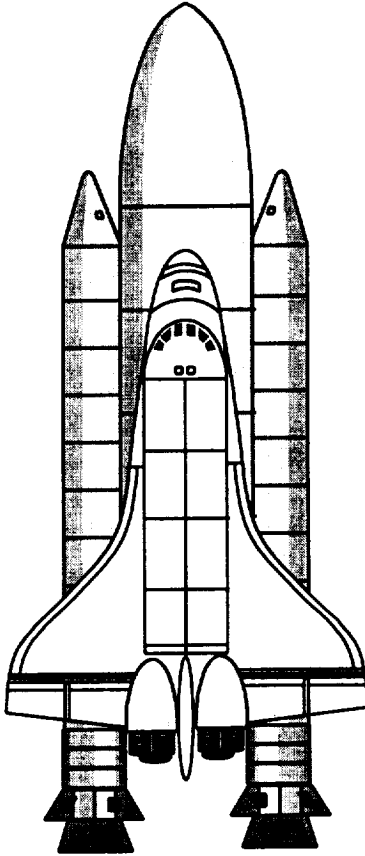
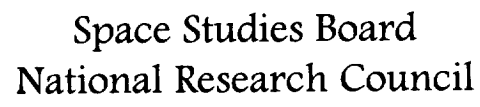


ARCHIVING MICROGRAVITY FLIGHT DATA AND SAMPLES



Space Studies Board
National Research Council



Space Studies Board
National Research Council

ARCHIVING MICROGRAVITY FLIGHT DATA AND SAMPLES

Committee on Microgravity Research
Space Studies Board
Commission on Physical Sciences, Mathematics, and Applications
National Research Council

NATIONAL ACADEMY PRESS
Washington, D.C. 1996

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This report has been reviewed by a group other than the authors according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

The National Academy of Sciences is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Bruce Alberts is president of the National Academy of Sciences.

The National Academy of Engineering was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Harold Liebowitz is president of the National Academy of Engineering.

The Institute of Medicine was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Kenneth I. Shine is president of the Institute of Medicine.

The National Research Council was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Bruce Alberts and Dr. Harold Liebowitz are chairman and vice chairman, respectively, of the National Research Council.

Support for this project was provided by Contract NASW 4627 between the National Academy of Sciences and the National Aeronautics and Space Administration.

Copyright 1996 by the National Academy of Sciences. All rights reserved.

Copies of this report are available from

Space Studies Board
National Research Council
2101 Constitution Avenue, N.W.
Washington, D.C. 20418

Printed in the United States of America

COMMITTEE ON MICROGRAVITY RESEARCH

MARTIN E. GLICKSMAN, Rensselaer Polytechnic Institute, *Chair*
ROBERT A. ALTENKIRCH, Washington State University
ROBERT J. BAYUZICK, Vanderbilt University
GRETCHEN DARLINGTON, Baylor College of Medicine
HOWARD M. EINSPAHR, Bristol Myers Squibb Company
J. GILBERT KAUFMAN, JR., The Aluminum Association
L. GARY LEAL, University of California, Santa Barbara
RONALD E. LOEHMAN, Sandia National Laboratories
MARGARET J. LYELL, West Virginia University
MORTON B. PANISH, AT&T Bell Laboratories (retired)
RONALD F. PROBSTEN,* Massachusetts Institute of Technology
JOHN D. REPPY, Cornell University
JULIAN SZEKELY, Massachusetts Institute of Technology (deceased)
FORMAN A. WILLIAMS, University of California, San Diego

SANDRA J. GRAHAM, Senior Program Officer
VICTORIA P. FRIEDENSEN, Senior Program Assistant

* Former member.

SPACE STUDIES BOARD

CLAUDE R. CANIZARES, Massachusetts Institute of Technology, *Chair*
JOHN A. ARMSTRONG, IBM Corporation (retired)
JAMES P. BAGIAN, Environmental Protection Agency
DANIEL N. BAKER, University of Colorado
LAWRENCE BOGORAD, Harvard University
DONALD E. BROWNEE, University of Washington
JOSEPH A. BURNS,* Cornell University
JOHN J. DONEGAN, John Donegan Associates, Inc.
ANTHONY W. ENGLAND, University of Michigan
DANIEL J. FINK, D.J. Fink Associates, Inc.
MARTIN E. GLICKSMAN, Rensselaer Polytechnic Institute
RONALD GREELEY, Arizona State University
BILL GREEN, former member, U.S. House of Representatives
HAROLD J. GUY,* University of California, San Diego
NOEL W. HINNERS, Lockheed Martin Astronautics
JANET G. LUHMANN, University of California, Berkeley
JOHN H. McELROY, University of Texas, Arlington
ROBERTA BALSTAD MILLER, Consortium for International Earth Sciences Information
Network
BERRIEN MOORE III, University of New Hampshire
MARY JANE OSBORN, University of Connecticut Health Center
SIMON OSTRACH, Case Western Reserve University
CARLÉ M. PIETERS, Brown University
JUDITH PIPHER,* University of Rochester
MARCIA J. RIEKE, University of Arizona
ROLAND W. SCHMITT, Clifton Park, New York
JOHN A. SIMPSON, University of Chicago
ARTHUR B.C. WALKER, Jr.,* Stanford University
ROBERT E. WILLIAMS, Space Telescope Science Institute

MARC S. ALLEN, Director

*Former member.

COMMISSION ON PHYSICAL SCIENCES, MATHEMATICS, AND APPLICATIONS

ROBERT J. HERMANN, United Technologies Corporation, *Chair*
STEPHEN L. ADLER, Institute for Advanced Study
PETER M. BANKS, Environmental Research Institute of Michigan
SYLVIA T. CEYER, Massachusetts Institute of Technology
L. LOUIS HEGEDUS, W.R. Grace and Co.
JOHN E. HOPCROFT, Cornell University
RHONDA J. HUGHES, Bryn Mawr College
SHIRLEY A. JACKSON, U.S. Nuclear Regulatory Commission
KENNETH I. KELLERMANN, National Radio Astronomy Observatory
KEN KENNEDY, Rice University
THOMAS A. PRINCE, California Institute of Technology
JEROME SACKS, National Institute for Statistical Sciences
L.E. SCRIVEN, University of Minnesota
LEON T. SILVER, California Institute of Technology
CHARLES P. SLICHTER, University of Illinois at Urbana-Champaign
ALVIN W. TRIVELPIECE, Oak Ridge National Laboratory
SHMUEL WINOGRAD, IBM T.J. Watson Research Center
CHARLES A. ZRAKET, The MITRE Corporation (retired)

NORMAN METZGER, Executive Director

Foreword

It is the preeminent role of objective, quantifiable data in scientific inquiry that most distinguishes it from nearly every other human endeavor. Hypotheses not firmly anchored to reproducible data become speculations that soon drift away from the realm of science. While all good experimenters support their deductions and inferences with published tables, plots, diagrams, and photographs, the huge quantities of data collected in a typical investigation generally preclude publication of anything beyond a tiny sampling of the partially digested information. After a short time, access to the original evidence, whether digital records or material samples, becomes impossible. This is particularly unfortunate in the space sciences, because of the cost and complexity of reproducing the original experiment or observation.

The easiest solution, to save everything, is clearly impractical and inefficient. The desire to preserve valuable data and samples must be balanced with the cost and difficulty of doing so. Moreover, saving raw records and materials is useless without the tools and recipes required to manipulate and understand them.

The community of microgravity researchers, a relative newcomer to space, is now confronting the question of what to preserve and how to preserve it. This report explores the complexities and ambiguities of the issue and describes a set of principles and processes to cope with them. It considers the differing practices of laboratory research and of the older space sciences. The conclusions of this report of the Committee on Microgravity Research and the Space Studies Board provide a sensible and responsible approach to reaping the largest return from the nation's sizable investment in microgravity sciences.

Claude R. Canizares, *Chair*
Space Studies Board

Preface

In the spring of 1995 officials at NASA's Microgravity Science and Applications Division (MSAD) requested that the Committee on Microgravity Research, a standing committee of the National Research Council's Space Studies Board, provide advice regarding the need for preserving and archiving microgravity data and samples. MSAD asked the committee for guidance on the types of microgravity data and samples requiring archiving; the location and duration for their preservation; the division of responsibilities shared between NASA and microgravity investigators for providing timely archiving of spaceflight results; access to data and samples; and archiving of flight data versus ground-based data.

During the course of this study, the committee was briefed extensively by MSAD personnel on MSAD's current archiving strategy and activities. The committee also examined a variety of large data archiving systems that have been developed and managed by other organizations. These systems included the Planetary Data System, the Space Telescope Science Institute database, the Global Change Archives, and the National Materials Property Data Network. Also examined were methodologies for archiving samples as varied as cosmic dust specimens and Antarctic meteorites. In addition, the committee discussed specific issues of microgravity data and sample archiving affecting the constituent scientific communities.

Among the previous National Research Council reports on data archiving, the committee took particular note of the following:

- *Preserving Scientific Data on Our Physical Universe* (National Academy Press, Washington, D.C., 1995) examined issues related to the archiving of scientific and technical data in the physical sciences and concluded that among the key criteria for determining what data should be archived are the difficulty and cost of replacing them.
- *Selected Issues in Space Science Data Management and Computation* (National Academy Press, Washington, D.C., 1988), a Space Studies Board report, emphasized that NASA should adopt and implement an explicit data management plan for all spaceflight investigations and should provide sufficient resources for archiving and continued protection of data. The report lays out a number of basic guidelines for archiving.
- *Networking of Materials Property Data* (National Academy Press, Washington, D.C., 1983), a National Materials Advisory Board report, documented the limited amount of archived data on materials and the need for a substantial increase in archiving of materials performance databases as well as for improved means of end-user access via menu-driven software that would help users to search for and understand the data. The

study emphasized the importance of careful documentation of the metadata (e.g., materials descriptions, definitions of test variables) associated with materials performance data.

- *Computer-Aided Materials Selection During Structural Design* (National Academy Press, Washington, D.C., 1995), a National Materials Advisory Board report, presented findings that confirm both the need for comprehensive databases covering a wide range of materials science phenomena and the importance of careful planning of the design of the archival databases used to collect the data and systematically store it for retrieval.

A number of people who assisted the committee during its preparation of this report deserve special thanks for their contributions: Robert Rhome, Roger Crouch, and Gary Martin of NASA Headquarters; Laura Maynard and Howard Ross of NASA Lewis Research Center; Charles Baugher and Robert Snyder of NASA Marshall Space Flight Center; J. Gilbert Kaufman of the Aluminum Association; Ray Walker of the Institute of Geophysics and Planetary Physics of the University of California at Los Angeles; Lola Olsen of the Global Change Master Directory of the NASA Goddard Space Flight Center; Roger Doxsey of the Space Telescope Science Institute; Judith Allton of Lockheed Martin; Michael Zolensky of the Planetary Institute at NASA Johnson Space Center; and Paul Uhler and Richard Hart of the National Research Council.

Contents

Introduction	1
The Need for Archiving	1
Purpose and Scope of This Study	2
Past Archiving Practices and Changing Needs	2
Current MSAD Archiving Strategy and Methods	4
Recommendations and Guidance for MSAD's Microgravity Data and Sample Archiving Strategy	5
Data to Be Archived	5
Samples to Be Archived	6
Location of Archives	6
Format and Accessibility of Archived Material	7
Proprietary Access and Submission to Archives	9
Retention of Data and Samples in Archives	9
Data and Samples from Ground-based Experiments	9
Further Recommendations	10
The EDMP	10
Future Assessment	11
Appendix: Input Form for the Experiment Data Management Plan	13

INTRODUCTION

The Need for Archiving

Experiments are conducted in microgravity primarily to add to our store of scientific knowledge and understanding. Their purpose is to acquire critical data and samples needed to test hypotheses and check specific theoretical predictions, and sometimes, to determine how certain phenomena are altered under a specified set of experimental conditions. As in every field of science, the primary method of archiving experimental data and results from microgravity experiments is through publication of the research in peer-reviewed journals. Experiments performed in a low-gravity environment should not be an exception to this practice, and NASA should continue to encourage its investigators to publish their results. However, it is also the case that today's journals are limited with respect to the amount of data they can publish from any one experiment. This limitation may pose a problem for other scientists interested in the results of experiments performed aboard orbiting spacecraft, because such experiments cannot be easily reproduced, owing in part to the high cost of conducting microgravity experiments, particularly those carried out in the long-duration microgravity environment of a spacecraft. The data and samples arising from such experiments are also costly, especially when compared with their terrestrial counterparts, which can usually be generated at a small fraction of the cost incurred for those returned from space. In addition, access to microgravity is limited through the stringent scheduling restrictions of space launches, making it either difficult or impossible for other scientists to plan repeat experiments and execute them in a timely manner.

A recent, informal survey distributed by MSAD to a limited sample of the microgravity user community suggested little demand for a formal process of data and sample archiving. Some of those scientists surveyed felt that traditional journal articles and other bibliographic sources sufficed for their personal archiving needs. However, it was also clear from their replies that some of the respondents were unaware of the opportunities for searching for and accessing data from an active microgravity archive.

The dual factors of high cost and limited access to space underscore the need for preserving at least some of the data produced in the course of conducting microgravity experiments. Indeed, an effective strategy for archiving data sets from flight experiments represents a prudent protection of NASA's and the nation's research investment. The justification for archiving data and preserving samples must consider their value, judged by both the cost of reproducing them and their future utility. Such judgments are, of course, subjective and difficult to quantify.

Purpose and Scope of This Study

To obtain help in evaluating its current strategy for archiving data and samples obtained in microgravity research, NASA's Microgravity Science and Applications Division (MSAD) asked the Space Studies Board's Committee on Microgravity Research for guidance on the following questions:

- What data should be archived and where should it be kept?
- In what form should the data be maintained (electronic files, photographs, hard copy, samples)?
- What should the general format of the database be?
- To what extent should it be universally accessible and through what mechanisms?
- Should there be a period of time for which principal investigators have proprietary access? If so, how long should proprietary data be stored?
- What provisions should be made for data obtained from ground-based experiments?
- What should the deadline be for investigators placing their data in the archive?
- How long should data be saved? How long should data be easily accessible?

As a prelude to making recommendations for optimum selection and storage of microgravity data and samples, the committee in this report briefly describes NASA's past archiving practices and outlines MSAD's current archiving strategy. Although the committee found that only a limited number of experiments have thus far been archived, it concluded that the general archiving strategy, characterized by MSAD as minimalist, appears viable. A central focus of attention is the Experiment Data Management Plan (EDMP), MSAD's recently instituted data management and archiving framework for flight experiments. Many of the report's recommendations are aimed at enhancing the effectiveness of the EDMP approach, which the committee regards as an appropriate data management method for MSAD. Other recommendations provide guidance on broader issues related to the questions listed above. This report does not address statutory or regulatory records retention requirements.

PAST ARCHIVING PRACTICES AND CHANGING NEEDS

Microgravity research has been identified as a distinct spaceflight activity within NASA's space science and applications programs since the mid-1970s. Some of the first microgravity experiments and low-gravity demonstrations were conducted by astronauts aboard the Skylab and the Apollo-Soyuz spacecraft. Early materials-processing spaceflight activities carried out under the auspices of the Marshall Space Flight Center (MSFC) led to the first recovery of microgravity spaceflight data and samples. These early experiments and demonstrations produced processed materials and biological samples as well as a wide range of data, including temperatures, pressures, and other physical measurements. The data were recorded in a variety of formats such as photographs, video tape, and cinefilm. Electronic data were sent to MSFC by telemetry or delivered there post-flight, often in the form of magnetic tape.

Principal investigators (PIs) who participated in early microgravity spaceflight activities submitted experiment implementation plans to NASA, the purposes of which were to describe their proposed flight experiments and their anticipated post-flight

analyses. After being given a specified period—usually 1 year—for post-flight data analysis, PIs prepared their final report, including a description of the disposition of the data and the microgravity samples. Unused portions of flight samples were to be returned to MSFC, cataloged by the archivist, and treated as “space artifacts.” Scientific results from individual spaceflight experiments usually were published in the form of NASA technical memoranda and as peer-reviewed journal articles. Such publications were tracked through a bibliographic archive also assembled and maintained at the Space Science Laboratory at MSFC. Summaries of experiments performed on the early microgravity missions were published by NASA as a series of Microgravity Materials Processing technical memoranda. Publication as a NASA technical memorandum of the combined scientific results of multi-investigator spaceflight missions continues to the present day.¹

Descriptive information on microgravity experiments and missions, accumulated by NASA and its investigators through 1990, began to be archived several years ago at MSFC. This set of experiment descriptions, originally a PC-based flat file with bibliographic references, was created by a NASA staff member at MSFC and then brought to its current status with the assistance of a contractor. This database has now been ported to the World Wide Web and renamed MICREX. The MICREX database is a useful source of historical microgravity information. Were the MICREX database updated to include descriptions and bibliographic information for all the microgravity experiments conducted after 1990, its value as a historical record would be greatly enhanced.

As NASA's microgravity research program expanded and matured beyond the early years, entirely new scientific components were added to microgravity research beyond the original emphasis on materials processing. For example, the growth of macromolecular crystals, low-temperature physics, fluid dynamics, combustion science, and the integration of international payloads were all gradually added. These new science activities have vastly complicated the requirements for archiving samples and data. For example, protein crystals cannot be archived because they degrade rapidly, and so a principal component of the archived data would be the x-ray diffraction data collected from such crystals. These data, when compared with similar data collected from laboratory-grown crystals, provide the primary means of analyzing the results of the protein crystal growth experiments conducted in microgravity. By contrast, mid-deck or glovebox microgravity experiments often yield photographic film and video tape data. These data are usually retained by the responsible NASA center, and the PI is provided the first copy.

International experiments, which are becoming more frequent, further complicate the archiving of data and samples because of the need for data sharing and negotiated sharing of resources and responsibilities among the participating national organizations. Experience has also shown that the methods employed in using, tracking, and returning space-processed materials are not uniform among the world's space agencies and PIs collaborating on international microgravity research. Contributing to the complications in the archiving of microgravity spaceflight results from international experiments is the fact that the lines of authority and the rules and responsibilities governing non-U.S. PIs, their

¹ See, for example, NASA Technical Memorandum 4283, 1991, “Microgravity Science and Applications Bibliography,” 1990 revision, Washington, D.C.; NASA Technical Memorandum 4348, 1991, “Microgravity Science and Applications Bibliography,” 1991 revision, Washington D.C.; and NASA Technical Memorandum 4469, 1993, “Microgravity Science and Applications Program Tasks and Bibliography for FY 1992,” Washington, D.C.

national funding sources, and the space agencies that develop and support their space facilities are usually different from those for U.S. microgravity investigators.

CURRENT MSAD ARCHIVING STRATEGY AND METHODS

Recently, MSAD has begun implementing a data management and archiving plan whose key feature is the requirement to file a formal Experiment Data Management Plan (EDMP) for each microgravity flight experiment currently funded and manifested for orbital flight by MSAD. MSAD expects each PI to file an EDMP prior to approval of an experiment for flight. An example of the input form now used by MSAD for the EDMP is shown in the appendix. The essential elements of this new procedure are the following:

- The PI and NASA sign an agreement—the EDMP—detailing the responsibilities of each with respect to management of data and samples after the mission.
- The PI delivers the agreed-upon data to NASA after the mission.
- The data are stored at the Johnson Space Center (JSC), the Marshall Space Flight Center (MSFC), the Lewis Research Center (LeRC), or the Jet Propulsion Laboratory (JPL), depending on the type of flight experiment.
- The data are stored at the NASA center in a variety of formats, with little subsequent processing of received data.
- No on-line access to the stored data is developed.
- The data stored at the four centers are actively managed by the MSFC (which also manages the JSC data) and LeRC (which manages the JPL data), and investigators interested in acquiring available data can contact the data manager at either center.

At present NASA has several main data archives covering different aspects of the MSAD research program. The MICREX is discussed above, and two of the archives located at LeRC are described here. The data produced by the two primary microgravity accelerometer systems developed by NASA are stored in archival form at LeRC. The first of these, the Space Acceleration Measurements System (SAMS), is capable of measuring the spectral power density of the acceleration environment along three axes aboard the space shuttle. The SAMS system permits on-board recording and near-real-time telemetry of the microgravity spectrum, including g -jitter over a spectral range from several hundred hertz down to relatively low frequencies of 1 Hz. The second accelerometer system, the Orbital Acceleration Research Experiment (OARE), is designed to measure the steady levels of the microgravity acceleration. Both SAMS and OARE data for the individual microgravity missions that carry these systems are currently available on CD-ROM with on-line access.

Observational data from fluids and combustion experiments are also stored, in the form of film and video, at LeRC. First-generation copies of these data are produced in CD-ROM format and are cataloged on the World Wide Web. Copies are available for distribution upon request.

To allow for analysis of the results of their microgravity experiments, PIs have traditionally been given exclusive access to the post-flight data and samples for 1 year following their delivery to the PI. Often it takes up to several months for the removal of microgravity experiment data from the spacecraft, intermediate processing, and final delivery to the PI. There are no uniform policies for distribution and retention of unaltered

space-processed samples, nor are there uniform requirements to preserve key microgravity facility components, such as sample cartridges and flight furnaces—even though some of these might be needed for replication of experiments in terrestrial runs performed after the mission.

RECOMMENDATIONS AND GUIDANCE FOR MSAD'S MICROGRAVITY DATA AND SAMPLE ARCHIVING STRATEGY

Every microgravity experiment conducted under the auspices of MSAD requires review to determine whether or not some of the data and samples warrant archiving and preservation. In this review, consideration should be given to the cost of archiving, the potential future utility of the data, and their intrinsic scientific value. The high costs and limited opportunities for reproducing microgravity data dictate the need for archival preservation² based on a careful selection process and an ongoing evaluation of archiving costs in relation to the scientific community's valuation and use of the microgravity data and samples.

Data to Be Archived

Flight experiments performed in the microgravity sciences vary enormously in size, scope, subject discipline, and data output. It is obvious that each experiment will have differing requirements for the archiving of its useful data and, therefore, decisions about what should be archived by NASA will have to be made on a case-by-case basis. The individual investigator, with the assistance of the NASA project scientist, is in the best position to recommend which data can be most usefully archived. Both the form and content of the archived data (such as video, numeric, photographic) will need to be considered prior to and after the flight. Although digital data should be encouraged because of ease of storage, copying, and access, in some cases much of the value of the original data may be lost if only the extracted numeric values are stored.

With proper implementation, NASA's EDMP process can serve as an appropriate tool for establishing the form and content of data to be delivered for archiving. The committee believes that the categories of information required by the EDMP for each experiment are appropriate and necessary to properly document the data and samples obtained from flight experiments. **The committee recommends that the process for establishing a mutually agreeable EDMP take place early in the mission planning process and that the list of data proposed for archiving by the PI be peer reviewed.** Several pre-mission science reviews are already in place, such as the Science Readiness Review and the Requirements Definition Review, during which decisions regarding the EDMP could be peer reviewed. If flight investigators were required to include at least a preliminary list of archival data and samples in the Science Requirements Document required by NASA for their experiment, the list could be peer reviewed and refined at one or both of these pre-mission science reviews.

² In the recently published National Research Council report *Preserving Scientific Data on Our Physical Universe* (National Academy Press, Washington, D.C., 1995, p. 34), the uniqueness of a data set and the cost of replacing it were both cited as primary criteria for deciding whether to archive data.

Samples to Be Archived

In the past, samples returned from some spaceflight experiments have been retained by the flight PIs, and some of these samples have been completely consumed during post-mission analysis. NASA does have a policy that calls for the archiving of unused portions of flight samples, and the committee encourages NASA to give greater consideration to how this policy should be carried out with respect to future microgravity missions. The same arguments justifying the need for archiving data from flight experiments also hold for samples. In addition, continuing advances in analytical instrumentation make it conceivable, if not almost certain, that new information could be obtained from a flight sample in the years following flight.

The great variation in experiments and sample types makes it impractical to develop a single set of decision rules regarding the disposition and archiving of flight samples. Nonetheless, decisions to preserve samples, and archive data, should ultimately rest on a cost-benefit estimate, wherein the potential costs for reproducing the results, their intrinsic scientific value to the user community, and the prospects for future utility are weighed against incurring present costs for storing the data. Some samples, such as protein crystals, degrade so quickly, or require such stringent storage conditions, that long-term archiving becomes impossible. Decisions about the archiving of samples should, therefore, also be made on a case-by-case basis. **The committee recommends that the EDMP process also be used to frame, answer, and then review the question of what portion of flight-generated samples will be retained by the investigator, and what portion, if any, should be transferred to NASA for archiving.**

Samples should be retained and made available according to accepted archiving practices. This means that samples of sufficient value to warrant archiving also warrant the expense and effort required to store them under conditions that preclude significant contamination or degradation of the material. Archiving of samples entails proper cataloging of samples so that sufficient information, including the necessary sample history, is available to interested scientists, allowing them to make intelligent queries for use of the sample.

NASA should also develop streamlined procedures by which decisions on applications for use of flight samples can be made expeditiously. Unless a situation arises in which there are a large number of conflicting demands for the use of flight-generated materials, loans of the samples to bona fide investigators can be left to the discretion of the responsible discipline scientist or project scientist. If a conflict arises, it might become necessary for the project scientist to consult with the appropriate MSAD discipline working groups (DWGs) to arrive at a decision.

Location of Archives

In general, samples will be more accessible to other interested scientists if they are stored at a limited number of locations, such as NASA centers. The most reliable method of ensuring future access to flight data and samples is for NASA to establish and maintain central archives. NASA centers are the obvious locations for maintaining these archives, and the committee sees no obvious problem with NASA's current plan to apportion management of its archives between MSFC and LeRC, as long as there are sufficient pointers guiding inquiries between the two. Each of these archives is now managed

independently, and NASA may wish to consider, after gaining further experience with the current system, whether greater coordination between the archives is needed.

In some instances, however, it may be decided that archived samples should reside at the laboratory of the flight investigator. In any case, the location of archived samples should be clearly indicated in the EDMP and in published references to the flight experiment. NASA should also consider other means of alerting the science community to the existence of repositories of samples, such as regular notices in science journals, in NASA newsletters and bulletins, and on Internet World Wide Web home pages.

Format and Accessibility of Archived Material

Archived data and samples serve no purpose if they cannot be subsequently retrieved and used. Access issues have been studied previously by a number of groups, and the committee agrees with the conclusions in *Networking of Materials Property Data*³ and *Computer-Aided Materials Selection During Structural Design*⁴ that menu-driven and intuitively understood search and retrieval interfaces are essential if the archived data are actually to be used in the future. The use of cryptic command-driven interfaces virtually assures that end-user scientists with limited time will rarely spend the time and effort to relearn those commands each time they want to search for information. In addition, data that are stored in inaccessible or unidentified physical formats will also go unused. Examples include digital data archived without the executable program required to read and organize the data, or a video tape that can be played back only on specialized equipment built by the original investigator. Access remains an issue of special concern for any database that is likely to contain many different types of data.

The committee assumes here that interest in the use of these data will be limited primarily to specialists in the same or related field of science in which the experiment was performed. Scientific users can be expected to be aware of the published literature, where the results of the flight experiment will in most cases be recorded. Such users will have an understanding and appreciation of the general types of data used in that scientific field, and even perhaps some familiarity with the specific data types collected on the archived flight experiment. Therefore, an archive can be designed with the specialized user in mind, as is common practice for many scientific and engineering databases. This approach, stressing specificity, serves both to lower the cost of maintaining the archive and to reduce the amount of ancillary interpretive information that must be created and stored with the experiment data.

The information that is required to interpret experimental results is generally referred to as metadata, and its importance is discussed in some detail in *Computer-Aided Materials Selection During Structural Design*.⁵ In designing an archiving capability, it is vital to provide for inclusion of information such as the experimental error and the various parameters that make specific data meaningful.⁶ It is important to know, for example, a particular material's composition and probably also the methods by which it was produced.

³ National Research Council, National Materials Advisory Board. 1983. *Networking of Materials Property Data*. National Academy Press, Washington, D.C.

⁴ National Research Council, National Materials Advisory Board. 1995. *Computer-Aided Materials Selection During Structural Design*. National Academy Press, Washington, D.C.

⁵ See footnote 4.

⁶ See footnote 3.

Parameters such as temperature, pressure, humidity, environment, and the like are key to understanding the limitations of the application of the data and how they may be compared with other data from older or future experiments. Metadata supporting the experiment data under evaluation are essential, and decisions concerning the metadata to be archived should be included in the EDMP.

It would not be practical to attempt to list all of the different types of data that might be collected from a microgravity flight experiment, or to try to indicate the information and physical devices that should be stored with primary data. The committee recommends, however, that NASA and its PIs consider the following general guidelines when making decisions on which data to archive and how to ensure their accessibility.

- All data selected for archiving should be accompanied by sufficient explanatory metadata to allow a knowledgeable scientist, with access to the published literature, to interpret the contents of the archive independently and with ease and accuracy.
- Digitally stored data must be accompanied by a copy of one or more computer programs that are capable of accessing, organizing, and properly displaying these data. Such programs should of course be chosen with ease of use and common platform compatibility in mind. Clear directions for the use of the program and data should also be included in the archive.
- Attempts should be made to convert data in rare formats (e.g., holographic film) to more accessible formats. In cases where conversion is not practical or the transfer would result in an unacceptable loss of information, NASA should decide, on the basis of cost, whether to maintain the equipment capable of accessing the data in its archived format.

The committee further recommends that NASA maintain running records of when and how often data and samples from a particular microgravity experiment are requested, in order to judge more accurately the awareness of, and demand for, these data and samples by the scientific community.

As with its archives of samples, NASA should make reasonable attempts to ensure that the scientific community is aware of archived microgravity data and has a means of gaining access to it. **The committee recommends that NASA take advantage of the growth in the Internet-based World Wide Web to post EDMPs on-line for all of its microgravity flight experiments.** On-line EDMPs ought to list and describe in sufficient detail (1) all of the data and samples that are or will be archived from a flight experiment, (2) the exact location and current status of the samples, and (3) the procedures required to gain access to both data and samples. Sufficient links from various NASA Web home pages should be set in place to allow individuals searching the Internet to locate the EDMPs readily. NASA should also consider using other effective means of alerting the scientific community to the availability of microgravity data and sample archives, such as placing timely notices in newsletters, bulletins, and journals. The committee concluded that so long as adequate mechanisms are in place to alert interested scientists and point them to the appropriate NASA contact from whom data can be requested, it is unnecessary to attempt to place all actual flight experiment data on-line. As a practical matter, many of these flight data sets are too large for on-line storage or access, and, in some cases, data are not in digital form, making on-line access to them difficult, if not impossible.

In addition to the EDMPs, MSAD should also maintain an easily accessible, on-line, central catalog of all of the flight experiments for which data and/or samples are archived. In the case of more recent experiments the catalog might merely contain a pointer to the

EDMP locations. For older experiments for which no EDMP was created, the catalog should list the various archived samples and data sets, their locations, and the procedures for gaining access to them.

Proprietary Access and Submission to Archives

It has commonly been NASA's practice to allow flight investigators exclusive access to their own flight data for 1 year following their receipt of these data. In general, a 1-year period of exclusive use should provide reasonable and sufficient time to allow a PI to analyze these data and initiate steps toward publication of the flight results without the concern of being preempted. The committee recognized that instances may arise in which a PI legitimately requires an extension of the period of exclusive use, and NASA should develop a petition process that allows such requests to be considered. At the end of the period of exclusive use, however, the PI should turn over the agreed-upon data to the NASA archives. NASA, in turn, should monitor PI compliance with this policy rather vigorously, because after the passage of 1 year investigators frequently shift their attention to other projects.

Retention of Data and Samples in Archives

In general, advances occurring in most laboratory sciences limit the utility of a data set to fewer than 10 years beyond the time it was collected. **The committee recommends that NASA maintain archived data and samples for 10 years, at the end of which period NASA should seek a recommendation from its internal scientific advisory groups as to whether further archiving is merited.** Such reviews could best be performed by the appropriate DWGs, and NASA should make available to them its records on the frequency of requests for the archived material over the preceding decade. Should the DWG determine that further retention of the archive is not needed, then it should recommend whether or not the material be turned over to some national archiving group for purely historical purposes. If neither archiving option is recommended by the review group, then the material should be offered by NASA first to the original flight PI, and then to collaborators. As a final option, NASA should consider utilizing these materials for their educational and outreach value. Space-flown samples, for instance, could be a valuable resource for schools and museums attempting to stimulate young people's interest in science.

It should be pointed out that in situations such as the on-line storage of digital data sets, the costs of retaining the data indefinitely may be trivial. In such cases NASA may wish to consider waiving the review and retaining the data set in perpetuity.

Data and Samples from Ground-based Experiments

The high cost of performing flight experiments and the limited opportunity to reproduce them have both been cited as reasons for archiving data. This argument does not generally apply to ground-based experiments performed by NASA. In most cases, the ease of reproducing ground-based microgravity experiments (such as those done in drop-

towers) and the additional cost incurred in archiving data from such experiments are likely to outweigh the benefits. However, the need for archiving does apply to baseline data collected on Earth that are a critical component of the flight experiment. Similarly, the EDMP should also contain references to publications derived from ground-based experiments that led to microgravity experiments conducted in flight.

FURTHER RECOMMENDATIONS

The EDMP

The EDMP approach is, in the committee's opinion, precisely the kind of data archiving and management policy needed by MSAD. It organizes and imposes discipline and uniformity on the highly diverse, sometimes conflicting, archiving requirements of the user community. Implementing a uniform and effective EDMP policy not only will impose order on the process of data archiving but also will help to ensure that NASA meets its responsibility to the public.

To enhance the effectiveness and utility of EDMPs, the committee offers several additional recommendations:

- The committee found that not all microgravity flight experiments have EDMPs, owing in large part to the relatively recent implementation of this policy. EDMPs should be required from all flight PIs early in the flight approval process and should become a part of the Science Requirements Document.
- The requirement for the submission of EDMPs ought to be described explicitly in each NASA Research Announcement calling for proposals to perform microgravity research. If the EDMP is specified as a requirement at the initiation of research proposals to perform microgravity flight experimentation, PIs will be encouraged to think about the need for archiving their results from the outset.
- EDMPs should be an item for discussion at each of the NASA science reviews at which peer review occurs. A tentative EDMP, developed jointly by the PI and the project scientist, should be presented as early as the Science Requirements Review stage, and then subsequently amended at each of the follow-on reviews. These reviews not only should consider the goals of the experiment in determining what should be archived but also should take a broad view of the potential value of the data to science.
- The EDMP should contain a current bibliography of all relevant reports in the public domain and of journal articles on related ground-based research authored by the PI prior to flight, and it eventually should be amended to include those published subsequent to flight.
- Completed EDMPs vary somewhat in form and content depending on the NASA center overseeing the preparation of the EDMP. For example, some finished EDMPs have cover pages signed by the PI and the project scientist, each agreeing to the content of the EDMP, whereas others do not. EDMPs should be uniform across all of the microgravity flight experiments sponsored by NASA.
- The PI and the project scientist should be given joint responsibility for ensuring on-line availability of their EDMP prior to flight. An Internet Web home page is currently recommended as a common access point for all active EDMPs. Subsequent amendment of

on-line EDMPs, both prior to and after a flight, should remain the joint responsibility of the PI and the NASA project scientist.

- The committee recognizes that EDMPs do not constitute a data archive per se. EDMPs can facilitate effective and convenient access to the actual archive and should be constructed so as to help serve that purpose.

Future Assessment

A number of important questions regarding the archiving of microgravity data and samples will remain unanswered until MSAD has built up sufficient experience in this activity. These questions include such issues as the likely demand for the archived materials and MSAD's level of success in developing appropriate archives. MSAD will need to monitor closely the implementation and performance of the archiving program in its early operational period in order to ascertain actual user needs and make any necessary adjustments. After a suitable period, perhaps 5 years, MSAD should formally reassess its microgravity archiving activities to evaluate their appropriateness and success.

APPENDIX

INPUT FORM FOR THE EXPERIMENT DATA MANAGEMENT PLAN

The following version of the input form for the NASA Experiment Data Management Plan (EDMP) was received from the NASA Lewis Research Center and is considered the most current version of this recently edited form. The EDMP form is formatted simply to allow easy access from the Internet.

Experiment Data Management Plan

<Experiment name>

<Mission Name>

<Date>

Prepared By:

Approved By:

<Name>	Date
<Experiment> Project Scientist	
<Affiliation>	

<Name>	Date
Chief, MSAD Science Branch	
NASA Headquarters	

Concurrence:

<Name>	Date
Principal Investigator/Experiment Program Scientist	
<Affiliation>	

EXPERIMENT DATA MANAGEMENT PLAN

<EXPERIMENT NAME>

PREFACE

The Experiment Data Management Plan (EDMP) provides the data management plans for the <Experiment Name> to be conducted on <Mission>. The EDMP will also be used to provide planning information to the MSAD Archive System and will provide a consolidated record of the experiment data and products.

The EDMP serves multiple purposes and is updated as necessary to reflect the current status of the experiment's plans. The initial version of the EDMP is developed before the mission to document the experiment's expected archival requirements, to help the MSAD Archive Centers prepare for accommodation of the experiment's data products, and to facilitate their transfer to the archive following the mission. After the appropriate data products have been delivered to the archive, the EDMP is amended to describe the actual data types archived, the quantities of each, and any necessary locating information. The amended EDMP is then incorporated into the Archive Center's local directory and is also used to develop entries for NASA's Master Directory. Some of the sections in the EDMP are pertinent only to the post-flight amended document and may be left blank in the pre-flight version.

This document has been prepared in accordance with the OSSA Program Directive on Data Management and the Microgravity Science and Applications Division Management Plan.

1.0 CONTACT INFORMATION

1.1 PRINCIPAL INVESTIGATOR

Name / Affiliation
Address / Phone / E-mail

1.2 PROJECT SCIENTIST

Name / Affiliation
Address / Phone / E-mail

1.3 EDMP AUTHOR

Name / Affiliation
Address / Phone / E-mail

1.4 ARCHIVE CENTER TECHNICAL CONTACT

Name / Affiliation
Address / Phone / E-mail

2.0 EXPERIMENT DESCRIPTION

2.1 EXPERIMENT NAME

2.2 MISSION

Mission name, STS flight, and launch date

2.3 INSTRUMENT USED

Instrument in which experiment was flown, e.g., Crystal Growth Furnace/CGF

2.4 PURPOSE

Short description of the purpose of the experiment. Should use existing write-ups from mission brochures, etc.

2.5 METHOD

Short description of experiment method. Should use existing write-ups from mission brochures, etc.

2.6 GENERAL EXPERIMENT SUMMARY

General overview of the experiment including objectives, data acquisition, special hardware requirements, etc.

2.7 SUMMARY OF RESULTS AND DATA

2.7.1 SUMMARY OF RESULTS

A short summary of experiment results to be included in the amended EDMP; can be taken verbatim from another source, e.g., Executive Summary of the Investigator's Final Report

2.7.2 SUMMARY OF DATA

A summary of the actual data that was collected during the experiment. This includes digital data (e.g., thermocouple readings, voltage readings, etc.), and visual data (e.g., the film/video showing the candle flame increasing/decreasing with voltage fluctuation, etc.)

2.8 KEYWORDS

Discipline, Subdiscipline, Parameter Group and Parameter are terms used in the Master Directory to aid in a user's search for data sets. There is a standard list of keywords provided by MSAD to use when filling out these sections. The first level of keywords, Discipline, will always be Microgravity

2.8.1 SUBDISCIPLINE

The name of the experiment's subdiscipline within Microgravity (e.g. Materials Science, Biotechnology, Fluids, or Combustion)

2.8.2 PARAMETER GROUP

Parameter Group should provide the next division under Subdiscipline to classify the experiment (e.g. Electronic Materials, Separation Sciences, Interface Dynamics, or Smoldering)

2.8.3 PARAMETER

Parameter should provide the next level under 'Parameter Group' needed to classify the experiment (e.g. Semiconductors, Electrophoresis, Surface Tension, or Flameless Combustion)

2.8.4 GENERAL KEYWORDS

Any keywords that do not fit within first four levels of keywords. Since there is no standard list for general Keywords, this section should contain all the specialized terms that apply to the experiment (e.g. Gallium Arsenide, Cells, Drops, Polyurethane Foam)

3.0 MEASUREMENT AND ANALYSIS DESCRIPTION

3.1 MEASUREMENT TECHNIQUES

Provides a brief description of instrumentation used, measurements taken, etc.

3.2 ANALYSIS TECHNIQUES

Provides a brief description of the types of analyses performed on-orbit, in the operations control center, and post-flight by the PI and team

4.0 ARCHIVING AND ACCESSIBILITY

4.1 DATA ARCHIVE CENTER

The pre-flight EDMP should state the MSAD Archive Center assumed to receive the data after flight; if there is more than one expected location for the data, please describe; the amended EDMP will state the location(s) of the data inventory

4.2 INVENTORY OF DATA TO BE ARCHIVED

The pre-flight EDMP describes data types and quantities expected to result from the investigation; the amended EDMP provides data types, media, formats, quality, quantity, location, special storage requirements, etc. for all archived data

4.2.1 VIDEO

Include quantity, tape format, tape numbers, description of tape content

4.2.2 FILM

Include quantity, film format, film numbers, description of film content

4.2.3 DIGITAL DATA

Flight digital data, documentation files, drawing files, etc.

4.2.4 SAMPLES

4.2.5 OTHER

4.2.6 PUBLICATIONS/REPORTS/ETC.

4.2.7 RELATED GROUND BASED EXPERIMENT DATA

4.2.8 DATA NOT ARCHIVED

If a decision is made not to archive specific data, give justification and description of that data

4.3 DATA ACCESSIBILITY AND AVAILABILITY

The pre-flight EDMP should describe the delivery schedules for the data to the archives and any expected accessibility restrictions; the amended EDMP should describe any unique accessibility or availability issues

4.4 POLICIES FOR PROPRIETARY DATA

State any proprietary policy agreements that will impact the availability of the data to the science community

