Memorandum of Understanding
between
The High Energy Astrophysics Science Archive Research Center
and
The National Space Science Data Center/
The World Data Center for Rockets and Satellites

Dated: April 18, 1997

Approved by:

Joseph H. King 4/18/97

Joseph H. King
Head, National Space Science Data Center
Director, World Data Center A for Rockets and Satellites
Goddard Space Flight Center

Nicholas E. White
Director, High Energy Astrophysics Science Archive Research Center
Goddard Space Flight Center

Günter Riegler

Guenter R. Riegler
Senior Scientist, Research Program Management Division
Office of Space Science

Joseph H Bredekamp

Joseph H. Bredekamp
Senior Science Program Executive/Information Systems
Office of Space Science

I. Prologue

Several NASA advisory boards have recommended that the most effective method to manage a data archive, is to co-locate it with active scientists who understand the contents and can provide expert advice and assistance. To this end NASA has established several discipline specific nodes for archiving astrophysics data, including in December 1990 the High Energy Astrophysics Science Archive Research Center (HEASARC) for X-ray and Gamma-ray astronomy data. This Memorandum of Understanding (MOU) redefines the roles of the National Space Science Data Center (NSSDC), the World Data Center A for Rockets and Satellites (WDCA R&S) and HEASARC in preserving and facilitating access to data from X-ray and Gamma-ray astronomy missions. It updates and supersedes a previous MOU concerning the NSSDC-HEASARC interaction made between the NSSDC and the Office of Guest Investigator Programs (November 17, 1992).

This revised MOU takes into account the lessons learned and computer technology advances over the past four years of HEASARC and NSSDC activities. The emergence of the Space Science Data System is also consistent with this MOU. It also uses as a guideline the report from the "Task Group on Science Data Management" (the Linsky report) which was produced in October 1996. In the original MOU the data management function (the physical archive) is the responsibility of the NSSDC, and the archive interface' the responsibility of the HEASARC. The dramatic reduction in the cost of mass storage, systems has made it possible for the HEASARC to provide a dedicated archive to support its users. This has enabled direct ftp access to the entire data holding, provided greater speed for data access and more flexibility in maintaining the archive. Simultaneously, the development of the world wide web has made it possible to integrate the data and services from various disciplines and remote data nodes in ways that were not anticipated when the HEASARC was created.

Over the past four years the HEASARC has effectively demonstrated that it can provide a high energy astrophysics node that effectively fulfills all the NASA requirements for archiving, maintaining and distributing data. This has led to some duplication in effort between the data services provided by the NSSDC and the HEASARC. This new MOU redefines the respective roles of the NSSDC and HEASARC. It transfers to the HEASARC full responsibility for all of NASA's high energy astrophysics archive and the community access to it. In this new arrangement the NSSDC (or any future organization so designated by NASA HQ) will hold a back-up copy of the HEASARC archive in a secure environment and provide broad data standards to ensure the HEASARC archive conforms to minimum NASA data standards.

II. Introduction

A. Purpose

The purpose of this document is to describe the roles of the NSSDC, the WDCA R&S and the HEASARC in acquiring, archiving, and distributing high energy astrophysics data (primarily consisting of X-ray and Gamma-ray astronomy data). This document sets the general constraints on the interfaces between the NSSDC, the WDCA R&S and the HEASARC. It supersedes a previous MOU between the NSSDC and the Office of Guest Investigator Programs (November 17, 1992). Additional implementation details will be specified in future documents which govern the details of the interface between NSSDC and the HEASARC.

B. Scope

The scope of this document covers areas of operation where there is joint responsibility between the HEASARC, NSSDC and WDCA R&S, or where, despite possibilities of ambiguity, there is sole responsibility of one or the other. Other documents exist which describe the full range of operations of each organization.

C. Audience

The following groups are the intended audience for the document:

1. The staffs of the NSSDC, WDCA R&S, and HEASARC.
2. The HEASARC, NSSDC and other NASA advisory groups.
3. Those who are preparing or restoring high energy astrophysics data such as Principal Investigators, and data restorers
4. Those who access high energy astrophysics data.
5. NASA Headquarters personnel who manage and provide the operating funds for the NSSDC and HEASARC.
6. Managers of institutions participating in NSSDC and HEASARC activities.
7. Members of other data centers.

D. Summary

The following are the key components of this MOU

(a) The HEASARC will serve as the point of entry for all high energy astrophysics (X-ray and Gamma-ray astronomy) data and software into the NASA archive environment.

(b) The NSSDC shall maintain and coordinate broad standards for data management and archiving which are appropriate to most or all disciplines and recommend, maintain and monitor implementation of minimum NASA data standards.

(c) The HEASARC shall be responsible for establishing and maintaining standards for data structures, data formats and software which are appropriate for use by the high energy astrophysics community and consistent with the broad data standards defined by the NSSDC.

(d) The HEASARC serves as the primary interface with the high energy astrophysics data producers, obtaining the data, analysis software, and documentation. The HEASARC will validate the delivered items for format, portability, completeness and content.

(e) The HEASARC will be the primary interface to the user community and will provide for all electronic distribution of the archival high energy astrophysics data sets, associated analysis software and documentation. The HEASARC will also distribute to its user community data and software on media such as CD ROM, DVD and-tape, where it is appropriate and cost effective to do so.

(f) The HEASARC will deliver at regular intervals a backup copy of the archive along with the associated catalogs and documentation. The NSSDC will maintain this backup copy as a "deep archive" to be used to restore the HEASARC primary archive in the event of a catastrophic loss of the primary archive. The back-up copy will be in a mutually agreed format, that will be primarily designed to rapidly restore, in a cost effective manner the HEASARC archive.

(g) The NSSDC will be the primary interface for bulk distributions of entire mission datasets and other large distributions. The NSSDC will provide distributions from its pre-HEASARC data holdings, restoration of old data tapes, and public/multidiscipline CD-ROM/DVD distribution.

(h) The NSSDC will maintain and operate the NASA master directory (which is to be inclusive of all NASA space science data, regardless of the discipline node where they reside). The HEASARC shall support the population of the master directory. In addition the HEASARC will maintain a catalog of its data holding and user interface to the online archive.

(i) The WDCA R&S will provide the auspices for international data requests from the HEASARC data holding (as currently defined for the NSSDC).

F. MOU Review Procedures

This MOU will be annually reviewed by the NSSDC, the WDCA R&S, the HEASARC and their advisory groups to determine each organization's compliance. Any statements in the MOU that conflict with the current policy and procedures will be identified. The reasons for these discrepancies will be reviewed and either the MOU modified, or the operations policies revised to resolve the discrepancy.

Any conflicts or discrepancies related to the roles of the NSSDC, WDCA R&S and HEASARC with respect to the MOU will be resolved by the directors of those organizations. In the event that resolution is not possible at this level, the matter will be elevated to NASA Headquarters, where the appropriate program sponsors will help resolve the issue.

III. Organizations

A. HEASARC

The HEASARC is the NASA discipline node that supports multi-mission X-ray and Gamma-ray astronomy archival research. Following the recommendation of several NASA advisory groups, the HEASARC was established in November 1990 in the Laboratory for High Energy Astrophysics at the GSFC with a charter to:

- maintain and disseminate data from previous and concurrent high energy astrophysics missions;
- provide software and data analysis support for these datasets;
- maintain and provide the necessary scientific and technical expertise for the processing and interpretation of the data holding;
- develop and maintain multi-mission analysis and support tools;
- provide catalogs of observations and ancillary information for the data holding; and

- coordinate data, software and media standards with other astrophysics sites.

B. NSSDC

Since 1965, the NSSDC has archived and disseminated data across the spectrum of Space Science and Earth Science disciplines. The NSSDC provides access to a wide variety of astrophysics, space physics, solar physics, lunar and planetary data from NASA space flight missions, in addition to selected other data and some models and software. NSSDC provides access to online information bases about NASA and non-NASA data at the NSSDC and elsewhere as well as the spacecraft and experiments that have or will provide public access data. NSSDC also provides information and support relative to data management standards and technologies.

C. WDCA R&S

The World Data Center system is an international confederation of data centers originally established in 1955 to support the exchange and archiving of data generated by the 1957-8 International Geophysical Year. WDCA R&S is hosted by NSSDC and has responsibilities for spacecraft launch and other descriptive information capture and dissemination. WDCA R&S serves as a conduit for requests from the international community for data from the NSSDC data archives.

IV. Project Interface

The HEASARC will serve as the point of entry for archival data from all NASA-sponsored satellite based high energy astrophysics missions and investigations (X-ray and Gamma-ray astronomy data). The HEASARC will maintain a mission interface which is responsible for negotiating the individual Project Data Management Plans (PDMPs) which satisfy the OSS Policy on Science Data Management in anticipation of HQ/SR signature of the PDMP's. The data, software and catalogs to be

delivered will be specified in the prelaunch PDMPs and may be augmented by agreement with HEASARC based on inflight experience. These data products will include raw science data, ancillary data, higher level data products, catalogs and relevant documentation. In all cases the HEASARC will have signature authority over the PDMP. All active mission data products that enter the NASA archive environment through the HEASARC mission interface will be validated and archived. Back-up copies will be passed onto NSSDC for safe keeping.

The HEASARC will be the archive entry point for all high energy astrophysics data restored from past missions. The NSSDC will continue to hold and distribute high energy astrophysics data from the pre-HEASARC era, until such time as the HEASARC, its user group and NASA headquarters decide otherwise.

The HEASARC and NSSDC will agree on and document the terms covering the transmission of archive products from HEASARC to NSSDC such as media type, volume and frequency of delivery. The HEASARC and NSSDC will also agree on the procedure for the restoration of the HEASARC archive in the event of a catastrophic failure.

The HEASARC will be responsible for the long term maintenance of the data archive, and its migration to new media before it becomes obsolete or deteriorates.

The HEASARC will support international requests for high energy astrophysics data in its holdings made to the WDCA (hosted by the NSSDC). Requests for data may be made directly to the HEASARC. The HEASARC will make a best effort to keep a record of all international data requests, and send a summary to the WDCA director if requested.

V. Catalog

The HEASARC will maintain a catalog of its data holding which allows users to conveniently locate and retrieve data. This catalog will contain links to data files in the physical archive

The NSSDC will maintain a multidisciplinary NASA master directory which will contain high level information about the HEASARC data holdings Appropriate high level information about the HEASARC data holdings will be sent to the NSSDC's master directory via the Directory Interchange Format (DIF). NSSDC will provide instructions to HEASARC in the appropriate use of the DIF format and will review HEASARC entries in the master directory. NSSDC will also maintain electronic links into the HEASARC catalogs from the master directory. HEASARC will provide spacecraft and experiment level information to the NSSDC Master Directory for new high energy astrophysics missions and experiments For each new data set sent to the NSSDC, HEASARC will also send a completed data set description template to be used by NSSDC in populating the NSSDC master catalog at the data set level.

The NSSDC shall coordinate the Master Directory design and evolution to meet mutual budget and system scope constraints. All changes made to the master directory design that will require HEASARC to develop and/or modify their operational system shall first be coordinated and agreed with HEASARC.

VI. Data

A. Data preparation

(i) Mission Data: As described in Section IV, all NASA high energy astrophysics data will enter the NASA archive environment through the HEASARC. The HEASARC will agree ahead of time with the relevant data provider on the products to be delivered, and these will be summarized in a PDMP. Upon delivery the HEASARC will verify the correct format, completeness, and continuity of the data and will monitor the validity and content of the incoming data. The HEASARC will also verify the completeness and contents of the associated catalog data and documentation.

(ii) Restored Data: The restoration of old high energy astrophysics data will be coordinated and overseen by the HEASARC. The priorities for data restoration will be set by the HEASARC users group and via proposals to NASA. As with mission data sets, the restored data sets will result in the update of the HEASARC mission catalogs and the relevant Master Directory entries.

(iii) International Data: Data from foreign missions may also be ingested into the HEASARC, as international agreements allow. The HEASARC will coordinate and cooperate with international agencies to ensure the data from foreign missions follow HEASARC standards and will seek to maximize the access to the complete world-wide high energy astrophysics dataset..

B. Data Storage

The final archive site for data from NASA sponsored high energy astrophysics missions will be at the HEASARC. In those cases where the mission is still actively modifying its data the physical archive may be remotely located at another site. In these cases the HEASARC catalog will contain entries and pointers to these data at the remote site. Every effort will be made to obtain a copy of the archive at the HEASARC as soon as is practical and cost effective.

The HEASARC will provide to the NSSDC a backup copy of the entire locally stored HEASARC data holding, software and associated catalogs. The NSSDC will hold this copy of the archive in a secure off-line environment, which will not be directly accessible by the outside community. The HEASARC will only send to the NSSDC back-up copies on a mission by mission basis of those data sets directly under its control. This backup copy will be optimized to be in a form that the HEASARC can use to quickly restore the primary archive in the event of a data loss. It will be the responsibility of the HEASARC to validate the correctness of these backup tapes before they are delivered to the NSSDC..

C. Data Distribution

The HEASARC will as resources and technology allows maintain its entire data holding online for immediate network access. Network distribution of the data will be the prime method of data dissemination from the HEASARC. Requests for moderate volumes of data on a distribution medium (tape, CD ROM, DVD) will be supported by the HEASARC on a case by case basis. Those cases where large volumes of data are requested, in particular an entire mission data set, the HEASARC may request the NSSDC to make a bulk copy from the backup tapes and send the copies to the requester. In these cases the NSSDC will apply its normal charging policy for each copy.

HEASARC will put the most popular datasets of a manageable size onto CD ROM, or DVD. When the initial procurement is made, the NSSDC will simultaneously procure a set of CDs. Both the NSSDC and HEASARC will fulfill requests for CD ROM or DVD media. The HEASARC will concentrate on distribution to its user community, the NSSDC requests from the general public, again following its normal charging policy.

VII. User support

The HEASARC has prime responsibility to support users of the high energy astrophysics archives. This will include providing expert advice on the contents and use of the data holding. The current NSSDC documentation (including web pages) that refer to high energy astrophysics data holdings will be modified to redirect users to the HEASARC services.