitted to the archive.

5.2.2 Supplimental Experiment Data Records (SEDRs)

The SEDRs provide viewing geometry and related spacecraft state parameters required for visual interpretation, computer enhancement, or computer search and retrieval systems to support image data analysis. SEDRs are produced by the same set of software which is used in planning image acquisition sequences and in the determination of sequence implementation commands. Predict SEDRs are generated during the sequencing activities and are used in the interpretation of images as they are received. Final SEDRs are created after the fact with the same system using refined instrument activity times and instrument pointing parameters, as well as the final planetary and probe ephemeris (PLET and PET) information, to more accurately determine the SEDR values. Final SEDRs exist in digital form for all JPL missions since Mariner 9. These are also on file for Lunar Orbiter, the Apollo Mapping, and Panoramic data sets.

The mission is responsible for producing one SEDR record for each image; however, the responsibility for verifying the SEDR -- establishing that the longitude and latitude measurements supplied are accurate -- has generally been placed upon the imaging team. This user validation system, conducted in an environment where thousands of images are taken but only a few are scrutinized in near real time, tends to delay identification of errors in SEDR generation for months or years -- sometimes so long that regeneration becomes impossible. Several per cent of the SEDR data for Viking and Voyager images are known to be inaccurate.

5.2.3 Picture Catalogs

The MARK IV File Management System has been used at JPL for the past 10 years to produce, maintain, and provide access to Picture Catalog files for each mission. The picture catalog combines parameters from the SEDR with information documenting digital and photographic products produced by the first-order processing system (MTIS) and by the Image Processing Laboratory (IPL). The most actively utilized parameters in search and retrieval operations include footprint latitude and longitude, sun or viewing angles, resolution of image data, EDR tape and file identifiers, and film roll and frame numbers. A printed catalog format is produced for each mission and distributed to NSSDC and to imaging team members. Auxillary listings are also produced using the powerful sorting and formatting capabilities of MARK IV; examples include images sorted by latitude and longitude and subsets of satellite, limb, or other target characteristics.

Ideally, in the generation of the picture catalog, all input data would be submitted in machine readable format. Input includes the SEDR tapes, processing summary records (PSRs) which document the digital and photographic first-order products, and the IPL library support file (LSF) which documents the second-order processing and products. Unfortunately, the complexity of compiling the various machine generated files into an accurate catalog has proved to exceed the budget allocated for cataloging the recent missions. Therefore, much of the product identification information is still input by hand.

5.2.4 Special Data Sets and Catalogs

There are at least two sources of consistent engineering and viewing geometry parameters for a variety of missions. These are the Planetary Image Data Base (PIBS), a subset of picture catalog and SEDR parameters in consistent format, and the Better Image Retrieval Program (BIRP) search system which utilizes the PIDB data set to provide an interactive search and retrieval capability featuring user-friendly, menu-driven processing. BIRP allows non-programmers to extract information from the data base with a minimum of difficulty, utilizing whatever programming or data management techniques that are available at the users institution. BIRP is a set of programs written in PDP-11 Fortran IV (compatible with VAX systems), and is designed for transportability and for implementation on other types of computers.

5.2.5 Ephemeris Archives

Ephemeris data are produced as part of the navigation, orbit determination, and trajectory operations for each mission. Planetary ephemerides describe the position of solar system bodies while the probe ephemeris describes the flight path of the spacecraft. Production of SEDR data requires the input of a set of measurement times (in Universal Time) along with refined instrument pointing angles. These are processed in conjunction with the probe and planetary ephemerides to provide instrument viewing geometry and related parameters.

Unfortunately, the production of ephemeris data sets, their maintenance, and documentation has not been a well organized activity. The projects do not recognize ephemeris tapes as archival products. It is, therefore, extremely difficult (potentially impossible) to locate accurate ephemeris data for a specific mission and time interval. A careful and systematic review of ephemeris archives which still exist should be commissioned. The task should include the collection of the digital ephemeris data into a single set of magnetic tape files with documentation of their epochs (time periods over which they are valid).

5.2.6 Photographic Products

The principle archive product from each mission consists of 70mm photographic negative transparencies of each image, produced by the MTIS and usually represented by two or three processed versions of each image. The selection of processing versions is the responsibility of the imaging team. The image format and the processed versions vary from mission to mission, as do the resolution and optical properties of the cameras. It is, therefore, not possible to draw other than general conclusions when comparing surface features imaged by successive missions to the same planetary body, such as Mariner 4, 6/7, 9, and Viking.

Additional products which result from the systematic processing are 70mm positive transparencies, 3 x 4 hard copy images, and 8 x 10 hard-copy enlargements. Products resulting from science processing and the mosaicing of images, either by hand or in the computer, provide a highly useful data set. Negative and hard-copy photographs of these are archived at the NSSDC for the Viking and Voyager missions. These products from prior missions reside at JPL. Photographs in 8x10 and enlargement format of all planetary images are resident at each of the Regional Planetary Library Facilities (RPIF).

5.3 Data Description

An overview of each planetary mission which includes an imaging instrument begins each discussion; the targets of the imaging experiment are described and a brief scenario of the imaging rationale is often included. Also a brief description of the imaging instrument is given in the overview. All instruments are then summarized in Table 5-1.

The overview is then followed by a description of the EDR and the SEDR files for each mission. The order proceeds from the most recent to the most distant past, and begins with the on-going Voyager Mission. If a portion of the discussion is omitted for a particular data set, that indicates a lack of information or non-existence of the data product.

Table 5-1

Description of Selected Flight Camera Systems

Mission	Focal Length (mm)	Field of view (deg)	Field of view Pixels (P) Lines (L)
Voyager			
Vidicon framing cameras			
Wide Angle	200	3.0 X 3.0	800P X 800L
Narrow Angle	1500	0.4 X 0.4	800P X 800L
Viking Orbiter			
Vidicon framing cameras			
Two cameras offset by 1.38 degrees	475	1.69 X 1.54	1204P X 1056L
Viking Lander			
Facsimile cameras	53.7	.04 X 20	512P X Var.
Mariner 10			
Videcon framing cameras			
Narrow Angle (2)	1500	.48 X .37	832P X 700L
Wide Angle Optics	62	14 X 11	832P X 700L
Mariner 9			
Vidicon framing cameras			
Narrow Angle	500	1.41 X 1.06	832P X 700L
Wide Angle	52	13.5 X 10.5	832P X 700L

5.3.1 Voyager

The Voyager mission has, so far, returned over 60,00 television images of Jupiter, Saturn, their satellites and their rings. More data are expected during the next decade as the Voyager 2 spacecraft flies past Uranus and Neptune prior to leaving the solar system. Many more objects have been observed and a richer variety of observations has been attempted with Voyager than with any other mission. In contrast with the Viking and Mariner 9 missions which acquired data steadily for many months or years, Voyager science activities were focused into a few weeks near the times of closest approach to each major planet. This intense and periodic saturation of information has resulted in a new set of problems in data reduction and documentation activities.

Each Voyager spacecraft carries identical imaging systems consisting of wide-angle and narrow-angle slow-scan TV cameras. Each is fitted with a square vidicon sensor, a mechanical shutter, and a rotating color filter wheel. A Voyager image is an 800 x 800 array of picture elements (pixels), and is coded as an 8 bit data number or DN which represents the brightness of each of the 640,000 picture elements read out during the slow-scan of the vidicon. Specifications for the Voyager cameras are given in Table 5-1.

The communications and on-board data storage capabilities of Voyager were important constraints on the amount and nature of imaging data acquired. Some television images were telemetered to Earth in "real-time", i.e., immediately after being read from the sensor. Other images were stored on a magnetic tape recorder and telemetered to Earth at a later time. Data were also edited in various ways in order to reduce the number of bits in each frame and to allow more frames to be recovered within the same communications budget.

Most of the photographic products generated from the digital data were produced by the Mission and Test Imaging System (MTIS). A comprehensive description of the computer programs used to process the Voyager data is given in the MCCC/MTIS Voyager Program Description Document, JPL 618-792. A final version of the Experiment Data Record (EDR) containing a minimum number of missing or defective pixels was generated for each Voyager image.

First order products were produced by the MTIS and are described in the appendix of MCCC/MTIS UPDD JPL 618-792. Second order processing at JPL's Image Processing Laboratory (IPL) was performed in support of research in many different discipline areas (meteorology, geology and ring dynamics). Second-order processing at the USGS Flagstaff facility focused on cartography of satellite surfaces. These second-order products comprise the primary data set for many of the scientific studies based on Voyager image data.

Dividing the complex Voyager imaging data set into subsets by target or resolution conditions is one way of assisting scientists to make more effective use of these data. For many of the sequences, it is useful to combine individual images to form mosaics, produce composite images to form color versions, and refer features visible in the images to object-centered coordinate systems. This is still underway for the Voyager data and many of the products are not yet ready for distribution. The status of the geodetic control nets for the satellites of Jupiter and Saturn is given in NASA TM 84211, 1981. Maps of the Galilean satellites based on this network include photomosaics of map projected images and maps of the Saturnian satellites. All-digital multispectral mosaics of the Galilean satellites, multispectral mosaics of Jupiter and geodetically-controlled photomosaics of Saturn's rings are either in preparation or contemplated. These data sets provide advantages over individual frames; they are radiometrically corrected, referenced to an object centered coordinate system, and they provide a synoptic view.

EDR Formats

Voyager EDRs are recorded on 9-track tape at a density of 1600 bpi. Each image comprises one file on the EDR tape, with as many as 24 files on a tape. One end-of-file marker (EOF) separates images. There are two EOFs after the last image on a tape. Each image file consists of a single header record of 1280 eight bit bytes, followed by 800 line records consisting of 1040 bytes each. An image file consists of 801 records. Both the header and line records include an MTIS Header block, a source data summary, and status/engineering data. The header record provides a source data histogram while the line record contains the 800 bytes of pixel data.

5.3.1.3 SEDR Format

A. Voyager. The Voyager SEDR data set is written on 9-track, 1600 bpi magnetic tape. Each tape contains approximately one command load of images. A single header record at the beginning of each tape identifies input data sets used to produce the SEDR (ephemeris tapes, etc.), and is followed by individual data records for each image ordered by FDS Count (spacecraft clock time). For simultaneous A and B camera shutterings the SEDR identifies images with the same FDS Count, however the photographs and digital data identify the wide angle camera image with an FDS Count from 1 to 5 counts higher than the narrow angle cameras. The Voyager SEDRs are processed at the Image Processing Laboratory to produce several other data sets. Each SEDR is used to update master SEDR files (one for Jupiter and one for Saturn) used by VICAR image processing programs stored on direct access devices in a special indexed format. The SEDRs are also processed to update the Voyager Picture Catalog which contains a subset of SEDR parameters, and to produce a SEDR stack tape containing all the SEDR data records.

5.3.2 Viking Orbiter

5.3.2.1 Overview

In the three year period from June 1976 to October 1979, two Viking Orbiter spacecraft returned more than 50,000 images of the Martian surface and satellites to Earth. Although the Voyager missions have returned a greater number of images, the larger image format of the Viking cameras makes this the largest planetary data set in existence. Descriptive material on this data set is dispersed through many scientific papers, formal reports, and informal memoranda. There is no single organized catalog of data such as those that exist for Mariner 9 or for the Viking Landers.

The Viking Orbiter Mission was developed primarily as a support function to survey the Martian planet and permit selection of a safe site to deploy the Landing Spacecraft carried by each orbiter. Daily sequences imaged the several landing targets in great detail. After safe deployment of the lander payloads, the orbiters continued to operate for many years after their predicted "life time." 90% of the surface was mapped at moderate resolution (ranges between 3000 and 18000 km from the surface), several high resolution survey strips of contiguous coverage were obtained between 1000 and 2500 km range; Martian seasonal changes were monitored over nearly two martian years; high altitude (10 to 14000 km range) color coverage was obtained of virtually the entire planet; and high resolution mapping was performed of the satellites Phobos and Deimas.

The two Viking Orbiter Spacecraft were each equipped with identical, paired telescopic cameras. Each camera consisted of a large-format vidicon sensor, selectable color filter-wheel near the focal plane, and a mechanical shutter assembly. Sensor readout was similar to that described for Voayger; however, the format of the image array was rectangular -- 1204 x 1056 pixels. Other particulars of the Cameras are included in Table 5-1.

Images were initially recorded on seven-track magnetic tape recorders on the spacecraft. The data were played back to earth one track at a time by operating the tape recorder alternately in the forward and reverse directions. Each raw data frame retrieved from the tracking station thus contains every seventh pixel arranged in either increasing or decreasing order. Image data reconstructed from these raw data frames by the MTIS form the Experiment Data Record (EDR).

Most of the photographic products generated from Viking Orbiter imaging data were also created in the MTIS facility as part of the Viking Imaging System Reconstruction and Processing (VISRAP) processor. For each image a shading corrected rectilinear image was produced and either a filtered rectilinear or an orthographic product depending on image viewing geometry. Some of the Viking Orbiter data were also subjected to second-order processing of a more specialized character. Processing performed at Flagstaff will be described in a document now in preparation, "Level 1 and Level 2 Digital Processing of Viking Orbiter Data by R. Batson.

The majority of Viking images were acquired as strips of contiguous and slightly overlapping images which are most useful when assembled into mosaics. In many cases multiple strips were assembled to form still larger mosaics which provide a synopic view of a regional surface and can support detailed geologic mapping. In many respects the mosaics are a more useful data set than the individual frames from which they were derived. However, individual frames used in detailed study of particular areas or features of the planet. Several families of mosaics exist or are in the final stages of preparation. The first set to be assembled were the "211" series of rectilinear images assembled without independent geodetic control. They served landing site selection and first order scientific interpretation. Until recently they have been the only reasonably complete compilation of the Viking images in mosaic form. A catalog of "211" series mosaics with reproductions and outline drawings identifying individual images is provided in the "Viking Orbiter Mosaic Catalog, Vol 1 and 2", NASA Cont. Rep. 3496. A more sophisticated product is a set of 198 "subquad" mosaics soon to be available in the "Catalog of Mars Subquad Mosaics", by R. Tyner et al. These mosaics are made from orthographically projected images and are compiled on standard mapping projections. The mosaics are accompanied by outline indexes and lists of all frames that lie within the mosaic - whether included in that particular mosaic or not. A final version is planned which will show the location of all Viking frames. Compilation of a series of 140 controlled photomosaics of Mars at 1:2 million scale is also underway. These mosaics use medium resolution images and are referenced to the control net defined by Davies and others in "The 1:2 Million Series of Controlled Photomosaics of Mars", NASA TM-84211, 487, 1981.

5.3.2.2 EDR Format

Viking Orbiter EDRs are recorded on 9-track tape at densities of either 800 or 1600 bpi (800 was used only during the first few weeks of the mission). A maximum of 14 images are contained on one tape recorded at 1600 bpi. Images are separated by single EOFs with two EOFs following the last image on a tape. Each image file consists of an EDR Header record of 1200 bytes, an SEDR record of 2000, bytes and 1056 line records, each 1600 bytes long. Records are not blocked. Pixel values are stored as eight bit integers with a range from 0 to 255. The values represented on the EDR tape are actually 2-times the value received in the telemetry stream. Because of the nature of the recording and playback system, strings of missing bits appear as vertical bars at 7-pixel intervals. Missing pixels are set to zero on the EDR tapes.

5.3.2.3 SEDR Format

Viking Orbiter SEDR's generally contain records representing two orbits of imagery. Each record on the SEDR tape represents a single image. Nearly all SEDR parameters are represented both in IBM floating point format and in EBCDIC (Extended Binary Coded Decimal Interchange Code) character format. The predict version of the SEDR tapes were processed by MTIS to provide annotation labels for processed images and the SEDR records (expanded to 2000 bytes) were also incorporated in the EDR tapes as the second record of each image file. The final SEDR's were processed by IPL to update the IPL SEDR file for access by VICAR application programs. A subset of the SEDR parameters was incorporated into the Viking Orbiter Picture catalog (a MARK IV File Mgt. System compatible format), and all SEDR data records were also stacked onto SEDR master files (3 tapes, VO-1, VO-2 and Survey mission).

5.3.3 Viking Lander

5.3.3.1 Overview

The Viking Lander imaging data set was acquired by the two Viking Lander spacecraft on the surface of Mars. Viking Lander 1 began operation on July 10, 1976 and continued to return data until early 1983, when communications were lost. Viking Lander 2 began operation on Sept. 3, 1976 and concluded transmission in Feb. 1980. The camera systems utilized by the landers are summarized in Table 5-1.

A comprehensive description of this data set is provided in a three document picture catalog. The first document ("Viking Lander Imaging Investigation: Picture Catalog of Primary Mission Experiment Data Record", R. B. Tucker, NASA REF. 1007, Feb., 1978) describes data from the primary mission covering the period from landing through the solar conjunction period in Dec. 1976 when data transmission ended. The other two volumes ("Viking Lander Imaging Investigation During Extended and Continuation Missions, Vol 1 and 2", K. Jones et al., NASA RP 1068, April 1981) describe the Lander 1 and 2 missions through Feb. 1979.

The introductory material in each of the three documents includes a description of the Viking Lander cameras, their calibration and the coordination of the Viking Lander cameras, and their image data. The Viking Landers were built with the capability for directly transmitting data to Earth at low data rates as well as for relaying data via the orbiters at much higher data rates. Those documents describe the implications of these data modes for the types of data recovered from the Landers as well as the extraction of image data from the telemetry to produce the EDR data set.

5.3.3.2 EDR Format

Viking Lander EDR's are recorded at 1600 bpi on 9-track tape. A variable number of images are contained on each tape with each image separated by a single EOF and two EOFs following the last image on the tape. Image files are written in standard VICAR format with a variable number of label records followed by 512 line records. The logical record length of each label or line record is equal to the number of vertical lines in the image (variable from image to image), or 360 bytes, whichever is larger. All label information is represented in EBCDIC format. Each pixel is an 8-bit binary value ranging from 0 to 255.

There is also another Viking Lander digital data set called the Team Data Record (TDR). This consists of sets of specially enhanced images and includes mosaic data sets, color images and other special products. Image data is stored on 9-track tapes in VICAR format recorded at 1600-bpi.

5.3.4 Mariner 10

5.3.4.1 Overview

Mariner 10 or Mariner Venus Mercury Mission (MVM) was the first multiplanet mission carried out by a single spacecraft. In one flyby of Venus followed by three flybys of Mercury, Mariner 10 returned more than 12,000 images to Earth. Images of the airless body, Mercury, were of interest to planetary geologists, while the images of cloud-shrouded Venus were of meteorological significance.

The Mariner 10 spacecraft was equipped with two vidicon framing cameras, whose characteristics are detailed in Table 5-1. The narrow angle camera was used to image Venus and Mercury from long range and to obtain the highest resolution pictures near closest approach to the two planets. The wide angle camera provided mosaicked mapping coverage of both planets during the periods before and after the closest approach during the four encounters.

Though provided with a selectable color filter wheel, no false color images were obtained as this technique had not yet been developed. A first, however, were the computer generated mosaics of the surface of Mercury which form the basis of the "Atlas of Mercury" discussed below. These are preserved in film format only. EDR data exist for Venus. EDR and radiometrically decalibrated data exist for Mercury.

The "Atlas of Mercury", M. Davies, S. Dwornik, D. Gaust and R. Strom, NASA SP-423, 1978 covers the Mercury portion of the Mariner 10 mission. The Atlas includes an outline of the mission, a description of the cameras, photomosaics, maps and a set of references to the scientific photomosaics, maps and a set of references to the scientific results. The Mercury observations of the Mariner 10 data set remain unique in that there have been no subsequent spacecraft observations of Mercury. The Venus images have recently been supplemented by the imaging type observations obtained by the Pioneer Venus Orbiter cloud photopolarimeter.

5.3.4.2 EDR Format

Mariner 10 images are recorded at 556-bpi on 7-track tape. Each image is recorded as a separate file on the EDR tape followed by an EOF marker. Image files are made up of 700 line records, each containing 832 pixels. The pixels, represented as data numbers, are encoded as eight-bit binary values ranging from 0 to 255.

5.3.4.3 SEDR Format

The Mariner 10 SEDR data set is written on 7-track magnetic tapes each containing a single header record describing the input data sets used to generate the SEDR data and followed by a variable number of SEDR records, each representing a single image. The SEDR tapes have also been stacked onto 5 tapes, STK001 thru STK005. A small subset of the SEDR data for each image has been incorporated into the Mariner 10 Picture catalog.

5.3.5 Mariner 9

5.3.5.1 Overview

Mariner 9 was launched on May 30, 1971. Between the time of Mars orbit insertion on Nov. 14, 1971 and Oct. 22, 1972 the spacecraft conducted an intensive orbital reconnaissance of the planet. The first complete geological map of Mars and the discovery of a host of hitherto unsuspected landmarks were among the spectacular results. A comprehensive description of the Mariner 9 mission and its imaging data set is provided in "Mariner Mars 1971 TV Picure Catalog", Vol 1, [585], 1974 and Vol 2, Sequence Design, by P. Koskela et al., JPL Tech Memo 33-585, 1972.

Mariner 9 was equipped with wide angle and narrow angle telescope cameras, small format selenium-sulphide vidicon sensors, and mechanical shutters. The wide-angle camera had a color filter wheel which functioned for only part of the mission. Characteristics of the cameras are given in Table I-1.

Two types of Mariner 9 data are available. The Experiment Data Record (EDR) data set consists of raw images in the form that they were read from the sensor; they still retain the photometric and geometric distortions of the camera system. The EDRs also contain telemetry errors introduced during transmission the data from the spacecraft to Earth; these are missing or defective pixels. The second product is the Reduced Data Record (RDR) generated from the EDR by a process known as decalibration. Camera calibration data were used to convert EDR pixel intensities to estimates of the light intensities that formed the image. The EDR pixels were also resampled to correct for camera geometric distortions. The RDR image file includes the electronic scanning raster coordinates of reseaus (mechanically fixed reference marks on the face of the vidicon), which were used for geometric decalibration. Because the raster drifted as a function of temperature and scene brightness, it was necessary to measure the reseau positions in each frame with a specially designed algorithm in order to make an accurate geometric correction.

The extensive observations of Mars acquired more recently by the two Viking Orbiters have rendered the Mariner 9 data, a data set of lesser interest. However, it is still of considerable value in studies of seasonal and temporal variation of the surface, the polar caps, and the atmosphere. The Mariner 9 data also contain surface observations of regions not covered by Viking at suitable resolutions for geologic analysis.

5.3.5.2 EDR Format

The Mariner 9 imaging EDR's are written at 556-bpi on 7-track tape. The tapes consist of a series of Recorded Science telemetry records (R7938) and engineering records (E124). There are no EOF markers between images, only a single EOF to signify the end of data on a tape. An image consists of 700 R7938 records each containing 832 9-bit pixel values.

A set of Mariner 9 RDR (Reduced Data Record) images also exists in digital form. The RDR tapes are written at 800 bpi on 7-track tapes. Each image is represented by two files on the RDR tape, the first file containing the digital image data and the second containing the location of each reseau mark in the image. Two EOFs follow the last file on a RDR tape. The RDR data are stored in VICAR format with a variable number of label records followed by 800 line records of 1900 bytes each. The 1900 byte logical records are blocked in sets of three to form 5700 byte physical records on the tapes. Each 1900 byte logical record is composed of 950 16-bit pixel values representing values from 0 to 511 (only 9 bits are utilized). The reseau files contain one record of reseau location data. Each reseau is represented by two 4-byte IBM floating point words indicating the line and sample number for that reseau. For the Mariner 9 A-camera 111 reseau points are given, while for the B camera only 63 are provided. Attempts to process existing RDR data tapes at JPL have been less than 40% succesful because of tape read errors.

5.3.5.3 SEDR Format

The Mariner 9 SEDRs each contain two orbits of image data and are recorded on 7-track tapes at 556-bpi. The entire SEDR record is also stored in the Mariner 9 Picture catalog and a complete stack of SEDR records is stored on COV001, in 9-track, 800 bpi format.

5.3.6 Mariner 4, 6, and 7

5.3.6.1 Overview

The Mariner 4 (Mars) mission acquired 20 images of Mars in 1965. By present standards the images are of very poor quality. A detailed description of the images is provided in "Mariner" Mars 1964 Project: Television Experiment", JPL Tech Rep 32-884, 1967.

Mariner 6 and 7 spacecraft acquired 200 images of Mars at a variety of resolutions in the summer of 1969. These data are described in the "Scientific Findings from Mariner 6 and 7 Pictures Final Report", JGR, Vol 76, 1971.

These sets of data are still available but little used execpt for studies of time variability.

5.3.6.2 EDR Format

Digital images from the early Mariner missions to Mars are recorded on about 220 7-track magnetic tapes recorded at 556-bpi. All of the images have been computer enhanced and actually represent a Reduced Data Record rather than an EDR. Mariner 4 images contain 200 lines each consisting of 200 6-bit pixels. Mariner 6 and 7 images consist of 704 lines of 945 pixels each.

5.3.7 Lunar Data

The lunar data set is not described in detail in this document. Other sources document Survey or, Lunar Orbiter, Apollo, and varies Societ image sets. Television images acquired by the Surveyor series of unmanned lunar landers in the late 1960's provided important information about the lunar surface prior to the Apollo landings. Although the subsequent Apollo data have diminished interest in the Surveyor data, only one of the Surveyor sites was visited by Apollo astronauts and so the remaining observations remain unique.

5.3.7.1 EDR Format

The Lunar Orbiter mission underscores the potential fate of non-archived data not updated with changes in technology. Although the original magnetic tapes exist and there has been interest in reprocessing the images, there is no mechanism for recoving the data owing to evolutionary changes in hardware.

5.3.7.2 SEDR Format

The Lunar Orbiter SEDR consists of a single 9-track tape recorded at 1600-bpi, ORBIT. There is one record on the SEDR tape for each low resolution image, and one record representing the combined area of three contiguous high resolution images. All numeric values are stored as IBM floating point numbers, and characters in EBCDIC format. Data included in this file were taken from the Boing Lunar Orbiter Photo tapes.

The Apollo Mapping SEDR is contained on two files of a single tape, APOLLO, recorded at 9-track, 1600-bpi. The first file represents the Mapping (or Metric) camera data and the second the Panoramic images. Data used to produce this SEDR were obtained from the Apollo Photographic Evaluation (APE) tapes produced by JSC.

5.3.8 Pseudo-Image Data

Instruments other than visual-wavelength optical sensors produce two-dimensional data which may be considered images in a broad sense. Several mission data sets which fall into this category are discussed.

5.3.8.1 Radar

Planetary radar data are considered in detail in Chapter 7. For purposes of this section it should be noted that radar images can be produced directly through delay-Doppler processing of an echo signal (e.g., Aricebo and Goldstone roughness/reflectivity images of Venus) or by resampling linear data traces (Pioneer Venus altimetry). In both cases two-dimensional arrays of relatively low resolution (by optical standards) result; these can be processed and analyzed in the same way as other images.

The quality and character of radar images expected from VRM and subsequent radar missions, as well as the existing Earth-based and Venera radar maps, bear many similarities to more traditional vidicon and CCD optical images. The reasons for this similarity include:

a) Data are inherently 2-dimensional, with highly-correlated systematic errors between neighboring pixels.

b) Radar images are ephemeris - dependent in much the same way as are optical images.

c) The images must be cataloged, smoothed, combined, enhanced, and compared in the same way.

d) Users seek the same types of information from radar and optical images.

5.3.8.2 Pioneer Venus OCPP Images

The Cloud Photopolarimeter (OCPP) on the Pioneer Venus Orbiter spacecraft has been obtaining spin-scan images of Venus since December 1978 during several imaging opportunities (Colin, 1980, Travis, 1978). The Pioneer Orbiter is expected to be able to continue acquiring data until about 1992. It is expected that a total of about 3000 images of Venus would be acquired over this period. The orbital elements (SEDR) and the raw imaging data need to be included in the central archive.

5.3.8.3 Pioneer 10 and 11 Image and Image Derived Data

Pioneer 10 and 11 imaging data for Venus represent a unique data set since they provide a photometric and polarization standard for the measurement sequence. In addition the Pioneer 11 trajectory provided observations of the polar regions of Venus which have not been measured by other spacecraft instruments. The observations were made at only red and blue wavelengths, and therefore provide limited spectral information. Since the observations were made from a line scan system, it is important that the reduced data (calibrated and geometrically corrected) data be available for distribution.

The following data sets are derived from or obtained in conjunction with the Pioneer 10 and 11 imaging observations at red and blue wavelengths.

- Polarization maps of Venusian clouds as a function of phase angle
- cloud parameters - particle size, refractive index maps

5.3.9 Imaging Data of the Future

Historically, planetary image data storage and transfer formats have been designed to emphasize (sometimes subtle) differences in experiment and instrument characteristics. Today, planetary science analysis often requires access to multiple data sets and tends to address their complementary aspects. This report recommends the establishment of standard formats for image data that recognize the differences while permitting common handling routines for all. It is of paramount importance that, when such standard formats are designed, they be sufficiently general to allow for the emerging data types as well as those of the past.

Several data types planned for flight experiments now in preparation differ from those used in the past. It would not be surprising if other differences were to appear later, some perhaps trivial and others of a more fundamental nature. Among the new data types will be those produced by the Venus Radar Mapper, the Galileo Near Infrared Mapping Spectrometer (NIMS) and the Charge Coupled Device (CCD) sensors to be used in framing cameras on the Spce Telescope and Galileo.

Raw imaging radar data must be transformed into a spatially correlated array before it is considered to be image data. Even then, the fundamental differences in the mode of acquisition lead to differences in reduction which are not yet fully mature. The most useful ways in which to represent the reduced data are probably not yet determined. NIMS data, also requires spatial correlation, although for a different reason, in order to be successfully dealt with in an imaging sense. Its spectral dimension, extending over hundreds of bands, being completely correlated as a byproduct of the initial spatial resampling, introduces more than a mere quantitative improvement over previous multi-band imagers. Ways to fully utilize the potential of this instrument are yet to be discovered. CCD imagers, replacing the classic vidicons in framing cameras, seem to resemble their predecessors in some ways and are much better in others so that we are inclined to overlook factors which will profoundly affect several aspects of the reduction of their data. The instrumental geometric signature, so annoying and difficult to reduce (not remove) with vidicons, is negligible with a CCD imager. On the other hand, radiometric responsivity of the new sensors, while totally linear, is quite variable on a pixel to pixel basis. Differences in device construction between the Space Telescope and Galileo CCDs will lead to other differences in both operation and data reduction. CCD's designed for enhanced X-ray response will have yet different qualities surely troublesome as well as wonderful.

5.4 Locations of Imaging Data Sets

Hardcopy of images from past planetary missions is available for viewing at the Regional Planetary Image Facilities. These are located at several Universities and research facilities in the United States and at several research centers in Europe.

5.4.1 JPL

The imaging EDRs and original "flight films" from all missions are stored either at JPL or are in the Federal Records Center (FRC) at Laguna Niguel. The converted image data in 6250 bpi format will be stored in the Science Data Library (Building 264-111A) to be available for processing on the Pilot Planetary Data System (PPDS) VAX computer currently located in Building 264-127. A duplicate copy of all converted data will be suppli9ed to USGS Flagstaff. Since Flagstaff currently does not have a 6250 BPI tape drive, copies of Mariner 9 and 10 data are being delivered to L. Soderblom at Caltech for verification and testing. Most processed IPL data tapes are stored either in Bldg. 231-B1 or FRC as are SEDR, catalog and calibration data tapes.

5.4.2 USGS - Flagstaff

Flagstaff has had the role in most missions of compiling mosaics and map products and thus has essentially a complete set of lunar and planetary data in the form of, magnetic tapes and photographic products, positives, negatives and photographic prints. These are available for Ranger, Surveyor, and Lunar Orbiter, as well as the many Mariner missions and the Viking and Voyager Missions.

Photographic records are available of ground based telescopic pictures, Apollo Hasseblad camera, metric camera, and panoramic camera pictures. These data are available as second generation positive, third generation negatives and fourth generation prints.

Magnetic tapes are available with raw data (EDR) position data (SEDR) and calibration files for the planetary missions: Mariner 4, Mariner 6 and 7, Mariner 9, Mariner 10, and Viking orbiter and lander data for all four spacecraft. Voyager 1 and 2 data from Jupiter and Saturn Satellites are available including pictures taken within 5 million kilometers from each object.

In addition there is a complete set of photomosaic and maps of the Moon, Mercury, Mars, Venus, and the satellites of Jupiter and Saturn. Both original compilations and photographic reproduction are on file. There is also a complete set of footprint charts indicating locations of pictures of all the above bodies.

5.4.3 Johnson Space Center

Original negatives and color transparencies of the moon from the Apollo missions are presently maintained in a temperature - and humidity - controlled vault. Access to these original materials is highly restricted. Color transparencies and films have been transferred to silver-based three-color black and white negatives with calibrations to permit color reconstruction. This is necessitated by the unstable nature of dye-based (color) emulsions.

5.4.4 Foreign Image Centers

Several Planetary Data Sets exist in various locations in Europe, such as:

- London, U.K. (Imperial College, University of London Observatory),
- Paris, France (Universite Paris-Sud at Orsay, Observatateire de Paris-Mendon),
- Toulouse, " (GRGS-CNES),
- Munich, Germany (University, IFVLR)
- Roma, Haly (Laboratorio de Astrofisica Spaziale, CNR)

These data are provided by NASA to the above mentioned group in the framework of NASA efforts to develop Regional Planetary Image Facilities (RPIF) and to facilitate, participation of European P.I., Co. I. guest Investigators in NASA space missions or data analysis programs. The data available in the different facilities are listed below:

Imperial College, London (UK):
- Voyager I & II digital data
- Viking " " " "
- Mariner 9, 10, imaging data

University of London:
- Viking I & II, " "
- Voyager I & II, " "
Observatory (UK)

Observatorie de Paris - Mendon (France):
- Lunar data (Ranger, Surveyor, Lunar Orbiter, Apollo)
- Mariner 4, 6, 7, 9, 10

University Paris-Sud, France:
- Mariner 4, 6, 7, 9, 10
- Viking Orbiter & Lander
- Voyager

Toulous, France:
- PVO (VRM)
- Viking > altimetric - gravimetric data

University Roma, Italy:
- Lunar Data
- Viking Orbiter & Lander
- Voyager > imaging data

Center for Remote Sensing
Imperial College London:

- Complete Voyager Data Set
  - All EDR Data
  - All photographic products
  - SEDR
  - Calibration files
  - Complete Vicar based software facilities for data analysis
- Selection of Voyager IRIS temperature and spectra - particularly N/S maps in digital form
- All IUE planetary observations
- Selected Viking digital images and photographic products.
- Considerable amount of Viking IRTM data (digital form)

At Imperial College, most of the imaging data are available under different formats: negatives, hard copies, mosaics. The digital data are original sets (Voyager, Viking). These data represent the milestone of an archiving center in Europe that will include the forthcoming data of the future US space missions (Galileo...) and of the planned European missions.

5.5 Archiving and Preservation of Image Data

5.5.1 Urgency and Nature of Rehabilitation Activities

Primitive image products including EDR digital data and "flight films" from the various missions currently are not in a central location, are not properly maintained, and are not systematically monitored. These materials must be identified and assembled at the earliest possible time. At present, imaging EDR's from all planetary missions are dispersed in three different locations in uncontrolled environments. In contrast, the original Apollo flight films are presently stored in temperature and humidity controlled vaults at JSC. Apollo color products recently have been transferred to three-color separation silver-based negatives.

The medium on which digital data are stored is not of archival quality and requires continuous supervision, maintenance, monitoring, and rehabilitation.

Magnetic tapes must be systematically rewound and monitored for bit errors and changes in physical state. A central archive should be established to perform these functions according well-established procedures for meeting industry standards. New storage media (e. g. video disks) should be explored, but back-up data sets on media of known (not projected) storage characteristics should be maintained until careful review demonstrates complete transfer to be safe.

5.5.2 Archiving Priorities

a) A systematic effort to document image data from all missions must be carried out. This includes resurrecting existing knowledge of instrument data set characteristics.

b) Data management plans for upcoming missions must be reviewed in light of the Planetary Data Workshop findings. Funding should be provided for proper planning and archiving of mission data sets.

c) Ephemeris data sets still in existance must be saved in some agreed upon format.

d) Locate and identify all image sets worthy of archiving and integrate them into central archive.

e) Copy existing tapes (e. g., IPL tapes) to a high-density format in order to extend their life and to reduce storage requirements.

f) Transfer existing data sets to new formats compatible with evolving hardware.

g) Transfer calibrated color products (e. g., Viking "real color" products) to proper storage conditions; essential products should be preserved on silver-based and stable emulsions as calibrated 3-color separates.

5.5.3 Rehabilitation of Digital Image Data

Rehabilitation of many of the planetary data sets must be approached as a two-phase operation. Large quantities of magnetic tapes are stored in less than optimum conditions and suffer from decay of the magnetic substrate as well as loss through miscataloging, and movement from storage area to storage area.

Rehabilitation, therefore, encompasses copying the old data tapes onto a higher-density tape to save the information, while scientists and data managers work to develop a more standardized means for encoding planetary data.

Imaging scientists have been in the forefront in developing both the interim salvage procedures and in developing requirements for standardizing the use of planetary image products from all missions.

5.5.3.1 Salvage Operations

Throughout the last two years the digital, imaging data base has undergone processes of organization and data compression as a result of JPL's Planetary Image Conversion task. By the end of FY84 the imaging data base will reside on 1000 hi-density tapes, stored in a "modified VICAR" format. Standard, mission dependent labels precede each file; are followed by the pixel data; and finally by the associated engineering parameters which are placed at the tail of each line. Any header records contained in the EDR file were placed after the last image line. Not only has the conversion resulted in a 12:1 reduction in the number of tapes necessary to contain the data set, but also the reorganization of the data has permitted the reading of all of the data with standard VICAR software.

The conversion activity has by no means resulted in the development of a common image data format. The goals of the task were compression, organization, and salvage of a data set rapidly undergoing loss through tape aging. Each mission data set retains the mission defined format characteristics in the header labels. Devising a universal, standard format that is sufficiently flexible to take into account the specific requirements of each mission is recommended, but development will be a lengthy and extremely difficult process.

5.5.3.2 Standardization Recommendations for Digital Data

A standard digital format should be adopted. This format should include a label record which has a data description in an agreed upon format which all types of machines can easily recognize. This would be followed by the data set and the ancillary data which are needed to process, define, or calibrate the data set. These would be packed in such a way that the "raw data" could be easily accessed independent of the ancillary data.

The astronomy community has for several years used an image data format called FITS (Flexible Image Transport System). It was designed to meet needs of all classes of users, from those having small computers to large, and those using 800 or 1600 bpi tape drives. FITS is inefficient for high density tapes (6250 bpi) because of the number of record gaps. Space Telescope Science Institute has adopted a modified version of FITS which allows larger record sizes called extended FITS. The planetary community in image processing has a similar format called VICAR. A third format is being provided by the pilot archiving program. A digital format might be established which includes the basic approach of all three formats.

Each digital image tape should begin with a label record, in standard format, which includes the key parameters, a basic description of the data, and its format. This would be followed by the data set packed with the ancillary data, arranged in an efficient format where each line starts with the imaging data and ends with the ancillary data.

5.5.4 Recommendations for Film Products

Photographic (silver-based) products represent a proven medium for preserving image-formated data. Past missions have and future missions will continue to provide such products for Team analysis and RPIF reference sets. The value of such sets to the geological community cannot be overstated. The cost to NASA of resurrecting high quality reference photographic products from the digital data by individual researchers would greatly exceed the costs incurred during the nominal mission phase. The proper preservation of first-generation and master photographic products (negative and positive transparencies, respectively) should be policy. For illustration, the digital data record for Lunar Orbiter images is inaccessable due to hardware changes over the last 15 years, but the photographic record is preserved. Although federal archives exist, the transfer of these products to such an archive involves the risk of inadvertent destruction and un-managed (with respect to the planetary community) access. Consequently, the following specific recommendation are made.

a. First-generation negatives produced in future missions should be considered "flight films" which are to be placed in a climate and access-controlled, dedicated facility. The first-generation negatives will be used to produce a second-generation positive transparencies from which third-generation negatives and subsequent prints can be produced for selected Team Members and the RPIFs. This sequence preserves the integrity of the first-generation products in case of damage to the "master" positive transparency. This second-generation "master" should be preserved once the mission is completed (perhaps in a Federal Data Center).

b. "Flight films" require proper handling, storage and controlled access. Silver-based emulsions represent the only proven medium for image preservation over a 100 year lifetime, provide they are processed and stored correctly. Well-defined specifications for film processing exist and should be followed. In order to insure proper storage conditions, access, and responsibility, the "flight" films and "master positives" should be maintained at and managed by a central facility.

c. Dye-based emulsions (color processes) are intrinsically unstable having lifetimes from 2-20 years. Irreplaceable color products should be preserved as three-color separation negatives (or positives).

d. Ideal storage conditions for first-and second-generation products include the following requirements.

- Climate: constant 65-75oF with 40%-50% relative humidity (0oF, 25-35% humidity preferred).

- Storage form: storage on stable (chemically inert) spools in chemically inert containers, e. g., black plastic (commercially available).

- o Access: First/second-generation products should be limited to successfully reviewed requests. Third-generation products residing in RPIFs can be accessed through RPIF management to qualified users.

5.6 Imaging Requirements on the Planetary Data System (PDS)

Utilization of the basic archived image data, both digital and hard-copy, is customized to the discipline requirements of the user. Two major discipline groups make up the principal users of image data; these are the atmospheric scientists and the geological scientists.

Scientists in the planetary atmospheres discipline tend to analyze the data in the temporal domain as well as the spatial domain. This involves the use of a variety of software tools which generate a time lapse sequence and support the subsequent analysis. Geologists analyze data primarily in the spatial domain, but require many images of the same surface feature taken under differing directional, phase, range, and lighting conditions. Primary analysis is performed using hard copy products. Digital data are employed in making photometric analyses or in multispectral studiés. Requirements for termporal support are primarily met by availability of data from repeated missions to the same planetary locale. It is therefore important to have imaging centers and facilities which can meet the needs of both planetary meteorologists and planetary geologists. It is not a foregone conclusion that these must be in the same place; however, access, organization, and workstation requirements should be similar for both groups.

5.6.1 Use of Existing Data Centers

The community which deals with the analysis of imaging data is unique in that it has been supported by the Regional Planetary Image Facilities which have acted as small-scale or prototype PDS operations handling hard-copy and (to a lesser extent) digital image data. Close examination of the data services provided and discussions with the various facility directors have provided insight into the support operations for imaging data and requirements necessary to meet future needs. Discussions are developed around current use of the RPIFs.

5.6.1.1 Nature of the Request

a. Digital data

Tapes: The user generally provides a PICNO or FDS and the librarian identifies the tape ID. Requests are of two types.

1) EDRF data for a given image.

2) The IPL JS tape which contains partially processing data.

b. Photographic materials and reproduction of photoproducts:

1) For use in a publication, generally an 8 x 10 print is produced.

2) If the image is needed for science anlaysis, the print is made to a desired scale in the photolab.

c. Access requirements

Searches can be made on the image data base for images that meet given parameters such as resolution or sun angle. This can be done by the librarian or by the investigator.

d. Prevalence: most requests fall within the following categories:

- Searches for specific images or products (e.g., mosaics, maps, color products)

- Reproduction of hardcopy (for publication, scaling, and cropping purposes and slides)

- Reproduction of magnetic tapes from various levels of processing (i.e., EDRs or second level processing)

- Documentation on data sets (i.e., tape formats, hardcopy enhancements, BIRP program)

- Information on relevant publications (e.g., science reports, special NASA publication (SPs, TMs)

- Cross-reference capabilities (i.e., figure numbers vs. image numbers, pre-release numbers vs. image number)

5.6.1.2 Most prevalent problems:

Certain problems which affect the quality of the data, ease of access and interpretation, and influence the ability of the user to interface with the system and acquire data must be addressed in planning a PDS. These are:

a. Definition of the desired product - as there is no standard identifier which is unique to the image as well as its processing version or its status (digital or hard-copy), how does the user define exactly and suscinctly what he wants?

b. How does the user know what is available either in coverage or in products without prior involvement with a project, etc?

c. How best are inaccuracies and incompletenesses in the data set or its catalog made known to the user? How are these notifications changed when the deficiency is corrected?

d. What means are available to expedite delivery of the designated products, both digital and hard-copy, to the requestor?

5.6.1.3 Accessibility of Imaging Mosaics

Imaging sequences which have been fitted-together by hand (mosaicked) to generate a scene are non-digital products which are but highly useful to both the atmospheric and the geologic scientist. The following suggestions are offered to enhance their availability to those who need them.

a. Make sure all mosaics are distributed to the RPIF's, universities and NSSDC.

b. Also send negatives to those facilities that have reproduction capabilities.

c. Update mosaic catalog frequently. It should be done 2 - 3 times a year, so users can keep abreast of most recently made mosaics.

d. Find out who really needs and uses the mosaics and adjust the distribution.

5.6.2 Planetary Scientist User Requirements

5.6.2.1 Planetary Image Data Requirements

Planetary images are a sophisticated data type which require auxillary information for proper use. Accurate and faithful archiving and ease of access to supporting data are crucial for successful utilization of imaging data sets. The lack of or incomplete nature of such supporting data has often been a significant impediment to full utilization for planetary research. These auxillary or ancillary data include the instrument engineering data; the spacecraft identification, time, and geometry; and the calibration factors or algorithms needed to correct the data to give useful parameters. The basic requirement on the archive would be to maintain a capability which can read, display, and process data from all missions stored in these formats.

a. Digital - Digital data are necessary for quantitative measurements on most data sets. Photometric and geometric analyses use data in digital form, and may utilize both the raw and calibrated or partially processed versions. Calibration files must be provided for reduction of raw data or, if the data are decalibrated when distributed, adequate documentation must be provided to the user to permit interpretation. If updates occur in decalibration files, this information must also be distributed. Some consideration may be given at a secondary level so alternative data formats for a subset of the most common home - institution processing systems, or to establishing a inter-institution standard. The SEDR must be available for quantitative analysis.

b. Hardcopy - Hardcopy format is generally necessary for photogeologic analyses. While arguments can be made that this is anachronistic, anyone who has been involved in photogeologic mapping knows the utility of being able to spread out 15 or 20 enlarged photographs and mosaics for comparison. This is a fundamental data format requirement for planetary photogeologic analysis.

   It may not be necessary to assemble a complete hardcopy data set for a given mission. Where multi-spectral data are involved, it may be adequate to only produce only the mid-spectral range in hardcopy (e.g. Voyager violet-frames for satellite data). Far-encounter approach imaging may be more useful as a digital data set (i.e. more appropriate for full disc photometry than for photo-interpretation) and thus may not need to be produced as hard-copy. In these cases subsequent research needs may dictate hard-copy requirements.

c. Analog Video Discs - Consideration should also be given to rapid dissemination of planetary photogeologic data by video disc. This would require some modest image processing of the raw data to provide appropriate contrast enhancement. The economy and ease of distribution of the discs is attractive, but in the future, video discs should be produced as part of the initial project data compilation, rather than as a later follow-on (e.g. Voyager/Viking images). Videodisc facilities are an essential part of the browse characteristics of Regional Planetary Image Facilities. The discs should also include time sequences of images for studies of atmospheric motions.

For photogeological research data is required in several formats. At the most basic level we require raw, digital data plus adequate and current radiometric and geometric updated decalibration files. Data must remain in digital form for quantitative studies (e.g., photometry, photolinometry, stereogrammetry). For qualitative studies (e.g. mapping, feature identification) hardcopy data in some convenient processed version should be available, not necessarily (but preferably) in all available spectral bands. In all cases, precise and updated current ephemerides need to be provided. All data should be in a format amenable to rapid comprehensive search for a range parameter values.

Images should be available in time sequences, and software should assess the effects of varying geometry. Display programs must be able to supply local planet time and season.

Analysis products derived from images should also be displayed in image format. Examples include slope, elevation, wind streak, channel, crater density, tectonic feature, and volcanic feature maps or displays.

5.6.2.2 Catalog Requirements

Perhaps as crucial as the data and its decalibration is the maintenance of adequate data catalogs. Catalogs should have not only identification of pictures by a unique picture number, but ideally should have breakdowns by geographic locations, time of day, camera parameters, spacecraft parameters, and geometric parameters These data should be accurate, timely, and available in formats appropriate to home-institution hardware, as well as in printed format. If processed versions of various image data are available, the processing history of each version should be annotated either in catalog form or in the processed image label.

5.6.2.3 Multispectral Analysis Requirements

The scientist must be able to locate frames forming a mosaic and/or a multispectral image, register these data or be able to obtain registered data, and obtain or produce cartographically controlled and rectified base map mosaics for a region of interest on a planet.

Geodesy is an important discipline in itself and one which supports other types of data analysis. For this type of work access to the digital image data base is required. A small, simple workstation is needed to support limited image processing, such as contrast enhancement and filtering. Geodesy does not require internal (camera distortions) or external (pointing) geometry information since these are solved independently. Generation of hardcopy is one of the biggest problems.

5.6.2.4 Archiving of Geophysical Data Derived from Imaging Data

In the analysis of imaging data for studies of planetary atmospheres geophysical information becomes a basis for meteorological studies often in conjunction with data from other experiments. These geophysical data are the result of a large amount of data processing which could not, and should not, be repeated.

Examples of geophysical data derived from imaging data are summarized below.

Atmospheric winds - from tracking of cloud features in time-sequence images.

Median Dust Carrying Capacity - from sequences of Viking Lander pictures.

Cloud heights/albedos - from multispectral imaging data

5.6.2.5 Direct Access to the Processed Data

The top priority should be for direct access to the data; that implies organization of RPIFs in Europe. A central archive must be funded and established to curate the digital image data. This center would be responsible for saving old data sets and updating and distributing data to regional centers. It will also establish requirements for future missions data sets. It would create accurate ancillary data (SEDR type info, navigation, calibration corrects) to place in the Planetary Data DBMS for search and browse capabilities. In the RPIFs, two types of data are necessary:

a. Complete sets of negatives and hardcopy of radiometrically calibrated and geometrically corrected images;

b. Digital data, that allow local image processing.

5.6.2.6 Browse/Retrieval System

The browse and retrieval of the planetary data imply a catalog access mode. This catalog must be in a digital format that allows selection of the data (images, mosaics, maps, altimetry, gravimetry) with a videodisc system or any other kind of display system.

5.6.2.7 Image Processing

The RPIFs do not produce hard copies of planetary images for distribution; and therefore do not replace the function of the NSSDC. Local centers should have a capability to produce their own copies for local users and to process digital data for special purposes such as image color processing and digital filtering, as well as reducing gravimetric data.

These requirements, and especially the requirements for digital reduced data and digital processing, should be made compatible with the data to be delivered from future planetary missions.

5.6.2.8 Requirements of general users with less sophisticated facilities

The small user is generally forced to a more limited operating framework. This means that resources, time, background, and requirements are also limited. For the planetary data system to be generally useful a major effort should be focused in the design requirements of the data structure, the data base structure, and data distribution system.

A catalogue system should be on-line accessible by modem interface from distant locations. Not only should the catalog indicate the availability of observations, but also descriptions of the data should be provided. One line descriptions are often inadequate. The small (or occasional) user is likely to be less sophisticated about data systems in general and less knowledgeable about PDS. An overly complete data description with cross-referencing would be very useful in helping to focus limited resources on achievable tasks.

The data which are requested from the distribution facility should be self-contained. The data-on-tape should be self-documenting. Header and trailer labels should give the history of the data. Most users can accept limitations of the data set, previous data processing, etc.; but they need to know what these are. Even if the header/trailer labels constitute a sizeable fraction of the whole, it is necessary for the small user since resources for requesting more documentation will be limited. The same explanatory material will also be useful to the frequent or large-scale user; after becoming familiar with the data set, the large user can simply skip over the documentation with little or no wasted effort.

5.6.2.9 Issues Regarding Data Release and Proprietary Data

In the past there has been a "proprietary" period when a principal investigator or team leader and associated investigators retained exclusive right to use of data. In many cases this was defined to be six months after receipt of data, except for image products that were released to the press or published in scientific articles. Publication meant that these data were immediately available to the scientific community. Presently there is some confusion about the rules for data access. Two aspects of this issue are relevant: 1) access to hard copy; and 2) access to digital records. Digital data include the EDR, the SEDR, calibration data that allows radiometric and geometric corrections to be made, and a time history of versions distributed as in many cases the data base improves with work and time. The proprietary interval allowed should be defined for each data type and the release ensured at the end of this agreed upon interval. The data must be stored in an available facility such as the National Space Science Data Center and that facility must be accessible by the general scientist. Delivery of these bases sets should be a project responsibility and funds must be designated to allow this delivery.

5.6.2.10 Research in which images provide context for other data types

Requirements may be placed on an image data system by users of other types of data. An example of this might be the scientist who wishes to compare IR radiometry data with imaging taken at the same season on Mars. It must be possible to register non-image data, which are likely to be of much lower resolution, to the digital image. To enable this, geographic coordinates must be associated with pixels. Graphics capabilities must allow the image and the other data set to be simultaneously displayed, and there should be a computer capability for binning the image data at the same resolution as the other data set. Also, the photometric information may be as important as the interpreted morphology in the image and it may be necessary to extract image intensities for a single pixel or averages over many pixels.

5.7 Summaries and Recommendations

5.7.1 Summary of Requirements

Image data will play an important role in all stage of future planetary missions including planning, implementation, operations, and data analysis. Users from Projects to individual scientists will need access to digital image data. The data must be available in raw or decalibrated form (with ancillary data) as appropriate for each need. In the immediate time frame it is suggested that an appropriate organization should prepare a study on how a unified central archive could be established and how a data distribution system could be implemented. Prototype EDR and SEDR standard formats should be developed as part of the PPDS activity. In addition, PPDS should develop standard decalibration schemes and investigate appropriate DBMS software to meet the needs of an image database. In the longer term it will be necessary to establish an image archive and a distribution system.

As new high-density digital storage media become available, crucial and relevant primitive digital data sets should be housed at discipline-oriented centers in addition to the central archive. Current technology permits condensing 5000 Viking EDR tapes to 500 tapes. Anticipated high-density storage on magnetic tapes in the next five years should permit further reduction to 50 tapes. Such advances in technology should permit important back-ups for the central archive.

5.7.2 Summary of Open Issues

The digital image data can potentially be put into good order and could be made readily available with few exceptions. However, three major concerns must be addressed relative to the archiving and use of the planetary imaging experiment data.

- Calibration of raw image data may present extreme problems. Instrument characteristics and certain data peculiarities have never been documented for many of the imaging systems. Raw calibration measurements, if preserved may be potentially impossible to resurrect.

- Supplementary Experiment Data Records (SEDR) and the catalogs generated from SEDR information are rarely complete. Means must be found to easily update these information sources in the post-project phases.

- Important questions remain unanswered regarding the hard-copy data. These involve storage life, central storage of the original flight negatives in a "safe" environment, what factors constitute the safe environment. In addition, consideration must be given to the handling of color photoproducts, as these may differ from the better understood black and white products.

5.7.3 Recommendations

Development of an administrative structure for archiving and distribution of all imaging data is recommended. This should include establishment of a central archive with overall responsibility for all raw digital data, flight film, and related calibration and spacecraft parameters. An administratively separate system should be established for the distribution of data.

It is also recommended that adequate resources be committed to preserving the existing spacecraft imaging data sets, since these represent the basis of current and future scientific investigations. The archive must include the spacecraft trajectory parameters (SEDR) that will enable quantitative investigations to be made. Future imaging data will be derived from CCD and image spectrometer instruments, which have different characteristics from vidicon systems used for past missions. We recommend that developments made at the central archive to support the reduction of all data to a common format unit which is compatible with both new and existing imaging data.

Specific Recommendations

a) The most primitive image formatted data, the EDR level data, should be saved in the central, archive.

b) Versions of the image data with camera signature removed (radiometric and geometric decalibration) should also be saved.

c) Software and calibration files necessary to "decalibrate" images, and the pointing and ephemeris data should be saved in the data base.

d) All data from current and future missions should be documented and prepared for archiving in a timely fashion, e. g. within a year after the data are obtained.

e) Each project should specify in its Project Data Management Plan the additional data products that would be most useful for future investigations - map products, filtered images, mosaics, etc.

6. NON-IMAGING REMOTE SENSING DATA

This Chapter addresses spectrometric, radiometric and polarimetric remote sensing observations of wavelengths from gamma-rays to microwaves. The basic form of the data is one dimensional arrays. At the high energy end of the spectrum, data are typically presented as pulse count versus energy, and at lower energies, as intensity versus wavelength. High spectral resolution measurements (better than 1% of wavelength) are particularly useful for identifying atomic, molecular, and ionic species while broader band measurements are adequate for identifying minerals and for determining total energy fluxes. Polarization data permit the study of finely divided material such as clouds and surfaces.

6.1 Criteria for Basic Observational Data and Catalogs

The heart of a Planetary Data System is its archive of basic observations and catalogs. Each set of basic observations consists of four parts: instrument description, calibration procedures, the basic data set, and documentation of the dataset. These four parts, as described below are to be provided by the experiment team and verified by the center.

6.1.1 An instrument description must contain, at minimum, all parameters pertinent to data interpretation which cannot be altered by command, such as field of view, spectral range and resolution, instrument or filter functions, and instrument noise characteristics. References should be given to more detailed descriptions in the open literature. (Consideration should also be given to depositing unpublished engineering details which might be of use for designers of future instruments.)

6.1.2 A description of calibration procedures and detailed tabulation of relevant calibration measurements is required. In conjunction with the instrument description, this should allow the data user to understand any algorithms included for the reduction of raw data (see 5.1.3a).

6.1.3 The basic dataset is composed of three types of information:

a) Data is required both in raw form and in physical units. It is important to archive the raw data accompanied by algorithms and software to enable calibration and reduction to physical units. This has the advantage of being easily updated as improved calibration procedures are devised and presented to the data center. It is necessary to archive data calibrated in physical units. In both cases records must be maintained specifying the appropriate update for the data calibration. All data should be included, with quality flags, since even marginal data can sometimes be of use with the development of new analysis techniques or upon correlation with other data sets.

The degree of processing which leaves data no longer in raw form is necessarily ambiguous and dependent on the nature of the data. Conversely, there is no point archiving data in a form so raw that Golay encoding used for the telemetry has not been deconvolved. It has been proposed that an appropriate operational definition of raw data is "data which have been processed only to the point that instrument calibration has not been included and on which no sampling, editing, averaging, etc. has been performed." It is the feeling of this group that the final decisions regarding the exact nature of raw and calibrated data sets rest with the experiment teams. For future missions where time sequential raw data may be available via an "end-to-end" data system, the archiving of such data should be a project responsibility.

A prime objective of many remote sensing experiments on spin stabilized spacecraft, e.g., Pioneer 10 and 11 IPP; PVO OCPP (see Appendix 1), is to obtain spin-scan "images". Though somewhat distorted by changes in perspective and by temporal changes in the target during the acquisition period, the image format permits extensive investigation of phenomena such as large scale cloud structure and motions. Such images, provided by the experimentor to the PDS in digital form with the rectification procedures appropriately documented, should be available to users with equivalently processed imaging data. Users can then utilize the image processing capabilities of the PDS to stretch, enhance, and otherwise analyze the data.

b) Detailed description of the instrument state for all observations is required. This consists of all commandable instrument parameters such as filter setting or grating position, field of view size (where programmable), gain state, and integration period as well as environmental parameters (such as temperature) which may affect performance.

c) Detailed information on the geometric and temporal state of the observations (pointing) is also necessary to the archived data set. This information is normally supplied to experiment teams as a Supplementary Experiment Data Record (SEDR). For each observation the SEDR contains: the time of each observation; target object; and detailed specification of the illumination and viewing conditions for instrument footprints (or specifications of the tangent ray if off the limb of the target). To provide direct access capability, it is normally necessary to have this information directly accessible in a header record. As with calibration, however, the determination of accurate pointing information often requires repeated revision. Storage of raw data (spacecraft and target ephemerides, with spacecraft and scan platform orientation parameters including positions within limit cycles) together with necessary algorithms and/or software will allow direct determination of pointing and will provide for updating. Dropouts in engineering data streams can provide gaps in coverage which require specialized algorithms and additional information to produce SEDR information which makes the process potentially highly complex. In the future, the determination of accurate pointing information (a perennial sore point with many experimenters and data users) requires much greater consideration at the project and spacecraft design levels than has been the case. For existing datasets, creation of software for the a posteriori determination of pointing through the use of images, where available, should be seriously considered.

6.1.4 Documentation of datasets should emphasize quality control criteria and known anomalies in the data or in ancillary information. All formats should be specified in detail.

Catalogs of the data, including concise data summaries, are critical adjuncts to the documentation. The basic catalog should contain the most important characteristics of the dataset; such as, type of observation, data storage medium, quantity, data quality, and level of processing. The catalog may be provided by the experiment team or may be compiled by personnel at the Data Facility.

The Data Facility, through scientifically qualified personnel, should review all incoming material for its adherence to PDS standards. To assure that all requirements of the facility are met and to keep each data set current, it may be appropriate to assign an "Archive Representative" to act as liaison with each experiment.

All data should be submitted in digital form. Data products should be available in a variety of forms, e.g., electronically, hard copy, magnetic tape, and disk; potential users should not find themselves unable to utilize the data simply because they lack the most up-to-date equipment.

6.2 Enhancements to Basic Observational Data and Catalogs

The usefulness of the data system can be increased significantly with the inclusion of the following types of additional derived and supporting information.

6.2.1 A catalog of cross references between experiments and experimental results should be provided. It should include lists of experiments, derived datasets, models, and available software, with indexing by results and by discipline.

6.2.2 Datasets derived from analysis of the basic data (e.g., atmospheric temperature profiles or maps of surface thermal inertia) should be accompanied by appropriate references and documentation on methods of analysis, algorithms, and discussions of uncertainties. This documentation should be actively solicited from the user community. These derived data sets should be reviewed by center personnel as carefully as (or perhaps even more so) the original experiment team submissions. Each data set should be accompanied by documentation indicating uncertainties.

6.2.3 Bibliographies of results obtained from the datasets should be included in a PDS. The small amount of total space taken up for reference lists argues that they should be fairly complete.

6.2.4 Models for various derived datasets, processes or phenomena, and baseline models for first order data fitting or analysis should be included in the archive. Beyond data manipulation, it is desirable to include scientific codes simulating physical processes in the archive together with appropriate documentation, literature references and statements of authorship. The placement of such codes should be encouraged and some review of such submissions should be provided. However, the data facility will have only limited responsibility as to the scientific reliability of the models. Authors may provide for protection of codes (which they feel are proprietary), via submission of compiled codes (e.g. load modules) as opposed to source codes, as long as basic documentation is supplied. Running of such codes on the central facility will place the greatest demands on the system and are regarded as lowest priority among the user needs to be provided by a PDS. Examples of appropriate physical models are aeronomical chemical composition, radiative-convective temperature equilibrium, surface/subsurface temperature profile and time-dependence, the basic versions of may have a broad utility.

6.2.5 Algorithms and software for data analysis vary from highly generic, modular codes (e.g. Mie scattering) to somewhat model-dependent codes (e.g. temperature or abundance sounding because their use is driven by assumptions made about the physical environment, especially regarding composition). General use of these techniques represents a potential danger. Any program, no matter how well documented, can be a source of severe problems. Not only may it contain undiscovered bugs, but much time may be wasted in trying to understand and use it. Appropriate programs are, however, extremely useful and save the user much development time. Documentation for all programs should include detailed elaboration of the coded algorithms; both to enable the user to assess efficiently the applicability of a given program to his needs, and in recognition of the fact that software is system specific.

6.3 Present User Community, Data Sets, and Archival Facilities

The membership of the American Astronomical Society alone contains over two hundred active scientists with direct or closely related interest in non-imaging remote sensing data. Large numbers of additional users, such as geologists and graduate students, could be expected to utilize such data if the datasets were more available and easier to use than at present.

Based on funding and facilities, the user community can be roughly divided into three groups. The first group of users is constrained by limited budgets or time; such investigators have not generally been directly involved in flight missions. Their facilities might typically consist of a terminal with graphics capability and a hard copy unit attached to someone elses' CPU, or to micro-sized CPUs with limited I/O capability and modest processing software. The second group of users includes rocket payload investigators, teams for some low data-rate spaceflight experiments, and some co-investigators on major mission teams. Their facilities may be represented by full-sized minicomputers with significant I/O, graphics, and hard-copy capabilities. The third group of data users are typically PI and experiment teams on major high data-rate flight experiments. Their facilities are designed to perform complex processing tasks and frequently cover the multiple functions of data formatting, calibration and extensive data analysis. These systems include large minicomputers (DEC 11/70, VAX, PRIME) or mainframe CPUs; they often represent small scale computing centers which serve both local and, through long distance communications channels, remotely located users.

A list of pertinent datasets residing in the NSSDC is given in Appendix 2. The quality, potential future usefulness and condition of these data remain to be assessed. Requirements for the refurbishment of datasets to be incorporated in the PDS should be established. Priority should be given to data presently heavily utilized (e.g., Viking, Voyager) and to data of critical use for near-term mission support (e.g., Voyager in support of Galileo).

A number of archival facilities presently exist and their approaches to data recovery should be examined. One outstanding example is the literature retrieval service at the Lunar & Planetary Institute. Another is the Mars Consortium containing a large number of map-type and other data which should be considered derived datasets and incorporated into the PDS facility. The Mars Consortium successfully allows intercomparison of mapped datasets by using a common latitude/longitude binning with bins sized in powers of two.

6.4 Entry into the Data Facility: Nature of a Request for Data

A user approaches a data facility with specific scientific objectives in mind. Additionally he usually has an idea of the types of data and conditions of observation which will be relevant to his problem. The primary communicative objective of the data facility should be to enable such a user to locate this information as rapidly and efficiently as possible. An efficient method of sifting through the mass of data is imperative. The starting point should be a general catalogs; and the system should provide comprehensible instructions to the user.

The efficient design of catalogs and database management or searching systems should be a high priority for software definition groups. An immediate difficulty arises because present datasets are archived under a variety of formats. The first requirement is to introduce a system of uniform formats for temporal, geometrical, and instrument state variables (headers), or a software system which makes actual format differences transparent to the user. This should result in the ability to sort quickly, at a relevant level of classification.

It would also be of value to address the data intensities themselves in the search process. At this point a host of additional problems arises. For example, one might wish to compare Voyager infrared spectrometer data with Galileo orbiter infrared PPR data, this requires that an intermediate "conformal" database to exist or to be created. Since the PPR provides broadband thermal data as a function of wavelength, it is necessary to integrate the IRIS data over the PPR passbands. This integration might be accomplished by: 1) the host institution providing computation facilities to permit this on line; 2) the host institution creating such a derived dataset independently in anticipation of need; or 3) leaving this type of comparison to the user by providing him with the two datasets (after which the user would be encouraged to contribute the derived dataset to the data facility). Clearly it would be desirable to have the derived dataset "available" at the data center, but which approach is most feasible depends on the data center configuration and organization.

Under the assumption that data intensities are available for use in the "browse" mode, a package of statistical routines available for on-line use on the host computer would be very desirable to investigate further properties of the datasets further (e.g. to look at variability or to work with low signal-to-noise ratio data by averaging). This type of activity we collectively call data "manipulation". The facility may require that this be restricted to subsets or averages of the entire data set in order to reduce the computational burden on the system. Such activity is distinguished from data "analysis" which is substantially more extensive and deals with a comparison between the data and the simulations of specific physical processes, i.e. "models," in order to derive various physical parameters from the data.

In any event, the successful creation of a catalog/browse facility requires reference to actual attitude information and to data in physical units. Unless extremely fast methods are available for determination of such quantities from basic spacecraft and target state vectors, and from raw data, reduced data records must be included. Unless these quantities are regenerated each time new attitude or calibration information is submitted, they will become out of date (derived datasets, such as might be contributed by outside data users, must clearly indicate the particular data calibration used, since it will be impossible to updated these as more refined calibration data or algorithms become available). While basic state vectors and raw data should be available for users, the bulk of the data center activity will ultimately deal with reduced data records.

In an advanced system, capability should be provided to permit video terminals to be utilized with plotting routines from the data facility. This would significantly enhance the browse capability, by permitting visual assessment of maps, spectral data, and correlations.

The result of any successful data search will be the identification of a non-empty dataset for the user. Beyond this point the contribution of the data center depends upon the user's analysis task and upon the data center configuration.

6.5 Role of the Data Facility in Data Analysis

The user should be able to obtain the data for analysis; and in order to accommodate the user's system these data should be available in hard copy, on tape, on disk, or through electronic transmission. An elaborate Data System might provide users with algorithms, software and computational facilities for data analysis. As indicated by the Planetary Data Survey, the processing capabilities most desired are calibration and attitude evaluation. Graphics capability are desired for generating maps, contour plots, etc. Subroutine libraries of frequently used algorithms were considered desirable by many, but required by few, perhaps reflecting caution over the use of "black boxes". With regard to more elaborate analysis software, such as radiative transfer codes, opinion within the splinter group is strongly divided (see item 4 under Enhancements to Basic Observational Data and Catalogs).

Several members of the group feel that a visit to the Data Facility is mandatory in order to learn to effectively understand and utilize the data through working with persons familiar with the data and the instrumentation. Concern exists that without such interaction and/or some basic screening mechanism, e.g., proposal submission or user charges, the computational capabilities will quickly become saturated with unproductive work. Even assuming that ground rules and procedures can be set up to avoid unproductive work, the potential risk of saturating the computational capacity of the Data Facility must be given very serious attention. It must be recognized that a substantial computational power for the host computer(s) at the Data Facility will be necessary, and that the priority should be placed in the following order: (1) quick and easy access to the catalog(s), (2) browse capability and limited data manipulation, (3) limited data analysis.

6.6 Configuration of the Data Facility

Given the state and trends of technology, an effective PDS should incorporate electronic access. Prototype systems already exist (Appendix 3).

It is agreed that the data facility should involve scientists who are actively working with the data. This will provide user oversight and curatorial management of the data, resulting in a higher quality and more responsive data system. In order to guarantee this it is imperative to incorporate investigators from ongoing experiments into the system. Active scientists should be present at all times.

The most certain way to guarantee active scientific involvement is to incorporate the intent of the data facility into the experiment teams. Funding for preparation of data would be allocated through the central data facility at a percentage of data analysis costs. A contractual arrangement, separate from the project data analysis plan, would be negotiated with the central data facility. Funds not spent by an investigation which successfully completes its contractual obligations would revert to the experiment team for data analysis. This assures that the most instrument-knowledgeable personnel are involved; should help, to assure availability of the most current datasets; and will engage the largest number of active personnel. It would also promote advanced planning on the part of the investigation team to provide data formats fully compatible with the data facility requirements from the outset.

Care must be taken to respect proprietary data rights and to avoid imposing an onerous burden of providing user services. For an experiment team failing to meet its contractual obligations to the data facility, the facility may then elect not to renew the contract but use allocated funds to provide its own personnel to complete the data transfer or aid the experiment team in this effort in the form of an Archive Representative. The experiment team may elect to acquire the aid of such a representative from the outset. As missions are completed and experiment groups dissolve, complete data transfer to a central facility would be made.

Another possibility is to set up networked Centers of Excellence which are formed around "instrumental techniques", e.g., UV spectroscopy. These would incorporate knowledgeable individuals from active experiment teams to yield much the same team-investigator oversight as the above. Additional support for the archive might be obtained by providing fellowships at the Centers under which investigators would split their time between research and archive activities.

A third possibility is to establish a network of Discipline Centers, the emphasis being on single large areas of scientific interest, such as meteorology or geology. Such centers would probably involve investigators from ongoing spacecraft missions, so that experimentor oversight would exist. This might provide the best environment for non-project scientists since focus would be on scientific rather than instrument issues. In addition, Discipline Centers might provide the best incentive for the construction and incorporation of derived datasets, as it would be in the Centers' vested interest to do so. Research/archive fellowships might also be offered.

The foregoing discussion suggests three levels of an evolving PDS, each to include preceeding ones:

Baseline System

Refurbish worthwhile existing datasets and catalogs (create catalogs, if not available). Engage experiment teams in maintenance of heavily used current datasets. Provide simple dial-up remote terminal access to catalogs (probably networked to specific catalogs at involved experimentor sites). Data to be available to users as hard copy, tape or disk. Create an oversight committee to review the operation of the PDS in light of users' needs.

Intermediate System

Actively solicit derived datasets from scientific community, to be subject to review by committee of active data users. Develop sophisticated cross referenced catalogs to handle data from experiment to discipline level. Provide calibrated data with pointing information to on-line browse facility. Incorporate statistical and plotting capabilities in host computer to support data manipulation and display from remote video terminals. Provide fellowships for investigators to conduct research and aid in data curation at the facility.

Enhanced System

Establish a network of research/archive facilities based on scientific disciplines or instrumental techniques to support use of data in interdisciplinary studies. Incorporate models, algorithms and software; provide data analysis capability on host computer to the extent that it does not diminish the capabilities of the primary storage and search functions.

Appendix 1: Pioneer Venus Orbiter Cloud Photopolarimeter Imaging

The two principal types of observations made by the Orbiter Cloud Photopolarimeter (OCPP) are 30-km resolution images of the clouds of Venus in the near ultraviolet and lower spatial resolution (250 to 500 km) maps of polarization over the disk in four spectral bands. Both the images and polarimetry maps are acquired by means of the spin-scan approach. With the OCPP telescope fixed at an appropriate look (cone) angle, the disk of the planet is scanned by spacecraft rotation, one scan line being generated each roll of the spacecraft. Polarimetry requires sixteen such rolls for one complete cycle, hence the lower spatial resolution. Since the spacecraft spin axis is oriented perpendicular to the ecliptic plane and the orbit is not far from polar (105^o inclination), the spacecraft orbital motion can provide the necessary cross scan translation. The completion of a full disk image or polarimetry map in this manner requires about four hours of continuous data.

The raw data for a typical full-disk image is composed of approximately 1000 scan lines, with up to 800 8-bit intensities, or pixels, for each scan line. Because of the nature of the spin-scan process, an image constructed simply by treating each pixel as a rectangular element and placing scan lines in a continuous sequence would be very distorted in appearance. Hardcopy image products are normally created from a digital data array which has been quasi-rectified using variable bilinear interpolation of the raw intensity data. This 'rectified' image has an appearance similar to that of a snapshot image; in particular, the bright limb is a circular arc. Note that this image is not equivalent to an orthographic projection. The spin-scan technique with the OCPP telescope fixed at a particular look angle during the acquisition of data for a full-disk image yields coverage of a full 180^o of latitude. However, the coverage for each scan line is limited to the fraction for the disk instantaneously visible from the finite spacecraft distance, i.e., a 'longitude' range equal to the supplement of the apparent angular diameter. Because the spacecraft orbit is not exactly polar, there is also a small change in the sub-spacecraft longitude during the acquisition of the image. As a result of these factors, along with the desire to produce an image that includes all of the data, the 'rectified' images are slightly distorted in comparison to the exact location of features in the hypothetical snapshot image. Further, there is not simple relationship between image coordinates in this version and body-fixed latitude and longitude.

For any applications which require geometric fidelity such as feature tracking, etc., a true geometric rectification of the image data must be performed. This involves the determination of body-fixed latitude and longitude for each raw image pixel, binning those pixels by latitude and longitude, and then reprojecting the image data into an appropriate projection such as the orthographic. Such an approach must distort the relationship between true illumination and limb geometry, or discard some disk coverage, or a little of both. However, regardless of the specific image 'rectification' employed, the resulting digital image data array can be processed using the same standard procedures (such as spatial filtering, contrast enhancement, etc.) applied to vidicon or CCD image data.

Appendix 2: Existing Datasets

Datasets considered to lie within the purview of non-imaging remote sensing observations (passive electromagnetic spectroscopic, radiometric and polarimetric observations from gamma-ray to microwave) are listed below. Suggestions for further inclusions are solicited.

Specifically excluded from consideration are the Apollo Laser Altimeter, Lunar Sounder, and Laser Ranging Retroreflector as being basically active ranging experiments for consideration by the Radio Science Splinter Group; and the Pioneer Imaging Photopolarimeter as falling under consideration of the Imaging Splinter Group.

Table A1. Datasets in NSSDC Considered Significant by the Non-imaging Remote Sensing Data Definition Splinter Group

Spacecraft	Experiment	Tapes	Mbytes	Requests
Mariner 2	IR Radiometer	---	---	1
Mariner 5	UV Photometer	---	---	---
Mariner 6,7	UV Spectrometer	4	1	0
	IR Spectrometer	---	---	1
	IR Radiometer	2	¼1	2
Mariner 9	UV Spectrometer	---	---	1
	IR Interferometer Spectrometer	5	159	6
	IR Radiometer	1	14	4
Apollo 15, 16	Gamma Ray Spectrometer	47	1,438	3
	X-Ray Fluorescence Spectrometer	2	21	6
Apollo 17	Far UV Spectrometer	5	33	5
	IR Radiometer	(10)	(24)	---
Pioneer 10, 11	UV Photometer	71	4,832	0
	IR Radiometer	---	---	---
Mariner 10	EUV Spectrometer	1 (150)	¼1	1
	IR Radiometer	1	¼1	1
Viking Orbiter	IR Thermal Mapper	30	600	4
	IR Water Vapor Detector	3	60	5
Pioneer Venus Orbiter	UV Spectrometer	10	44	1
	Cloud Photopolarimeter	19	127	3
	IR Radiometer	3	154	1
Voyager 1, 2	UV Spectrometer	2 (500)	4 (104)	0
	Photopolarimeter	2	¼1	0
	IR Interferometer Spectrometer	41	2,261	6
	Totals	249(650)	9752 (10^4)	½51

--- Splinter group did not obtain information
( ) Numbers in parentheses represent estimates for data outside of NSSDC

Appendix 3: Pioneer Venus Orbiter UVS Data Facility (D. Anderson)

The PVOUVS operations are a good example of a data facility that provides a bibliography, browse, data selection and transmission, and analysis capability.

The PVOUVS data base is accessed through reference to off-line indices. These indices consist of a collection of the command plans for each orbit, orbit summary images and command summaries. Additionally, there are files on tape which summarize the disk observations for wavelengths of general interest, e.g., OI 1304, [OI] 1356 and HI Lyman alpha.

The command plan consists of a timeline with the grating position or operating mode recorded. For the Pioneer Venus Orbiter, with an orbital period of 24 hours, this is still manageable even as we approach the end of the 5th year of operation.

In practice a data set is selected by choosing a wavelength of interest and scanning the command plan listing for orbits of interest. The orbits have been written to magnetic tape and can easily be transferred to a scratch area on disk for analysis.

The orbit summary images which are stored on magnetic tape consist of a summary of the data collected over an orbit collected in such a way that the data are displayed as an image on the color graphics terminal. The image is divided into three parts. The "apoapsis" image represents data from the first view period (from about 2 hours before periapsis to about 30 minutes before periapsis), the second section, the "periapsis" image, represents data collected from the second view period (about 8 minutes before periapsis to 4 minutes after). As a final indication of the data acquired for that orbit, a "summary" image contains all the data collected. On each part of the image tic marks are recorded which indicate the time. Data dropouts are also delineated. Each orbit summary image also records the orbit number and instrument operating parameters. In practice, the images allow the appropriateness of the data collected on that orbit. These orbit summary images are also reproduced as hard copies of three dimensional plots and are stored in a binder.

Command summaries are listings of relevant instrument and spacecraft parameters as a function of block number on the data tapes. Among others, the time from periapsis, integration period, wavelength of observation, and spacecraft altitude, are given. This listing enables the user to check the time and altitude range of observations at a desired wavelength. It is generated as a computer printout during the initial processing of the raw data tapes.

The wavelength orbit image summaries are files stored on magnetic tape which contain the apoapsis images for useful wavelengths. This allows the user to compare the images as a function of time or viewing/illumination geometry.

To enable the data base user to access these files in an easy and straightforward manner a subset of IDL (Interactive Data Language), known as

PVIDL has been developed. IDL, currently in use at NRL, the University of Colorado, etc. is a high level interactive vector oriented data processing language.

PVIDL enables the user to select data sets and obtain orbit attitude information such as instrument look direction, spacecraft location and altitude, timing information, sun position, and direction angles. Vectors containing the minimum ray height altitude and solar zenith angle are easily generated. The data can be plotted, using IDL's graphics capability, as a function of any parameter (usually a geometric one such as the minimum ray height altitude of the observation). Due to IDL's extensive analytic capability data analysis is facile. A data vector can be smoothed, filtered, extracted, fourier transformed, data dropouts and spikes can be removed, all with one line commands. Fortran programs can also be run from IDL if the results of computationally intensive calculations are to be compared to the data. At any point, processed or "raw" data can be saved on a mass storage device such as a disk or magnetic tape.

7. IN SITU DATA

7.1 Introduction

In-situ planetary data include all measurements made by a wide variety of instruments whose sensors are exposed directly to the planetary environment. These instruments are mounted upon the spacecraft and collect data along its trajectory. Also included in this category are measurements of the local conditions on a planetary surface made from landed instruments. Typical measurements include those of magnetic fields, energetic particle densities and energy distributions, plasma wave characteristics, ionospheric and neutral gas densities, temperatures, composition and motions. These measurements support a wide range of scientific disciplines including the study of atmospheric, magnetospheric, and solar wind phenomena. In-situ data are especially valuable in interdisciplinary investigations involving such phenomena as solar wind interaction with planetary magnetospheres (Jupiter, Saturn and Earth) and with planetary atmospheres and ionospheres (Venus, Mars, comets, etc). Thus many investigators require in-situ data from a wide range of instruments, obtained by different spacecraft, and of the many planetary targets and interplanetary spaces.

A large body of scientists in the United States, and the world, are involved in research using in-situ data. Many of these scientists are associated with Mission investigations groups which provide specific types of in-situ data. These groups, in the course of data analysis, use data provided by similar groups. The interdisciplinary nature of these investigations fosters exchange and reliance upon one another for data is heavy. To a lesser extent, correlative data sets are obtained from the National Space Science Data Center (NSSDC). However, the data entries often have limited temporal and spatial resolution, are of uncertain quality, and are usually available only for measurements that are several years old.

7.2 The Nature of In-situ Data

In-situ measurements are made using sensors whose outputs are transformed electronically into voltages suitable for digital telemetry to the Earth. The telemetry bit stream is decommutated to produce magnetic tapes containing the raw data from each instrument. These tapes are shipped to individual instrument groups who reduce the data to physical parameters of the planetary medium itself, such as the local magnetic field strength, the electron density, or velocity of the solar wind. These are referred to as processed data or derived parameters. Some in-situ techniques involve intermediate stages in which the actual electrical output of the instrument is an important parameter to be stored as a data base. An instrument calibration must be applied to convert the data to physical parameters. Since these calibrations may be variable with time, or are subject to revision, they are supplied as ancillary information with the data. Thus the nature of the planetary data base differs from instrument to instrument. Only processed data are archived by the NSSDC. Raw data are usually available only from the instrument groups themselves. These groups also have processed data with higher spatial and temporal resolution than is available at the data center itself; and if the spacecraft mission is still active, they will also be able to provide processed data for more recent periods that are currently available from the NSSDC.

7.3 The Need for Processed Versus Raw Data

The vast majority of requests for planetary in situ data involve data that have been processed to physical parameters of the environment rather than raw data or uncalibrated data, as most data users do not have the expertise to convert the raw data to a useful form and are happy to have that service performed for them by the instrument groups. Furthermore, the computer codes for performing the reduction to physical parameters are usually computer specific and may not be adequately documented. In some cases, however, later discoveries about the planetary environment or the instrument operation itself require reprocessing of the raw data (or portions of it) to extract additional information or to improve the accuracy of data reduced earlier. It is important, therefore, that the raw data be archived indefinitely. Often those portions of raw data that may conceivably have future value have already been converted to a more compact form, and these major density tapes may represent the best form for raw data storage.

7.4 Ancillary Data

Ancillary data required for the use of in-situ data include the following:

- orbital or trajectory information such as the position of the spacecraft in latitude and altitude or, in the case of interplanetary spacecraft, the heliocentric position.

- attitude of the spacecraft, instrument pointing information

Catalog, Document¬tion and Bibliography

These topics are covered in detail in the Chapter 3 on user requirements.

7.5 Key Parameter Data Sets

The interdisciplinary nature of in-situ data places special emphasis on the need for the ability to browse through simultaneous data from a number of instruments, usually from the same spacecraft. For purposes of discussion, we identify such data as key parameter data. There are fixed sets of data in a common compatible format, probably with fixed spatial or temporal resolution along the orbit. Such sets are already available at the NSSDC for certain Earth and planetary missions such as the Atmosphere Explorer and the Pioneer Venus orbiter. We envision that the content of these key parameter data sets would continue to be defined by each mission science team in the normal course of mission planning. During the lifetime of the mission, key parameter data would be assembled by the project and delivered to the NSSDC. We recommend that such data sets, for future and even current active missions, be defined and assembled. Recommended standards should be defined to encourage compatibility for browse purposes. When electronic networking becomes widely available to the planetary science community, key parameter data sets should be placed on line permanently for remote access and browsing. Such data, in the meantime, should be available from the NSSDC in the forms of magnetic tape and/or microfilm or microfiche, and by direct computer dial-up.

7.6 Means of Data Access

Planetary data currently in the NSSDC are available only in the form of digital tapes and, in some cases, as microfilm or fiche. Data are also available from Principal Investigator groups in the form of tapes and hardcopy and by direct computer transfer in cases where the user and provider are connected to a common network or the user has dial-up access to the provider's on line data base. The latter are informal arrangements that should be encouraged and supported by NASA. In the near term, we recommend that that NSSDC be supported to provide dial up access to the existing key parameters, and in the longer term, to assemble additional key parameter sets for future dial-up access.

Regarding the more complete data sets that may reside at the NSSDC or at the PI sites, data access will for some time continue to be in the form of tapes and hardcopy, except as noted above when it is possible for the user to dial-up an online data set at the PI site. As soon as possible, node to node transfers of data should be provided for those heavy data users and providers that are expected to benefit most by this technology.

7.7 Levels of Accessibility of Various Types of Data

The catalog data described in Chapter 3 (User Requirements) and its associated documentation should be readily available by computer dial up on a 24 hour per day, 7 day per week basis. The key parameter data sets also should be available on line for dial-up access. User specific requests for other in-situ data sets should be handled on a case by case basis by the NSSDC or the PI groups whose data are requested. Currently most such requests will be met by the mailing of tapes. As node to node transfer becomes feasible and online storage capabilities increase, more rapid transfer of data will be possible.

7.8 Data Manipulation/Analysis

In the design and evolution of a computerized support system for planetary data activities, it will be important to make a distinction between the functions of data delivery and data analysis. Data delivery involves providing the ability to search the data base for a subset of data that satisfies the request. This may require sorting the data by time, or location or range of physical conditions and perhaps manipulating the data to permit various plot displays. Certain standard mathematical transformations or correlations may also be desired by the user. In general, the in situ splinter group felt that the system should be designed primarily to aid the user in selecting a data set that is appropriate to his investigation. The system should then deliver that data as quickly as feasible to the user for analysis at his own facility. The dividing line between data manipulation and data analysis prior to delivery should be established by a more detailed study than is possible here. However, an important factor in the definition will be the cost impact on the computer system and the potential for over-use or abuse of analysis capabilities at the expense of data delivery.

7.9 Quality of the Existing NSSDC Data

The personal experiences of many data providers, and many users, indicate that the data currently archived at the NSSDC is of variable quality. Some data sets have questionable accuracy, while others exist only is limited amounts or have much lower spatial and temporal resolution than the data potentially available from the instrument. The reasons for this are many, but generally they involve the desire of the investigators to obtain the maximum amount of scientific understanding within a constrained budget. After the end of each project, funds are generally unavailable for the routine processing of data for submission to the NSSDC. Seldom can the PI group complete all processing within the period in which project funding is available.

A number of steps will be necessary to upgrade the data base and to assure that future data will meet a standard that will justify the additional expenditures required to improve the data delivery system. This should involve the identification of data sets that are in greatest demand, followed by efforts to remove existing errors and add more recently acquired data. Appropriate funding will be required over a several year period to perform this upgrading. The actual extent of such a program is not easily estimated as it involves assumptions concerning the likelihood of conducting new planetary missions which might provide even better data for some areas of existing data. The availability of knowledgeable groups to perform the upgrading of the in-situ data base demands that any such upgrading activity be conducted within the next few years. The opportunity to assemble a first rate planetary database will have been lost, if reconstruction is not begun quickly.

7.10 Rules of the Road for Users

Many science teams have found it desirable to adopt a list of rules that govern the use of data that are provided by individual PI groups but are shared by the entire team. Some of these rules were devised to avoid errors in the data base which result in misinterpretations and publication of faulty concepts and results.

These rules are suggested as being highly important:

- The data user should contact the provider regarding the accuracy of any data he plans to publish as part of his investigation. This rule is intended to give the data provider the opportunity to verify the data accuracy if he should choose to do so.

- The data user should acknowledge the source of the data and quote references to papers describing the instrument and methods of data analysis.

Although the upgrades to the data base recommended in the previous section will reduce the incidence of error, these rules for data users will encourage users to continue to discuss the database with the data originators in order to become better informed about the potential and the limitations of the data.

7.11 Standards for Future Data Submissions

The standards for the submission of data to the NSSDC are now left to the individual PI groups or, in some cases, to the science teams of the flight projects. The diverse organization of the various data sets makes more the electronic node to node transfer of such data more difficult. It will be important in the future to devise guidelines for future projects and individual data providers to follow when preparing their data for archiving. Detailed definition of such standards is beyond the scope of this document. However, the submission of a key parameters data base in a specified format which allows flexibility for differences in spacecraft and instrument design constraints must be required.

8. RADIO SCIENCE

This chapter is the result of an extensive canvas of the radio science and radar communities. In order to accommodate the findings within the body of the PDS report, it has proved necessary to eliminate several lengthy tables and appendices:

- o Institutions conducting radio science and radar investigations, their hardware, software and telecommunications capabilities. (36 entries)
- o Partial list of members of the planetary radio science community, their affiliations and interests. (84 names)
- o Preliminary catalog of planetary radio science data sets listing investigators, wavelength, site or spacecraft, data product, epoch, reference publication(s), and data set status. (140 references; 27 bodies and solar wind).
- o Current catalog of NSSDC radio-science datasets.
- o List of observatories which might have taken planetary data -- to help future data archival searches.
- o Report by Ray Jurgens on the status of the JPL Radar facility -- as a pilot project in the archival/retrieval of radio-science data.

8.1 Nature of the Data and Data Sources

Radio science experiments use electromagnetic waves to probe or study the solar system. We identify three major research areas within this discipline: radio astronomy, radar astronomy, and celestial mechanics. Attributes which separate the three are given below. These should not be considered rigid separations, but they will aid in the discussion of actual data sets later.

8.1.1 Interest Areas within Planetary Radio Science

Radio astronomy (or radiometry) is the detection and measurement of naturally produced radio frequency emissions. Sources include surfaces, atmospheres, rings, and plasmas. Either "continuum" or "spectral" measurements can be made. Measurements can be of "bulk" properties or, if the source can be spatially resolved, may be presented as maps. Maps are a common product when interferometric (multiple antenna) techniques are employed. Most radio astronomy has been conducted from earth-based facilities (such as the VLA and NRAO at Kitt Peak), but there are several data sets resulting from spacecraft observations (Pioneer Venus and Voyager).

Radar astronomy is the observation of man-made signals after their interaction with a "target". Targets include surfaces, atmospheres, rings, and plasmas. Both imaging and non-imaging data may result. Data may be presented in the time domain, frequency domain, or both. Monostatic radar systems (such as at Aricebo or Goldstone, or on board Pioneer Venus) use the same antenna for both transmitting and receiving. Biostatic systems use separated transmitters and receivers, most commonly with the transmitter on board a spacecraft and the receiver on earth (examples: Venera orbiter transmissions scattered by Venus' surface and received at earth stations in the USSR, or Voyager 1 transmissions through the solar wind to receiving stations of the DSN). We include Apollo laser altimetry and lunar laser corner reflector studies in the radar astronomy category even though they have been conducted at optical wavelengths.

Celestial mechanics is meant here to include all studies related to the motions of (and gravity fields of) bodies within the solar system. Data from radar experiments (for example, Arecibo ranging to Mercury) and from radio transponders (e.g., on the Viking landers) are used. Observations of relativistic effects (such as Voyager at Jupiter) are included in this category. Gravity wave studies by spacecraft are included here, though the observations to date have led to no detections.

8.1.2 Basic Measurements and Calibrations

Raw Data: Most raw data from radio science experiments are in the form of voltage samples taken periodically from one or more receiving antennas. These are rarely saved in their original form, however, some averaging being necessary to reduce the data volume. Whether phase information is retained throughout the averaging process depends on the eventual use of the data. In certain experiments relatively low data rate outputs from radio receivers (such as ranging data in spacecraft tracking) are the most primitive data form.

Calibrations: Proper interpretation of the radio data requires that experimental parameters be understood. There is interplay among some parameters (tracking data must be used to determine spacecraft positions before an atmospheric occultation may be analyzed, for example), but the basic requirements are as follows:

a) antenna gains, polarizations, radiation patterns, and pointing;
b) system functions (filter responses and noise figures);
c) frequencies, times, and sampling rates;
d) reference calibrations (sky sources or hardware devices);
e) transmission power and waveforms (if any); and
f) experimental geometry (positions, velocities, orientations).

8.1.3 General Comments on Levels of Data

Beyond the raw data samples, four levels of processed data have been defined which would be of interest to users. The first level, uncalibrated data, includes data with which the experimentor initially works. As an example, these might be range-Doppler arrays from a planetary radar observation, identified by a time tag and (perhaps) some ancillary information giving operating parameters of the radar system. These data could be of interest to other radar astronomers but would probably be of little use to workers in other areas of planetary science. In H. Kieffer's H22:PDSFLO.DOC 1983 Oct 31 (Generic Data Flow) terminology, these data would be somewhere around level 3.

The category Calibrated Data, groups data which have been converted to standard formats or units. The power spectrum of microwave emission from Venus given in terms of watts/Hz would fall into this category. These data represent the response of an ideal instrument operated in the same experiment given in physically meaningful units.

Derived Data are corrected for certain observational effects: they may be either model independent or the result of the application of a generally accepted model or simple set of assumptions. Radar-determined topographic elevations with respect to the 6.1 millibar surface of Mars would be one example, as would the temperature-pressure profile from a radio occultation. These data could be used by anyone familiar with the general behavior of radio signals and the conduct of such experiments. In Kieffer's terminology, these data would be about level 7 data; if consolidated in a database, they could be considered level 8 data.

The category Model-Dependent Data, includes results which have been extensively massaged (export ephemerides available from JPL or SAO) or which result after correlation with another data set (a map showing the ratio of radar topography to Bouger gravity anomalies). Such data could be of immediate use in planetary studies without the user's needing to know in detail how the results were obtained. These are at least level data in Kieffer's terminology.

The size of a data set generally decreases, while the potential audience for it increases, as one moves from the raw data to the Model-dependent data. The risk of misapplication of data sets also increases as remoteness from the original experiment increases; more and better documentation (and more caveats) is needed even though the data volume is less.

8.1.4 Data Attributes

The quantities derived from radio science data vary with the experiment. The chart below shows some of these, grouped by sub-discipline and using level of processing as a parameter.

Calibrated Data*	Derived Data*	Model-Dependent Data
	RADIO ASTRONOMY	
Amplitude#	Radiance	Particle size(s)
Polarization	Emissivity	Surface roughness
	Temperature	Density
	Opacity	Rotation rate
	Electron density	
	Spectra	

Calibrated Data*	Derived Data*	Model-Dependent Data
	RADAR ASTRONOMY	
Amplitude#	Reflectivity	Particle size(s)
Polarization	Opacity	Density
	Electron density	Roughness
	T-P profiles	Rotation rate
Time of arrival	Position/elevation	
Frequency/phase	Velocity	
	CELESTIAL MECHANICS	
Range residuals	Position/elevation	Mass
Doppler residuals	Velocity	Gravity field
Time of arrival		Relativity
Frequency/phase		

*usually given as a function of time and/or frequency
#may be power or voltage

8.2 Current Situation -- Active Research Groups and Summary of Extant Data Sets

The number of sites equipped for observational research programs in planetary radio science is limited, but researchers who are active in the field can be found at a number of institutions.

Data sets in the field are diverse. Some of these presently reside at NSSDC, including most of the planetary mission data sets. Most NSSDC data are in fairly primitive form ("raw", "uncalibrated", or "calibrated") and would be of limited interest to users. Other data sets (such as those supporting planetary missions) reside with the original Principal Investigators. Some are more readily recovered than others, the parameter determining usually being age. Data sets more than ten years old are considered lost. Data sets which have been acquired as independent scientific research programs (many radio astronomy data sets are in this classification) would be difficult to recover even if the observations were made recently because of non-standard formats and lack of calibration data.

8.2.1 Active Research Groups

The radio community is scattered among many institutions -- the celestial mechanics group is, for practical purposes, limited to the Smithsonian Astrophysical Observatory and the Jet Propulsion Laboratory. Radar data are generated at the Aricebo and Goldstone facilities; additional investigators reside at MIT, JPL, Stanford, and Cornell while "users" come from a variety of universities and government centers. The radio astronomy community is more diverse and includes all of the above as well as many more.

Present capabilities for information processing vary widely within the radio science community. In terms of hardware the most common manufacturer is Digital. FITS (Flexible Image Transport System) is the only data format being used by more than a handful in the community. Most users are not connected to a network (such as TELEMAIL).

8.2.2 Major Data Sets of Interest

1. Celestial Mechanics

Two groups have almost entirely dominated this field. The JPL group supports mission operations and works closely with project navigation teams. Miles Standish has led this effort for many years. A JPL Development Ephemeris is available on magnetic tape upon request and has been used by many both inside and outside JPL. The Ephemeris is updated periodically. John Anderson and Bill Sjogren conduct scientific investigations, determining planetary masses and mass distributions which can then be fed back into solar system models for ephemeris improvement.

The Smithsonian group (previously associated with MIT) includes Irwin Shapiro and Bob Reasenberg. Their independently maintained data set has been used to produce an export ephemeris which (along with selected software) is also available upon request. Shapiro and Reasenberg carry out scientific work on gravity fields and mass distributions.

It would be relatively easy to incorporate both ephemerides into a PDS data bank (with some modest software development).

Raw tracking data as well as range and Doppler residuals presently reside in NSSDC for most planetary missions. These might be of interest to individuals familiar with procedures used in celestial mechanics and gravity studies. It is unlikely they would have wide appeal.

2. Radar Astronomy

a) Earth-based: Most earth-based radar observations have been conducted at Arecibo Observatory and Goldstone Tracking Station (JPL). Observations at Haystack Observatory were discontinued in the early 1970's when deteriorating equipment could not be refurbished.

Arecibo observations since the mid-1970's have been at 12.5 cm wavelength and in more or less standard formats: range-Doppler maps of slowly rotating large bodies (e.g., Venus), range delay observations for topography and absolute position (altimetry on Mars, range to asteroids), and power spectra for detection and measurement of difficult targets (comets, Galilean satellites). Some solar wind scintillation data have also been acquired. Prior to 1975, Arecibo radar observations were at 70 cm wavelength. That system remains functional and continues to be used routinely for ionospheric observations, but incorporating those planetary data into PDS would be more difficult.

Data from Goldstone are of similar types but are in somewhat different formats than those at Arecibo. Observations have been conducted at 12.5 cm wavelength since the 1960's and at 3.5 cm since the mid-1970's. Much of the Goldstone and JPL hardware is in the process of being replaced, so new formats can be expected in the future. Ray Jurgens is presently attempting to consolidate the Goldstone data on a new rapid-access disk system at JPL. Those data would then be available through dial-in ports on the new computer to anyone maintaining an account there. The system is patterned after BIRP, developed at Washington University. No major difficulties would be expected in making this system compatible with PDS.

Most Haystack data would be difficult to recover though some still exist. The major data set worth attention is the atlas of 3.8 cm lunar backscatter images produced by Stan Zisk. Pam Clark is presently generating digital mosaics from those images.

Reports in the literature of Soviet experiments should be noted. Details, however, have been sketchy, and it is unlikely that Soviet data could be easily acquired for PDS.

b) Spacecraft Radar: The major spacecraft radar data set is from Pioneer Venus. Maps of altitude, reflectivity, and small-scale surface roughness have been produced and are available from the PI, Gordon Pettengill, in digital form.

Apollo laser ranging data (W. Kaula, PI, and Apollo lunar radar sounder data (R. Phillips, PI) are listed by NSSDC. They are also available (in "derived" form) through the lunar remote sensing consortium in Flagstaff and (being the most accurate lunar topographic data now in existence) have been used by a number of scientists.

Several users of radar data have expressed the hope that SEASAT or SIR-A radar images could be included within the PDS data base. Those data do not fall within the PDS charter as presently written but the opportunities for Venus-analog studies and VRM comparisons are obvious.

c) Bistatic Radar: Most surface studies have been conducted at Stanford University (G.L. Tyler and others). Some data have been forwarded to NSSDC. Most are available at Stanford on tapes. With a small effort (to reconcile different formats) these could be incorporated into PDS. Targets include the moon and Mars.

Limited surface experiments have been reported at JPL (Mars). There have been many Soviet experiments (moon and Venus) reported, but data would be difficult to acquire. In neither case does a significant PDS effort appear warranted.

Ionospheric/atmospheric occultations have been (or will soon have been) conducted on all the planets except Pluto. Titan occultation data and solar corona/wind data can also be considered part of this data set. Most data reside at JPL (Arvydas Kliore and Gunnar Lindal) or at Stanford (Eshleman, Tyler, and others). Some have been sent to NSSDC. Most recent data would be easily incorporated by PDS; older data would be more difficult.

Ring occultation measurements from Voyager 1 at Saturn are an active data set at Stanford; preliminary results have been sent to NSSDC. Similar data will be obtained from Voyager 2 at Uranus in 1986.

Soviet spacecraft (particularly from the Venera series) have been used for atmospheric/ionospheric occultation experiments. Those data would be difficult to acquire for PDS.

Occultation studies have also been used to determine topography on planets with thin atmospheres. The most extensive of the data sets are from Mariner and Viking orbiters of Mars and reside at JPL.

8.2.3 Radio Astronomy

1. Earth-based radio-astronomy: The major observatories used for planetary radio astronomy include the VLA, NRAO, and Owens Valley. Observations are usually conducted by a small number of scientists and support personnel; procedures vary significantly from experiment to experiment. Most in the radio astronomy community concede that it would be difficult (if not impossible) to recover and archive most of these data, which are seldom retained by the observatories themselves.

VLA data are a noteworthy exception. Those are recorded in a limited number of standard formats and original data tapes are stored in an archive maintained at the VLA. Experimenters use standard software packages to convert the raw data to forms suitable for transport to their home institutions. The raw data tapes are, in principle, transferred to the public domain twelve months after creation. Several observers indicated they would not look kindly on outside use of those data (even after the twelve month hiatus for proprietary work) unless permission from the original investigator were obtained. There was a consensus that the novice would find working with the raw data difficult. Only limited cataloguing of observing programs (such as appears in the annual Bull. AAS reports) is available.

NRAO millimeter-wave observations at Kitt Peak may be archived and a terse catalogue may exist. Several other observatories conduct millimeter wave observations, often to obtain molecular spectra of comets. Snyder's review (Icarus, 51, 1) may be helpful in locating comet data sets; their condition is unknown.

Major planetary data sets at centimeter wavelengths include those of DePater, Klein, Epstein, Muhleman, Kenderdine, Dickel, Caldwell, and Jaffe on the outer planets and Muhleman, Janssen, Schloerb, and Wilson on the terrestrial planets.

Decametric observations of Jupiter have been made and stored for many years. Joe Alexander (GSFC) has produced a thick summary of those data (NASA Tech. Mem. 80308 by J.R. Thieman, 1979). The data base spans 15 years, but funding cuts stopped the summary program in 1979. Major data sets include those from U. of Florida and Chile (T.D. Carr), Colorado (J. Warwick), and Texas and Yale (J. Douglas).

Data from the International Halley Watch will be reduced to the FITS format (Flexible Image Transport System; see Wells et al., Astron. Astrophys. Suppl., 44, 363-370, 1981) which is emerging as a standard in astronomy (IRAS, VLA, etc.). There is thus some hope for standardization in the future.

2. Spacecraft radiometry: The most extensive spacecraft radiometry data set is from the Voyager radio astronomy experiment; it covers 1.2 KHz to 40 MHz and includes Jupiter, Saturn, and solar observations. The data set remains active; digital data through at least 1981 are at NSSDC.

A radiometric map of Venus has been compiled from Pioneer Venus observations at 17 cm wavelength. This data set is in a form similar to that used for P-V radar images and can be obtained from the PI (Gordon Pettengill).

Several Soviet spacecraft (particularly in the Mars series) have carried radiometers; without extraordinary efforts, it is unlikely those data could be obtained for PDS.

3. Preservation of Primary Data: Sources of funding and objectives of sponsors and scientists are varied. As a result there are relatively few standards for data collection and storage. In some cases the primary data are physically lost almost as soon as they are collected; in others, they are meticulously logged and archived, only to be lost in a practical sense when time passes them by. The details vary from case to case; we note one example here.

There is a large body (an estimated 7000-8000 tapes) of raw data representing radio occultation measurements made over the years that is insufficiently characterized for useful cataloging and that is on the verge of being lost due to its age. These data have been recorded using equipment on the ground; the history of that ground equipment is volatile and is not included in spacecraft records. If these raw data records are to retain their full value in an archive, they should be annotated to reflect the known effects of equipment and operator anomalies. The capacity to do this task is rapidly disappearing, as fewer than a handful of the people involved in the original data collection remain available to provide assistance in carrying out this documentation work.

To provide a scale for the magnitude of this problem, we estimate that 1-2 man-years of data technician support would be required simply to copy the 7000-8000 occultation tapes noted above. An additional 2-4 man years would be required to annotate and inventory satisfactorily those same tapes.

The above situation also applies in other branches of the radio community. For example, several thousand earth-based radar data tapes also need attention.

8.3 Description and Requirements Of PDS

8.3.1 The Preliminary Inquiry

Most producers and users of planetary radio science data indicate that a comprehensive catalog of existing data sets would be the most valuable first step in constructing PDS. Parameters required for effective use of such a catalog include most of those listed in Section 8.1.2 under "Calibrations". In addition, tags identifying coverage in a relevant planetocentric or planetographic coordinate system (latitude, longitude, for example) would be essential if the catalog is to be generally useful. Resolution, formal estimates of uncertainties, and instrument sensitivities should also be included.

Users of data sets requested that pointers to review articles be included in an early version of the catalog. Later, when actual data sets come on line, tutorials, glossaries, and help files should be added. Two major hazards for those just starting to use radio data are not understanding terminology and misunderstanding the limits on data validity. A "smart" sign-on procedure (to separate unsophisticated from sophisticated users) would keep the true novice from becoming lost at the beginning.

Search parameters and system responses for the three interest areas within radio science are as follows.

a) Celestial mechanics -- An early version of PDS should include one or both of the export ephemerides (JPL or SAO) now available. Development of a versatile software package suitable for all users could require a major documentation effort since proper understanding of terminology is crucial in this field. Users would want to know visibility of solar system bodies from the earth, spacecraft, and other bodies. In a well-developed implementation, such a program could be used to predict conditions for future observations as well as reconstruct geometries for past experiments. An ephemeris package would likely have more initial appeal for data "producers" than data "users", but there is considerable potential for all.

Celestial mechanics data themselves are limited to range and Doppler measurements. As noted earlier relatively few people would have an active interest in these data. Searches would be limited to a small number of parameters such as spacecraft name, a time interval, or a three-dimensional spatial window. A successful search should return information on sampling rates.

Some work has been done toward construction of consortium type data sets. Binned topographic data in one degree latitude-longitude cells would be useful in this work. Presumably gravity maps on similar scales would result.

Scientists involved in detailed studies of planetary gravity fields have noted that the ability to display spacecraft ground tracks on images or maps (particularly topographic maps) would be of considerable value in locating and interpreting anomalies.

b) Radar Astronomy -- Active radio systems usually produce data which can be tagged to a particular point in space such as a surface resolution cell or a latitude-longitude-altitude triplet in a planetary atmosphere. Preliminary searches would thus be directed toward whether data exist within such a time and/or space window. A successful search should return the name of the spacecraft and/or earth station, spatial and time coverage, wavelength, type of data (maps, spectra, T-p profiles, etc.), some indication of the accuracy of the data or the sensitivity of the system, and a pointer to the actual data set.

Parameters derived from surface scattering studies (altitude, surface roughness, reflectivity and depolarization) are applicable to a broad range of geological studies. As such, these data need to be available in a usable fashion to investigators with only a minimal background in radio science. Physical values of the parameters need to be in interpretable units (e.g., relative or absolute elevation, RMS slope) with the history of the processing used to derive the values readily available. For most surface studies, data need to be accessible from an atlas-style catalog whereby the user can identify data through a latitude-longitude search (i.e., BIRP type archive/retrieval). The display format of the data should be a function of the data type. Data obtained along ground tracks (e.g., altitude) should be available as individual linear arrays as well as two-dimensional arrays (e.g., projections onto maps) with resolution being a function of the data resolution. Where binned data are provided they should retain information regarding frequency of occurrence, smoothing and/or filling. Data suitable for an image format (e.g., Earth-based range-Doppler maps, SAR) should be available to both analog and digital images. Analog versions of images would allow for browsing the data sets. These analog images should be treatable in a manner similar to that of optical system images. A notable distinction, however, is that the user should be alerted to the differences in image production. Digital format of image data would allow for more detailed analysis and should be provided in two formats; range-Doppler and planetary coordinates. Data handling would be comparable to that for visual wavelength images.

For atmospheric work refractivity and absorption vs. altitude would be fundamental data sets; temperature-pressure profiles would be simple adaptations. Searches of large data sets for seasonal and latitude-longitude effects could be expected.

Display of data in terms of power spectra of the amplitude and the phase of waves transmitted through plasmas (planetary ionospheres or solar wind) has been common recently. These data would be searched by altitude (or distance from the sun) for the effects of turbulence and/or magnetic fields. Dual frequency measurements would be scanned for differential phase changes as a function of time and/or position of the ray-path closest approach.

c) Radio astronomy -- Passive radio data sets are generally of coarser spatial resolution than those derived from radar observations. Though users may hope to search for emissions by latitude and longitude on the body, those searches are likely to be less revealing than in the radar case. An exception is the Pioneer Venus radiometry data set which has footprints as small as 7000 sq km and covers most of the planet.

Radio maps would presumably be searched according to observing wavelength, portion of target visible, and observing date. Radio spectra would be searched by frequency and time of observation. A table of known emission lines might facilitate this process.

8.3.2 Nature of an Analysis Task

a) Need for locally available data sets: Since the JPL and SAO ephemerides are now available on magnetic tape, there is no compelling need for immediate storage of such a data set in an electronic PDS. It is clear that inclusion of ephemerides would be very useful, however. Widespread use of the ephemerides could be expected if a comprehensive and readily available software package were developed. In that case it would probably be necessary to maintain both data and software in a single location. Users would do whatever work was needed remotely and transfer the results to their own facility.

Many respondents expressed a need to have visual images readily accessible. If technology permits, laser disks may be available, providing at least a browse capability locally. Actual processing of these images would require a non-trivial software package. Ray Jurgens is adapting BIRP to handle radar data at JPL. Present storage capabilities are limited but it does show promise. If BIRP or a similar system is sufficiently portable, then much work with images could be done locally. If not, a high data rate (and expensive) electronic system may be the only alternative. Even within the radio science community there is a significant interest in working with images.

Images or consortium-format data constructed from radio science data represent a relatively small number of bits when compared to those in the planetary visual image data set. If visual images were available locally on laser disks, then an electronic network could be used to transfer other data. Storage of those data sets locally would then not be required, though it would be convenient if external data sets could be copied into the local system for extended analysis. During browse phases or preliminary analysis (such as with the producer's own on-line software package) manipulation within the source machine would be most efficient.

A major advantage of an all-electronic PDS would be the relative ease with which data sets could be modified or corrected by the producer. This would not be possible if all planetary data were written onto laser disks, then shipped to users--unless updated disks were distributed on a frequent basis. We would hope that data entered into PDS for archiving would be in a fairly complete form, but perfection cannot be expected. A rate of one change per data set per year might be high enough that some users would lose confidence in the currency of data distributed through a laser disk system (for certain data sets, such as visual images, this rate might be acceptable, however).

Corrections or updates should be the responsibility of the data source. For user protection the system might include an automatic mail feature which would check a log of recent users and advise them that a change had been made. A management requirement that papers based on PDS data sets be reviewed by producers of source data sets might also guarantee the same result, albeit at some potential cost to the author.

b) Data Processing Level - Data entered in PDS first should be that which has the widest audience. The data catalog mentioned earlier is the obvious first step; pointers can be included within that catalog for an upgraded system which would eventually handle data. Images and polished (final) data sets which are already being distributed among planetary scientists should be entered next -- the so-called "derived" data. These might also include tabulated data and figures from published papers (or the papers themselves). These data would be the key to increased interdisciplinary work in planetary studies. Data sets should be adequately documented with easily understood explanations of their meanings, limitations, and processing histories. Third priority would be the transfer of archived "uncalibrated", or "calibrated" data from storage to the active data base. These would be the most massive of the radio science entries, but they would provide "producers" with raw material for detailed comparisons. There should be clear pointers to calibration and documentation files in the headers to each of these entries.

Most in the community felt that augmentations to the radio/radar data base from extended analysis or reanalysis of uncalibrated or calibrated data would not significantly increase the volume of the database. Of more importance would be the "model-dependent" data since these could represent synthesis efforts incorporating several diverse elements from within PDS. "Model-Dependent" data would be the most speculative. There might be no need to make a formal, separate entry of these data since analyses of data already within PDS would probably produce more than enough of these additional files.

Most models used in radio studies are mathematical and would be described in tutorial or help files. There would be, at most, only limited need to store numerical models in PDS.

c) Need for Basic Analysis Software - In radio science, basic analysis software would most effectively be provided by the investigator supplying the data. By default it will be the investigator who decides which software to provide. If a distributed data base is adopted, programs could be run on the resident computer with only results being transmitted to the requestor. This structure simplifies the need for major revisions in software to make programs portable, substituting the much simpler requirement (which will be necessary anyway) that the data transmission/reception formats be compatible.

If a more centralized system is developed, then several special needs can be identified.

1) All image processing software should handle radar and other two-dimensional data sets with equal facility.

2) It should be possible for multiple users at separate remote locations to operate cooperatively on the same data set or display.

3) It would be desireable if researchers could store back-up copies of proprietary data sets in secure areas of PDS.

In the area of more general software support, the radio/radar community would find the following useful:

a) Fitting routines;
b) Statistical packages;
c) Graphics packages;

d) FFT, correlation, and other signal processing packages (one- and two-dimensional);
e) Resampling and filling routines (for images-type formats, etc.);
f) Standardization routines (to produce consortium-format and other arrays, for example);
g) Ephemeris package; and
h) Algebraic and vector manipulation

If radio data are to be generally useful to the planetary community, some effort must be mounted to develop tutorial and help files. Some of these (explaining radio science concepts and defining terms) could be centralized with the data catalog. Those peculiar to a specific data set would have to be supplied by the data source.

d) Future needs - In the 1990's PDS, or a similar system, may be used for distributing and analyzing data in near real time -- for example, with MGCO. Under those conditions the need to update data sets rapidly will be essential. Also, the distinction between quick-look and long-term analyses will be less (or at least different). Both points argue for networking at multi-kilobit data rates. The evolution from an initial 1200 baud catalog network should be smooth. PDS should begin looking at Code T and TDRSS style high speed networks so as to anticipate those changes. At present only Voyager radio astronomy data (on the NEEDS network) and the JPL prototype retrieval system even come close to high-speed data transmission systems within the radio science area.

8.4 Summary, Recommendations, and Concerns

8.4.1 Major recommendations for PDS from the radio science community are summarized in an expanded format.

- PDS should be developed in phases

  - Overview

    Scientists should be involved at all levels of system design (NASA should encourage maximum use and feedback from scientists as the project develops).

    Performance (rather than design) criteria should be emphasized at all stages

  - Implementation

    Design should use state-of-the-art engineering, management, and library science skills.

    Design should make maximum use of existing technology

- Phases of development

Catalog	earliest time
Inventory	
Data access	to
Data manipulation	
Data analysis	distant future

- Levels of data (user community is diverse)

  Data should be keyed to geophysical and other vocabulary as well as to radio terminology

  Data included should appeal to users over a wide spectrum of sophistication

o PDS development environment

- Funding for system development should not seriously impact ongoing science efforts

- A design for present and future missions and experiments would be most palatable; old data can be incorporated as resources and interest dictate

- Fading data should be preserved now (at least to the extent of copying decaying magnetic tapes)

8.4.2 Rehabilitation of Existing Data Sets

As a precursor to developing a PDS data base, it is essential to take steps to preserve the information in existing data records. Within most groups, these are the only data that are generally accepted as uncontaminated by other users. Consequently, these are the primary and most important data within that community. First priority should be given to, at a minimum, making fresh copies of the older data tapes and preparing a rough inventory. It would be preferable, at the same time, to examine all of these records and annotate them, providing a more detailed inventory. The revised tapes could then be transferred into a "preservational" archive for safekeeping, and pointers to the detailed inventory could be included in a PDS catalog.

8.4.3 Requirements of PDS - implementation sequence and emphasis

A step-wise development for PDS is highly desirable. A simple system with limited (but useful) capabilities should be produced first. Evolution to more sophisticated systems should proceed over a period of years. At each stage performance (rather than design) requirements should be set to ensure maximum utility to the user community. Use of the system at each stage should be widely encouraged and feedback from the users fully understood (and hopefully incorporated) before the next stage is designed. Prototype systems already exist and their performance and limitations should be reviewed before a new effort is mounted. Where possible PDS should take advantage of networks (or their lessons) developed in other branches of NASA.

A "data catalog" allowing anyone with a "dumb" terminal and a modem to locate potentially useful data is a reasonable first step in development. Pointers to published papers and tutorial documentation should be included. On-line access to the radio data is desirable but a longer term objective. Most in the radio community feel that rapid access to visual-wavelength images (or mosaics or maps) would be more useful than rapid access to radio data in the near future. Local laser disks for rapid image browsing might suffice for visual images. Radio data converted to consortium formats (e.g., quarter by quarter degree binned data) are valuable for regional studies, but these should be backed up with original linear data sets. There is but limited interest in having primitive radio/radar data sets on line, except as a very long range objective, a fact consistent with NSSDC records of requests for radio data. Depending on costs of data storage, manipulation, and transmission, costs of computation, and the availability of truly general software, there may be distinct advantages in the long term of retaining only primitive data (and calibration files). This option should be kept active but it does not represent a viable approach to a generally useful first-order system.

Data obtained as part of planetary missions (e.g., Voyager radio astronomy and Pioneer Venus radar) would be most easily incorporated into PDS. In that sense it would be more practical (and cost effective) to design PDS for future missions, defining formats and establishing protocols now. Existing data, which (in the long term) will represent a relatively small fraction of the total, would be added to PDS as time and interest dictate. Independent observing programs (typical of earth-based radio astronomy) would be the most difficult to integrate into PDS. Data more than ten years old (with a few exceptions) are effectively lost; data taken more recently are recoverable with the recovery effort roughly proportional to age.

Funding and manpower limitations probably will dictate a distributed system for PDS. Effort should be directed toward establishing standard formats for transmission of data now.

The system should be compatible with (or transparent to) most present and anticipated hardware. On the other hand, designing toward the most common equipment now in use.

Responsibility for generating, updating, documenting, and accessing data sets goes almost de facto to those actively working with data sets -- unless there is a separate, major influx of new funding for PDS. As a result, these same scientists will be responsible for deciding what software will be necessary and how it should be provided. Most data, in a distributed system, will be manipulated in the computer of residence, only results will be transmitted to the requestor. Exceptions include image and consortium-format data, which might be transmitted once, then stored for later use. A one-week turn-around on requests for hard-copy images from NSSDC (rather than the now typical six weeks) would facilitate much work with images.

8.4.4. Continuing Concerns

The planetary radioscience community is generally supportive of any effort to improve the usefulness and accessability of planetary data sets. The producers of radio data (as opposed to users - those whose interests are primarily interdisciplinary) tend to have reservations about the implementation of any real system.

The most commonly expressed concern is that funding for a new archival/retrieval system should not impact actual analysis of existing data. Most radio scientists are unable to analyze, fully, the data already in hand. PDS implementation geared toward future missions would have the least impact on current research. Considerable foresight is required, however, the needs of scientists who would like enhanced interdisciplinary capabilities now are not addressed. Proposals to conduct limited archiving have not been received favorably by NASA in the past. New emphasis on the end-to-end progression of data must come from NASA and projects.

A second major concern is that scientists be fully involved in development of PDS -- from the beginning and at every intermediate stage. Historically, data systems lacking this interaction have turned out to be cumbersome, little used, and/or expensive. On the other hand, scientists should probably not be charged with the actual development of the system; well-defined assignments of responsibility and authority are necessary to make the implementation effective and efficient. A careful blend of science guidance and review with engineering expertise and management efficiency will be needed to create a viable and useful system.

9. EARTH-BASED PLANETARY DATA

9.1 General Nature Of Data Available

NASA sponsors Earth-based planetary astronomy to provide data needed for:

a. Planning future missions,
b. Supporting missions already in progress, and
c. Enhancing the scientific analysis of data returned by past missions.

Such data complement spacecraft data by providing observational coverage of important planetary phenomena at epochs before and after spacecraft encounters, with time scales longer than a spacecraft encounter, and at wavelengths and spectral resolutions not covered by spacecraft instruments. These criteria fit a wide variety of Earth-based data, acquired by a number of techniques, which are discussed in the sections that follow. Moreover, for some future spacecraft targets (such as asteroids, comets, and the planets beyond Saturn) Earth-based data are the only kind that now exist.

Earth-based planetary data are defined here as: observations made with ground-based telescopes, balloon-borne instruments, rocket-borne instruments, and Earth-orbiting spacecraft. Data in the last category differ somewhat in nature and availability from the other three. In all categories, planetary data are only a subset of the total astronomical observations on record. As a result part of the task is to identify and extract (or separately catalog) those observations which are planetary. The data addressed in this chapter covers wavelengths in the ultraviolet, the visual, and the infrared. Earth-based radio and radar data and particles-and-fields data are addressed in other sections of this report. The instrumental techniques, addressed are: imaging, polarimetry, photometry, spectrophotometry, and spectroscopy.

A large percentage of the ground-based observations (as distinct from other Earth-based observations) are made by individual investigators (and their collaborators) acting largely on their own initiative. The investigators who are supported by federal grants, have an obligation to publish their results in the scientific literature but none to archive their original data in the public domain. In most cases, no accumulation of data of archival interest has been created, and no one but the original investigators could properly reduce of the raw data. There are, of course, exceptions; and those are the cases we discuss in this report.

Earth-orbiting observations tend to be organized with some degree of similarity to planetary missions. For the early phases of an Earth-orbiting mission (such as Space Telescope), Investigation Definition Teams are selected by NASA on the basis of proposals. These teams function in collaboration with NASA, and are under contractual obligation to put their data into the public domain after a prescribed proprietary interval. Earth-orbiters of current interest here include OAO, IUE, IRAS, ASTRO, ST, and SIRTF. For all of these, planetary observations constitute (or will constitute) only a modest fraction of the total data obtained and archived.

9.2 Ground-Based Data

9.2.1. Imaging

The brighter planets (Mercury, Venus, Mars, Jupiter, and Saturn) have been photographed frequently with ground-based telescopes for nearly a century. Features on Mars and Jupiter have been monitored with some regularity. Mars is accessible to Earth-based instruments for periods of several months at intervals of about 26 months, and Jupiter at intervals of about 13 months. Owing to the broad variation of observing conditions, the resolution of detail varies greatly. For features of medium contrast, it is rarely better than 25 cycles per Martian diameter or 50 cycles per Jupiter diameter, and it is often much poorer than that. Even so, the historical record is rich in information about Martian clouds and dust storms and about changes in Jovian belts, zonal features, and spots.

In recent years, photographic imaging of planets has most commonly been done with broad passband filters for ultraviolet, blue, green ("visual"), and red. Photographs of a particular planet through a particular filter are typically taken in a contiguous sequence of ten or more images, all within a time span of a minute or two, and therefore all representing (within the limits of ground-based image resolution) the same face of the planet. The individual images in this sequence vary in resolution due to turbulence in the Earth's atmosphere, so the sequence provides a user with the opportunity of selecting the best images and rejecting the rest. When several images of a sequence are good, they can be combined into a single image of improved signal-to-noise ratio. In the past this has been done by photographic "compositing," which is analogous to the stacking of digital images. Among the sequences currently in the Planetary Research Center catalog, several thousand have been composited (mainly at Lowell, New Mexico State University (NMSU), and Meudon) and are alternatively available in that form.

Electronic imaging with vidicons, CCDs, and spectracons has been used to attack specific scientific problems, particularly utilizing the gain in quantum efficiency to permit short exposures through narrowband filters. Results have largely been extracted and published, though there is no organized archival collection of original data. In the future, however, electronic imaging may eventually displace photography for synoptic monitoring. The value of ground-based synoptic monitoring, to a continuing planetary exploration program in NASA, tends to be underestimated. It will not become superfluous until planetary orbiters can see all faces of a planet every day or until a dedicated Earth-orbiter is available to take over the task.

Prior to 1969, ground-based photographs of planets were obtained at only a few observatories, notably Lowell, NMSU, Lick, Mount Wilson, University of Arizona (UA), and Meudon. In 1964, two centers were established by the International Astronomical Union for archiving and cataloging planetary photographs, one at Lowell Observatory (with NASA support) and the other at Meudon Observatory near Paris. The Meudon center is no longer active, but the Planetary Research Center at Lowell has either originals or copies of virtually all good-quality planetary photographs from all sources predating 1969.

Starting in 1969, the International Planetary Patrol Program was established with NASA support to monitor Mars and Jupiter on an uninterrupted hourly schedule with identical telescopic cameras distributed in longitude around the Earth (Baum, Planetary and Space Science 21, 1511-1519, 1973). Film processing, editing, and cataloging were all carried out at the Planetary Research Center (PRC), Lowell Observatory, where the Patrol films are now archived. From 1969 through 1973, the Patrol multiplied by fivefold the number of high-quality images in existence. Sharp cutbacks have reduced coverage since 1974.

In addition to the Patrol network coverage archived at the PRC, planetary photography has been continued at NMSU, where images accumulated since about 1970 have not yet been systematically copied or integrated into the PRC archive. Photographic copying of the post-1970 NMSU material would probably not be cost effective, but we do recommend listing those observations in the PRC catalog, with the idea that the NMSU images themselves continue to be archived only at NMSU.

The PRC catalog already lists most of the existing ground-based planetary image sets of usable quality from most observatories through 1983, currently about 140,000 of them. Post-1970 NMSU images will add a few percent. Each image set is identified by planet, U.T. date and time, and filter; it is accompanied by all useful ephemeris information plus codes specifying the observatory of origin, telescope aperture, observer, image scale, length of exposure, image quality, present location of the original, and whether a composite had been made. The PRC catalog amounts to about 15 Megabytes. It is available on magnetic tape and can be accessed by planet, date/hour, central meridian, illumination geometry, color (wavelength), resolution, location of archive, etc. The complete catalog and associated software will be directly available to PDS to put on-line. The amount of NASA funds needed for maintaining the archive and annually updating the catalog is quite modest.

A useful on-line adjunct to the catalog would be an atlas of selected images characterizing the significant phenomena observed on each planet year by year. These images should be digitized, calibrated, rated in terms of data quality, and be available in rapid access mode. The purpose of this atlas would not be to include all historic details but to provide a historical overview readily available to mission scientists during planning and encounter periods. This atlas would also provide a concentrated briefing for general users of the ground-based archive. It would require about 1 Megabyte per year of coverage and might usefully start with a total volume of 20 Megabytes. An information proctor will need to be given responsibility and support for preparing and periodically updating the atlas and for supplying tutorial material associated with it.

9.2.2 Photometry and Polarimetry

Photoelectric photometry data sets of solar system objects were obtained at least as early as the 1950s. Most of the early data sets are on the UBV system. More recently, the ubvy, 8-color (0.34 to 1.04 micrometer) and 24-color (0.32 to 1.04 micrometer) systems have been used extensively in observations of asteroids, comets, planets and planetary satellites. Most of these data sets have been published in the open literature, primarily the Astronomical Journal and Icarus. To the best of our knowledge, there are no extensive unpublished data sets in this category. Some of the post-1975 data sets are available in machine-readable form.

Earth-based observations provide the longest available observational record of Solar System bodies, are useful in providing calibration of spacecraft instruments, and can follow up discoveries made by fly-by missions. Advantages of including data of these types in the PDS are: (1) concatenation of data sets which were originally published in many separate papers, (2) having these data available in machine-readable form, and (3) putting all observations (especially infrared) on a common calibration system. The fact that these types of observations continue to be made means that provisions for the updating of data bases need to be incorporated into the PDS.

As an example of an existing planetary photometry data base, we cite the asteroid compilation called TRIAD (Tucson Revised Index of Asteroid Data). This data base was established in 1979, and initial publication was in the book Asteroids (T. Gehrels, editor, University of Arizona Press, Tucson, 1979). TRIAD consists of eleven files, each created and maintained by a "contributor." Each file contains the "best" parameters and is limited to one 80-character record per asteroid. Basic data (UBV colors, 24-color and polarization parameters, osculating orbital elements, etc.) and derived data (polarimetric and radiometric albedos and diameters, proper orbital elements, taxonomic classifications, etc.) are both included. The contributor is responsible for selecting or assigning the "best" parameter for each asteroid. The TRIAD data base is available on cards, magnetic tape, and in hard copy. It was intended to be used by mission planners and others interested in having, in one place, canonical parameters for all asteroids for which they are available.

The PDS asteroid data should be organized differently from TRIAD because the new data base would be on-line (it is now about 250k in size), would be kept current (most of the files in TRIAD have never been updated), would have less restrictive formats, and would contain more original data rather than only selected "best" parameters. It will still be necessary to have a designated individual oversee each distinctly different data file, but the on-line nature of the data base will greatly facilitate file maintenance. Finally, the distribution formats will be of wider use (e.g., via modem directly into the user's computer), and some form of data-base management system could be made an integral part of the retrieval system.

Photometric parameters (together with dimensions and orbital data) should similarly be listed on-line in PDS for all other bodies (except comets) of the Solar System. Although such data are readily available in the published literature, having that information handy on-line would be a convenience. The Halley Archive, which will result from the International Halley Watch, provides a good guideline for contents and format for the incorporation of comet data files into PDS. Such files would facilitate the choice among candidate objects for a comet mission.

An on-line catalog needs to be generated for photopolarimetry data, which lists objects together with information on observing periods and observational constraints. New data sets should be entered as they become available. The historical Venus data set is an example of a desirable entry into this archive. There will be ongoing interest in spatially resolved frequency-dependent photopolarimetry. Venus displays a large range of phase angles to an earth-orbiting observer and is bright enough to be observed with instruments developed for other projects. Historical integrated disk photopolarimetry obtained by Gehrels et al. at the University of Arizona should be available for planning and comparison with new data.

The table below gives an overview of the categories of photometric and polarimetric data sets arranged according to technique. Under each technique we have indicated the types of objects for which data are available.

9.2.3. Photometry and Polarimetry Data Sets Ordered by Technique

1. Visual Filter Photometry Data Sets
   (UBV, ubvy, 24-color, 8-color, and specialized filter sets)

   Asteroid, comet, and natural satellite lightcurves, magnitudes, and colors.
   Standard star lists.
   Planets.
   Zodiacal light.

2. Infrared Filter Photometry Data Sets:
   (JHKLL', specialized filters?)

   Asteroid, comet, and natural satellite lightcurves, magnitudes, and colors.
   Standard star lists.
   Planets.

3. Circular Variable Filter (CVF) Photometry Data Sets
   (0.7 to 14 um)

   Asteroid, comet, and natural satellite lightcurves and fluxes.
   Standard star lists.
   Planets.

4. Thermal Infrared Filter Photometry Data Sets
   (MNQ, 30 um, "silicate" filters)

   Asteroid, comet, and natural satellite
   lightcurves, and magnitudes.
   Standard star lists.
   Planets.

5. Submillimeter Filter Photometry Data Sets
   (350 um, 800 um, 1000 um)

   Asteroid, Comet, and natural satellite fluxes.
   Calibration details.
   Planets.

6. Filter Polarimetry (Linear and Circular) Data Sets
   (UBV, other?)

   Asteroid and natural satellite lightcurves and
   polarizations; polarizations of comets,
   planetary rings, and planets.
   Standard star lists.

9.2.4 Spectrophotometry

The definition of spectrophotometry, as used here, includes only work done with a dispersive element (grating, prism). Further broken down, spectrophotometry can be divided in two categories: absolute (flux-calibrated) and relative (normalized to some convenient level). For each category, attempt has been made to identify relevant data sets.

Absolute spectrophotometry of planets include data sets from approximately a half-dozen investigators. Each data set consists of I/F or geometric albedo spectra, generally in digital form, and contains a few hundred to a few thousand bytes. All of these data could be conveniently stored in an on-line archive for maximum accessibility.

Absolute spectrophotometry of small bodies (comets, asteroids, satellites) are relatively recent data, and all are in digital form on tape though many have not been published at this time. A larger quantity of such data exists, especially for comets, than for outer planets. Nonetheless, they could all be conveniently archived on-line as they become available.

Relative spectrophotometry (low-resolution spectroscopy). These data are in the form of relative intensity versus wavelength. Venus and the small bodies mentioned above are targets of these measurements. All such data could be stored conveniently in an on-line archive (i.e., a few tens of kilobytes).

9.2.5 High-Resolution Spectroscopy

There are two principal categories of high-resolution spectroscopic data: photographic (mostly older than ten years) and photoelectric (including observations with solid-state detectors).

Photographic spectra exist for all of the planets except Pluto. Similar data exist for a number of comets. These data cover specific spectral features and, in some cases, specific spatial features. There is a large quantity of such data. For the purpose of publication results have generally been reduced to a single number (e.g., an equivalent width). It is not clear that digitizing raw spectra for inclusion in an on-line archive would be useful or cost effective. However, a comprehensive on-line catalog of these spectra should be compiled.

Spectra recorded with photoelectric and solid-state detectors utilize a variety of dispersive techniques (e.g., gratings, Fourier transform spectrometers (FTS), and Fabry-Perot etalons) and detector formats (e.g., photomultipliers, one-dimensional arrays such as Reticons, and two-dimensional panoramic detectors such as CCDs). The various permutations of dispersive techniques and detectors which have thus far been used can be represented succinctly by the matrix below. Some elements have been given proper names.

Dispersive Technique	Grating	Fabry-Perot	FTS
Detector			
Single-channel	Spectrum scanning	Yes	Yes
1-dimensional array	Yes	Planned	No
2-dimensional array	"Long-slit" spectroscopy	Spectral imaging	SPIFI*

*Spectrophotometric imaging Fabry-Perot interferometer.

A large quantity of raw photoelectric data (of the above types) is extant. While it would be technically feasible to store the raw data in an on-line archive, because they generally exist in digital form, it is not clear that this would be cost-effective or useful. For example, in the case of FTS observations, the raw data consist of very large numbers of interferograms which must be individually calibrated, then co-added and transformed to produce usable spectra. Data reduction procedures of similar complexity must be applied to data acquired by other techniques. The investigator who performed the observations is best equipped to do the reductions; consequently, only reduced data should be considered for inclusion in an on-line archive.

The reasons for putting spectroscopic and spectrophotometric data sets into an on-line archive include their use in constraining physical models of the atmospheres or surfaces being studied, and their value in supporting studies of time-variable (e.g., seasonal) planetary phenomena. The most useful form of these data is that which permits direct quantitative comparisons, for example, one may want to fit band or line profiles to spectroscopic data. Important ancillary data, to be included in the data set, would include pointing information such as the spatial location spectra of extended objects, as well as information concerning the resolution and quality of the original data.

9.3 Balloon And Rocket Data

Ultraviolet observations of planets and comets from balloons and sounding rockets represent only a very small data base, resulting from short observing times (typically five minutes for a rocket, a few hours for a balloon) and infrequent flights. The available observations do, however, extend back to 1967 and provide very useful data concerning a small number of problems, such as the variation of Jovian HI Lyman-alpha emission over a complete solar cycle. (Skinner et al., Astrophys. J. 265, L23, 1983).

Most of these observations were made with instruments specifically designed and built for a given flight. No two instruments are used in exactly the same configuration and each instrument was calibrated individually (usually both pre-flight and post-flight when possible). The data quality was often dependent on the performance of the rocket attitude control (pointing system). Successful flights were usually the basis of a Ph.D. dissertation, so extensive experiment descriptions, calibration, and data reduction descriptions can be found in theses and in the published literature.

It is recommended that the archive contain a catalog of solar system observations by rocket and balloon and that a description, as complete as possible, of the instrumentation be included. Researchers desiring to obtain and analyze unpublished data from these experiments whould, however, be advised to contact the individual Principal Investigators.

9.4 Earth-Orbiter Data

9.4.1 Orbiting Astronomical Observatories (OAO-A2 and OAO-C)

The use of OAO-A2 for planetary astronomy consisted of broadband photometry (1900 to 3200 A) of Venus, Mars, Jupiter, Saturn, and Comets Bennett and Tago-Sato-Kosaka. Special observational constraints and instrumental problems were encountered which made the data very hard to interpret. The primary reason for archiving these data would be for investigating long-term seasonal changes in the atmospheres of the planets. Any interpretations are very dependent on the calibration procedures and on the solar spectrum used. The archival data set should consist only of the final calibrated data.

During the long lifetime of OAO-C "Copernicus" (1972-1979), high-dispersion profiles of the Lyman-alpha resonance line were obtained for Mars, Jupiter, Saturn, Io, Titan, and several comets. These data have been calibrated and reduced using concurrent observations of the geocoronal Lyman-alpha profile, and the results have been published. Raw data files exist in the NSSDC and at Princeton University. The archival data set at NSSDC lacks the critical pointing information through the Earth's geocorona, which exists only in the form of a separate computer listing and in notebooks. The published data consist of a few calibrated scans that could easily be made available for a data archive. Of the raw data, only the Jupiter Lyman-alpha scans have enough S/N to be used for additional analysis, such as a study of variations of Lyman-alpha intensity with Jovian longitude.

9.4.2 International Ultraviolet Explorer (IUE)

The International Ultraviolet Explorer satellite was launched in January 1978 and has proved to be a valuable tool in many areas of solar system astronomy. The IUE with two unique spectrographs used for observations of a wide variety of astronomical objects, has required a large amount of attention to the evaluation of instrumental response and its variation with time. Thus, the photometric accuracy of the IUE spectrographs has been established with a high degree of reliability. This feature enhances the usefulness of IUE for synoptic studies of the major planets, many of which are tied to near-simultaneous IUE observations made during the Voyager 1 and 2 flybys of Jupiter and Saturn. Approximately 15% of the total U.S. IUE observing time has been devoted to planets, comets, and satellites during the first five years of operation.

IUE images are readily available from the NSSDC. Each image is available in several forms:

a) A raw image of the spectrum as recorded on the vidicon detector;

b) Processed line-by-line spectra (low dispersion) parallel to the direction of dispersion, with the optical transfer function of the instrument and the wavelength calibration included; and

c) Integrated fluxes with the background subtracted, in physical units.

All of the processing algorithms are also available, as are the sensitivity factors.

A catalog of all IUE images taken to date is available in several formats, including one in which the entries are sorted by target or by class of object. The two spectrograph apertures are a 3" diameter circle and a 10"x20" rectangle (i.e., both smaller than many planetary disks or comets) so that detailed pointing information is required for proper interpretation of the recorded spectra. The motion of the object and (sometimes) a parallax correction are compensated for by slowly slewing the spacecraft, and the position of the aperture is often checked at the end of an exposure or at regular intervals during a long exposure. The errors in pitch and yaw are generally recorded only in the observer's notebook, not in the science header of the archived image, and that lack of information in the archive poses a problem which is particularly acute for comets within 1 AU of the Earth.

There should be no difficulty in copying the IUE data products from the NSSDC into PDS. Care should be taken in making the conversion of given flux values into surface brightness for extended objects. Spatial imaging of particular spectral features with a resolution of about 5" is possible with data taken with the large spectrograph aperture and has been discussed in several different contexts in the literature. Instrumental problems, such as long-wavelength scattered light, limited dynamic range, and false features due to detector hot spots, have also been discussed in the literature. It is recommended that a copy of the IUE archive be included in the PDS. It is anticipated that another copy of the IUE archive will exist at the Space Telescope Science Institute.

9.4.3 Infrared Astronomy Satellite (IRAS)

The Infrared Astronomical Satellite was launched on 25 January 1983. Its primary objective was to conduct an all-sky survey in four infrared bands centered near 12, 25, 60, and 100 micrometers. Due to their strong infrared fluxes, the planets themselves (and their satellites) were avoided. Asteroids and comets, however, were observed in the survey mode. A special group, called the Asteroid Advisory Group (AAG) and chaired by Dennis Matson, has been set up at the Jet Propulsion Laboratory (JPL) to aid the IRAS Project in producing data products which will be useful to asteroid and comet specialists. These data will be archived at the NSSDC, but it is expected that they can be put in a form suitable for inclusion in PDS.

9.4.4 Space Telescope (ST)

There are to be five scientific instruments aboard Space Telescope: two camera systems (WF/PC and FOC), two spectrographs (HRS and FOS), and a high-speed photometer (HSP). There is also a fine-guidance system (FGS) that additionally has some astrometric capability. In its long-focus mode, the ST Wide-Field/Planetary Camera (WF/PC) will provide a spatial resolution of 0.1 arcsecond over a field 68x68 arcseconds, with an assortment of filter passbands available from 1200 A to 1.0 micrometer. Thus, WF/PC's advantages over ground-based imaging of planets are a factor of 10 in resolution, together with access to the ultraviolet. The High-Resolution Spectrograph (HRS) covers 1100 to 3200 A with spectral resolutions ranging from 2,000 to 100,000. Thus, the HRS can be used for investigating comets and planetary atmospheres. The fainter parts of comets might also be studied at lower resolution in the ultraviolet with the Faint Object Spectrograph (FOS). The High-Speed Photometer (HSP), with sensitivity from 1200 to 7000 A, may be useful for improved occultation observations.

All of the Space Telescope data will be archived at the Space Telescope Science Institute (STSI) in a 32-bit/pixel format, using a data-base management system. Proprietary data will be accessible only to certain users. Other researchers may be granted access to released (non-proprietary) holdings, based on a proposal selection process similar to that for the allocation of ST observing time. The STSI plans to have the capability of accommodating ground-based as well as ST planetary data for both real-time and non-real-time research.

A periodically updated catalog of Space Telescope planetary observations should be created for the PDS archive. A tutorial, which characterizes the nature of observations and provides typical sample data sets, should be available as an on-line component of the planetary archive. This will allow a potential user to determine whether useful Space Telescope observations are available within the confines of the planetary science data files.

A copy of ST planetary data themselves should be transported on a continuing basis to the PDS archive and integrated into the data base, in order that a complete chronological record will be available to present and future users. Currently, Space Telescope is commissioned to archive observations for only 17 years. Each full-size WF/PC image will occupy 10 megabytes, but other ST instrument outputs very small in comparison.

9.4.5 Shuttle Payloads

This area represents a new generation of planetary missions that will encounter archival problems similar to those currently experienced by imaging teams. These missions can have high data rates and, in the case of free-flying packages may last for several years. It is not possible to estimate the extent of the demands of these missions on the PDS archive, but significant data sets will become available and should be included.

There are a number of payloads currently being built or designed for flight aboard Spacelab on the Space Shuttle, and they are capable of making unique observations of Solar System objects. The data from these instruments will, in most cases, be deposited with NSSDC due to contractual obligations. It is important to recognize at this time the desirability of ensuring that the pointing and tracking details be recorded in the archives, including, when possible, images from any devices used for acquisition and tracking that are integral to the instruments. Any Shuttle events that may affect the data quality should also be flagged in the archival record.

One of the early Shuttle payloads with potential for planet and comet studies will be ASTRO. It consists of three boresighted telescopes to be mounted on the Instrument Pointing System (IPS): (1) The Hopkins Ultraviolet Telescope (HUT) will provide spectrophotometry from 850 to 1800 A at 2-A resolution. For solar system objects, in second order the spectrograph will be sensitive to below 500 A. (2) The Wisconsin Ultraviolet Photo Polarimeter Experiment (WUPPE) will obtain spectrophotometric and polarimetric data in the wavelength range 1300 to 3200 A. (3) The Ultraviolet Imaging Telescope (UIT) will obtain 40'x40' images in selected ultraviolet bands.

The first flight, ASTRO-1, is planned for March 1986 to include P/Halley as a prime target. Digital images from a slit-jaw camera on HUT will be transmitted every 20 seconds and will provide the requisite pointing and tracking information. For this purpose, filters to isolate the gas and dust emissions from a comet have been included in the design.

Another Shuttle payload of interest will be the Spacelab Infrared Telescope. This instrument is a small liquid-helium-cooled telescope which is attached to the Spacelab pallet. Its goal is to measure Shuttle contamination and to survey the sky at far-infrared wavelengths. Planetary data expected will include photometric data on comets, asteroids, and the outer planets. New information on the spatial distribution of the zodiacal light is also expected.

SPARTANs are rocket-class payloads that will be released from the Shuttle orbiter bay and will operate autonomously in orbit for a period of about five days. Potential SPARTANs for planetary astronomy are yet to be determined.

Finally, there is to be the Shuttle Infrared Telescope Facility (SIRTF). SIRTF is expected to be a free flyer at the time of this writing, but funding has not been initiated. It is to be a cryogenically cooled 0.85-meter telescope with a lifetime of over one year. Considerable imaging, spectroscopic, and photometric data on solar system objects can be expected.

9.5 Ancillary Data

There are a number of planetary studies that require the knowledge of external parameters. Such parameters include solar ultraviolet full-disk fluxes, solar-wind composition and velocity, and the occurrence of solar events. In some cases this information is available from measurements by instruments on board planetary missions, but in others it may be necessary to acquire the needed data base from other sources. This data needs to be located, digitized, and entered into PDS.

9.6 Requirements

It is believed that the present ground-based planetary data sets recommended in this document would amount right now to about 40 Megabytes of storage. The updating of those files might add 5 Megabytes per year. IUE may add 150 Megabytes plus annual increments. Future missions such as Space Telescope could produce precipitous rises in the Earth-based data collection. However, we assume that only a small fraction of the total Earth-based archive needs to be continuously accessible on-line, the remainder being transferable on-line upon request. High data rates will be important only for ST images and data of similar format.

The amount of labor required to digitize analog data as proposed, to update existing archives (e.g., the TRIAD file, the PRC film collection, etc), and to prepare initial tutorials may total as much as 5 man-years. File and archive maintenance is estimated at 2 full-time equivalents. It is expected that the fractional time of many individuals would be utilized rather than the full time of a few.

9.7 Summary Recommendations

Earth-based planetary data are used as an aid in planning future missions, conducting current missions, and interpreting results of earlier missions. So it is recommended that selected Earth-based data sets included in the PDS data base.

Certain Earth-based data sets (a small fraction of the total) have been identified for on-line access through the PDS. Other data sets, of less general interest, should be catalogued and be available on-line upon request. It is recommended that a continuing mechanism be established to select appropriate Earth-based data sets for incorporation into the PDS.

A large portion of the Earth-based data which should be included in PDS currently does not exist in digital or machine-readable form. These data must be digitized and documented before inclusion in the PDS.

Documentation of each data set should include a tutorial that is detailed enough to enable non-experts to make correct use of the data. Scientifically knowledgeable experts should be identified and supported to construct this documentation and assist in creating and maintaining the data base.

If NSSDC does not become a partner in PDS with on-line interactive capability. It is recommended that the existing Earth-based planetary data bases in NSSDC be copied for PDS, to be available in on-line format.

Holders of NASA research grants should be urged (even required) to ensure that their future data, if appropriate, are produced in formats convenient for inclusion in PDS.

Data produced under sponsorship of other agencies (e.g., NSF) and from foreign sources (e.g., Spacelab) should be actively sought out for inclusion in PDS.

In order that users will know what Earth-based planetary data sets exist in PDS, a listing of them with suitable keywords in the PDS directory is recommended.

10. LABORATORY

10.1 Definition of Laboratory Data

Laboratory data are traditionally the measurement of one parameter as a function of one or several other parameters. Some of the most common are intensity, transmission, or reflectance versus wavelength and are commonly referred to as spectra. Other common lab data are intensity versus phase angle, line strength versus wavenumber, band strength versus temperature sputtering yield versus impact energy, photoionization cross section versus photon energy, secondary electron emission yields from surfaces and many more. Less common data include such things as viscosity versus temperature. For the purpose of this report, we consider all (x,y) paired data that might be obtained in the laboratory and of interest to the planetary science community. All such data sets will be referred to as spectra. Some data are tabular in nature, but can be thought of (or catalogued as) multiple paired data sets, where the abscissa has the same values. Some data (e.g., high resolution spectra of a gas over a large wavelength range) contain many tens of thousands of points (a few cases containing over a million points are known to exist), while other data contain only a few points (e.g. low resolution spectra of a mineral). An extension of this concept is 3-dimensional data such as Pressure, Volume, and Temperature plots (PVT) for which the data can be treated similarly to (x,y) paired data.

10.2 Sources of Data and Data Volume

Laboratory data are acquired in numerous research laboratories around the country and world. Such data are vital to NASA Planetary programs as they provide the basis for the interpretation of planetary data. Perhaps more important is the need for laboratory data in all scientific research areas, of which planetary science is a small part. The archiving of many laboratory data sets is vital to other NASA and non-NASA activities outside the planetary community. For example, virtually all reflectance spectra of minerals and rocks of interest to planetary science are also of interest to the US Geological Survey, mining and oil companies. In fact, many investigations of basic reflectance data on rocks and minerals have been funded by the Air Force, and the USGS (e.g. data by Graham Hunt and colleagues). Transmission data of gases are important for terrestrial as well as planetary research, and research in this area is funded by the National Oceanic and Atmospheric Administration (NOAA), the Environmental Protection Agency (EPA) and the Department of Defense (DOD). The National Science Foundation funds basic research in many areas. The Department of Energy funds electron, ion, and photon - collisional work for laser and high-temperature plasma (fusion) applications. The Department of Defense funds cross section measurements for gaseous dielect and fast-opening-switch research. Data from all these sources are vital to planetary sciences and need to be archived along with NASA-funded data. Thus, NASA must fund archiving of those data which are of interest to the planetary community, preferably, or seek the assistance of Congress and/or other agencies in a much more general archiving effort.

Laboratory data are obtained by individual Principal Investigators (PI's) and Co-Investigators (CoI's) and are not funded in the same manner as data obtained via spacecraft. Thus, there has been no requirement for archiving data as with spacecraft mission data. Currently, with the tightening of the NASA Planetary budget, the incentive and ability to archive data in individual research groups has also decreased. NASA may have to require each PI in the future to submit data, which are the products of the funded research, to an archive. Such a requirement must include adequate funding for obligatory tasks.

The following tables list selected sources and data volumes of laboratory data relevant to planetary research. The compilation is certainly not complete. However, it represents a good estimate of what exists and what is projected. The basic need is to archive in digital form, the physical properties of all materials of interest to the planetary community. These data include densities, compressibilities, thermal and electric properties, etc. This type of data is implied in the miscellaneous entries in the tables and CRC Handbook entry in Table 4.

10.3 Required Support Documentation for Lab Data

The quality of laboratory data must be assured before they are added to the PDS archive. The PDS laboratory data archive must have extensive documentation to ensure and convince users that the data are of adequate quality for their purpose and that any quirks in the data are fully explained. For published data, a reference to the publication may be all that is required, however, additional supporting documentation would be required if not in the publication (e.g. such as manufacturer, purity, lot number for a gas; x-ray diffraction and microprobe analysis for a mineral). If adequate documentation is not available for published data, this fact should be specifically stated in the archive. If the sample still exists (e.g. a mineral) the last known location could be given. This would ensure, for instance, that if a spectrum of such a material were important for the interpretation of some planetary data, the sample might be found and appropriate analyses could be undertaken to validate the archived data. The peer review engendered by publication is one way to help ensure the quality of data considered for archiving. In cases where a published result has later been shown to be in error, a flag should be added to the data set indicating such, as well as the location of the newer data and a description of the error. In the case of unpublished or preprint-form data, a notation should be appended as to the stated accuracy, anticipated journal and date of publication, and authors' telephone numbers; however, such data must be scrutinized before qualifying for addition to the archive. This review could be done by the steering committee for the discipline.

Table 1
Reflectance Spectra and Associated Data For Solids

Location and People	Description	Volume #spectra	#megabytes
Brown U. C. Pieters	Reflectance data of various minerals. From Re-Lab formerly at JSC	1000	1.5
Cornell J. Veverka J. Gradie + others	Reflectance and phase data on various minerals	1000	1.5
JPL A. Goetz	Reflectance data on various minerals and terrestrial locations	1000	1.5
JPL B. Smythe D. Nash	Reflectance data on various minerals, including ices, irradiated species.	1000	1.5
U. Hawaii R. Clark M. Gaffey B. Singer	Reflectance data on minerals including ices. Analysis products such as curves of growth, temperature shifts of Bands.	5000	7.7
U. Mass. B. Huguenin	Reflectance data on minerals	1000	1.5
U. Wash (Seattle) J. Adams	Reflectance data on minerals and lunar material. Analysis products such as pyroxene band 1 vs. band 2.	3000	4.6
USGS G. Hunt R. Clark	The basic data set of the reflectance spectra of most major rocks and minerals	10,000	15.
A. Lane	UV data on solids	?	?
Miscellaneous from various locations		4000	6.
TOTAL		27,000	41

Notes: The spectra in this table were estimated by a survey done in fall 1982 and were "normalized" to 256 point segments for a spectrum. The volume in megabytes includes 1024 bytes per spectrum (256 real numbers) and 512 bytes of description, on the average. It was estimated that the data volume would grow about 20 percent per year for the next 5 years.

Table 2
Representative Laboratory Spectra of Atomic and Molecular Species

Location and People	Description	Volume #spectra	Volume #megabytes
AMES L. Giver	Newsletter indicating spectra obtained by groups all over country. Primarily gases	<2000	300-500
Sadtler Research Labs., Inc. Philadelphia	Compilation of infrared and UV spectra of organic and inorganic gases.	8,000	24
U. Denver D. G. Murcray A. Goldman	Editors of Chemical Rubber Publishing Co. handbook of high resolution infrared spectra of atmospheric interest.	~500	1.5
U.C. Berkeley S.R. Davis- N. Phillips	Newsletter describing, in part, atomic and molecular spectra and cross sections	6,000	60
A. Lane	UV spectra	?	?
TOTAL		10,000	300-500

Notes: The Giver newsletter indicates a current volume of about 50-100 megabytes per year. This data comes from many research labs, so the data from the other sources indicated in the table may be included. The extent of any overlap needs to be determined. The constraints on accessing commercial proprietary data (i.e., Sadtler Labs) are unknown and need to be investigated.

Sadtler has available on magnetic tape some 8000 infrared spectra of gases and is adding 600-900 more per year to its collection (or about 2 megabytes/year). These tapes include relatively low resolution (4 cm-1) spectra over 600-3600 cm-1 range and a search data base for identifying unknown gases given positions of features. This data base costs $11K for a 5-year lease, with annual updates included. NASA should negotiate a price if the data base is to be implemented on a distributed system.

Subsets of the Sadtler data base have been incorporated into various manufacturers' spectrometers, including Perkin-Elmer, Digilab, Nicolet and IBM. For example, a subset is available on a Perkin-Elmer 3600 at GSFC. One Digilab system implemented on a 5 megabyte disk includes 4500 full 4 cm-1 spectra, with capability for displaying a 32 cm-1 deresolved spectrum. The

Nicolet system is somewhat abridged compared to the Digilab system. IBM's FTIR systems 85 and 95 are 32 megabyte versions with the same kind of capability, generally speaking. Mattson makes a spectrometer called Starlab which incorporates the UNIX operating system to provide similar capabilities. The specific sizes and capabilities of the system would be available from Mattson.

Some spectra may not be included in the above newsletters but are published in journals such as the Journal of Molecular Spectroscopy, Applied Optics, etc. This data volume is unknown and a careful survey would need to be undertaken before any estimates are given.

It was estimated that about one half the data indicated in the U. C. Berkeley newsletter is of use to the planetary community, and that the current growth rate is about 1200 spectra or 12 megabytes per year.

Table 3

Molecular Line Atlas Compilations

Location and People	Description	Volume #transitions	Volume #megabytes
AFGL* (Rothman et al. see below)	Air Force Geophys. Lab IR, millimeter, microwave terrestrial and trace gas parameters.	278,000	22
GEISA* (N. Husson and A. Chedin)	Gestion et etude des information spectroscopiques atmospheriques IR terrestrial and outer planets' gas parameters	228,000	18
GSFC* (W. Maguire V. Kunde)	NASA/Goddard Space Flight Center (Code 693.2) IR terrestrial and outer planets' gas parameters	222,000	18
JPL (R.L. Poynter H. Pickett)	Microwave transitions for 40 molecules, atoms, ions and 140 isotopic species. Data are in the frequency range 0-100 cm-1, and include actually observed and theoretically-calculated spectral lines. Each listed transition has an identifier which flags the particular species. The tabulation is available in tape or microfilm form.	160,000	~8 increasing by ~15% annually for next decade
TOTAL			66

*There is considerable overlap among these atlases, each of which contains aboug 30 gases comprising about 50 isotopic species over 0-17,900 cm-1. Numbers do not sum to total for this reason. For description of AFGL catalog, see Rothman et al., Applied Optics 22 (1983) 1616 and 2247. Plots of the logarithm of absorption line strength (cm-2 atm-1 at 296K) vs. wavenumber for 23 gases in the 0-17,900 cm-1 range, based on the AFGL catalog, are available. See Jae H. park et al., NASA Reference Publication 1084, Dec. 1981. Finally, part of the JPL catalog is contained in the AFGL catalog.

Table 4

Collisional and Reaction Rate Data

Data Base	Description	Volume #megabytes
IAEA JILA Oak Ridge	Photoionization, photoabsorption photodissociation cross sections as a function of wavelength for various species (e.g. He, CO, O, NO, N2, H2O, etc.).	<1.0?
IAEA, IPP, JILA, Oak Ridge	Electron-ion collision cross sections	<2.0?
IAEA, IPP JILA, LASER TRC	Electron swarm, and cross sections as functions of energy and scattering angle for various inelastic processes (ionization, dissociation, electronic excitation, vibrational excitation, and rotational excitation).	<2.0?
IAEA, LASER JILA, TRC	Elastic electron impact cross sections as functions of both energy and scattering angle for various species.	<2.0?
IAEA, JILA LASER TRC	Energetic electron impact cross sections as functions of energy and scattering angle for various species.	<2.0?
JPL and various	Chemical reaction rates of ion-neutral and neutral-neutral for various reactants as a function of temperature (and pressure where relevant).	<1.0?
IAEA, Oak Ridge, Bell Labs, U. Va.	Sputtering yields and products as a function of energy and ion for energetic ion impacts into solid surfaces (e.g. ices).	<2.0?
CRC	A computerized version of the CRC Handbooks.	??<10
TOTAL		~<22

Notes: The growth factor is estimated to be not more than a factor of about 3 over the next decade or so for all of the above, except the CRC.

The casual user needs extensive help menus and users guides to effectively access the data base. Experienced users need to have the ability to turn off extensive help features to speed-up access.

A data base management system is essential due to the diverse nature of the data and the search requirements. Such software systems exist commercially and one might be adapted to such an archive. Most are oriented around business related activities (e.g. find all clients older than 32 with blue eyes who live on the west coast and have ordered widget x3). It is a very different matter to search for a spectrum with an absorption band. By the nature of the data base, the spectra will be of different resolutions, covering different wavelength regions, and be transmittance, emittance, or reflectance. In reflectance, an absorption band can occur on a steeply sloped continuum so that there is no real minimum. It becomes difficult to analytically decide whether or not an absorption band exists. Any commercial data base management system would need extensive additions to handle such cases.

10.6 Analysis of Data

Basic analysis routines must be available to perform simple manipulations of data, display data in various forms (e.g., most commonly as x-y plots), including overlaying various data sets for proper comparison. A history of commands performed should be available, as well as a history of any manipulation of data. The user interface must be in a free format and user input should be analyzed before being executed so that data cannot be destroyed. No user should be able to destroy archive data intentionally or accidentally. Workspace for manipulation of archive data should be available, and software should protect data in the workspace from being inadvertently deleted.

More sophisticated routines should be part of the archive software if well documented and of general use, such as a general synthetic spectrum generating program. User programs should also be allowed.

The required software for such simple but very diverse analysis tasks probably does not exist. There are two systems which might serve portions of requirements:

- specpr : developed at the U. Hawaii as a general one dimensional array processing system and is currently handling all non-imaging data reduction and analysis in the Planetary Geosciences Division as well as some processing for researchers as far away as the eastern US coast. Implementation is imminent at the USGS Denver Geophysics Branch. See Clark (1980), Pub. of the Astron. Soc. of the Pacific, v.92, 221-224.

- IDL : An interactive data language developed at LASP by David Stern. It has been used in Analysis of IUE spectra and Voyager Saturn photopolarimeter ring occultation data.

10.7 Location and Access to the Data Archive

One central facility needs to coordinate addition of data to the archive, and ensure that all documentation needs are met. Due to the many researchers desiring access to such a system, and the computational load (a relational data base searching for absorption bands puts an extreme load on even a large VAX-like computer), individual researchers need at least major portions of the archive online at their home institutions. The cost of individual computers (many adequate systems already exist) is modest compared to constant remote access via commercial telephone lines. It is questionable whether even a very large computer could adequately serve many users online querying a relational data base (see Chapter 12: Data Base Management). The central archive could distribute new laboratory data to regional centers or centers of excellence.

Since the data volume is modest, the laboratory data base would best be served by a distributed system. This would off-load the computational burden to the research centers that need constant access to such data. A distributed system would have the requirement that software be directly or at least easily transportable to several different computer systems. Such standardization is extremely difficult to achieve with the different computer architectures, operating systems, and various versions of language compilers (e.g. FORTRAN) common today. One possibility might be the use of Unix, an operating system available on many machines, including home computers that are approaching VAX-like performance.

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA CP-2343 (Part 1)		

4. Title and Subtitle	5. Report Date
PLANETARY DATA WORKSHOP	October 1984
	6. Performing Organization Code EL
7. Author(s) Dr. Hugh H. Kieffer, Chairman	8. Performing Organization Report No.
9. Performing Organization Name and Address Office of Space Science and Applications Solar System Exploration Division	10. Work Unit No.
	11. Contract or Grant No.
	13. Type of Report and Period Covered Conference Publication
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D.C. 20546	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

This is the proceedings of the Planetary Data Workshop, held November 29-30 through December 1, 1983, at Goddard Space Flight Center, Greenbelt, Maryland. The community of planetary scientists addresses two general problems regarding planetary science data: (1) important data sets are being permanently lost; and (2) utilization is constrained by difficulties in locating and accessing science data and supporting information necessary for its use. These proceedings, in two parts, explore a means to correct the problems, provide science and functional requirements for a systematic and phased approach, and suggest technologies and standards appropriate to the solution.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
planetary science data science data management data base management discipline centers	UNCLASSIFIED - UNLIMITED primary distribution to conference participants; secondary distribution by request Subject Category 91

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
unclassified	unclassified	130 (Part 1)	A07

National Aeronautics and
Space Administration

Washington, D.C.
20546

Official Business
Penalty for Private Use, $300

NASA

National Aeronautics and
Space Administration

Washington, D.C.
20546

SPECIAL FOURTH CLASS MAIL
BOOK

Postage and Fees Paid
National Aeronautics and
Space Administration
NASA-451

Official Business
Penalty for Private Use $300

U.S.MAIL

7 26 1U,J, 840928 S00008ASR
NASA
LANGLEY RESEARCH CENTER
ATTN: TECHNICAL LIBRARY, MS 185
HAMPTON VA 23665

NASA

POSTMASTER: **If Undeliverable (Section 158 Postal Manual) Do Not Return**