FINAL

IN-45-CR

O CIT

P. 7

FINAL REPORT OF WORK
PERFORMED UNDER NCC8-21
FOR THE PERIOD

May 26, 1993 - May 25, 1994

Submitted to:

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
MARSHALL SPACE FLIGHT CENTER, ALABAMA

Submitted by:

SCIENTIFIC INFORMATION SYSTEMS LABORATORY
COLLEGE OF SCIENCE
UNIVERSITY OF ALABAMA IN HUNTSVILLE
HUNTSVILLE, ALABAMA

Dr. Sara J. Graves
Principal Investigator

August 15, 1994

Final Report for
Work Performed Under Cooperative Agreement NCC8-21

May 26, 1993 to May 25, 1994

INTRODUCTION

UAH was awarded Cooperative Agreement NCC8-21 to perform the second year of research, originally awarded as three-year grant NAG8-904, entitled "A Data and Information System for Processing, Archival, and Distribution of Data for Global Change Research". The third year of this research project is to be covered under the IGCRE Cooperative Agreement. The focus of UAH's work on this project is on information management techniques for Marshall Space Flight Center's EOSDIS Version 0 Distributed Active Archive Center (DAAC). The centerpiece of this effort has been participation in EOSDIS Catalog Interoperability research, the result of which is a distributed Information Management System (IMS) allowing the user to query the inventories of all the DAACs from a single user interface. UAH has provided the MSFC DAAC database server for the distributed IMS, and has contributed to definition and development of the browse image display capabilities in the system's user interface. Another important area of research has been in generating value-based metadata through Data Mining. In addition, UAH has provided information management applications for local inventory and archive management, and for tracking data orders. Activities performed during the term of this cooperative agreement, are summarized below.

INFORMATION MANAGEMENT SYSTEMS

DISTRIBUTED IMS DEVELOPMENT -- SERVER

Research on the EOSDIS distributed IMS prototype has continued through the term of this agreement. MSFC DAAC is responsible for an inventory Server, which communicates with the user interface, or Client, through a well-defined message protocol. UAH's participation in this

prototyping effort during the past year has focused on enhancements to the MSFC IMS Server software necessary for the Alpha 1 release of the entire IMS distributed system, on October 12, 1993, and for the final testing period before operational release of the prototype system in July, 1994.

One of the biggest challenges of the IMS prototyping research effort has been the distributed nature of the project. UAH has participated in consensus decision making to agree on protocol changes, and has led discussions on standardizing descriptions of data packaging. UAH has also done extensive testing of the distributed system. Problems with the local Server are tracked and corrected, while Client and message protocol problems are reported to the distributed IMS development group. At the end of this cooperative agreement, the MSFC DAAC IMS Server was in good shape for a July release of the system. Continued development will be necessary, however, to correct any problems found in the final system test period, and to respond to user requests during the early release period.

DISTRIBUTED IMS DEVELOPMENT -- CLIENT

While the IMS Client is developed by a system-level team at GSFC, MSFC DAAC has played a lead role in development of the Client's browse facility. This feature allows the user to preview image data before ordering by browsing a reduced-resolution image. Under this cooperative agreement, UAH has designed and prototyped a method to attach a legend to each browse image. The design will allow either continuous or discrete legend data with labels to be described in a common text format. The Client will interpret this text to display the legend graphically.

DAAC-UNIQUE IMS DEVELOPMENT

UAH obtained a satellite orbit model and display software package from the University of Colorado. It plots actual orbits as well as orbit projections using satellite ephemeris data , such as the NORAD two-line element sets. This software be used as the basis for the planned orbit display feature of the DAAC's local IMS user interface. Under this cooperative agreement, a user interface for the orbit model display was developed, with the look and feel of the distributed IMS Client

DATA MINER

The version 1 Data Miner was completed under this cooperative agreement. This initial version provides limited functionality to prove the data mining concept. It will take and HDF file, extract a channel, and identify the data holes that exist. This miner code was integrated with IMS access code and jukebox access code to provide an end-to-end system that will take as input a dataset name (initially limited to SSM/I Pathfinder data) and time range. It will query the IMS for the file names, then proceed to request each file from the jukebox and mine it for holes. The version 1 miner's data structure, an equal degree grid, has the necessary capability to support the more complex mining tasks planned for the version 2 Data Miner. UAH plans to continue research in this area under IGCRE and other contracts.

DAAC METADATA

IMS inventory tables contain a variety of metadata describing the MSFC DAAC data. UAH has worked closely with MSFC DAAC Data Acquisition and Distribution System (DADS) personnel, as well as science data producers, to assure that the IMS inventory accurately describes current datasets, and to plan for metadata generation for future datasets. We have updated the ORACLE database with information on new datasets, and also contributed to the development of a series of User Guide documents for the MSFC DAAC.

DATA DICTIONARY

The distributed IMS Data Dictionary uses a common format to provide a description of the diverse metadata schema at each of the DAACs. UAH has played a lead role in some aspects of this distributed research effort. Under this cooperative agreement, UAH has continued participation in the data dictionary task, and has generated reports as required to reflect changes in IMS metadata database schema and content.

DADS DEVELOPMENT

JUKEBOX FILE LOCATION DATABASE

The MSFC DAAC archive uses an optical disk jukebox system for mass storage. The jukebox's file management system provides for storage and access of files, but requires the user to specify the disk as well as the file name, and for file access, the disk offset as well. UAH has developed a database application for storing the file location information. Read and write routines were developed to interact with the database, and subsequent enhancements were added. These routines are now used for all DAAC archive operations.

DATA ORDER TRACKING SYSTEM

UAH is developing an automated Data Order Tracking System (DOTS) for the MSFC DAAC. Under this cooperative agreement, the database schema for both user and order information was finalized. UAH transferred the data in an existing user information database to the order tracking database schema, and developed a prototype user interface, using Oracle's SQL*Forms tool, for DAAC User Services personnel. Design of a graphical user interface for this database, to be implemented using JAM, was begun.

IMS/DADS INTERFACE

Prototype development of interface software for the IMS and DOTS also continued. The interface automates the entry of the IMS product request into the database. While the initial design included logging the message, the user address, and each item ordered, enhancements to the design were made including a comparison check to determine if the user was currently listed in the user database. Another enhancement is a statistics counter, which tracks the number of times any specific dataset is ordered. Error handing features were also incorporated into the design; if an error occurs, the software will notify User Services of the type error that occurred during the logging of the order and forward the product request message for manual entry in the database. This software has been linked to the IMS Server for testing, to assure that all product requests received through the IMS will be correctly logged.

DATA ORGANIZATION AND ACCESS FOCUS TEAM

UAH is also involved in the transition of EOSDIS from the Version 0 prototype to Version 1, to be developed by the EOSDIS Core System (ECS) contractor. Several Focus Teams, composed of Version 0 developers and engineers, have been tasked with reviewing ECS requirements and designs, to assure that EOSDIS goals are met. UAH provides an IMS developer to the Data Organization and Access Focus Team (DOAFT). This task involves reviewing documents, participating in telecons, and attending system-wide meetings.

EQUIPMENT PURCHASES

Several equipment purchases were made under this cooperative agreement. A Macintosh Powerbook was purchased for shared use by all members of the development team. Also, a Silicon Graphics Indy workstation was purchased for off-site Data Miner development.

PRESENTATIONS AND DEMONSTRATIONS

Marilyn Drewry gave an overview presentation of the IMS, and its ordering capability to Congressional Staffers on September 17, 1993. The presentation also included an interactive, distributed demonstration of the search process and results interpretations, discussion and display of the browse capabilities, and access to the Global Change Master Directory (GCMD).

Similar presentations and demonstrations of the IMS distributed system were given for two local MSFC groups, including the MSFC Repository. These groups are interested in developing similar systems, based on our research.

MEETINGS, TRAINING, AND TRAVEL

In addition to informal working meetings with MSFC science and DAAC personnel, UAH has participated in weekly DAAC meetings, bi-weekly Earth Science Information Systems (ESIS) Configuration Control Board (CCB) meetings, and regular IMS, IMS Visual Aids, Data Dictionary, and DOAFT telecons.

Susan McCoy took advantage of several local MSFC training opportunities. She attended a two week (40 hour) Advanced C class, a one day (4 hour) Networks class, and a one week (20 hour) Operating Systems class.

Helen Conover attended the ECS System Requirements Review (SRR) and subsequent DOAFT meeting, which took place in Greenbelt, MD. The SRR was held at Goddard Space Flight Center on September 14 and 15, while the DOAFT was at the University of Maryland on September 16 and 17, 1993.

Evans Criswell and Susan McCoy attended a JAM training class in Wellesley, MA, September 27 - October 1, 1993. Subsequently, the JAM product was used for local user interface development at MSFC DAAC.

Marilyn Drewry attended the Consortium for International Earth Science Information Network (CIESIN) Second User's Workshop Meeting in Atlanta, October 11-13, 1993. CIESIN's focus is similar to that of the IMS, i.e., distributing Earth Science data to a wide range of users.

Helen Conover and Marilyn Drewry attended the IMS Development Team meeting at Langley Research Center, in Hampton, VA, on November 30 through December 2, 1993. At the meeting, they took the action to standardize data distribution media options presented to the user through the IMS.