The NASA STI Program ... in Profile

Since its founding, NASA has been dedicated to ensuring U.S. leadership in aeronautics and space science. The NASA Scientific and Technical Information (STI) Program plays an important part in helping NASA maintain its leadership role.

The NASA STI Program provides access to the NASA STI Database, the largest collection of aeronautical and space science STI in the world. The Program is also NASA's institutional mechanism for disseminating the results of its research and development activities.

A number of specialized services help round out the Program's diverse offerings, including creating custom thesauri, translating material to or from 34 foreign languages, building customized databases, organizing and publishing research results.

For more information about the NASA STI Program, you can:

- **Phone** the NASA Access Help Desk at (301) 621-0390
- **Fax** your question to NASA Access Help Desk at (301) 621-0134
- **Send us** your question via the **Internet** to help@sti.nasa.gov
- **Write to:**
  NASA Access Help Desk
  NASA Center for AeroSpace Information
  800 Elkridge Landing Road
  Linthicum Heights, MD 21090-2934
The first Applied Information Systems Research Program (AISRP) Workshop provided the impetus for several groups involved in information systems to review current activities. The objectives of the workshop included: (1) to provide an open forum for interaction and discussion of information systems; (2) to promote understanding by initiating a dialogue with the intended benefactors of the program, the scientific user community, and discuss options for improving their support; (3) create an advocacy in having science users and investigators of the program meet together and establish the basis for direction and growth; and (4) support the future of the program by building collaborations and interaction to encourage an investigator working group approach for conducting the program. For individual titles, see N93-29223 through N93-29241.
APPLIED INFORMATION SYSTEMS RESEARCH PROGRAM

WORKSHOP PROCEEDINGS

JULY 22-24, 1991

BOULDERADO HOTEL
2115 13th Street
BOULDER, COLORADO
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td>1</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>2</td>
</tr>
<tr>
<td>Workshop Overview</td>
<td></td>
</tr>
<tr>
<td>• Synopsis</td>
<td>3</td>
</tr>
<tr>
<td>• Follow-on Actions</td>
<td>5</td>
</tr>
<tr>
<td>Appendix</td>
<td></td>
</tr>
<tr>
<td>Agenda</td>
<td>A</td>
</tr>
<tr>
<td>Attendee List</td>
<td>B</td>
</tr>
<tr>
<td>Programmatic Overviews</td>
<td>C</td>
</tr>
<tr>
<td>Principal Investigator Presentations</td>
<td>D</td>
</tr>
<tr>
<td>Program Presentations</td>
<td>E</td>
</tr>
</tbody>
</table>
APPLIED INFORMATIONS SYSTEMS
RESEARCH WORKSHOP

PREFACE

The first Applied Information Systems Research Program (AISRP) Workshop provided the impetus for several groups involved in information systems to review current activities. Investigators representing fifteen of the twenty-two Office of Space Science and Applications (OSSA)/Information Systems Branch (ISB) NASA Research Announcement awards were in attendance. Attendees also included representatives from the Science and Technology divisions of NASA, directors of NASA's Centers of Excellence, specific research institutes, and members of the academic and remote sensing arena.

The objectives of the workshop are outlined:

• To provide an open forum for interaction and discussion of information systems research activities.

• To promote understanding by initiating a dialogue with the intended benefactors of the program, the scientific user community, and discuss options for improving their support.

• Create an advocacy in having science users and investigators of the program meet together and establish the basis for direction and growth.

• Support the future of the program by building collaborations and interaction to encourage an investigator working group approach for conducting the program.
APPLIED INFORMATIONS SYSTEMS
RESEARCH WORKSHOP

ACKNOWLEDGEMENTS

We wish to gratefully recognize the following persons for the skillful support given towards the success of this first Applied Information Systems Research Program (AISRP) workshop. The capable efforts and contribution of each individual are greatly appreciated. Thank you.

- Elaine Hanson/Colorado Space Grant Consortium, University of Colorado, for hosting the workshop
- Susan Solari/Colorado Space Grant Consortium, University of Colorado, for hosting the workshop and making local arrangements
- Karen Friedman/National Center for Atmospheric Research (NCAR), for giving assistance with local arrangements, especially for the workshop reception held at NCAR
- Anne Novomy/NASA HQ, Code SM, and Leslie Emerson/Center for Space and Advanced Technology (CSAT) for providing registration and on-site logistical support
The workshop proceeded according to the agenda provide herein, (see Appendix, attachment A). The first day was given to programmatic presentations. Joseph Bredekamp/NASA HQ, ISB, provided the overall context for the program. Representatives of OSSA Science Disciplines and of the Office of Aeronautics, Exploration and Technology reviewed current as well as planned data and information systems activities. Presentations were also made on the Center of Excellence in Space Data and Information Sciences (CESDIS) and the Center for Aerospace and Space Information Sciences (CASIS). These Centers work to encourage collaborations between government, academia, and industry.

Following this, presentations were made by the participating principal investigators of the Applied Information Systems Research Program. In attendance to these presentations were scientists, software developers, program managers, technologists and computer systems personnel. This diversity of audience allowed for evaluation of the broad issues regarding use, development, and maintenance in information systems.

Presentations topics included visualization and associated analysis, data management including distributed databases, and software tools for modeling. Several presentations highlighted remote sensing, software development, and high performance computing. Many of the activities reported on are based on existing and commercial products, but all included advances in the field.

The final day of the workshop was devoted to discussions centered toward three key topics; technology transfer options, technical issues and future interaction.

Technology Transfer Options

This discussion yielded three different directions for information diffusion, specifically across disciplines within OSSA, from OAET to OSSA, and across the broader science community. Issues and options for each direction area differed.

Across disciplines within OSSA, consensus was made that the at-large science community needs to be made aware of the tools and techniques under development for the Applied Information Systems Research program. Several attendees suggested that the transfer activity be made a formal, explicit part of the program. Some encouraged OSSA to provide the infrastructure to support and disseminate results. Others suggested that the technology transfer responsibility for results be made part of the NRA award.
OSSA to OAET

Workshop attendees agreed that a closer interaction between the two organizations was needed, but the mechanism was unclear. Attendees agreed that OSSA should be the 'implementing' code.

Broader Science Community

All agreed that outreach is important. There was also agreement that many options are already available, and need to be fully utilized. The group considered several other options. It was suggested that the Computer Software Management Information Center (COSMIC) at the University of Georgia be considered a candidate for functioning as the software distribution mechanism for the agency.

Discussions of technical issues emphasized a variety of interchange issues. Further discussion of these topics and development of additional topics, was delegated to future splinter group discussions. Overall, the discussion emphasized the need to consider user requirements, both current and future, during the tool development.
FOLLOW-ON ACTIONS

Two near-term actions were identified. The first is to poll investigators to identify current computing platforms being used, preferred mode of operation, etc. The second action is to establish a set of test dataset suites, to help compare and evaluate effectiveness of various software tools.

Since this was the first meeting of the Applied Information Systems Research activity, many of the continuing logistics aspects were discussed as well. The preferred format seemed to be an annual meeting, with 'splinter group' sessions for specific topics or subsets of the group to be included at the same meeting. As additional topics warrant, interim 'mini workshops' should be convened as well.

Future meetings will continue to include both the science and project communities, as well as investigators in the program. Future meeting sites will continue to require infrastructure for demonstrations of tools (workstations, network access, etc.) In the interim, a network bulletin board will be established for rapid and timely communications.
APPLIED INFORMATION SYSTEMS RESEARCH WORKSHOP

JULY 22-24, 1991

BOULDERADO HOTEL
Boulder, Colorado

AGENDA

MONDAY, JULY 22, 1991

11:30 a.m. Registration

1:00 p.m. Welcome and Logistics  Elaine Hansen

1:15 p.m. Program Overview  Joe Bredekamp

2:00 p.m. Office of Space Science and Applications  Randy Davis
            Data Environment and Future Plans
            -Earth Observing System Data  Martin Ruzek
            and Information System
            -Planetary Data System
            -Astrophysics Data Systems  Alice Bertini

4:00 p.m. Office of Aeronautics, Exploration and  Paul Hunter
           Technology Information Systems Program

4:20 p.m. Center of Excellence in Space Data and  Ray Miller
           Information Sciences (CESDIS)

4:40 p.m. Center for Aerospace and Space  Mike Flynn
           Information Sciences (CASIS)

5:00 p.m. Adjourn

PRECEDING PAGE BLANK NOT FILMED
APPLIED INFORMATION SYSTEMS RESEARCH WORKSHOP

JULY 22-24, 1991

BOULDERADO HOTEL
Boulder, Colorado

AGENDA

TUESDAY, JULY 23, 1991

8:30 a.m.     Principal Investigator Presentations
12:00 p.m.    Lunch
1:00 p.m.     Principal Investigator Presentations
5:30 p.m.     Reception at the National Center for Atmospheric Research

WEDNESDAY, JULY 24, 1991

8:30 a.m.     Flight Project Office Information Systems Testbed (FIST)  Patricia Liggett
8:50 a.m.     Intelligent Data Management       Bill Campbell
9:15 a.m.     Open Discussion

-Where we go from here to work together on program: potential collaborations, subgroups, future workshops, etc.

-Technology Transfer and Infusion

12:00 p.m.    Adjourn
**APPLIED INFORMATION SYSTEMS RESEARCH WORKSHOP**

**P.I. PRESENTATIONS**  
(Morning Session)

<table>
<thead>
<tr>
<th>Title</th>
<th>Presenter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment's Laboratory for Visualized Interactive Science</td>
<td>Ms. E. Hansen</td>
</tr>
<tr>
<td>Grid Analysis and Display System (GRADS): A Practical Tool for Earth Science Visualization</td>
<td>Dr. J. Kinter</td>
</tr>
<tr>
<td>A Distributed System for Visualizing and Analyzing Multivariate and Multidisciplinary Data</td>
<td>Dr. A. Jacobson</td>
</tr>
<tr>
<td>Advanced Data Visualization and Sensor Fusion: Conversion of Techniques from Medical Imaging to Earth Science</td>
<td>Dr. R. Savage</td>
</tr>
<tr>
<td>Development of an Expert Data Reduction Assistant</td>
<td>Dr. G. Miller</td>
</tr>
<tr>
<td>System of Experts for Intelligent Data Management (SEIDAM)</td>
<td>Dr. D. Goodenough</td>
</tr>
<tr>
<td>Construction of an Advanced Software Tool for Planetary Atmospheric Modeling</td>
<td>Dr. R. Keller</td>
</tr>
<tr>
<td>Knowledge-based Assistance for Science and Analysis Using Large Distributed Databases</td>
<td>Mr. T. Handley, Jr.</td>
</tr>
<tr>
<td>Multi-Layer Holographic Bifurcative Neural Network Systems for Real-Time Adaptive EOS Data Analysis</td>
<td>Dr. Hua-Kuang Liu</td>
</tr>
</tbody>
</table>

**LEGEND:** * P.I. Presentation not given/ ** P.I. not in attendance
<table>
<thead>
<tr>
<th>P.I. PRESENTATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Afternoon Session)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Title</th>
<th>Presenter</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Distributed Analysis and Visualization System for Model and Observational Data</td>
<td>Mr. M. Arrott</td>
<td>Univ. of Illinois</td>
</tr>
<tr>
<td>An Interactive Environment for the Analysis of Large Earth Observation and Model Data Sets</td>
<td>Professor K. Bowman</td>
<td>*</td>
</tr>
<tr>
<td>A Land-Surface Testbed for EOSDIS</td>
<td>Dr. W. Emery</td>
<td>Univ. of Colorado</td>
</tr>
<tr>
<td>Geographical Information System for Fusion and Analysis of High-Performance Remote Sensing and Ground Truth Data</td>
<td>Mr. A. Freeman</td>
<td>JPL</td>
</tr>
<tr>
<td>Development of a Tool-Set for Simultaneous, Multi-Site Observations of Astronomical Objects</td>
<td>Dr. Chakrabarti</td>
<td>*</td>
</tr>
<tr>
<td>SAVS: A Space Analysis and Visualization System</td>
<td>Dr. E. Szuszczewicz</td>
<td>Science Applications International Corp.</td>
</tr>
<tr>
<td>Planetary Data Analysis and Display System: A Version of PC-MclADS</td>
<td>Dr. S. Limaye</td>
<td>Univ. of Wisc., Madison</td>
</tr>
<tr>
<td>Interactive Interface for NCAR Graphics</td>
<td>Mr. R. Lackman</td>
<td>National Center for Atmospheric Research</td>
</tr>
<tr>
<td>The Development of Generic and Extensible Software to Support the Study of Space Science Data</td>
<td>Mr. G. Goucher</td>
<td>*</td>
</tr>
<tr>
<td>VIEWCACHE: An Incremental Database Access for Autonomous Interoperable Databases</td>
<td>Assoc. Prof. Nick Roussopoulos</td>
<td>Univ. of MD</td>
</tr>
<tr>
<td>A Spatial Analysis and Modeling System for Environmental Management</td>
<td>Mr. C. Vermillion</td>
<td>NASA/GSFC</td>
</tr>
<tr>
<td>Introduction to CADET: Center for Advanced Data Evaluation Technology</td>
<td>Ms. Cathy Schulbach</td>
<td>NASA/ARC</td>
</tr>
<tr>
<td>Topography from Shading and Stereo</td>
<td>Professor B. Horn</td>
<td>*</td>
</tr>
<tr>
<td>Multivariate Statistical Analysis Software Technologies for Astrophysical Research Involving Large Data</td>
<td>Mr. G. Djorgovsici</td>
<td>JPL</td>
</tr>
<tr>
<td>High Performance Compression of Science Data</td>
<td>Dr. J. Storer</td>
<td>*</td>
</tr>
</tbody>
</table>

**LEGEND:**  * P.I. Presentation not given/ ** P.I. not in attendance
# ATTENDEE LIST

<table>
<thead>
<tr>
<th>Name</th>
<th>Electronic Address</th>
<th>Affiliation</th>
<th>Phone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matthew Arrott</td>
<td><a href="mailto:arrott@ncsa.uiuc.edu">arrott@ncsa.uiuc.edu</a></td>
<td>NCSA</td>
<td>(217) 244-6829</td>
</tr>
<tr>
<td>Alice Bertini</td>
<td><a href="mailto:bertini@coma.colorado.edu">bertini@coma.colorado.edu</a></td>
<td>CU/CASA</td>
<td>(303) 492-6422</td>
</tr>
<tr>
<td>William S. Boyd</td>
<td><a href="mailto:boyd@mmm.ucar.edu">boyd@mmm.ucar.edu</a></td>
<td>NCAR</td>
<td>(303) 497-8159</td>
</tr>
<tr>
<td>Joe Bredekamp</td>
<td><a href="mailto:jbredekamp@nasamail.nasa.gov">jbredekamp@nasamail.nasa.gov</a></td>
<td>NASA</td>
<td>(202) 453-1506</td>
</tr>
<tr>
<td>Bill Campbell</td>
<td><a href="mailto:campbell@nssc.gsfc.nasa.gov">campbell@nssc.gsfc.nasa.gov</a></td>
<td>NASA-GSFC</td>
<td>(301) 286-2785</td>
</tr>
<tr>
<td>Randal Davis</td>
<td>aquila::rdavis@nasamail</td>
<td>Univ. of CO</td>
<td>(303) 492-6867</td>
</tr>
<tr>
<td>Gitta Domik</td>
<td>domik%<a href="mailto:cosmos@vaxf.colorado.edu">cosmos@vaxf.colorado.edu</a></td>
<td>CU/CASA</td>
<td>(303) 492-4062</td>
</tr>
<tr>
<td>Bill Emery</td>
<td><a href="mailto:emery@orbit.colorado.edu">emery@orbit.colorado.edu</a></td>
<td>Univ. of CO</td>
<td>(303) 492-8591</td>
</tr>
<tr>
<td>M. J. Flynn</td>
<td><a href="mailto:m@sierra.stanford.edu">m@sierra.stanford.edu</a></td>
<td>Stanford</td>
<td>(415) 723-1450</td>
</tr>
<tr>
<td>Anthony Freeman</td>
<td><a href="mailto:freeman@jplmsrs.jpl.nasa.gov">freeman@jplmsrs.jpl.nasa.gov</a></td>
<td>JPL</td>
<td>(818) 354-1887</td>
</tr>
<tr>
<td>David Goodenough</td>
<td><a href="mailto:dgg@ccs.cdnl">dgg@ccs.cdnl</a></td>
<td>CCRS</td>
<td>(613) 952-2760</td>
</tr>
<tr>
<td>Charles Goodrich</td>
<td><a href="mailto:cgg@astro.umd.edu">cgg@astro.umd.edu</a></td>
<td>Univ. of MD CP</td>
<td>(301) 405-1516</td>
</tr>
<tr>
<td>George Hagedorn</td>
<td><a href="mailto:hagedorn@hugo.gsfc.nasa.gov">hagedorn@hugo.gsfc.nasa.gov</a></td>
<td>GSFC</td>
<td>(301) 286-4058</td>
</tr>
<tr>
<td>Tom Handley</td>
<td><a href="mailto:thandley@spacemouse.jpl.nasa.gov">thandley@spacemouse.jpl.nasa.gov</a></td>
<td>JPL</td>
<td>(818) 354-7009</td>
</tr>
<tr>
<td>E. Hansen</td>
<td>ehansen@nasamail</td>
<td>Univ. of CO</td>
<td>(303) 492-3141</td>
</tr>
<tr>
<td>Paul Hunter</td>
<td>phunter%<a href="mailto:nasamail@ames.arc.nasa.gov">nasamail@ames.arc.nasa.gov</a></td>
<td>NASA/HQ</td>
<td>(202) 453-2704</td>
</tr>
<tr>
<td>Allan Jacobson</td>
<td><a href="mailto:budj@scnl.jpl.nasa.gov">budj@scnl.jpl.nasa.gov</a></td>
<td>JPL</td>
<td>(818) 354-0693</td>
</tr>
<tr>
<td>Paul D. Jensen</td>
<td><a href="mailto:pdj@vap.span.nasa.gov">pdj@vap.span.nasa.gov</a></td>
<td>Consultant</td>
<td>(303) 449-7029</td>
</tr>
<tr>
<td>Chuck Jorgensen</td>
<td><a href="mailto:jorgensen@pluto.arc.nasa.gov">jorgensen@pluto.arc.nasa.gov</a></td>
<td>NASA</td>
<td>(415) 604-6725</td>
</tr>
<tr>
<td>Rich Keller</td>
<td><a href="mailto:keller@ptolemy.arc.nasa.gov">keller@ptolemy.arc.nasa.gov</a></td>
<td>NASA-ARC</td>
<td>(415) 604-3388</td>
</tr>
<tr>
<td>Jim Kinter</td>
<td><a href="mailto:kinter@umbc.umd.edu">kinter@umbc.umd.edu</a></td>
<td>Univ. of MD BC</td>
<td>(301) 405-5384</td>
</tr>
<tr>
<td>Allison Kipple</td>
<td><a href="mailto:kipple@audra.colorado.edu">kipple@audra.colorado.edu</a></td>
<td>Univ. of CO</td>
<td>(303) 440-6091</td>
</tr>
<tr>
<td>Margi Klemp</td>
<td><a href="mailto:margi@rana.colorado.edu">margi@rana.colorado.edu</a></td>
<td>CU/LASP</td>
<td>(303) 492-7289</td>
</tr>
<tr>
<td>Joseph B. Kelemp</td>
<td><a href="mailto:klem@ncar.ucar.edu">klem@ncar.ucar.edu</a></td>
<td>NCAR</td>
<td>(303) 497-8902</td>
</tr>
<tr>
<td>Steve Koch</td>
<td><a href="mailto:koch@vap.span.nasa.gov">koch@vap.span.nasa.gov</a></td>
<td>NASA-GSFC</td>
<td>(301) 286-7188</td>
</tr>
<tr>
<td>Robert Lackman</td>
<td><a href="mailto:rll@ncar.ucar.edu">rll@ncar.ucar.edu</a></td>
<td>NCAR</td>
<td>(303) 497-1224</td>
</tr>
<tr>
<td>Pat Liggett</td>
<td><a href="mailto:pat@puente.jpl.nasa.gov">pat@puente.jpl.nasa.gov</a></td>
<td>JPL</td>
<td>(818) 354-1414</td>
</tr>
<tr>
<td>S. S. Limaye</td>
<td>madraf::limaye</td>
<td>Univ. of Wisc,</td>
<td>(608) 262-9541</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Madison</td>
<td></td>
</tr>
<tr>
<td>H.K. Liu</td>
<td><a href="mailto:hkliru@nieron1.jpl.nasa.gov">hkliru@nieron1.jpl.nasa.gov</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vance McCullough</td>
<td><a href="mailto:miller@stsu.edu">miller@stsu.edu</a></td>
<td>Hughes</td>
<td>(303) 344-6146</td>
</tr>
<tr>
<td>Glenn Miller</td>
<td><a href="mailto:miller@cs.umd.edu">miller@cs.umd.edu</a></td>
<td>Space Telescope</td>
<td>(301) 338-4738</td>
</tr>
<tr>
<td>Raymond E. Miller</td>
<td><a href="mailto:miller@cs.umd.edu">miller@cs.umd.edu</a></td>
<td>CESDIS</td>
<td>(301) 405-2675</td>
</tr>
<tr>
<td>Lynne Norikane</td>
<td><a href="mailto:norikane@jplmsrs.jpl.nasa.gov">norikane@jplmsrs.jpl.nasa.gov</a></td>
<td>JPL</td>
<td>(818) 354-1235</td>
</tr>
<tr>
<td>Martin N. Orton</td>
<td></td>
<td>JPL</td>
<td>(818) 354-0728</td>
</tr>
<tr>
<td>Michael Robson</td>
<td><a href="mailto:robson@ccrs.cdn">robson@ccrs.cdn</a></td>
<td>CCRS</td>
<td>(613) 952-2759</td>
</tr>
<tr>
<td>Paul Rotar</td>
<td><a href="mailto:rotar@ncar.ucar.edu">rotar@ncar.ucar.edu</a></td>
<td>NCAR</td>
<td>(303) 497-1277</td>
</tr>
<tr>
<td>Martin Ruzek</td>
<td>omnet/m.ruzek</td>
<td>NASA</td>
<td>(202) 453-8137</td>
</tr>
<tr>
<td></td>
<td>m.ruzek@nasamail</td>
<td>Hughes</td>
<td>(303) 344-6176</td>
</tr>
<tr>
<td>R.C. Savage</td>
<td>mcl::szusz</td>
<td>JPL</td>
<td>(818) 354-3467</td>
</tr>
<tr>
<td>E.P. Szuszewicz</td>
<td></td>
<td>JPL</td>
<td>(818) 354-3467</td>
</tr>
<tr>
<td>Amy Walton</td>
<td>awalton@nasamail</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPLIED INFORMATION SYSTEMS
RESEARCH PROGRAM WORKSHOP

PROGRAMMATIC OVERVIEWS
APPLIED INFORMATION SYSTEMS
RESEARCH PROGRAM WORKSHOP

Joe Bredekamp
NASA
CODE SMI
Washington D.C.
(202) 453-1505
NOTES:

THIS PAGE LEFT INTENTIONALLY BLANK
Applied Information Systems Research Program

Workshop Objectives

- Exchange information on status and plans of research efforts
- Initiate dialog with OSSA science disciplines and other related NASA efforts
- Establish basis for conducting program
  - Means to facilitate communications and collaborations
  - Advocacy building
  - Technology transfer and infusion
  - Future directions

Office of Space Science and Applications

Information Systems Strategy

- Science discipline divisions provide primary focus for discipline-specific data management systems that integrate project data plans and on-going research needs into total research capability
- Provide robust, multi-discipline infrastructure
  - Architecture, policies, standards, practices, etc. to promote interoperability and resource sharing
  - Access to high performance computing
  - Network services
  - Information services (directories, catalogues, etc.)
- Apply and exploit advances in technology to evolve and enhance systems capabilities

Applied Information Systems Research Program

Workshop

Boulderado Hotel
Boulder, Colorado

July 22-24, 1991

Joe Bradekamp
Information Systems Research and Technology
and Systems Evolution

**Objective:** Apply advanced information systems technology as appropriate
to improve support to OSSA science programs

**Elements:**
- Applied research (tools and capabilities, etc.)
- Testbed demonstrations and insertions

**Approach:**
- Broaden participation through open solicitations
- Provide as part of infrastructure capability to support testbeds,
rapid prototyping, etc.
- Leverage with related efforts in NASA/OAET; NSF, etc.

**Status:**
- Applied research program initiated (you are here!) and well-received
- Next solicitation targeted for Summer 1992

**Issues/Opportunities:**
- Need to develop systematic process for technology transfer and
  infusion
- Sharpen research agenda to focus future solicitations
The Earth Observing System Data and Information System (EOSDIS)

Applied Information Systems Research Workshop
Boulder, Colorado

Sara J. Graves
July 22-24, 1991

EOSDIS and Earth System Science

EOSDIS is a pivotal part of the U.S. Global Change Research Program and the international effort to understand how the Earth functions as a complete system.

Earth system science objectives require a data and information system that will:
- Encourage multidisciplinary and interdisciplinary investigations
- Combine data from the EOS platforms with data from other agencies and nations (other satellites, aircraft, in situ operational and experimental data)
- Integrate EOS information with models of environmental processes and global change

Major System Drivers of EOSDIS

| Mission life | 15+ years |
| International partners | Japan, European Space Agency (ESA), and Canada |
| Interdisciplinary research | Climate, hydrologic system, biogeochemistry |
| Multidisciplinary research | Land, oceans, atmosphere |
| Multiple instruments | MODIS, HRIS, ITR, SAFIRE... |
| Multiple platforms | EOS-A, EOS-B, EPOP-M, JP01, EPOP-N |
| Distributed system architecture | GSFC, EDC, LARC, JPL, ASF, MSFC, NSIDC, investigator sites |
| Large number of users | >500 AO scientists and thousands of users |
| Massive data volume | 50,000 Tbytes |
| Massive data processing | As much as 60 Mbps input, 500 Mbps output |
| Metadata | Comprehensive directory, catalogs, inventories, and browse products |

EOSDIS Major Functions

- Mission planning, scheduling, and control
- Instrument planning, scheduling, and control
- Effective resource management
- Communications
- Computational facilities at Investigator sites to support research
- Production of standard data products
- Production of special data products
- Archiving and distribution of data and research results
**EOSDIS Strategy**

- Work with the users to meet their needs
- Make data easily accessible for research users
- Begin development immediately
- Use lessons learned from current efforts
- Build on existing infrastructure
- Utilize open, distributed architecture
- Evolve with advances in technology and Earth science

**Early EOSDIS Objectives**

- Provide increased utility to scientists
- Improve access to existing data sets
- Produce new research-quality geophysical products
- Provide better metadata/data services
- Maintain existing services
- Build EOSDIS infrastructure/unifying capabilities
- Build working relationship between DAACs, Project, and users
- Implement first communications links, standards, etc.
- Prototype/test/evaluate DIS elements and standards
- Produce results to guide continuing development
- Adopt Earth system science vs. "Earth sciences" view

**EOSDIS Research Users**

Research users must abide by the stipulations outlined in a "Research Agreement":

- Publish in the open literature results of research based in whole or in part on data obtained from EOS
- Make available to the research community the derived data, algorithms, and models at time of acceptance for publication
- Data used only for the researcher's bona fide research purposes
- Data may be copied and shared among other researchers provided that they are covered by a Research Agreement or the researcher who obtained the data is willing to take responsibility for compliance

**EOSDIS Design Implications**

EOSDIS is one system even though distributed
- DAACs are the institutionalization of EOSDIS
- Each new step should advance us towards the Integrated EOSDIS
- The EOSDIS users are the researchers, not the general public
- Results as well as data must be archived
- System must be evolvable to survive
- Meeting user needs must be paramount
- Respond to DAAC User Working Group
- The payoff is in improved efficiency/effectiveness of the research

**EOSDIS**
**What is the EOSDIS Version 0?**

Version 0 is the first manifestation of EOS that anyone will see
Version 0 is the working prototype of NASA's Earth science data system
Version 0 is SCFs, DAACs, networks, and IMS
Version 0 is populated by:
- Existing data
- Pathfinders
- In situ, aircraft, and field campaign measurements
- Precursor missions
- Modeling and other research results

**Why is EOSDIS a Distributed System?**

- Recognizes the distributed nature of Earth science community
- Recognizes the existing distribution of assets and capabilities
- Provides for diversity of styles in data use
- Provides for multiple prototypes
- Ensures capability to interface with external systems

**Early EOSDIS Pathfinder Data Sets**

- For existing data important for global change
- Level 1 data to active working storage/archive
- Scientists (academic, agency) develop/select community-consensus algorithm/products
- Generate and validate the derived products
- Make data and products accessible through Distributed Active Archive Centers (DAACs) and metadata/browse services by Information Management System (IMS)
- Working with NOAA, start with four data sets
- AVHRR—SST and vegetation products
- GOES—Products TBD
- TOVS—Vertical profile products
- SSMI—Sea ice, precipitation, etc.
- Others TBD—Science needs to drive selection
Version 0 - Version I

Time scale of Version 0 is from now until
Our needs require Version 1
Our means are adequate to implement Version 1
Transition from Version 0 to Version 1 should feel like a continuous activity
EOSDIS is an evolutionary system
Hardware and software will be coming and going forever
Service to the community is preserved by being institutionalized
Continuity of DAACs and network
Meet user expectations, listen to user advice, and obtain user support
AN OVERVIEW OF THE PLANETARY DATA SYSTEM

Randy Davis
Univ. Of Colorado
Boulder, Co
(303) 492-6867
An Overview of the Planetary Data System

Old-Style Planetary Data Delivery

Modern Planetary Data Access
PDS Project Organization

- **Project Manager**: S. McMahon, JPL
- **Science Manager**: M. Maran, JPL
- **Baseline Manager**: D. Maran, JPL

**Data Engineering Team**
- **Manager**: Y. Fletcher, JPL

**Discipline Nodes**
- **System Engineering Team**
- **Software Team**

The PDS Discipline Nodes

- **Atmospheres**
  - University of Colorado
  - Geosciences
    - Washington University
  - Small Bodies
    - University of Maryland
  - Imaging
    - U.S. Geological Survey
- **Plasma/Particles**
  - UCLA
  - Rings
  - Ames Research Center
- **Navigation and Ancillary Information**
  - JPL
  - Jet Propulsion Lab

Sample Discipline Node (Atmospheres)

- **U. of Colorado**
  - Planetary Atmospheres Database and air-quality modeling
  - Standard Space Flight Center
    - Infrared astronomy and spectroscopy
  - Jet Propulsion Laboratory
    - Spacing spectroscopy
  - U. of Arizona
    - Ultraviolet spectroscopy
  - How Mexico State U.
    - Imaging of atmosphere and long-term climate data

How Users Access the PDS

1. Users can access HSSDC and PDS data holdings through the HSSDC Home Directory
2. Users can access the HSSDC central node using the PCN located on the desktop home for HSSDC
3. Users can access the PDS central node through the HSSDC Home Directory
4. Users can access HSSDC and the PDS central node through the Discipline Node
Standards Being Embraced or Examined

- Operating Systems
  - VMS
  - Unix
  - Macintosh OS
  - MS DOS

- Networks
  - DECNET
  - TCP/IP

- Database Management Systems
  - Commercial RDBMSs
  - Homegrown Systems

- Data Analysis Environment
  - IDL
  - Fortran
  - C

- Programming Languages
  - Fortran

Concluding Remarks

- Thanks to CD-ROM, PDS is moving from distributed databases to distributable databases.
- Planetary CD-ROM disks are available for use in your work.
- The common denominator computer hardware configuration for planetary science is changing:
  - From: A VAX and a VT-100
  - To: Workstations with sophisticated graphics capabilities
- Macs and MS DOS machines may become more important since good tools are available on these platforms.
- The planetary community is striving for consistent and robust information interfaces between scientists, flight project systems, and the discipline archives (PDS).
  - We in the PDS welcome discussions on how to make your research more useful and available to planetary scientists.
  - PDS is a working system, with stringent quality and configuration control, but good new technology will always be considered.

- Object Description Language (ODL) for identifying and describing data
- Limited use of CCSDS Standard Format Data Units (SFDUs)
- ISO 8859 CD-ROM for importing/exporting large volumes of data.
- User Interface
  - X Windows for basic user interface functions
  - Motif for look and feel

ADS PROJECT REVIEW
INTRODUCTION & ADS OVERVIEW

Alice Bertini
CASA
Univ. Of Colorado
Boulder, Co
(303) 492-6422
ADS PROJECT REVIEW

Introduction & ADS Overview

Alice Berliti
CASA/University of Colorado

ADS Project Review

Introduction

What is the Astrophysics Data System?

- Response to Needs Recognized by Various Studies
  - CODMAC Reports I (1982) and II (1985)
  - Astrophysics Data Systems Study (1988)
  - Information Systems Strategic Planning Project (1990)

- Common Themes of Studies
  - Large influx of Data: 10's Gigabytes to 10's Terabytes
  - Knowledge of Data Holdings: What? Where?
  - Access to Data and Meta Data: Catalogs, Data Products, Archives, Documentation
  - Applicable into 21st Century: Flexible, Extensible, Evolvable
  - Science Driven System: Embrace or Enable Science Research e.g., "Fan Chromatic"

ADS Project Review

Status of System Now

- Operational System
  - 100 Registered Users
  - 16 Databases from 8 nodes
  - 600 Queries 1st Month
  - 30,000 Records 1st Month

- Transition from Development to Operations
  - Internal Re-organization of Project in Progress
  - Emphasis on Operation of System
  - Project Plan Based on Providing Science Services to Users
Astrophysics Data System Node Locations

**ADS PROJECT PLAN**

- **IPAC IRAS**
- **IPAC Account**
- **IPAC Software**
- **CASA Account**
- **CASA Software**
- **PSJ Account**
- **PSJ Software**
- **SDO Account**
- **SDO Software**
- **FISDEC Account**
- **FISDEC Software**
- **Einstein**

**TRADITIONAL ACCESS METHOD**

- Many Accounts
- Many Different Software Packages to Learn
- No Means of Integrating Data

**ADS PROJECT PLAN**

- **IPAC IRAS**
- **Einstein**
- **ADS Catalog**
- **ADS Access Method**
- **SDO Catalog**
- **FISDEC Account**
- **FISDEC Software**
ADS Project Review

Top Level Organization Chart

Major Milestones of ADS

- Astrophysics Data System Workshops: 1987, Report 1988
  Working Group Formed 1988
- Internal Alpha Release 1 - Q2 1989
  User Interface Prototype
- Internal Alpha Release 2 - Q4 1989
  Distributed Data Access Prototype
- External Data Release 1 - Q2 1990
  Distributed Data Base Access via MPS
  Prototype Command Language
- External Data Release 2 - Q1 1991
  Database Server
  Host Environment
  Operational Testing
- Operational System Release 1 - Q2 1991
  16 Catalogs On-Line with Documentation
  100 Registered Users

Science Objectives of ADS

- Provide Information on Existence and Location of Data Holdings
- Provide Access to Data Holdings
- Provide Sufficient On-Line Information on Data for Scientific Use
- Provide Tools for Selection, Retrieval, and Manipulation of Data
- Provide Tools for Information Retrieval (Textual Data Access)
- Provide Access to Data Processing Services
- Provide Tools for Data Processing, Visualization

ADS Project Review

Overview

Scientific Based Scenarios Imply Functional Requirements

- Cross Correlation of Catalog Data
  Coordinate Conversion
  Sky Finding, Distance Test (Search in Cone)
  Combining Data
  Export Results

- Location of Data/Objects
  Name Resolution
  Browse Data Base
  Order Data Set
  Immediate Transmission
  Visualization

- Data Processing (Examples)
  Source Detection with Custom Parameters
  Deconvolution of Images
  Spectral Fitting
  Flat Fielding
ADS Project Review

Overview

Scientific Based Sensors Imply Functional Requirements

- Simple User Scenario
  What sources are seen in both X-ray and IR spectra observations?
  Is there a correlation of X-ray and IR spectra for a class of objects?

- Actions
  Get X-ray sources over some range of parameters, e.g., RA, DEC, Class
  Get IR sources in same range of parameters
  Compare source lists
  Plot flux vs flux

- Functions
  Access Remote Databases
  Convert Coordinates
  Match Objects by Position
  Extract Interesting Measurement Data
  Plot Results
Future Functional Milestones

- Catalog Access Tools
  - MEO Interface Q1 '92
  - Name Resolver Q2 '92
  - Locate Data on an Object (P Q3 '93) Q2 '94
  - SIMBAD Interface 2

- Data Set Access
  - Data Archive Capability (P Q1 '92) Q4 '92
  - Image Display (P Q2 '93) Q2 '94
  - Graphics Display (P Q2 '93) Q3 '94
  - Proprietary Data Access Q1 '95
  - Data Analysis Tools ?

- Test Retrieval
  - Distributed Documentation Q4 '92
  - Text Database Indexing Scheme Q3 '95
  - Keyword Augmentation to Indexing Q3 '95
  - Text Retrieval System Q4 '95

Scientific Project Enabled by Functionality of ADS

- On-Line Astronomical Literature and Documentation
  - Scanned Literature from Astronomical Journals
    - Data Archive
  - Document Retrieval Technologies
    - Indexing, Retrieval
  - Transmission and Viewing of Documents
    - Display of Graphs and Images
    - Display of Scanned Text is a New Function
OAET COMPUTER SCIENCE & DATA SYSTEMS PROGRAMS

Paul Hunter
NASA/OAET
CODE R
Washington, D.C.
(202) 453-2704
DISTRIBUTED ACCESS VIEW INTEGRATED DATABASE (DAVID) SYSTEM

Due to the diversity of computers, operating systems, management systems, network protocols, etc., NASA space scientists have to learn many different access methods in order to obtain data. For example, in NASA's Space Astrophysics Program, Astrophysics observations, such as the International Ultraviolet Explorer (IUE), Hubble Space Telescope, Gamma Ray Observatory (GRO), etc., generally have one or more data centers for the analysis and distribution of data. The heterogeneity of its data centers makes Astrophysics multi mission research almost impossible.

The DAVID approach to the heterogeneous distributed systems problem is at four levels. At the lowest level, we develop universal object management systems to provide uniform access to heterogeneous database, images, spacecraft, nanosatellite, etc. management systems. At the second level, we develop "book" and "file" management systems to provide uniform access to the aggregate sets of data objects. At the third level, we develop libraries to provide uniform access to local area networks of computers containing "books" and "files". At the fourth level, we develop libraries to provide access to sets of libraries.

The DAVID software is being used on the NSSDC Data Archive and Distribution System (NADAS) as a front end and to distribute IUE observational data.

Technical Contact: Barry E. Jacobs, GSFC, (301) 344-5661.
**Problem:**
NASA space scientists have to learn many different access methods in order to obtain data/software.

**Solution:**
Distributed Access View Integrated Database (DAVID)

**DAVID Concept**
- Universal object type management systems: database, spreadsheet, tool, index, image, etc.
- Aggregate Sets of Related Objects: books, series, kits
- Libraries: Collections of holdings

**DAVID Approach**

**OBJECTIVES**

The objective of the intelligent data management project is to develop a fast and efficient knowledge base and database management system for access to a large set of heterogeneous data sources. This approach contains the provision of a set of tools to access and manage the data stored in various databases and to provide a user-friendly interface for accessing the data. The project will develop a system that allows users to search through the data stored in the databases and extract the information they need.

**TECHNICAL CONTACT**
Wilson J. Elsperman, O.S.C., (909) 256-8785

**CS F1**
**INTELLIGENT DATA MANAGEMENT PROCESSES**

**CS F2**
**Automatic Image Data Encoding and Analysis**

**CS F3**
**Intelligent Data Management**

**楮**

**STAFF**
This work has been presented at various conferences and in refereed journals, including a workshop titled "Data Management and Open Systems."
SOFTWARE ENGINEERING PROGRAM

- Includes: LoRC, GSFC, JPL, & JSC
- Supports: RICIS
- Program Drivers
- Very reliable software
- Solvable productivity
- Very large complex software systems
- Legacy systems
- Process management

The Portable Common Execution Environment (PCEE) is a framework designed to support the development and execution of software systems in a portable and common environment. It facilitates the integration and interaction of software components across different platforms and environments, ensuring compatibility and interoperability. PCEE aims to address the challenges of legacy systems, process management, and the need for robust software solutions. It provides a robust foundation for development, ensuring efficiency and effectiveness in software engineering projects.
Machine Learning Experiment in the KBSEE

**Description**: A machine learning experiment is conducted to validate an algorithm that can be used to analyze large sets of data. The experiment is designed to test the effectiveness of a specific machine learning model in predicting outcomes based on historical data. The algorithm is trained on a dataset containing past project information, and its performance is evaluated using metrics such as accuracy, precision, and recall.

**Knowledge Based Software Engineering Environment (KBSEE)**

- **Significance**: The KBSEE is an approach to software development that integrates machine learning and knowledge management to improve software quality and reduce development costs. By incorporating machine learning techniques, the KBSEE aims to automate parts of the software development process, such as code generation and defect prediction.
- **Static**: An experiment is conducted to evaluate the effectiveness of the KBSEE in practice. The experiment involves applying the KBSEE to a real-world software development project and comparing the results with those of a traditional development process. The KBSEE is found to improve development efficiency and reduce the number of bugs.

**Technical Contact**: [Contact Information]

---

**The Software Management Environment**

The Software Management Environment (SME) is a tool designed to assist software project managers in monitoring, analyzing, and controlling ongoing software projects. The SME integrates data from various sources, including project management systems, cost tracking tools, and quality assurance tools, to provide a comprehensive view of project status. It allows managers to make informed decisions based on real-time data and predictive analytics.

**During Fiscal Year 1919**, the SME was enhanced to include new features such as real-time alerting and automatic data synchronization. The enhanced SME was implemented across multiple projects, leading to improved project management and reduced costs.

**Technical Contact**: [Contact Information]
ESC, or Encyclopedia of Software Components

**TASK DESCRIPTION:**
- A hypertextual software cataloging and retrieving system
- An electronic metaphor of an encyclopedia
- Organizes software into a dynamic, linked knowledge structure
- Multiple access modalities
  - Goal-directed searching
  - Browsing with hypertext tracing
- Distributed contents
- Contents in many programming languages and of many levels of granularity
- Contents in many forms: code, documentation, graphics, etc.
- Encyclopedia Construction R&D supports user contributions and specialized handbooks

**NASA NEED AND SIGNIFICANCE:**
- An interchangable software parts technology is greatly needed
- Such a technology has existed for hardware for over a century
- Reuse will not be widespread until it is easier and cheaper to find software than to write new

**CURRENT STATUS:**
- Conceptual prototype: June 1990
- Prototype runs on color Macintosh systems
- Prototype is available for 1990
- Collect feedback from interface and contents
- Update the conceptual prototype
- Engineer insertion and retrieval mechanisms
- Conduct preliminary investigation of issues of AI (Artificial Intelligence)

Technical contact: Dr. Brian Beckman, JPL, (818) 354-1222

---

**APPLICATION AND ASSESSMENT OF INDUSTRY-STANDARD GUIDELINES FOR THE VALIDATION OF SOFTWARE**

**D'S 11**

**S. PHILIPS**

**SYSTEMS VALIDATION THEORETICAL GROUP**

**Research Triangle Institute**

**Rept. 1990**

**January 1990**

**Code SC**

**NASA**

**OBJECTIVES**

- Establish industry standards for the development of real-time software into an experimental test bed for evaluating the failure points of mission software and assess the effectiveness of mission software with these standards.

**RESULTS**

- Develop validation activities along with configuration management and software quality assurance policies for the Department of Defense and the National Aeronautics and Space Administration (NASA) in accordance with the NASA Technical Instruction 8805.2. The objectives of the software validation plan are achieved using the System of Systems (SOS) validation activities as roadmap for the development of the SOS code. Collect all activity data to evaluate the effectiveness of software development and validation methods and to provide the basis for software reliability model development.

**ACHIEVEMENT**

- Validation and verification activities in accordance with the DO-178 standards for mission software have been defined and implemented in the development process at the OSS level. Software data is being collected throughout the development cycle.

**SIGNIFICANCE**

- Although many software reliability experiments have been conducted, these experiments have not considered complex mission software that is critical to mission success. The OSS experiment provides a realistic mission for investigating the failure behavior of mission software. Engaging the DO-178 policies will yield data along with real-time software developed according to industry standards.

**RELEVANCE**

- The ability of the experiment is to be implemented using the philosophy of system-oriented software. The data from the OSS experiment for dependent estimation of reliability is unique in that it can provide a comprehensive and independent measurement of software reliability. The experiment further provides a technique for investigating the effectiveness of development methods, such as those prescribed by the FAA, for mission software.

**OUTSIDE**

- Complex linking of the three OSS versions and analysis of the resulting software error data. Develop user interface software models and others effective adaptive software development procedure based on this information.

**PRODUCTION OF REALISTIC SOFTWARE ERROR DATA**

- Independent Versions of Guidance & Control Software
- OSS version 1
- OSS version 2
- OSS version 3

**Software Error Data**

**Reliability Modeling**

**Method Assessment**
SPACE DATA SYSTEMS PROGRAM

- Includes: GSFC, JPL, LaRC, ARC

- Supports: VLSI Design @ U. Idaho

Program Drivers
- Ambitious future space instruments
- Long-duration, high-radiation requirements
- Time delay for insertion of commercial technology
- Transition activity for externally developed technology (e.g., DoD)

SPACE DATA SYSTEMS
STATE-OF-THE-ART ASSESSMENT

<table>
<thead>
<tr>
<th>MISSION</th>
<th>EPOCH</th>
<th>PROC. TYPE</th>
<th>MISSION</th>
<th>MEASURED AS OF 01/01</th>
<th>PERFORMANCE</th>
<th>CPU STATE %</th>
<th>RAM USE CHARACTERISTICS</th>
</tr>
</thead>
<tbody>
<tr>
<td>70.6</td>
<td>1982</td>
<td>DAELEO</td>
<td>FLOWN</td>
<td>1000 W 1800</td>
<td>8.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>70.5</td>
<td>1986</td>
<td>MADELLAN</td>
<td>FLOWN</td>
<td>350 W 730</td>
<td>8.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>70.4</td>
<td>1986</td>
<td>ESPI</td>
<td>LEVEL 6</td>
<td>340 W PS 64</td>
<td>8.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>70.4</td>
<td>1986</td>
<td>SSF</td>
<td>LEVEL 6</td>
<td>340 W PS 64</td>
<td>8.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>70.4</td>
<td>1986</td>
<td>EIV</td>
<td>LEVEL 5</td>
<td>10.13 W PS 64</td>
<td>3.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>70.4</td>
<td>1986</td>
<td>CAFM</td>
<td>CASSIS</td>
<td>10.13 W PS 64</td>
<td>3.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>70.4</td>
<td>1986</td>
<td>ACS</td>
<td>LEVEL 5</td>
<td>2.4 W PPS 64</td>
<td>5.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>70.4</td>
<td>1986</td>
<td>MAFC</td>
<td>LEVEL 5</td>
<td>350 W PS 64</td>
<td>8.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>70.4</td>
<td>1986</td>
<td>SSF</td>
<td>LEVEL 6</td>
<td>26.25 W PS 64</td>
<td>4.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>70.4</td>
<td>1986</td>
<td>SSF</td>
<td>LEVEL 6</td>
<td>26.25 W PS 64</td>
<td>4.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NOTE: ALS is currently underdeveloped and may be infeasible.
In 1990, the MA5 program continued implementation of fault-tolerant features and all software developments. A MA5 system with three computing modules has been delivered to JPL for independent evaluation. The latest version of special application codes with 90% of results has been implemented and tested. A simplified implementation of the MA5 system processor communication controller has been developed. A demonstration application which includes an autonomously operated rover is being developed. JPL and JRC have developed a fault plan and specification for the CASAS system which will enable it to operate with multiple computer systems. The system will have the capability of handling multiple computer systems. The system will enable the integration of multiple computer systems with the MA5 system.
CONFIRMABLE HIGH RATE PROCESSOR SYSTEM (CHIPS)

- Possible future use of chip in flight data system
- University of Idaho–Space Engineering Research Center for VLSI (Very Large Scale Integration) System
- Developed by the University of Idaho–Space Engineering Research Center for VLSI (Very Large Scale Integration) System
- One of nine centers selected in mid-1984
- Perform more operations per second, contains 25,000 transistors
- Major improvement in performance/reliability over existing system
- SPECIFICATION:
  - 1.0 Mbit chip
  - 2.0 Mbit chip
  - 4.0 Mbit chip
  - 8.0 Mbit chip

USAIRFUCI

DESIGNER DESIGNED VLSI CHIP BEING TRAINED FOR USE IN COMPUTER DETERMINATION BEGUN IN 1983

Possible Future Use of Chip in Flight data system...
ADSP CHARACTERISTICS

- The Engineering Model contained 73 Total Boards (excluding interfaces) which included 26 unique designs.
- The Boards contained a total of approximately 28,000 ICs, nearly filling two electronics racks and consuming about 15 KW of power.
- At the 20 MHz maximum clock rate the processor could sustain about seven billion floating point operations per second, the approximate compute rate required to process SEASAT data at real time rate.
- Magellan added five additional boards for special radiometric compensation and MultiLook Control resulting in a processor capable of processing the Magellan SAR data at least four times faster than the real time acquisition rate.
- The V/0 computer system limits the actual processing rate to approximately real time. This rate is still about four times faster than back up (engineering SAR processor), which is built from the most efficient commercial hardware available when the ADSP was complete in 1986.

WHAT IS THE ADSP?

- FLEXIBLE SAR PROCESSOR
- Over 6 Gigaflops peak compute rate
- Real time rate capability for missions such as SEASAT & SIR-B
- 150 Megabytes of memory
- About 35,000 integrated circuits

HIGH PERFORMANCE COMPUTING & COMMUNICATIONS PROGRAM

- Includes: ARC, GSFC, LaRC, LaRC, JPL
- Supports: CESDIS, RIAS, ICASE
- Program Drivers
  - NASA mission needs for high-performance computation in
    - Aeronautics/Astronautics
    - Space & Earth Science
    - Exploration
  - Support for nationally & globally networked collaboration
  - Highly trained scientists & engineers for future missions
GOAL:
- ACCELERATE THE DEVELOPMENT AND APPLICATION OF HIGH PERFORMANCE COMPUTING TECHNOLOGIES TO MEET NASA SCIENCE AND ENGINEERING REQUIREMENTS

OBJECTIVES:
- DEVELOP ALGORITHM AND ARCHITECTURE TESTBEDS CAPABLE OF FULLY UTILIZING MASSIVELY-PARALLEL CONCEPTS AND INCREASING END-TO-END PERFORMANCE
- DEVELOP MASSIVELY-PARALLEL ARCHITECTURES SCALABLE TO SUSTAINED TERAFLOPS PERFORMANCE
- DEMONSTRATE TECHNOLOGIES ON NASA RESEARCH CHALLENGES
  - INTEGRATED AEROSPACE VEHICLE SYSTEMS
  - EARTH AND SPACE SCIENCE PHYSICAL PHENOMENA
  - SPACEBORNE APPLICATIONS
REMOTE EXPLORATION AND EXPERIMENTATION

**GOAL**
Develop spaceborne computing technology which will enable high performance, fault-tolerant, adaptive space systems for a new generation of missions to explore the Earth and Solar System.

**OBJECTIVES**
- Demonstrate the feasibility of 1000-fold increase in spaceborne computer systems performance.
- Demonstrate a parallel, scalable architecture.
  - Low weight and limited power
  - Affordability
  - Fault tolerant
- Demonstrate a solid prototype-to-flight system engineering path.

**CANDIDATE APPLICATIONS**
- Analysis of data from the troposphere emission spectrometer, an Earth observing system instrument.
- Integration of robotic algorithms for sensing and manipulation astrophysics missions.
- Robotics.

NATIONAL RESEARCH & EDUCATION NETWORK
CESDIS
THE CENTER OF EXCELLENCE IN SPACE
DATA AND INFORMATION SCIENCES

Raymond E. Miller
CESDIS Director and
Professor of Computer Sciences
University of Maryland
College Park, MD
(301) 286-3805
**DUKE UNIVERSITY**

Parallel Compression of Space and Earth Data

**PERIOD OF PERFORMANCE:** October 1989 - September 1991

**GOALS**
- Develop parallel algorithms and architectures for lossy and lossless data compression.
- Design compressions applicable to large variety of data.
- Design good parallel vector quantization algorithm.

**ACCOMPLISHMENTS**
- Devised several algorithms, investigated their theoretical aspects, and evaluated performance against USC Image database to compare results with existing schemes using same images.
- Developed multiresolution lossy methods capable of controlling amount of information lost by trading off between compression rates and distortion.

**GEORGE WASHINGTON UNIVERSITY**

A Knowledge-Based Advisory System for General Scientific Data Visualization

**PERIOD OF PERFORMANCE:** September 1988 - August 1991

**GOALS**
- Investigate and specify visualization vocabulary to embody essential concepts required for scientific visualization.
- Incorporate user interface for visualization - naive users.
- Develop system design extendable for knowledge refinement and new visualization techniques.

**ACCOMPLISHMENTS**
- Compiled 145 rules and principles pertaining to expressiveness of visualization techniques.
- Implemented primitive rendering algorithms.
- Designed VISTA (Visualization Tool Assistant) and completed first prototype.
- Conducted preliminary test of prototype using data supplied by GSFC NSSDC.

**Stanford University**

Computer Assisted Analysis of Auroral Images Obtained From High Altitude Polar Satellites

**PERIOD OF PERFORMANCE:** January 1989 - December 1991

**GOALS**
Develop, implement, evaluate, and utilize advanced computer software tools whose purpose is to automate the analysis of global auroral images obtained from DE-1 and Viking satellites.

**ACCOMPLISHMENTS**
- Completed implementation and testing of "snakes" algorithm for finding auroral oval inner boundary.
- Have made progress on extending technique to regions where daylight overlaps aurora and to simultaneously find both inner and outer boundaries.
- Project personnel met with 13 GSFC and NASA HQ scientists to discuss potential collaborative efforts.

**University of North Carolina at Chapel Hill**

Image Pattern Recognition Supporting Interactive Analysis and Graphical Visualization

**PERIOD OF PERFORMANCE:** September 1988 - August 1991

**GOALS**
- Develop research software base to support interactive image pattern recognition;
- Develop artificial visual systems to solve measurement and detection problems;
- Develop multi-resolution image descriptions as a language for human-computer communication.

**ACCOMPLISHMENTS**
- Have explored nature of image regions formed by some coherence of intensity, edge surround, orientation, or texture, across scales based on 2 image description techniques: Multiscale Orientation Fields and Spatial Spectrum Classification.
- Beginning to apply to Landsat images obtained from GSFC.
- Applied algorithm to Hubble data in effort to reduce blurring of images without destroying frequency spectra.
ADDITIONAL RESEARCH TASKS

Brown University: Jeffrey Vitter and Paul Howard
Data Compression Algorithms

Penn State University: Eric Fargason and Michael LaValley
Advanced Statistical Methods for Improved Data Analysis of NASA Astrophysics Missions

University of Maryland: Noah Friedland
Simulated annealing, neural networks and Markov random field methods to solve inversion problems in atmosphere sounding

George Washington University: Burt Eclason and Herman Helgert
Concepts, plans and experiments to utilize the Advanced Communications Technology Satellite in supercomputing networking

George Washington University: John Silber and Cindy Starr
Computer graphics tools and techniques for scientific visualization

George Washington University: Rainald Lohner
Computer codes for simulation of 3D compressible magnetohydrodynamics flows

Stanford University: Philip Schwerer and Richard Bogert
Astrodict: Electronic mail for the Astrophysics Community

Carnegie Mellon University: Douglas Smith (Cray Fellow) and K.T. Kang (Undersea) Intermediate Language and Virtual Architecture for High Performance Image Processing

WORKSHOPS

AUGUST 1989
Computing Challenges in Managing Future Massive Image Systems

OCTOBER 1990
The Role of Computer Science in Mission to Planet Earth

TECHNICAL REPORT SERIES

Technical reports from CESDIS research projects.
Currently have 74 reports.

FUTURE PLANS

- Develop industrial associates program
- Fund additional research associated with HPCC and EOS
- Run small specialized workshops
- Continue to cooperate in organizing and running conferences e.g., Data Compression Conference Workshop on Parallel Algorithms
- Build more collaborative efforts
NASA TRENDS AND CESDIS ROLE

- NASA missions are moving from using state-of-the-art computing techniques and technologies to missions requiring major advances in computing hardware, software, algorithms, and communications.

- Advanced computing needs lie at the center of the NASA missions.

- Many areas of computer science research have a direct bearing on NASA needs.
  Database systems, distributed computing systems, networking and communication, AI/Expert Systems/Neural Nets, Human/Computer Interfaces, Computer Graphics, Visualization, Pattern Recognition, Algorithms, Software Engineering, Robotics, etc.

CESDIS PROPOSED HPCC ACTIVITIES

FY 92

- Hire Associate Director for HPCC.

- Help manage Goddard NRA peer review.

- Staff and run evaluation activity for space and Earth science grand challenge problems on various high-performance test-bed architectures.

- Prepare and issue CESDIS call-for-proposals in HPCC research.

FY 93 - 97

- Build research activities for HPCC in:
  Relative evaluation for high performance computers in space and Earth science applications.
  Numerical techniques for parallel computation.
  Management of massive amounts of data in distributed systems.

- Support education of Ph.D.'s in areas of HPCC research funded through:
  peer reviewed projects
  post-docs
  graduate fellowships

OVERALL CESDIS PLANS

1992 - 2000

Build a community of computing science researchers who collaborate with space and Earth scientists on problems of interest to NASA, through peer-reviewed proposal research at universities, through collaboration with industrial researchers, and by building a small core of about 5 full time researchers at CESDIS, plus a similar size University of Maryland group of faculty and graduate student researchers.
CASIS BACKGROUND

Professor Michael J. Flynn
Departments of Electrical Engineering and Computer Sciences
Stanford University
Stanford, CA
(415) 723-1450
Some special facilities/centers:
- Center for Integrated Systems
- Center for Telecommunications
- Center for Reliable Computing
- Center for Concurrency Studies
- Program in Manufacturing Science

CASIS — Past
- 10 years
- Supports about 20 research students per year
- About 110 total
- About 5 supported students receive Ph.D. each year
- About 35 total
- Over 300 supported reports, publications, etc.
Data Visualization
Data Recognition
Data Transmission
Data Processing

Data Visualization

Ice Flows Using SAR
Aurora
Terrain Elevation Mapping Using SAR

Basic Studies in 3-D
Data Representation

INFORMATION EXTRACTION FOR IMAGE DATABASES AND VISUALIZATION

1) Image analysis and computer vision tools may be used to extract information and parameters from image data.

2) The extracted parameters from the data may be used as automatically generated indices of image databases.

3) The extracted information may also be used as automatically generated geometric primitives for visualization.

EXAMPLES OF CASIS RESEARCH IN INFORMATION EXTRACTION FROM DATA

1) Generating sea ice motion vectors from radar images.

2) Extracting area and boundary information from UV satellite images of aurora.

3) Finding curvilinear features in radar images.
Data Recognition

Sparse Memory

Neural Nets

ORIGINAL PAGE IS OF POOR QUALITY
**EXPRESSNET**

A Unidirectional Broadcast Bus Local Area Network with A conflict-free round-robin Media Access Protocol

Basis for many network Concepts used in High Speed, Fiber Optic Local Area Networks, including the Metropolitan Area Network Standard DQDB (IEEE802.6)

Described in Most Important Textbooks
Tanenbaum; Stallings; Gallager and Bertsekas; etc.
Fiber Optic Communications

- Impairment: Dispersion distorts pulses resulting in Intersymbol Interference (ISI)
- Goal: Find simple signal processing techniques to compensate for ISI, allowing for improved data rates and distances

Application Specific Systems Development
- avoiding the software bottleneck for key applications
  - gas dynamic/fluid flow
  - pattern recognition/neural nets
  - finite element analysis

Computing
Concurrent Studies
Optimizing Compilers
  for Large Scale Multiprocessors
Parallel Architect's Workbench

Parallel Processor Software support
  - GEM
  - scheduling optimization
Architectures for Lattice Gas Cellular Automata

- Main application domain: fluid flow simulation
- Approach One
  - Until recently, most fluid flow models are based on partial differential equations (e.g., Navier-Stokes equations).
  - Solution usually involves numerical techniques with large numbers of floating point operations.
- Approach Two
  - Lattice gas cellular automata are new discrete models of fluids.
  - The fluids are made up of idealized particles that move according to simple rules.
  - Collections of large numbers of these particles can show overall continuous behavior which agrees with results predicted by Navier-Stokes equations.
  - This approach involves only logical operations and does not require floating point arithmetic.

The Computer Architect's Workbench

- The AWB is a set of software tools to predict the relative performance of alternative computer and system architecture features
  - Actual applications are used as benchmarks
  - High quality compiler support
  - Low requirements on hardware descriptions
- Trace Driven at basic block granularity

AWB: A General Research Tool and a Problem Solving Tool

- Research Tool
  - Architecture and Instruction Caches
  - Architecture and Data Buffering
  - Optimization and Instruction Architecture
- Design Tool
  - ASIC alternatives
  - Cost/Performance Analysis
  - Specific Workload Evaluation

Current Research

- The Effects of Programming Paradigms on the Instruction and Data Streams of a Shared Bus Multiprocessor
- Increasing the Performance of Shared Bus Multiprocessors
  - Effects of shared caches
  - Compiler techniques
    - Prefetch of Data
    - Write and Flush
  - Utilizing relationships between variables
    - Processor Architecture
NOTES:
PRINCIPAL INVESTIGATORS PRESENTATIONS

PROGRAM PLAN FOR AN EXPERIMENTER'S LABORATORY FOR VISUALIZED INTERACTIVE SCIENCE

Elaine R. Hansen/P. I.
Colorado Space Grant Consortium,
Univ. of Colorado
Boulder, CO
(303) 492-3141

Marjorie K. Klemp/Co.I.
Sally W. Lasater/Co.I.
Marti Szczur/Co.I.
Joseph B. Klemp/Co.I.
I. PROGRAM GOALS AND OBJECTIVES

GOALS

- Provide a capability that will help scientists of the '50s to interactively visualize data in order to better understand the large, complex, and multidimensional data sets of our future space missions
- Provide an Interactive Visualization Environment to support science research within and across NASA science disciplines
- Provide an Interactive Visualization Environment that can be easily tailored by the scientists themselves to best fit their individual research problems and display preferences
- Enable general members of the Space Science Community to use advanced visualization tools at an affordable price
- Capitalize on existing information systems techniques, technologies, and tools

OBJECTIVES

- Provide a multidimensional, visualization capability based on the research needs of NASA scientists
- Support interactive, exploratory analysis
- Handle multiple, simultaneous, and diverse data sets (ingest, present, manipulate)
- Provide a user interface that is intuitive for and responsive to the needs of the general science user
- Provide software that will run on a variety of popular, affordable workstations
II. CONCEPT FOR ACHIEVING OBJECTIVES

- First and Foremost, product must have a suitable acronym!

IVIE (Interactive Visualization Environment)

IVAN (Interactive Visualization Analysis Node)

ELVIS (Experimenter’s Laboratory for Visualized Interactive Science)

II. CONCEPT FOR ACHIEVING OBJECTIVES

- Process

- Early and Continual Focus on the Science Users through interviews, observations, and participative design

- Parallel Design of all aspects
  - The user interface design must start at the same time or before software design

- Early and continual user testing with prototypes to enable feedback to the design process
  - User testing of user interface, functions and services, help, documentation, training, etc.

III. BUILDING BLOCKS

1. 3D Data Generation and Rendering, “PolyPaint”
   - Foundation system developed by NCAR
   - Key individuals are Joseph Klemp (PolyPaint Originator and Lead), William Boyd, (Systems designer/programmer), Matt Irvin (Programmer, summer student, Physics grad at MIT in Academic Year)

   - “Poly” Features
     - Polygon generation for contour surfaces within any specified subset of a three-dimensional grid volume
     - Contour surfaces that enclose either higher or lower values
     - Coordinate transforms for 3D data in non-Cartesian coordinates
     - Line normal vectors that are calculated either separately for each polygon vertex or averaged among polygons sharing that vertex
     - Polygon generation for planar surfaces cutting through the data domain
     - Polygon generation for surfaces defined by the values of a two dimensional data array
     - Polygon generation for two-dimensional contour lines (computed as ribbon surfaces) with a specified contour interval in any coordinate plane
     - Polygon generation to convert a three dimensional curve into a ribbon like surface
III. BUILDING BLOCKS

- "Poly" Features (cont'd)
  - User specified limitation on the maximum number of vertices in polygons
  - Wire frame plots to preview the polygon data
  - File storage for polygon data in formats that can be utilized within PolyPaint or other surface rendering facilities

- "Paint" Features
  - Either index color (8 bits) or true color (24 bit RGB)
  - Depth queuing using a z buffer stored in memory
  - Display of multiple objects that can be superimposed sequentially within the same view domain
  - Polygon shading using either color (Gouraud) or surface normal interpolation
  - Reversal of unit normal vectors to view the inner side or both sides of surfaces
  - Display of 3D real time stereo objects
  - Volumetric rendering of three dimensional data
  - Wire frame rendering with shaded and antialiased lines
  - Combined display of solid surface, wire frame, and volumetric rendering
  - Procedures for storage and display of images to and from disk
  - Multiple color table partitions for displaying different colored objects in index color systems
  - Lighting options that include:
    - Up to 10 light sources

III. BUILDING BLOCKS

2. Management of Data Objects, "Object Work"
   - A software tool, developed at LASP
   - Key individuals: Sally Lassale and Randy Davis
   - Manages data objects (store, retrieve, update, delete)

3. Interactive Interface Objects, Services and Designer bench
   "Transportable Applications Executive — TAE"*
   - Developed by GSFC's System Technology Division
   - Key individual: Mark Siclaire
   - Enhanced for realtime applications by LASP at JHU
   - Key individuals: Merg Klenk and Eric Hales
   - Features
     - An easy to use workbench for design and layout of new user interfaces with "interaction objects"
     - Run-time services to display and control workbench designed user interfaces
     - A run-time interpreted command language to control an application's user interface

III. BUILDING BLOCKS

4. Data Access Interface — Network Common Data Form (NetCDF)
   - Concept originated by GSFC's NSSDC
   - Enhanced as general purpose tool by Undata
   - Features
     - General tools for storing and retrieving range of science data types
     - Enables access software to treat all data types in the same way
     - Treats data as collection of self describing, network transparent objects, accessed through simple interface
     - Supports random access of multi dimensional variables, and direct access to subset data values, to a hyperslice of data, and to records

  * Program Plan
  * Page 3
IV. PLANS

- Plans have been reduced and stretched to fit available funding levels

(hoping for more funding!)

MILESTONES

- System data flow in 1/2 years
- Complete system with limited users, data types, handlers in 3 years
PRINCIPAL INVESTIGATORS PRESENTATIONS

GRID ANALYSIS AND DISPLAY SYSTEM (GrADS): A PRATICAL TOOL FOR EARTH SCIENCE VISUALIZATION

James L. Kinter III/P.I.
Univ. of Maryland
Dept. of Meteorology
College Park, MD
(301) 405-5384

Brian E. Doty/Co.I.
THE GRID ANALYSIS AND DISPLAY SYSTEM (GRADS):
A PRACTICAL TOOL FOR EARTH SCIENCE VISUALIZATION

JAMES L. KINTER III
BRIAN E. DOTY

CENTER FOR OCEAN-LAND-ATMOSPHERE INTERACTIONS
DEPARTMENT OF METEOROLOGY
UNIVERSITY OF MARYLAND
COLLEGE PARK, MD 20742

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
APPLIED INFORMATION SYSTEMS PROGRAM WORKSHOP
BOULDER, COLORADO

22-24 JULY 1991

GrADS - Design Goals

* INTEGRATE | Data access, manipulation and display
  - ACCESS
    - Four dimensional data use
    - Concentrated along one or even two of three dimensions
  - MANIPULATION
    - Data operations through command line and buttons
  - DISPLAY
    - All standard techniques
    - Four graphs
    - Contouring (contouring)
    - Grid functions
    - Axes labels
  - All standard map properties and marks

* INTERACTIVE
  - Interactive response mode
  - Time sensitive
  - Control of data/display characteristics
  - Language
  - Programmability

* EASE OF USE
  - Steep learning curve
  - User acceptance

* HARD COPY

SCIENTIFIC VISUALIZATION = 3:

I SEE

SCIENTISTS DATA
- QUANTITATIVE
- INTERACTIVE
- VIEW + MANIPULATE

YOU SEE

RESULTS JOURNALS
- QUANTITATIVE
- PUBLICATION QUALITY
  - LEGIBLE
  - INFORMATIVE
  - COMPARABLE

WE SEE

IMAGES PUBLIC
- QUALITATIVE
- AESTHETIC

ORIGIANL PAGE IS OF POOR QUALITY
The generalized view of the data is 4-Dimensional. Spacing may be non-linear except in time.

- X, Y, Z, T = Longitude, Latitude, Height, and Time
- Any number of variables

- Data resides on disk in binary format. Data sets are easily created or read by FORTRAN programs.
- X varies fastest, then Y, then Z, then each variable, then time.

- Variables may have different numbers of levels, but must have same number of X, Y, and T elements
- Use separate data set for different scaling

GrADS Station Data Set

- Each data element arbitrarily located in space and time.
- Data resides on disk in a structured format. Subroutines are provided for creating the data set from FORTRAN programs.
- A utility is run to create a 'map' of the station data set, allowing efficient I/O access for any data request.

GrADS Data Description File

- Name of binary data set
- Defines scaling between grid coordinates and world coordinates
- Describes each variable
- Abbreviation to be used for the variable within expressions
- Number of levels provided
- Units value (not currently used)
- Long name of variable
Example 1

Examples using the data sets from the descriptor files shown earlier (files have been opened).

SET LON -90
SET LAT 40
SET LEV 850
SET TIME 0026DEC1982 18Z10DEC1982
DISPLAY TV
DRAW TITLE 850mb Virtual Temperatures at 90N, 40N
DRAW YLAB Degrees Kelvin

Notes:
- A 1-D 'slice' of TV (a basic variable) is being displayed
- We get a graph by default
Example 2

CLEAR
SET LON -140 -60
SET LAT 15 65
SET LEV 500
SET T 1
DISPLAY Z
DRAW TITLE 500mb Heights
DRAW XLAB 002 December 8, 1982

Notes: • If X and Y are varying, a map is automatically drawn.

Example 3

CLEAR
SET LON -90
SET LAT -90 90
SET LEV 1000 50
SET T 1
DISPLAY U
SET CSTYLE 3
DISPLAY TV
DRAW TITLE U Winds / Virtual Temperatures

Notes: • Vertical Cross Section
• Multiple DISPLAYS without CLEARing result in overlaid contours
PRINCIPAL INVESTIGATORS PRESENTATIONS

A DISTRIBUTED SYSTEM FOR VISUALIZING AND ANALYZING MULTIVARIATE AND MULTIDISCIPLINARY DATA

Allan S. Jacobson/P.I.
Jet Propulsion Laboratory
Pasadena, CA
(818) 354-0693

Mark A. Allen/Co.I.
Michael J. Bailey/Co.I.
Ronald G. Blom/Co.I.
Leo Blume/Co.I.
Lee S. Elson/Co.I.
The Linked Windows Interactive Data System (LinkWinds)

1. A visual data exploration/analysis environment with data displayed in a series of interdependent windows. Interdependence is established at user's discretion by "linking" visuals and controls.

2. A standard graphical user interface with additional "linking" rules. Results in an intuitive interface accessible with a minimum of training.

3. A user accessible applications prototyping environment.

4. Implemented with an object-oriented programming model, with "links" establishing message flow paths. There is an underlying command language (Lynx) based upon scheme.

5. A multi-user science environment (MUSE) requiring a minimum of network band-width.

The Linked Windows Interactive Data System (LinkWinds)

Development Approach

1. Employ an incremental development process using rapid prototyping of applications.

2. Work with users and developers in a tight loop throughout the process.

3. Provide software and workstations to users to stimulate product use in research activities.
The Linked Windows Interactive Data System (LinkWinds)

Application Development Cycle

User Interface Design Philosophy

1. Users are impatient and want to get started quickly on productive work. They are discouraged by large manuals.

2. Users learn from self-initiated exploration, making mistakes and correcting them.

3. Users refer to documentation only when the software doesn't conform to their expectations. Then they skim around in manuals or online help to find the answer to the current problem.

Future Plans

1. Port to X-Windows, using OSF Motif toolkit, for expansion to other platforms.

2. Implement an applications generator to support user application development.

3. Expand MUSE capabilities:
   - Session Management, Floor Management,
   - Telepointers, Network clipboard

4. Implement hardcopy and video output.
PRINCIPAL INVESTIGATORS PRESENTATIONS

DATA REDUCTION EXPERT ASSISTANT

Glenn E. Miller/P.I.
Space Telescope Science Institute
Baltimore, MD
(301) 338-4738

Mark D. Johnston/Co.I.
Robert J. Hanisch/Co.I.
Data Reduction Expert Assistant
AISR Workshop
22-24 July 1991

Glenn E. Miller
Mark D. Johnston
Robert J. Hanlach
Space Telescope Science Institute
3700 San Martin Dr.
Baltimore, MD 21218

Data Reduction
The process of converting raw instrumental output into physical measurements

Data Analysis Systems
- IRAF - Image Reduction and Analysis System (NOAO)
- STSDAS - Space Telescope Science Data Analysis System (STScI)
- MIDAS - Munich Interactive Data Analysis System (ESO)
- IDL - Interactive Data Language
- ... Many other systems

Very successful approach
- widespread distribution of these systems
- systems written for one wavelength extended to serve others
- incorporation of independently developed packages

Philosophy of These Systems
- Modular operators which work on standardized types of data files
- Command Language to execute single commands or scripts of commands (in interactive or batch mode)

Advantages:
- Flexibility for the user:
  - Individual commands can be chained (or "pipelined") to construct powerful, customized procedures
- Ease of development:
  - Well-defined methods for adding new modules. Thus many programmers and scientists may independently contribute to the growth of a system.
- Standardization
Disadvantages

- Learning a system isn't easy
  Commands can be complex with many parameters and even experts don't
  know the entire system. Users may have to learn more than one system,
  especially if they work at different institutions or their work is multi-
  disciplinary/multi-spectral

- Difficult to capture expert knowledge
  Manuals, on-line help, local gurus have drawbacks

- Data management problem
  A few nights observations can result in hundreds of data files which must
  each pass through many reduction steps

Expert Assistant

An alternative approach which builds on the foundation of these systems

- Gather information about the available data (typically from header
  information in the data files).

- Develop a plan for data reduction based on the user's goals, actual
  properties of the data and on limitations of available resources (e.g. disk
  space).

- Translate the plan into explicit reduction commands for a specific data
  analysis system.

- Monitor the plan and its execution for problems (e.g. missing calibration
  files) and alert the user.

- Be extensible to incorporate new types of data reduction, new analysis
  modules and new data analysis systems. The Expert Assistant will provide
  users with tools for this purpose.

- Present powerful and effective user interface including mouse-and-menu
  graphics (which is also found in non-expert systems) and natural language
  interface.

Useful Goals - But can it be done?

Yes:

- Expert system/artificial intelligence technology in routine use at the
  STScI to support operations

- Prototype data analysis assistant developed by one of us in 1987

- Doesn't require specialized hardware - the same workstation which runs
  IRAF (e.g. a Sun 4) can run expert system software
Implementation Considerations

- Choose language suited to symbolic processing, expert systems, and other artificial intelligence paradigms
- Ability to construct powerful user interface
- Provide flexibility for porting to other workstations
- Low cost for our users (existing data analysis workstation, plus minimal if any software licenses)

Choices:
- Common Lisp
  - Well-tested Lisp utility library at STScI
- Object Oriented (Common Lisp Object System)
- X-based window system
  - Mold? CLIM? GINE

Summary

Expert Assistant will significantly enhance research by removing much of the burden of routine data reduction from scientists and freeing them to focus more attention on the physical interpretation of the results. It will also alert the user to problems encountered in the reduction process.

Technically feasible

Innovative in use of expert systems/Al technology but built on foundation of existing data analysis systems

Project oriented towards developing a useful tool that will run on commonly used systems

An explicit design feature of this work is independence of any particular scientific discipline or data analysis system, which allows it to be applied to multiple scientific disciplines.

How to Best Involve the Scientific Community?

Lead Users:

- Involve scientists in use and independent evaluation of expert assistant as soon as possible (in addition to scientific input of PI and CoIs)
- We feel that it is important that this tool be used with real data reduction problems as early as possible.
- Even the initial versions of the Expert Assistant will be sufficiently powerful to pay back the Lead Users for their investment of time.
PRINCIPAL INVESTIGATORS PRESENTATIONS

OVERVIEW OF THE SYSTEM$ OF EXPERTS FOR INTELLIGENT DATA MANAGEMENT (SEIDAM)

David G. Goodenough/P.I.
Canada Centre for Remote Sensing
Ottawa, Ontario
(613) 952-2760

Ko Fung/Co.I.
Joji Iisaka/Co.I.
OVERVIEW OF THE SYSTEM OF EXPERTS FOR INTELLIGENT DATA MANAGEMENT (SEIDAM)

by

David O. Goodenough, Ko Fung, Joji Ikeda, Mike Robson, Cornelius Kushigbor

CANADA CENTRE FOR REMOTE SENSING
1547 Merivale Road, Ottawa, Ontario K1A 0Y7
Tel: 613-952-2760 OR 613-952-0500
Fax: 613-952-0872
E-mail: DGG@CCRS.CA Boulder, Colorado

NASA I CCRS I PFC

WHICH AGENCIES ARE INVOLVED?

ENERGY, MINES AND RESOURCES CANADA
- CANADA CENTRE FOR REMOTE SENSING

FORESTRY CANADA
- PACIFIC FORESTRY CENTRE

U.S. NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
- APPLIED INFORMATION SYSTEMS RESEARCH PROGRAM

INDUSTRY, SCIENCE AND TECHNOLOGY CANADA
- STRATEGIC TECHNOLOGIES BRANCH

NASA I CCRS I PFC

SEIDAM

OUTLINE

- PROPOSAL SELECTED BY NASA IN NOV. 1991 THROUGH COMPETITION FOR THE APPLIED INFORMATION SYSTEMS RESEARCH (AISR) PROGRAM.
- INTRODUCTION
- MULTIPLE DATA SOURCES INTEGRATION
- SEIDAM STRUCTURE
- MACHINE LEARNING
- USER INTERFACE
- SPATIAL KNOWLEDGE EXPERT SYSTEM
- DIGITAL TERRAIN MODEL (DTM) AND SEGMENTATION EXPERT SYSTEMS
- SYSTEM VALIDATION
- TASKS AND PLANS

NASA I CCRS I PFC

SEIDAM

SEIDAM - OTTAWA TEAM

CCRS
David Goodenough - Principal Investigator
Ko Fung - Co-Investigator
Joji Ikeda - Co-Investigator
Mike Robson, Cornelius Kushigbor
Alain Menard, Jean-François Maunier, Karl Steenz

University of Ottawa
Stan Matek, Dan Charles

NASA I CCRS I PFC
SEIDAM OBJECTIVE

TO CREATE A SYSTEM OF EXPERTS FOR INTELLIGENT DATA MANAGEMENT (SEIDAM) WHICH WILL INTEGRATE REMOTE SENSING DATA FROM SATELLITES AND AIRCRAFT WITH GEOGRAPHIC INFORMATION SYSTEMS AND MANAGE LARGE ARCHIVES OF REMOTELY Sensed DATA FOR DYNAMIC SELECTION OF DATA SOURCES AND SENSOR CHARACTERISTICS FOR RECOGNITION OF FOREST OBJECTS APPROPRIATE FOR ENVIRONMENTAL FOREST MONITORING.

SEIDAM

WHY USE MULTIPLE DATA SOURCES?

- A SINGLE DATA SOURCE MAY NOT BE ABLE TO CAPTURE ALL THE SIGNIFICANT CHARACTERISTICS NEEDED TO IDENTIFY AN OBJECT.
- A SINGLE DATA SOURCE MAY BE SUBJECT TO SYSTEMATIC ERRORS AND NOISE.
- MULTIPLE DATA SOURCES MAY PROVIDE COMPLEMENTARY INFORMATION.
- REDUNDANT INFORMATION FROM MORE THAN ONE DATA SOURCE INCREASES THE ACCURACY AND CERTAINTY OF OBJECT IDENTIFICATION.
- ONE IS MORE LIKELY TO HAVE DATA AVAILABLE AT THE DESIRED TIME.
- HISTORICAL GIS INFORMATION CAN GUIDE RECOGNITION PROCESS.

SEIDAM

INTEGRATION OF INFORMATION

- Levels of data fusion
- Recognized object
- Derived components
- Derived attributes
- Sensor measurements
- Environment
- Symbols
- Illumination source
- Signals

SEIDAM

PROBLEMS WITH MULTIPLE DATA SOURCES

- COSTS OF ACQUISITION
- COMPLEXITIES OF HANDLING MULTIPLE SENSORS:
  - VARIABLE SPECTRAL CHARACTERISTICS OF SENSORS.
  - VARIABLE SPATIAL RESOLUTIONS OF SENSORS.
- HIGH RESOLUTION SENSOR MAY RESOLVE CLOSELY SPACED OBJECTS, THUS GIVING TOO MUCH DETAIL.
- LOW RESOLUTION SENSOR MAY NOT RESOLVE OBJECTS OF INTEREST.
- OBJECTS WILL VARY IN SHAPE AND TEXTURE FROM SENSOR TO SENSOR.
- DIFFERENT SENSORS MAY HAVE DIFFERING VIEWING GEOMETRIES, MAKING MATCHING AND REGISTRATION OF SPATIAL OBJECTS DIFFICULT.
- VARIABILITIES DUE TO DIFFERENT RESPONSES TO THE ATMOSPHERE.
- VARIATIONS DUE TO DIFFERING RECORDING CRITERIA, FORMATS AND STANDARDS.

NASA | CCRS | PFC
MISMATCHES BETWEEN GIS DATA AND REMOTE SENSING DATA

- Sensor resolution may not be sufficient to record spatial features used for polygon delineation.
- GIS features may not be visible in images due to non-stationarity of objects (e.g., crops, water boundaries, burns, etc.).
- Generalizations used for base map may create spatial displacements with respect to image data.
- Different GIS sources may have spatial errors between their respective base maps.
- Image objects reflect seasonal variations.

Object Recognition Approach

1. CONVENTIONAL STACK APPROACH
   - 1.1 maximum likelihood with or without prior probabilities.
   - OR 1.2 logistic classifier.
   - OR 1.3 feature selection followed by classification.

2. SYMBOLIC APPROACH
   - 2.1 derive attributes from the minimum number of data sources
   - AND 2.2 object identification based on the attributes.
   - OR 2.3 dynamically select new data source if partially successful in object identification.

SYMBOLIC INTEGRATION APPROACH

- Knowledge-Based Methodology
- Object-Oriented approach
- Objects are grouped into an hierarchy according to the ease of recognition from imagery
- Available data sources are ranked and selected automatically to derive attributes required for object recognition
- Attributes are derived from the optimum data sources on an as-needed basis
- Object instances are distinguished by the use of derived attributes from selected data sources
- The search space is limited by pruning unlikely instances
DATA SOURCE SELECTION
- Select data source with best coverage of the ground area.
- Select data source with lowest inherent errors such as geometric and sensor errors.
- Select data source with optimum spatial resolution. We don't want too coarse a resolution so that desired objects cannot be identified. Also, we don't want too fine a spatial resolution so that desired objects become too complex in the image.
- Select best wavelength intervals of the sensor measurements for the desired objects.
- Select data sources for which there exist implemented algorithms to derive the required attributes.
- Select data source with lowest cost pre-processing required.
- Select data source with the least computational cost of the analytic algorithms.

SEIDAM

COMBINATION OF ATTRIBUTES FOR OBJECT RECOGNITION
- Combination through weights on the attributes
- Heuristic Methods (e.g. majority voting)
- Bayesian Approach
- Dempster-Shafer rules of combination
- Fuzzy Logic
- Endorsement Methods

SEIDAM

CONFLICT RESOLUTION FOR THE ENDORSEMENT METHOD

<table>
<thead>
<tr>
<th>Endorsements</th>
<th>Endorsements against</th>
<th>Inferred</th>
<th>Remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definite</td>
<td>Definite</td>
<td>Definite Yes</td>
<td>Unknown</td>
</tr>
<tr>
<td>Definite</td>
<td>Contingent</td>
<td>Definite Yes</td>
<td>Definite No</td>
</tr>
<tr>
<td>None</td>
<td>Definite</td>
<td>Definite No</td>
<td>Unknown</td>
</tr>
<tr>
<td>Contingent</td>
<td>None</td>
<td>Weakly Yes</td>
<td>Unknown</td>
</tr>
<tr>
<td>Contingent</td>
<td>Contingent</td>
<td>Weakly No</td>
<td>Unknown</td>
</tr>
<tr>
<td>Contingent</td>
<td>Definite</td>
<td>Definite No</td>
<td>Unknown</td>
</tr>
<tr>
<td>Contingent</td>
<td>Contingent</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

SEIDAM
Objectives of work on applying learning and planning in SEIDAM:

- to facilitate the development of new experts
- to re-use previous designs, particularly in the areas requiring specialized AI-related expertise
- to evaluate usefulness of Machine Learning technology in deployment of Expert Systems, with a Remote Sensing focus

Programming-by-example helps build interfaces

LEAR: LEarning Advisor Rules

NASA  |  CCRS  |  PFC
SEIDAM Human-Computer Interface (HCI)

Purpose of the HCI
- simplify prompt-interactive procedures
- generate level of knowledge
- arrange the information on the screen
- augment the capability of the worker

Benefits of the HCI
- complex set of tasks become less intimidating by relating tasks to the goal and to other tasks
- communicate at the level of the users
- adapt to the level of the user
- event driven dialogue rather than predictable

Development Considerations
- communication strategy
  - level of abstraction (target user)
  - consistent grammatical rules and terminology
  - consistent presentation
- amount of freedom available to the user
- inform the user
- what has happened, what is happening
SEIDAM - Implementation of Spatial Knowledge for Operations

Spatial Knowledge Expert System

- Background Demands:
  - Conventional Image Analysis for Remote Sensing:
  - Little integration and utilization of spatial knowledge
  - Multi-sensor data integration for different resolutions
  - Need more intelligent data abstraction

- Emerging of supporting technologies:
  - Artificial Neural Network Computing
  - CCRS' developed method "Pixel Swapping" to treat image spatially and spectrally in a unified way.

- Experience:
  - CCRS' long term experiences with image analysis and expert system development.

SEIDAM - Spatial Knowledge OBJECTIVES

1. Develop spatial information acquisition methods and system:
   - Investigation of Neural Networks capability:
     - Spatial information extraction, Multi sensor data fusion
     - Data abstraction methods for high resolution spectral data, image indexing methods.

2. Develop an expert system for spatial information analysis system
   - Embedding or linked with neural networks and pixel swapping functions.

3. Collect spatial knowledge observed in remote sensing data and create pilot knowledge bases.

4. Validate spatial knowledge acquired for remote sensing.

SEIDAM - Spatial measurements knowledge

Geometrical measurement:
- Classical geometrical measurements:
  - Area, class, moment, direction and ellipticity length, direction, additional geometrical measures:
    - Fractal Dimension, Statistical and Structural Texture measures

Geometrical shape characterization:
- Point like, line like, region like.

Spatial structure:
- Tree-like, network like, spatially clustered...

Association:
- HU, overlapped, included...
- Positively associated, negative associated or randomly associated.
SEIDAM

SPATIAL KNOWLEDGE EXPERT SYSTEM

Preprocessor
- Linked with N-N
- Feature conversion
- Spatial knowledge acquisition
- Spatial measure estimation
- Spatial knowledge manager

Interface manager:
- Spatial query handling
- Spatial knowledge presentation
- Knowledge validation
- Spatial knowledge description, indexing and retrieval
- Post processor:
  - classifier
  - noise remover
  - predictor
- Interfaces to other expert system:
  - Import and Export relevant information to other expert system.

N A S A  C C R S  P F C

SEIDAM

DTM EXPERT FLOW

1. Estimate the Terrain type
2. Estimate Satellite image accuracy
3. Estimate DEM accuracy
4. Do Terrain Correction
5. Resample DEM

N A S A  C C R S  P F C

SEIDAM

DTM EXPERT STRUCTURE

- DTM Expert Structure
- Corrected Image
- Slope & Aspect

N A S A  C C R S  P F C

SEIDAM

CONCLUSIONS FOR DTM SYSTEM

- Elevation, slope and aspect from DTM improve significantly object recognition accuracies.
- Expert system amplifies use of DTM
- Expert systems offer the possibility of distributed, national access to DTMs from provincial and federal sources.
- DTMs are essential for multisource integration.
- Expert prototype successfully tested with image and DEM of mountainous terrain in B.C.
- In the future generate revised accuracy estimates for each point in the DTM

Adaptive grid size with DEM
for Slope & Aspect generation

N A S A  C C R S  P F C
**SEGMENTATION Introduction**

- SEGMENTATION is a complex process with knowledge distributed over several levels of the recognition procedure.
- Users are unfamiliar with the many issues to be considered in choosing the best edge operator, the best segmentation method, sensor characteristics, and the most appropriate spatial and spectral features.

→ Expert Systems can simplify access to this powerful but complex process!

---

**Prototype Flow Chart**

GIS polygons updates:

- Use labelled segments to update GIS polygons:
  - leave polygons that match segments
  - split polygons covering multiple segments classes
  - redraw polygon limits to match segments classes
  - update GIS database with new polygon attributes

- Difficult cases can be submitted to forest inventory experts and/or local foresters for human labeling.
SEGMENTATION CONCLUSIONS

- The expert system supplies and speeds up the segmentation process.
- Great advantage over standard programming techniques.
- Prototypes being incorporated into an operational system for segmentation.
- "Intelligence" of our expert is easily improved with new knowledge as soon as it becomes available.

CONCLUSIONS #2

- We still need to work on GIS updating process from segments of remotely sensed images:
  - Segmentation parameters?
  - Improve segment classification
- Establish RULES for labeling segment classes from GIS
- Rules to modify GIS polygons from classified image segments?
- Expert to be tested on several forest inventory maps

WHAT RESEARCH WILL BE DONE?

1) AUTOMATED METHOD:
   - Automatic method is employed to set up the segmentation process.
   - Prototypes being incorporated into an operational system for segmentation.
   - "Intelligence" of our expert is easily improved with new knowledge as soon as it becomes available.

WHAT RESEARCH WILL BE DONE? #2

2) IMPROVE EXISTING METHODS FOR FOREST INVENTORY:
   - Geographic correction of satellite data
   - Expansion of GIS data sets
   - Estimation of GIS label and boundary accuracies
   - Interactive creation of new areas
   - Update GIS
WHAT RESEARCH WILL BE DONE? #3

SEIDAM

HOW DO THESE RESULTS GET COMBINED WITH FOREST CANOPY MODELS TO MAKE IMPROVED ESTIMATES OF FOREST STAND
PRODUCTIVITY?

- CHARACTERIZATION OF LAND SURFACE FROM DTM AND ESTIMATION OF SITE QUALITY FROM SOILS, GEOLOGY,
  HYDROLOGY, AND DTM ANALYSIS

- INTEGRATION OF GIS - RS RESULTS WITH FOREST CANOPY MODELS FOR SHORT AND LONG TERM PRODUCTIVITY
  ESTIMATION ON A POLYGON BASIS

- EXPERIMENTATION WITH AIRBORNE IMAGING SPECTROMETER DATA FOR CHEMICAL INPUTS INTO MODELS

NASA I CCRS I PFC

WHAT RESEARCH WILL BE DONE? #4

SEIDAM

WHAT COMPUTING SCIENCE AND ENGINEERING RESEARCH WILL
BE DONE?

- CREATION OF EXPERT SYSTEM SHELL IN COMITUS PROLOG
  WITH G41 (GRAPHICAL USER'S INTERFACE) FOR UNIX AND VMS
  - RETAIN EXISTING SOFTWARE INVESTMENT
  - DISTRIBUTED PROCESSING WITH INTELLIGENT AGENTS
  - OBJECT-ORIENTED (FRAMES)
  - ABILITY TO PAUSE AT USER'S CONVENIENCE

- MACHINE LEARNING FOR CREATION OF EXPERT SYSTEMS
  TO CONTROL TASKS AND FOR LEARNING FROM EXPERIMENTS

- 150 EXPERT SYSTEMS WILL BE CREATED WITH WINDOW
  INTERFACE

- DISTRIBUTED GIS AND RELATIONAL DATA BASE ACCESS

NASA I CCRS I PFC
HOW WILL SEIDAM BE TESTED?

THE PACIFIC FORESTRY CENTRE AND CCRS ARE SELECTING THREE TEST SITES, THREE NEAR INVERMERE, BRITISH COLUMBIA AND ONE ON THE WEST COAST. DATA COLLECTION FROM AIRCRAFT AND SATELLITES IS BEING PLANNED FOR THE SUMMER OF 1993. CCRS HAS PREVIOUSLY COLLECTED DATA ON THE INVERMERE SITES. THESE DATA HAVE INCLUDED SEVERAL YEARS OF TM, HRV, MOS-1, AND NOAA AVHRR AIRCRAFT DATA PREVIOUSLY COLLECTED INCLUDE CCRS AND AVHRR. MORE THAN 150 FOREST INVENTORY GIS FILES HAVE BEEN ACQUIRED. OTHER DATA CORRESPONDING TO THE TEST SITES HAVE BEEN OBTAINED FROM FEDERAL AND PROVINCIAL SOURCES.

SEIDAM WILL BE TESTED AT CCRS, PFC, AND POSSIBLY NASA-ARC FOR THESE SITES AND FOR OTHER REMOTE SENSING DATA. SOME INDUSTRIAL CONTRACTORS MAY ALSO CHOOSE TO EVALUATE THE SEIDAM METHODOLOGY.

NASA | CCRS | PFC

HOW WILL SEIDAM BE TESTED? #2

CCRS HAS CONDUCTED FIELD WORK IN THE ROCKY MOUNTAIN TRENCH PREVIOUSLY. FOR AIRCRAFT CALIBRATION, OPTICAL MEASUREMENTS OF LAKES AND HOMOGENEOUS TARGETS ARE MADE. A GRID OF REFLECTORS AND ACTIVE RADAR CALIBRATORS ARE DEPLOYED. A GRID OF MARKERS IS LAID OUT FOR MULTI-SENSOR REGISTRATION. ECOLOGICAL SAMPLES WILL BE TAKEN WITH THEIR LOCATIONS DEFINED BY GPS GROUND EQUIPMENT. MEASUREMENTS WILL BE MADE OF CHLOROPHYLL, NITROGEN, LIGNIN, AND OTHER CHEMICALS. THE WEST COAST SITE HAS AN EXTENSIVE HISTORY OF ECOLOGICAL MONITORING. WHERE POSSIBLE, GPS POSITIONING WILL BE USED TO DEFINE AIRCRAFT LOCATION.

NASA | CCRS | PFC

Artificial Intelligence Research Projects

<table>
<thead>
<tr>
<th>SHERI</th>
<th>PIKES</th>
</tr>
</thead>
<tbody>
<tr>
<td>(System of Hierarchical Experts for Resources Inventories)</td>
<td>(Photo Interpretation Keys Expert System)</td>
</tr>
</tbody>
</table>

SEIDAM

(System of Experts for Intelligent Data Management)

(January 1991 - January 1995)

NASA | CCRS | PFC

SEIDAM SOFTWARE STATUS

- More than 45 expert systems created; implementing Motif interface for user interface
- Ported Shell and Expert Systems to Unix
- Ported Shell and Expert Systems to other platforms (Sun, Macintosh...)
- Research, develop and integrate new expert systems
  - (Segmentation, Digital Terrain Model, Texture, Spatial Knowledge, Ecosystem modeling with fuzzy methods, GIS input/output, intelligent land information system, forest productivity and monitoring, etc.)
- Interface Expert System and RDBMS (through INGRES, C, Pick and SQL)
- Plan knowledge acquisition experiments for SEIDAM validation
- Integrate Object Oriented Database
- Use Machine Learning to accelerate the development of new experts

NASA | CCRS | PFC
**SEIDAM**

## TASKS AND MILESTONES

<table>
<thead>
<tr>
<th>Project Approval by NASA</th>
<th>Nov. 1, 1990</th>
</tr>
</thead>
<tbody>
<tr>
<td>Submit Preliminary Proposal to ISTC</td>
<td>Mar. 13, 1991</td>
</tr>
<tr>
<td>Submit Final Proposal to ISTC</td>
<td>June 29, 1991</td>
</tr>
<tr>
<td>SEIDAM Overall System Design</td>
<td>Sept. 30, 1991</td>
</tr>
<tr>
<td>Remote Sensing data EB</td>
<td>Apr. 30, 1992</td>
</tr>
<tr>
<td>GIS EB</td>
<td>May 30, 1992</td>
</tr>
<tr>
<td>Field Data EB</td>
<td>June 30, 1992</td>
</tr>
<tr>
<td>Aerial data acquisition</td>
<td>July 31, 1992</td>
</tr>
<tr>
<td>Data Fusion EB</td>
<td>Nov. 30, 1992</td>
</tr>
<tr>
<td>Annual Report to NASA/CCRS</td>
<td>Dec. 30, 1993</td>
</tr>
<tr>
<td>Data Source Selection EB</td>
<td>Feb. 3, 1993</td>
</tr>
<tr>
<td>Neural Network development for data fusion</td>
<td>Mar. 31, 1993</td>
</tr>
<tr>
<td>Method Selection EB</td>
<td>May 30, 1993</td>
</tr>
<tr>
<td>Data Interpretation Model ES</td>
<td>July 31, 1993</td>
</tr>
<tr>
<td>Solution Planning ES</td>
<td>Sept. 30, 1993</td>
</tr>
<tr>
<td>Final SEIDAM Validation</td>
<td>Nov. 1, 1993</td>
</tr>
<tr>
<td>Final Report to NASA/CCRS</td>
<td>Dec. 30, 1993</td>
</tr>
<tr>
<td>Project Completion for NASA portion</td>
<td>Mar. 31, 1994</td>
</tr>
</tbody>
</table>

**SEIDAM**

## FINAL REMARKS

WE ARE DELIGHTED TO HAVE THE OPPORTUNITY TO JOIN OUR COLLEAGUES IN THE APPLIED INFORMATIONS SYSTEMS RESEARCH PROGRAM.

PLEASE FEEL FREE TO CONTACT US AT THE NUMBERS GIVEN ON THE FIRST VIEWGRAPH.

IF YOU ARE INTERESTED IN INTERACTING WITH THE SEIDAM PROJECT, PLEASE CONTACT DR. DAVID GOODENOUGH.

WE THANK THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION FOR SUPPORTING THIS PROJECT.

WE ALSO THANK THE CANADA CENTRE FOR REMOTE SENSING, OF ENERGY, MINES AND RESOURCES CANADA, AND THE PACIFIC FORESTRY CENTRE OF FORESTRY CANADA.
PRINCIPAL INVESTIGATORS PRESENTATIONS

THE SCIENTIFIC MODELING ASSISTANT: AN ADVANCED SOFTWARE TOOL FOR SCIENTIFIC MODEL BUILDING

Richard M. Keller/P.I.
Michael H. Sims/P.I.
NASA Ames Research Center
Moffett Field, CA
(415) 604-3388
The Scientific Modeling Assistant:
An advanced software tool for scientific model-building

Richard M. Keller
Michael H. Sims
Principal Investigators

David Thompson
Michal Rimon
Project Staff

Christopher McKay
Jennifer Dungan
Science Collaborators

NASA Ames Research Center
Artificial Intelligence Research Branch
 Moffett Field, CA 94035-1000
(650) 604-5284  joseph@goerge.ames.nasa.gov

Methodology

- Study existing scientific modeling software
- Design tool to meet scientists' needs
- Reimplement models with our new modeling tool

Initial Focus Areas

- Planetary Atmospheric Modeling: "Titan Greenhouse Model" (C.P. McKay, NASA Ames)
- Earth Ecosystem Modeling: "Forest BGC" (S.W. Running, U. of Montana)

Project Goal: Build a specialized software tool to assist in Scientific Model-building

Analysis & visualization facilities
Interactive graphical interface
Intelligent assistance
High-level modeling language
Libraries of data sets, equations, subroutines, physical quantities

TECHNIQUES
Object-oriented programming
Artificial intelligence
Symbolic manipulation

FEATURES

Overview of the Scientific Modeling Process

1. Conceptualization
   \[ \frac{\partial}{\partial t} n + \frac{\partial}{\partial x} \left( d n \right) = 0 \]
   \[ \frac{\partial}{\partial t} T + \frac{\partial}{\partial x} \left( d T \right) = -gb \]

2. Model-building

3. Experimentation

4. Analysis & Interpretation
### Outline

1. Project Introduction
2. Problem & Approach
3. Case Study: Modeling planetary atmospheres
4. Summary

### Problems and Opportunities

#### Problems:
- Labor intensive
- Difficult to understand, share, and modify

#### Root causes:
- Scientist ≠ programmer II
- Wrong level of abstraction
- Implicit knowledge buried

#### Potential payoffs:
- Improve scientific productivity
- Foster scientific communication
- Disseminate modeler's expertise

---

**The scientists' plea: “Spare me the programming details!”**

- Variable declarations, storage allocation
- Scientific units bookkeeping & conversion
- Data structure manipulation & management
- Documentation
- Symbolic manipulation
- Interfacing with scientific datasets
- Numerical programming:
  - ODEs, integration, interpolation, convergence
1. Project Introduction
2. Problem & Approach
3. Case Study: Modeling planetary atmospheres
4. Summary
Prototype Modeling Tool Provides:

- Library of relevant:
  - data sets
  - physical variables
  - equations
- Mechanisms allowing user to select and apply equations to data and compute new physical variables
- Model/data display features:
  - data dependency network
  - plotting/graphing facility
  - tabular display
- Model management environment:
  - multiple models and multiple scientists

Outline

1. Project Introduction
2. Problem & Approach
3. Case Study:
   Modeling planetary atmospheres
4. Summary
Status

- Completing development and evaluation of first prototype
- Redesigning system interface and internals
- Adding additional domain knowledge
- Moving toward more portable system
- Investigating other domains:
  - Ecosystem modeling

Major Milestones

- FY 1991: Completion and evaluation of initial "alpha" prototype
- FY 1992: Design and implementation of a "beta" version prototype that will be used by a small group of atmospheric scientist-users
  
Begin development of ecosystem model
- FY 1993: Enhance system and make available to a group of Cassini Mission scientists who are involved in instrument design
  
Make system available to ecosystem scientists
PRINCIPAL INVESTIGATORS PRESENTATIONS

DATAHUB
KNOWLEDGE BASED ASSISTANCE FOR
SCIENCE VISUALIZATION AND ANALYSIS
USING LARGE DISTRIBUTED DATABASES

Thomas H. Handley Jr./P.L.
Jet Propulsion Laboratory
Pasadena, CA
(818) 354-7009

Donald J. Collins/Co. I.
Richard J. Doyle/Co.I.
Allan S. Jacobson/Co.I.
DataHub

Knowledge-based Assistance for Science Visualization and Analysis Using Large Distributed Databases

Team
Data Management: T. Handley (PI), P. Li
Science: D. Collins (CO-I), A. Tran
Visualization: M. Wade
Research: T. Maddox (Summer Faculty)

Applicable Technologies

- Data Discovery and Management
- Management and Analysis of Complex Information
- Inference and Reasoning Assistance
- Man-machine Interface
- Exploratory Data Analysis

Objectives

- Scientific Data Models
  - Data-driven analysis
  - Data transformations
  - Data semantics
  - Analytical-related knowledge about data
  - Data discovery, ingestion, extraction, ...
  - Self-describing data structures
- Intelligent assistant system(s) with some knowledge of data management and analysis built-in.
- Use of expert system technology to aid exploratory data analysis, i.e., expert systems, neural nets, classification systems.
- Capture and encode knowledge about the data and their associated processes. Encode scientific knowledge into the routines, processes and procedures.
- Provide data management services to exploratory data analysis application(s) i.e., LINKWINDS.
Data Hub Functional Architecture

Logical View
- Uniformly gridded n-dimensional data
- Examples:
  - Ocean Temperature Data
  - HRB Data

Physical View
- HDF
- VICAR
- FITS
- CDF

Logical Access Methods
Selection Operators
- Range selection for one or more independent variables. Required conjunctive conditions for n-k dimensions (n-dimensional data, k-dimensional display).
  - Depth < 62, where Time = 11
  - Time = 12, where Depth = 4
- Enumerated selection for one or more free variables
  - Time = [11, 12, ... 1n]
- Selection for the dependent variables
  - Temperature > 30 degree-C for Temperature Dataset
  - Frequency > 25 for Flag Dataset

Data Representation

Meta Data Representations

Data Classes
- Rest Data
- Scientific Data

Logical File Structures
- Mapping Between Logical Access Methods and Physical Access Methods
- Semantic Structure
Logical Access Methods
Transformation Operators

- Data format conversion operator
  - MCSST DSP → MCSST HDF SDS
- Raster conversion operator
  - MCSST HDF → MCSST HDF Raster 8-bit
- Resampling operator
  - MCSST HDF Raster 8-bit Global → MCSST HDF 360x180
- Mapping between collected data and known facts
  - Ocean data versus LandMask
- Filling in missing data
  - Use local data to fill
  - Use data selected from other time frame
  - Use data selected from other correlated dataset
- Dataset registration
  - Warping
  - regliding
  - Registration
- Other map projections

Logical Access Methods
Update Operators

- Transaction
  - Atomic action
  - Traceable unit
  - Recoverable Unit
  - Result datasets stored in user area

Example

Begin-Xin
DSP → HDF SDS
HDF SDS → HDF Raster 8-bit
HDF 8-bit → HDF 360x180
End-Xin
An Intriguing Solution
(Summer Research Fellow)

- The development of intelligent assistant systems with some knowledge of data management and analysis built-in.
- LinkWinds: allowing users to select data sources and link them with various transformation procedures to analyze and display data.
- Graphical Representation of Knowledge (GROK): nodes and links (graphs) are a useful visualization of the grouping and ordering of general knowledge.
- Science Analysis Assistant (SAA): an embedded formalism for advising scientists about the salience of data.
- The goal is to reduce the difficulty of managing or exploring complex or large datasets or knowledge bases.
Our Hypothesis
(Summer Research Fellow)

- A high level of satisfaction with LinkWind: maybe because it is a comfortable reflection of the data analysis skills and processes of the development team.
- The combination of procedures linked by the user can lead to the identification of interesting features of the original dataset and the possible discovery of new knowledge.
- Emergent behavior seems to be a result of the linking process which is driven by the user's domain specific and general knowledge and experience.
- The process of data analysis is well-represented by a graph of connected data, analysis and display procedures.

Possible Outcomes
(Summer Research Fellow)

- Saliency measures which are locally interesting (from data perspective) and globally interesting (from the domain specialist's perspective).
- Concurrent data exploration paths.
- Experimental trials to determine if algorithms faithfully represent user preferences and goals.
- Machine learning techniques used to generalize data analysis algorithms.

Research Objectives
(Summer Research Fellow)

- Investigate how information (from 1) the original data, 2) analysis and display procedures and 3) the user may be combined to facilitate knowledge discovery in a dataset.
- Explore the development of algorithms which represent how information regarding analysis goals and hypotheses, transformation procedures, and metadata may be combined to facilitate the discovery of knowledge.
- Seek to develop dataflow representations of the information between these sources.
- Characterize interaction dynamics of information top-down in a goal-driven manner from the user towards the data and the analysis and display procedures and bottom-up in an event-driven manner.
- Use resulting metadata to characterize and classify the analysis and display procedures.
- Classify the cooperative or competitive information flow dynamics.
Science (cont'd)

The Pacific studies require a much broader context in which to interpret the measurements provided by the time-series and process-oriented studies. To address the broader context, satellite imagery can provide a description of the spatial and temporal scales of the biological and physical processes and their variability, including changes in water mass, incident radiation, nutrients, and the consequent formation of blooms of different species of marine phytoplankton and bacteria. We will use a time series of CIGS images of the atmospheric aerosol and pigment distributions for the north Pacific Ocean, obtained to include the equatorial region, extending to 2008, to study the seasonal variability of primary productivity. This data will be combined with time series of the CIGS product forming the air-sea surface contact to determine chlorophyll concentration of the equatorial and sub-tropical north Pacific Ocean. These time series will be examined to determine the spatial and temporal patterns of productivity, including the interannual variations that occur in the productivity caused by variations in the physical environment.

The regional imagery from the north Pacific will provide an historical context for the continuing study of this region. For the description of the seasonal patterns of primary productivity, composite maps will be generated from the individual scenes produced from the CIGS archives. The goal of the composite images is to provide a seasonal description of the productivity of the region and to present the temporal character of data that is fundamentally episodic in nature. Two such maps will be produced, including maps of the total productivity, computed from the surface temperature and the net column pigment content through a model described by Cullen et al., 1993, and the carbon flux through a model under development by Mitchell et al., 1991. The model of total productivity permits the computation of the total carbon assimilation by phytoplankton in the ocean, while the model for the carbon flux estimates the fraction of the total carbon assimilated that is removed from the surface mixed layer through processes which contribute to the transport of CO2 from the atmosphere.

Science (cont'd)

This task will address the issue of the assessment of pigment concentration, and of the radiance leaving the sea surface, quantified by 100 pigment concentration, for data collected in the equatorial region. The problem is fundamentally different from data collected in the equatorial region. The statistical description provided by the use of data analysis will enhance our ability to predict the development of these fields through the use of these techniques to assess the contribution of each of the physical processes in the development of the pigment fields in space and time.

With the development of properly interpolated and extrapolated fields, estimates of the net and total primary productivity will permit an assessment of the carbon flux in the equatorial Pacific Ocean. These concepts will be extended to provide an assessment of the net flux of carbon from the atmosphere into the ocean on a yearly basis. Our preliminary estimates of this quantity are 3.4 Gt CO2 per year based on an estimate of the carbon leaving the ocean and a representation of the water leaving radiance and carbon flux based on techniques developed by Mitchell et al., 1991. These estimates will be refined through a better estimate of the pigment concentration and flux of the water leaving radiance.
Expert Systems (cont'd)

Interpretation problems present challenges across scientific disciplines. The oceanographic domain presents a rich set of data parameters that must be clustered and combined to build an interpreted image. The time required to construct this interpretation is inhibited by the computational resources needed to find a comprehensive correlation over the entire set. Artificial intelligence (AI) will support this effort by combining the data sets using state-of-the-art machine learning techniques. Leveraging of this task is provided by the Scientific Analysis Assistant (SAA), funded under Data RC. Together these projects are investigating how machine learning techniques can assist scientists with efficient data ingestion.

Proof of Concept Near Term Goals
- Evaluate and compare machine learning techniques
  - GID2
  - Parallel distributed processing (PDP) networks
  - AutoClass
- Apply selected techniques and test performance
  - Apply the selected machine learning algorithm to two-thirds of the data set and test the algorithm performance on the remaining one-third of data.

Solution Long Term Goals
- Apply proven technique
- Produce the interpolated image using selected data parameters.
- Compare performance
  - Measure the performance of machine learning algorithm application to conventional techniques.

Data Management (cont'd)

Currently available database technology is largely designed for business data processing applications, and seems inadequate for scientific applications. This R&D is addressing the following issues in scientific data management:

1. Scientific Database Models
   1. Data format (usage of data formats, data types, data formats for different analysis methods, etc.)
   2. Data schema (mapping of data values, relationship between datasets, discipline-dependent data access/analysis methods, etc.)
   3. Analysis-related knowledge about data
   4. Data quality assurance (identification of missing, and missing data, i.e., NULL value representation, data quality data validity after data transformation, etc.)

2. Resource Sharing Environment for Science Databases
   1. Data exchange protocol to facilitate data ingestion and data visualization
   2. Data abstraction and storage of extracted data
   3. Tracking, logging, and synchronization of data access activities

Figure 1 depicts today's world of science data management. There are many data formats from the data sources in different science communities. There are many data conversion tools available to convert from the suppliers' formats to the data formats desired by the data consumers. One of the objectives of the ongoing project is to build a resource sharing environment depicted in Figure 1. A system, named D'azur, will be a server between the data suppliers and the data consumers to facilitate data exchanges, to assist science data analysis, and to provide a systematic approach for science data management.
**Figure 1. Science Data Management: Today**

**Figure 2. Science Data Management: Tomorrow**

---

**KNOWLEDGE-BASED ASSISTANCE FOR SCIENCE VISUALIZATION AND ANALYSIS USING LARGE DISTRIBUTED DATABASES**

**GOAL / OBJECTIVES**

- Provide an environment for supporting development, demonstration and testing.
- Use machine learning techniques to ingest management data and infer knowledge.

**APPROACH / BUDGET**

- Design and develop the system along with the internal data sets for a knowledge-driven system.
- Use the data to provide algorithms and interfaces for developing science applications and enabling science-driven interface design for the system. This data can be used by the knowledge system to support data and interactions.

**FY91 ACCOMPLISHMENTS**

<table>
<thead>
<tr>
<th>Task Description</th>
<th>Start Date</th>
<th>End Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Data Set</td>
<td>09/91</td>
<td>09/91</td>
</tr>
<tr>
<td>Data Set Select &amp; Analysis</td>
<td>09/91</td>
<td>09/91</td>
</tr>
<tr>
<td>Knowledge Caching &amp; Mining Data</td>
<td>09/91</td>
<td>09/91</td>
</tr>
<tr>
<td>Data Set 1 (Example)</td>
<td>09/91</td>
<td>09/91</td>
</tr>
<tr>
<td>Data Set 2 (Example)</td>
<td>09/91</td>
<td>09/91</td>
</tr>
<tr>
<td>Data Set 3 (Example)</td>
<td>09/91</td>
<td>09/91</td>
</tr>
<tr>
<td>Data Set 4 (Example)</td>
<td>09/91</td>
<td>09/91</td>
</tr>
<tr>
<td>Data Set 5 (Example)</td>
<td>09/91</td>
<td>09/91</td>
</tr>
</tbody>
</table>

**FY92 MAJOR MILESTONES**

<table>
<thead>
<tr>
<th>Milestone</th>
<th>Start Date</th>
<th>End Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analysis Assistance</td>
<td>05/92</td>
<td>05/92</td>
</tr>
<tr>
<td>Data Set &amp; Retrieval</td>
<td>05/92</td>
<td>05/92</td>
</tr>
<tr>
<td>Additional Data Sets (Example)</td>
<td>05/92</td>
<td>05/92</td>
</tr>
<tr>
<td>Data Management Assistant</td>
<td>05/92</td>
<td>05/92</td>
</tr>
</tbody>
</table>
PRINCIPAL INVESTIGATORS PRESENTATIONS

MULTI-CHANNEL HOLOGRAPHIC BIFURCATIVE 
NUERAL NETWORK SYSTEM 
FOR REAL-TIME ADAPTIVE EOS DATA ANALYSIS

Hua-Kuang Liu/P.I. 
Jet Propulsion Laboratory 
Pasadena, CA 
(818) 354-8935

J. Diep/Co.I. 
K. Huang/Co.I.
Presentation To
APPLIED INFORMATION SYSTEMS RESEARCH WORKSHOP
July 22-24, 1991
Boulder, Colorado

MULTI-CHANNEL HOLOGRAPHIC BIFURCATIVE NEURAL NETWORK SYSTEM
FOR REAL-TIME ADAPTIVE EOS DATA ANALYSIS

Hua-Kuang Liu
Jet Propulsion Laboratory
Pasadena, CA 91109

OBJECTIVE

To research and develop a novel optical bifurcating neuromorphic pattern recognition system for making optical data array comparisons and to evaluate the use of the system for EOS data classification, reduction, analysis and other applications.

OUTLINE OF THE TALK

OBJECTIVE
TECHNICAL APPROACH
FUNDAMENTAL, PHYSICAL CHARACTERISTICS
IMPORTANCE
PROGRESS SUMMARY
FUTURE RESEARCH WORK PLAN

IMPORTANCE OF THE WORK

The bifurcating optical data and pattern recognition and classification system is based on the theory of the nonlinear wave scattering and interaction in photorefractive crystals and the nonlinear neuromorphic interconnection and activation. Theoretical understanding of the system is important to nonlinear science and technology.

The system combines advanced spatial light modulator technology with holographic and photorefractive material and devices. These elements function together to allow massively parallel processing at the speed of light. Experimental results of the system can lead to inventions important for technology breakthrough.

The multi-channel version of the system allows the input of data simultaneously from many different sensor sources and thus offers sensor analysis and fusion capability with high capacity and throughput. The system has important potential applications include parallel database search, image and signal understanding and analysis, natural language processing, in addition to real-time multi-channel adaptive EOS data analysis.
TECHNICAL APPROACH

- The face-space holographic interconnection and optical parallel processing capabilities with multi-channel input capacity.
- The multi-channel input is realizable by holographic optical elements (HOL) and high-speed switchable SLM's.
- Photorefractive crystals will be used for the formation of the bifurcating decision-making process.
- Input data are applied via optical spatial light modulators into the optical system.
- Output are naturally classified into multiple channels of pairs of branches.
- Separately controlled or integrated and organized decisions of the out information can be made based on the output.

PROGRESS SUMMARY (Fy91 Accomplishments)

(1) Completed a simplified model for the dynamic beam fanning phenomenon via inhomogeneous scattering centers in photorefractive crystals.

(2) Developed collaborative analysis activities with JPL colleagues including J. Gimenez, W. Tal, W. Fang, and Prof. K. Hwang of USC.

(3) Found interrelationship and areas of potential applications of the system in data format standardization program of NASA.

(4) Investigated basic neuromorphic associative retrieval processing that include the terminal attractor based Hopfield model in the aspects of the data format and optical implementation challenges.

PROGRESS SUMMARY (Fy91 Accomplishments) (Continued)

(5) Investigated hardware issues including key holographic optical elements and new electronically addressable spatial light modulators.

(6) Recent theoretical and experimental results on the spatial and spectral effects on noise fanning for photorefractive bifurcating process was submitted for publication at the OSA annual meeting in Nov. 1991.

J. R. Lin

(7) Presented an invited talk on nonlinear photorefractive optical processing at the Electrical Engineering Department of the City College of New York on June 24, 1991. He also discussed the projects on the new wavelet theory and its optical implementations with the researchers led by Prof. Y. Li at CCNY. They explored the wavelet RADAR signal processing. The technique may be important to NASA's SETI program.

PROGRESS SUMMARY (Fy91 Accomplishments) (Continued)

(8) A U. S. patent No. 5,015,354 entitled "Method and Apparatus for Second-Stage Beam Generation" was received. This invention is useful for the generation of control and interconnection matrices in real time and is important for future optical computer networking applications. The photorefractive crystal used in the patent was the same as the one used in the current NASA RTOP work. The invention is useful to the current project.

J. R. Lin

(9) Attended the Gordon Research Conference in the week of June 16, 1991 at the Plymouth College, Plymouth, NH. He presented a talk on the topic of self-amplification in optical pattern recognition. The talk was well-received by about 100 invited experts in the field of holography and information processing from various countries including Japan, France, Israel, Canada, Germany, USSR, China, Sweden, and USA. Many interesting and important techniques were discussed in a very informal atmosphere.
FUTURE RESEARCH WORK PLAN (FY92 Major Milestones)

(1) Experiment with the photorefractive crystals including BaTiO$_3$ and KNbO$_3$, on their capability of bistatic diffraction via the studies of beam polarization, incidence angle of the input beam, crystal orientation, and spatial light modulator characteristics.

(2) Investigate the analogy between the neuromorphic processing and the nonlinear dynamic wave coupling phenomenon in the photorefractive crystals for gaining a deeper understanding of the fundamental building block of the system.

(3) Investigate the data format requirement and the output utilization algorithm of a multichannel system for the preparation of the design of an application-oriented architecture design.
APPLIED INFORMATION SYSTEMS
RESEARCH PROGRAM WORKSHOP

PRINCIPAL INVESTIGATORS PRESENTATIONS
(P. M. SESSION)
PRINCIPAL INVESTIGATORS PRESENTATIONS

NATIONAL CENTER FOR
SUPERCOMPUTER APPLICATIONS

Matthew Arrott/P.I.
NCSA
University of Illinois
Urbana, IL
(217) 244-6833
Prototype
Investigate the issues and mechanics for possible integration of atmospheric data across scales and use of modelling and observational data

Divide and conquer
- Storm (supported by NCSA and NSF)
- Regional (supported by NASA)
- Global
- Ocean

Computational Application Model
- Data Generation
- Observation
- Modelling
- Analysis
- Inspection

Develop
- Integration of successful prototype efforts into a series of robust applications
NASA MEETING

NCSA GOAL
NCSA TEAM
PROJECT STRATEGY
COMPUTATIONAL RESEARCH ENVIRONMENT

NCSA GOAL

To develop comprehensive computational research environments through the use of evolving software technology.
Computational Research Environment

Base criteria
Interactive
Visual
Distributed
Extensible
Vendor supported

Silicon Graphic's "Explorer" Software
PRINCIPAL INVESTIGATORS PRESENTATIONS

COLORADO CENTER FOR ASTRODYNAMICS RESEARCH (CCAR)

William Emery, P.I.
Univ. of Colorado
Boulder, CO
(303) 492-8591
EOSDIS TESTBED: AN AVHRR DATA SYSTEM FOR SNOWPACK AND VEGETATION STUDIES
BIL H. EMERY, JEFF DOZIER, AND PAUL ROTAL

OBJECTIVE: TO PROVIDE AN END-TO-END DATA SYSTEM FOR THE COLLECTION, PROCESSING AND ANALYSIS OF AVHRR DATA FOR TWO SPECIFIC DISCIPLINES

1. COLLECT AND PROCESS AVHRR IMAGERY OF THE WESTERN U.S. AT CUCAP
2. STORE PROCESSED DATA AT NCAR ON MASS STORE
3. MAKE DATA AVAILABLE TO RESEARCH USER COMMUNITY ON OR OFFLINE AS REQUIRED
4. DEVELOP USER COMMUNITY INTERESTED IN AVHRR APPLICATIONS IN SNOW COVER AND VEGETATION
5. DEVELOP APPLICATION SOFTWARE ON MACS AND UNIX FOR THIS USER COMMUNITY; DISTRIBUTE SOFTWARE TO USERS (I.E., IMAGE PROCESSING, IMAGE NAVIGATION, IMAGE CALIBRATION, ETC.)

ORIGINAL PAGE IS OF POOR QUALITY
SATellite DATA PROCESSING SOFTWARE
AT CU/CCAR

IMAGE PROCESSING SOFTWARE

1. SISO: A generic image processing program that runs on SUN workstations under SUNTOOLS/SUNVIEW (also on DEC 110)
2. DECSSO: A version of SISO that runs on DEC WINDOWS
3. QTIFSISO: A version of SISO that runs on any X11 workstation
4. IMAGIC: A similar image processing program that runs on the MACINTOSH-II family of computers

AWARD DATA COLLECTION AND HANDLING

a. Antenna Operation (scheduling, tracking, ingest, archival)
b. Image Navigation and IR Calibration
c. TOYS profiles (ITT or 3)
d. Applications (skin SST, SST motion, ice edge and concentration, ice motion, vegetation, cloud top temperature, cloud height, 3D perspective, etc.)

OTHER SATellite DATA HANDLING

a. SSM/I (ice and water vapor), LANDSAT & SPOT (3D mapping), SAR (ice mapping)
NCAR Mass Storage Systems (MSS)

**Usage Data**
- 105,000 tape cartridges in use
- Total data stored 19 Tbytes
- Over 722,000 files
- Average file length 26.3 MB
- < two-minute delivery

---

**Mass Storage System Hierarchy**

- **MSS Controller and Directory**
  - Disk Farm
    - 120 GBytes
    - 10 sec access
  - STK 4400 ACS
    - 6K cartridges
    - About 1 minute access
  - Open Shelves
    - 61K cartridges
    - 2-10 minute access

**Control Path**
- Y-MP8/64 UNICOS

**Fastpath**

---

**Remote Access**

- **Front-end Computer**
  - Shavano
  - Divisional Computer
- **Blitz-Up Access**
  - Including 13,400 Numbers
  - MASnet
  - NSFNET/Internet
  - BITNET
  - NASA SN
  - NASA SPAN

---

**Wide Area Network Connections at NCAR**

North America Map showing connections at NCAR.
As of this date, there are 499 days of data starting on March 3, 1989 – July 18, 1991. The first images consisted only of channels 1 and 2. Now all channels are being archived.

Current Number Of Images
Images on line: 1,200
Size of Image: 1 Megabyte
Storage Size: 1.2 Gbytes

Monthly Number Of Images
Images processed per month: 120/mth
Processed Data Size: 120 Megabytes/mth

Daily Number Of Images
Images Processed Per Day: 4/day
Size of Image: 1 Megabyte
Storage size: 4 Megabytes/day
Channels 1, 2, 3, 4, 5

Image Projection
Images at End Of Project: 4,200
Image Storage Size: 4.2 Gbytes
PRINCIPAL INVESTIGATORS PRESENTATIONS

GEOGRAPHIC INFORMATION SYSTEM FOR FUSION AND ANALYSIS OF HIGH-RESOLUTION REMOTE SENSING AND GROUND TRUTH DATA

Anthony Freeman/P.I.
Jet Propulsion Laboratory
Pasadena, CA
(818) 354-1887

Pascale Dubois/Co.I.
Franz Leberl/Co.I.
L. Norikane/Co.I.
JoBea Way/Co.I.
SCIENTIFIC OBJECTIVES

Conceived with the potentially massive volumes of data from remote sensing instruments and ground data collection for this site, the application scenario might ask the following questions:

i) How do I manipulate such a large data set?

ii) How do I assess its information content?

iii) How do I find the optimum combinations of data to study changes in a given biogeochemical parameter?

iv) How do I visualize the results?

v) How do I validate my models relating in-situ measurements to the remotely sensed data?

vi) What scientific generalizations can be made from this study, even though there are only two case studies?

vii) What important issues emerge which would impact EOS GIS and Forest/Resource research with highly multi-dimensional data?

viii) To what extent do the GIS and Image Processing systems, both of which were designed to work with scientific data in raster format, and which are essentially state-of-the-art, show significant weaknesses in handling large volumes of high-resolution data? This is a practical issue of great importance to the remote sensing community.
<table>
<thead>
<tr>
<th>Task</th>
<th>Completion Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>7. Data Preparation for Case Study II</td>
<td></td>
</tr>
<tr>
<td>SAR data geocoding (Vexcal)</td>
<td>07/92</td>
</tr>
<tr>
<td>AVORS data (JPL)</td>
<td>08/92</td>
</tr>
<tr>
<td>Ground Truth data formatting (JPL)</td>
<td>09/92</td>
</tr>
<tr>
<td>Simulated radar data generation (UCSB)</td>
<td>09/92</td>
</tr>
<tr>
<td>8. Data Entry (Case Study II) (JPL)</td>
<td>10/92</td>
</tr>
<tr>
<td>9. Data Analysis (Case Study II) (JPL, UCSB)</td>
<td></td>
</tr>
<tr>
<td>Extract radar parameters from data</td>
<td>11/92</td>
</tr>
<tr>
<td>Cross-correlate data sets</td>
<td>12/92</td>
</tr>
<tr>
<td>Principal components analysis</td>
<td>1/93</td>
</tr>
<tr>
<td>Produce Classification maps</td>
<td>2/93</td>
</tr>
<tr>
<td>Sensitivity analysis</td>
<td>3/93</td>
</tr>
<tr>
<td>Radar model validation</td>
<td>05/93</td>
</tr>
<tr>
<td>Optimize visualization</td>
<td>06/93</td>
</tr>
<tr>
<td>Report/journal paper</td>
<td>10/94</td>
</tr>
<tr>
<td>10. S/W Documentation (JPL, VEXCEL)</td>
<td></td>
</tr>
<tr>
<td>Write S/W Description Document</td>
<td>11/94</td>
</tr>
</tbody>
</table>

**GEOGRAPHIC INFORMATION SYSTEM**

- **STATUS OF UNIX VICAR/IBIS S/W:**
  - Starting point was a set of UNIX VICAR routines translated by Arizona State University.
  - Currently have 45 VICAR applications programs running.
  - 8 of them are IBIS routines (need ~30 IBIS routines in total).
  - Current capabilities:
    - Can read radar image data into the VICAR format.
    - Can set up - image files (RASTER), graphics files (VECTOR), VP files (TABULAR).
    - Perform image rotation, stretching.
    - Image rubber-sheeting (for registration).
    - Draw polygons on display and convert to image files.
    - Fill polygons with data attributes.

**DEVELOPMENTS:**

1. Case Study II: Report/Journal paper
2. Case Study II: Report/Journal paper
3. S/W Description Document
JPL GEOGRAPHIC INFORMATION SYSTEM

- DATA PREPARATION
  - CASE STUDY 1: BONANZA CREEK EXPERIMENTAL FOREST, AK
    - 2 SEASAT IMAGES
    - 2 SPOT IMAGES
    - 10 AIRSAR FRAMES (~ 50 IMAGES)
    - 1 DIGITAL ELEVATION MODEL (USGS)
    - GROUND TRUTH SURVEY DATA
    - AERIAL PHOTOS
    - FURTHER AIRSAR DATA (5/91), GROUND TRUTH DATA AND ERS-1 SAR DATA WILL BE RECEIVED IN NEAR FUTURE
  - CASE STUDY 2: BELIZE RAIN FOREST
    - 5 AIRSAR FRAMES (~ 45 IMAGES)
    - AVIRIS DATA
    - SOME GROUND TRUTH DATA

JPL GEOGRAPHIC INFORMATION SYSTEM

- UCSB RADAR MODEL
  - UCSB RADAR MODEL(S) ARE INSTALLED ON THE SPARC STATION AT JPL
  - MODELS WERE DEVELOPED FOR STUDIES OF MT. SHASTA FORESTS
  - CURRENT EFFORTS ARE FOCUSED ON:
    - ENSURING THAT THE CONTINUOUS AND DISCRETE FOREST CANOPY MODELS AGREE IN THE LIMIT
    - ADAPTING THE MODEL FOR THE TYPE OF TREES FOUND IN ALASKA
    - GENERATING SIMULATED RADAR MEASUREMENTS FROM BONANZA CREEK GROUND TRUTH DATA
    - SIMULATED RADAR MEASUREMENTS FROM THE MODEL WILL BE USED TO FILL POLYGONS IN THE GIS TO GENERATE SIMULATED IMAGES

JPL GEOGRAPHIC INFORMATION SYSTEM

- UCSB RADAR MODEL - BACKSCATTER COMPONENTS

ORIGINAL PAGE IS OF POOR QUALITY
CLASSIFICATION ALGORITHMS (FOR RADAR DATA)

- FIRST STEP IS THE REDUCTION OF POLARIMETRIC RADAR IMAGE DATA TO FIVE BASIC QUANTITIES FOR EACH FREQUENCY (HH, HV, VV, ARG(HHV)) AND (HHV*)
- THIS GIVES 15 'CHANNELS' FOR EACH RADAR IMAGE SET
- THEN USE LINEAR DISCRIMINANT ANALYSIS TO COME UP WITH A SUPERVISED CLASSIFICATION OF THE IMAGE
- APPROACH WAS TESTED ON AN AIRSAR IMAGE OF AN AGRICULTURAL TEST SITE IN THE NETHERLANDS WITH 15 DIFFERENT CLASSES OF TARGET

VISUALIZATION S/W AND ANALYSIS TOOLS

- PVWAVE IS INSTALLED ON THE SPARC3 STATION (GRAPHICS, RGB, PERSPECTIVES, ETC.)
- NO VF WITH VICAR YET
- POLTOOL S/W - FOR POLARIMETRIC RADAR DATA ANALYSIS HAS BEEN INSTALLED, ALLOWS PLOTS OF POLARIZATION SIGNATURES, SYNTHESIZED IMAGES OF ARBITRARY POLARIZATIONS
- NO VF WITH VICAR YET
- 'LIGHT TABLE' S/W DEVELOPED BY VEXCEL - ALLOWS USER TO 'ROAM' THROUGH LARGE IMAGES USING MOUSE
PRINCIPAL INVESTIGATORS PRESENTATIONS

SAVS
A SPACE DATA ANALYSIS AND VISUALIZATION SYSTEM

Edward Szuszczewicz/P.I.
SAIC
Laboratory for Atmospheric and Space Science
McLean, VA
(703) 734-5516

Charles C. Goodrich/Co.I.
Alan Mankofsky/Co.I.
APPLIED INFORMATION SYSTEMS
RESEARCH WORKSHOP

NOTES:

THIS PAGE LEFT INTENTIONALLY BLANK
SAVS
A SPACE DATA ANALYSIS AND VISUALIZATION SYSTEM

Presented by
E. P. Szuszczywicz

Laboratory for Atmospheric and Space Science
Science Applications International Corporation
McLean, VA

Applied Information Systems Workshop
July 22-24, 1981

CONTRIBUTING ORGANIZATIONS

SAIC
Laboratory for Atmospheric and Space Science

STARDENT
Computer Inc.

UNIVERSITY OF MD
Department of Physics and Astronomy

SAVS: A DATA ACQUISITION, MANIPULATION, ANALYSIS AND VISUALIZATION SYSTEM

MAJOR COMPONENTS:

INNOVATIVE VISUALIZATION SOFTWARE (AVS)

ADVANCED DATABASE TECHNIQUES (DAVID)

SET OF MATHEMATICAL, STATISTICAL, ANALYTICAL AND IMAGE PROCESSING TOOLS

STRONGLY DEVELOPED SENSE OF THE SCIENTIFIC REQUIREMENTS
**THE NASA VISION FOR THE COMING DECADE**

**The Vision:** This vision includes a comprehensive attempt to understand the flow of energy into and through the geospace environment, with the goal that "quantitative study of the geospace environment, created by the interplay of solar and terrestrial processes, will be progressing toward a full-scale predictive stage" [from OASS's Strategy Document (1989)]

**Implementation:** Coordinated multi-spacecraft missions (e.g. ISTP) to collect "in situ" plasma and field measurements throughout the magnetosphere, the ionosphere and the interplanetary medium.

- Requires imaging of the thermosphere, the ionosphere and the Sun, and coordinated ground-based magnetometer, ionosonde, radar, and photometric observations
- Requires interactive database handling and visualization and analysis techniques
STATEMENT OF THE PROBLEM

Increased focus on large-scale system phenomena

The cross-disciplinary nature of many investigations

Higher data rates and projections of increased volumes of data

Enhanced measurement capabilities (need for cross-correlation of global images with "in situ" and ground-based observations)

Increasing number of large-scale 3-D numerical models available as analytical tools for data synthesis and interpretation

CHALLENGES FACED

Interactive data analysis and graphics environment

Ability to cross disciplinary boundaries with ease and understanding

Ability to "compress" data into a visually-organized form optimized for analysis and interpretation

Easy-to-use mathematical, statistical and image processing tools

Tools to obtain data sets from remote archives

Access to empirical and numerical model results to correlate with the data and assist in data analysis and interpretation

An integrated user-friendly system they can afford
**PROJECTED CAPABILITIES**

**Visualization:** 1-, 2-, and 3-dimensional displays, along with animation, compression, warping and slicing

**Analytical Tools:** Generic mathematical and statistical techniques along with the ability to use large-scale models for interactive interpretation of large-volume data sets

**Platform:** Implemented on Sun and DEC UNIX workstations and on the Stardent graphics supercomputer

**Immediate Application:** Solar, heliospheric, magnetospheric, ionospheric and thermospheric physics, with extensions to a broader range of planetary, interplanetary and astrophysical environments

---

**THE VISUALIZATION SYSTEM (AVS)**

AVS is designed for a distributed network environment...single system or a network of systems.

**Complete Image Display Capabilities:**
- Real-time pan and zoom
- Rotation and transformation
- Flipbook animation
- Support for 8-bit, 24-bit and floating point images

**Image Filters Include:**
- Look-up table operations (pseudo-coloring, histogram balancing, data resizing, interpolation, cropping and sampling)

**Provides a variety of tools for rendering volume data; a real-time iso-surface generator; etc.**

---

**Start customization of the AVS Interface**

Address the porting issues

Adapt AVS to NASA applications of 1-, 2-, and 3-D displays

Implement basic mathematical and statistical functions

Support locally-resident NASA data sets

Begin remote data handling capabilities

Implement relevant large-scale numerical models
PRINCIPAL INVESTIGATORS PRESENTATIONS

A PLANETARY VERSION OF PC-McIDAS

Sanjay S. Limaye/P.I.
Univ. of Wisconsin-Madison
Madison, WI
(608) 262-9541

Mike Martin/Co.I.
R. S. Saunders/Co.I.
L. A. Sromovsky/Co.I.
APPLIED INFORMATIONS SYSTEMS
RESEARCH WORKSHOP

NOTES:

THIS PAGE LEFT INTENTIONALLY BLANK
A Planetary Version of PC-McIDAS
Applied Information Systems Workshop

Sanjay S. Linaye (UW-Madison), Principal Investigator
I.A. Smolensky (UW-Madison), R.S. Saunders (JPL), Mike Martin (JPL), Co-Investigators

- WHAT IS PC-McIDAS?
  A CONCEPT SYSTEM FOR ANALYSIS OF ATMOSPHERIC IMAGING DATA

- WHY A PLANETARY VERSION?
  THE NATURE AND VALUE OF THE PLANETARY DATA AND UNIQUENESS

- WHAT
  - UBS AND OS/2 VERSIONS
  - UBS
  - INITIAL DEVELOPMENT AT SSEC. TESTING AT JPL FOR MAGELLAN APPLICATIONS
  - FUTURE: COULD SUPPORT NASA OBSERVERS AND CASSINI MISSIONS

McIDAS Workstations
- INTEL 8088 based "dump" terminals
  - used on the Harris versions
  - Analog disk and digital cam display memory
- McIDAS - OS/2 Version - first "dump" terminal
  - Dual mode operation: local mode and host mode
  - Digital image display memory
- McIDAS-OS/2 Version - first multi-tasking implementation
  - EGA/VGA support for local display
  - used during the Voyager Neptune encounter
  - portable for field work

Wide Workstation (WWS) support
- Additional image display memory
- 16 bit display memory
- dual channel video support
- McIDAS-UNIX Version - X-Windows support

McIDAS Hardware Developed at SSEC

SATELLITE DATA INGESTION

FM MOUSE WORKSTATION
DIGITAL CASSETTE ARCHIVE FOR GOES DATA

Real Time Data Available on McIDAS

GOES satellite images 24 hours/day
- McIDAS visible and infrared images
- NOAA-10 polar orbit AVHRR or "L" data (NEX, 150)
- NOAA-11 polar orbit AVHRR or "L" data (NEX, 150)
- GOES-Iband full disc images
  - resolution: 700 km
  - GOES-IR band full disc images
  - resolution: 700 km

Cassini and PHSKMC cameras
- Harris, Minolta
- Black and white lenses

McIDAS and PC-McIDAS

DATA COMPUER INTERACTIVE DATA ACCESS SYSTEM
AN ANALYSIS TOOL FOR INTEGRATION OF WEATHER DATA FROM SATELLITES AND CONVENTIONAL SOURCES.
SOFTWARE AND HARDWARE COMPONENTS DEVELOPED AT SPACE SCIENCE AND ENGINEERING CENTER (SSEC)
INSTALLED GLOBALIY IN OPERATIONAL ENVIRONMENTS. EFFORT BEGAN IN EARLY 1980'S

- HARRIS DATACUBE ACCESS VERSION
  384K memory, 120 No Peripherals Dish Storage
  3 807/1600 bpi tape drives
  2 user image display terminals
  used during national 10 Venus Flyby

- HARRIS 16 NETWORK VERSION
  1-2 Database Managers and 6 Applications Computers
  768K core memory on each
  16 No peripheral storage on "main" and "data"
  140 No peripheral storage on each "sub" or "daughter"
  2 807/1600 tape drives
  1 707/1600 tape drive
  2+4 display terminals on each (as well as Remote)
  - use during the Voyager Flyby of Jupiter and Saturn

- IBM MVS PSS 70
  - Memory, 32 No real mes
  - 76 peripheral storage
  - 3 807/1600 tape drives
  - 2 707/1600 tape drives
  - communication support
  - 15 terminals globally
  - use during grand and great, voyager encounters

- MSW Working Stations
  - used on McIDAS Workstations
  - DIGITAL Cassette Archive for GOES Data
  - Real Time Data Available on McIDAS
  - GOES satellite images 24 hours/day
  - McIDAS visible and infrared images
  - NOAA-10 polar orbit AVHRR or "L" data (NEX, 150)
  - NOAA-11 polar orbit AVHRR or "L" data (NEX, 150)
  - GOES-Iband full disc images
  - GOES-IR band full disc images
    - resolution: 700 km
    - GOES-IR band full disc images
      - resolution: 700 km
  - Cassini and PHSKMC cameras
    - Harris, Minolta
    - Black and white lenses
  - McIDAS and PC-McIDAS
  - DATA COMPUER INTERACTIVE DATA ACCESS SYSTEM
  - AN ANALYSIS TOOL FOR INTEGRATION OF WEATHER DATA FROM SATELLITES AND CONVENTIONAL SOURCES.
  - SOFTWARE AND HARDWARE COMPONENTS DEVELOPED AT SPACE SCIENCE AND ENGINEERING CENTER (SSEC)
  - INSTALLED GLOBALIY IN OPERATIONAL ENVIRONMENTS. EFFORT BEGAN IN EARLY 1980'S

- HARRIS DATACUBE ACCESS VERSION
  - 384K memory, 120 No Peripherals Dish Storage
  - 3 807/1600 bpi tape drives
  - 2 user image display terminals
  - used during national 10 Venus Flyby

- HARRIS 16 NETWORK VERSION
  - 1-2 Database Managers and 6 Applications Computers
  - 768K core memory on each
  - 16 No peripheral storage on "main" and "data"
  - 140 No peripheral storage on each "sub" or "daughter"
  - 2 807/1600 tape drives
  - 1 707/1600 tape drive
  - 2+4 display terminals on each (as well as Remote)
  - use during the Voyager Flyby of Jupiter and Saturn

- IBM MVS PSS 70
  - Memory, 32 No real mes
  - 76 peripheral storage
  - 3 807/1600 tape drives
  - 2 707/1600 tape drives
  - communication support
  - 15 terminals globally
  - use during grand and great, voyager encounters
Archived Data available on McIDAS

- Continuous WVR digital imagery data going back to 1977, some to 1976
- High-resolution images of Venus
- Voyager 1 and 2 images of Jupiter, Saturn (Digital tapes)
- Voyager 2 images of Uranus and Neptune (Digital tapes)

McIDAS Installed Base

Mainframe Facilities Installed at:
- NASA/MSFC facilities (HGF, HSC, JSC, MSC, NASA) for shuttle operations
- NOAA (NMC, MSC, NASA)
- Private Companies (Federal Express, MFI)
- Universities (USU, UC-Chico)
- Weather Services facilities in Spain, China, Australia, CANADA

Remote Workstations:
- SME-Albany
- UW-Madison
- University PC McIDAS at 500 atmospheric science departments at colleges and universities

Planetary PC McIDAS

- Many different kinds of data and spacecraft failure data
- Nature of analysis different from operational weather satellite data
- Different research goals and disciplines
- Not a solution for all planetary data
- Build on the core McIDAS applications and utilities
- Differences between McIDAS and other similar systems

Enhancements/Modifications

- Bridge to Speculator kernels for planetary navigation
- Direct support of CD-ROM database
- Improved program product observation reporting
- New applications for geologic data analysis from planetary images
- Multicolor composites and animations
- Documentation of planetary programs/algorithms
- User guide
COMPARISON OF MCIDAS AND VICAR

- DIFFERENT OPERATING ENVIRONMENTS
- MULTITASKING VS. SINGLE APPLICATION MODE
- MAJOR FRAME VS. SINGLE FRAME DISPLAY
- DIFFERENCES IN HANDLING OF USER FILES (NATIVE OPERATING SYSTEM VS. MCIDAS OPERATING ENVIRONMENT)
- USER APPLICATIONS TYPICALLY BUILT MODULAR RATHER THAN MONOLITHIC
- USER PROGRAMMABLE CHAINING OF OPERATIONS ('MACROS')
- USER APPLICATION 'PROFILING' FOR APPLICATION PROGRAM INPUTS
- DIFFERENT SYSTEM ARCHITECTURE PHILOSOPHIES
- SIMPLIFIED FILE STRUCTURE ('IN' FILES)
- GRIDDED DATA STRUCTURE ('GRID' FILES)
- SCHEMA DATA STRUCTURE ('MO' FILES)
- UNIFIED SINGLE DATA NAVIGATION ACCESS
INTERACTIVE INTERFACE TO NCAR GRAPHICS

Robert Lackman, P.I.
NCAR
Boulder, CO
(303) 497-1224
NOTES:
Interactive Interface
to
NCAR Graphics

Bob Lackman
Scientific Visualization Group

NCAR Graphics Features

- Standards conforming
  - ANSI C and FORTRAN 77
  - GKS and CGM
- Portable
  - UNIX and VMS
- Scientifically oriented
- International distribution
  - University, government, and industry
  - Since 1971
  - Approximately 1500 sites
- Evolving
  - UNIX Version 3.1.1

NCAR Graphics Current State

- Features
- Components
- Interface

NCAR Graphics Components

- Test drivers for all utilities
- Over two dozen high-level utilities
- An intermediate interface library (SPPS)
- NCAR GKS-0A (Optional)
- Fortran and C translators (Optional)
- Miscellaneous tools and utilities
- Documentation
NCAR Graphics Utilities

- Contouring
  - Regular and irregular grids
  - Various quality levels
- Maps
  - Outlines or filled
  - Ten popular transformations

NCAR Graphics Utilities (continued)

- X-Y plotting
- Surfaces and isosurfaces
- Streamlines and velocity vectors
- Histograms
- Stroked-character generators
- Others

NCAR Graphics Programmatic Interface

User application programs
NCAR Graphics utilities
  - X-Y graphing
  - Contouring
  - Maps
  - etc.
  - SPPS library
  - GKS (NCAR or other)

CGM

NCAR Graphics

NASA Proposal Enhancement

An Interactive Interface
A Typical Application

GUI

X Toolkit

Text-based UI

NCAR Graphics

Design Considerations
- Ease of use
- Portability
- Enhanced functionality
- Standards
- Interoperability
  - Module compatibility with AVS, mpf, khoros, IDL, ...
- Distribution and support

Overview of NCAR Interactive

User Interface and CDL Interpreter

Stand Alone CDL Interpreter

Stand Alone Postscript Interface

GUI

CDL file

CDL Interpreter

ASCII Editor

CDL file

CDL Interpreter

Data file

Data file

Data file

CDL + Command Description Language
PRINCIPAL INVESTIGATORS PRESENTATIONS

VIEWCACHE: AN INCREMENTAL DATABASE ACCESS METHOD FOR AUTONOMOUS INTEROPERABLE DATABASES

Nick Roussopoulos/P.I.
Univ. of MD
College Park, MD

Timoleon Sellis/Co.I.
VIEWCACHE: AN INCREMENTAL DATABASE ACCESS METHOD FOR AUTONOMOUS INTEROPERABLE DATABASES

Principal Investigator:
Nick Roussopoulos

Co-Principal Investigator:
Timoleon Sellis

Department of Computer Science
and
Institute for Advanced Computer Studies
University of Maryland
College Park, MD 20742

July 5, 1991

Submitted to the First Workshop of the
Applied Information Systems Research Program
Boulder, Colorado
July 22-24, 1991
PROJECT SUMMARY

The objective of this work is to illustrate the concept of incremental access to distributed databases. An experimental database management system, ADMS, which has been developed at the University of Maryland, in College Park, uses VIEWCACHE, a novel database access method based on incremental search. VIEWCACHE is a pointer-based access method that provides a uniform interface for accessing distributed databases and catalogues. The compactness of the pointer structures formed during database browsing and the incremental access method allow the user to search and do inter-database cross-referencing with no actual data movement between database sites. Once the search is complete, the set of collected pointers pointing to the desired data are dereferenced.

One of the most attractive features of VIEWCACHE is its versatility in providing External Multi Gateway Access to commercial database servers, such as INGRES, ORACLE, and SYBASE, supporting existing and independently maintained databases. We have designed and implemented a Client-Server Database Management System Architecture which utilizes powerful workstations for managing inter-database queries and cached data. The workstation environment provides the "glue" for interoperating otherwise foreign environments. Gateway database access methods are enhanced by the incremental techniques of VIEWCACHE for caching local subsets of useful data.

VIEWCACHE is especially suited for distributed scientific databases maintained on commercial database systems. Scientists who search such large catalogued databases, cannot write complete queries because search is defined during browsing. Current computer and database technology allows a user only to browse a single database at a time. It is, however, extremely useful to be able to browse and correlate data sets from multiple data servers. During browsing, there is no need to move data from one database to the next to compare and do cross-referencing; instead, VIEWCACHE creates a working set of pointers to the data sets or records that are candidates for the final retrieval. The working set is continuously refined until the user finds the exact data sets he needs. Furthermore, VIEWCACHE allows the creation of mixed breed views from several data sets and caching of them onto the users workstations. Such new data sets provide value added to the exiting data.

Another feature that ADMS supports is spatial search on image data sets. This capability is not offered by commercial DBMS but it was absolutely necessary for searching for objects in a giver "view
field.” Spatial search is based on R+-trees, an established access method. The indexing technique we are planning to support, provides for very efficient search over large databases of spatial objects.

The concept of VIEWCACHE has been demonstrated with the various traditional database benchmarks (like the Wisconsin benchmark); we propose to investigate and develop extensions to ADMS to make our system suitable for accessing very large space data sets. Clearly, NASA’s huge numbers of distributed data sets collected from space and ground stations cannot be supported by existing distributed commercial distributed database systems because they require unnecessary and bulky data movements. VIEWCACHE, on the other hand, provides an inexpensive processing without interrupting the independence of existing and autonomous databases.
TECHNICAL OBJECTIVES

Our objective is to illustrate the potential of the incremental access of VIEWCACHE in a real NASA environment of distributed databases, in particular on a collection of Astrophysics databases. The compactness of the VIEWCACHE pointer structures formed during database browsing and the incremental access method allow the user to search and do inter-database cross-referencing with no actual data movement between database sites. Once the search is complete, the set of collected pointers pointing to the desired data are cached. This will provide a uniform user interface to a large number of databases using the VIEWCACHE concept, and the Client-Server Architecture. We will also examine ways to interface VIEWCACHE with the DAVID system and its library-based access methods for providing access to Heterogeneous and Distributed Databases.

In more detail, the following technical objectives will be sought during the next three years:

1. An extension of the physical pointer structure of VIEWCACH to a logical one. This will permit VIEWCACHE to do inter-machine pointer referencing across a communication network (Local or Wide Area Networks). Access of heterogeneous DBMSs will be done by gateway software based on the SQL language.

2. Investigate features and characteristics of various Client-Server DBMS architectures and measure their performance under a wide variety of key parameters, including speed of the storage media, speed of the CPUs, network transfer rates.

3. Incorporate a spatial access methods and accordingly extend the SQL query language to include spatial search and operators which deal with imagery databases.

4. Design and develop a user interface and the appropriate tools to facilitate handling and distribution of data sets and documents. The Astrophysics environment will be the first to target this interface. Other environments will also be investigated and appropriate tailoring of the interface will be attempted.

5. Investigate ways to interface VIEWCACHE with DAVID, the Distributed Heterogeneous Database System developed at Goddard.

Work on this project has already started on the items 1, 2, and 3 from above. We have been experimenting with Oracle DBMS on the subject of logical pointers. We have also developed simulation
packages for evaluating alternative Client-Server Architectures for DBMSs. Experiments are currently being conducted with very large simulation runs.
PRINCIPAL INVESTIGATORS PRESENTATIONS

SPATIAL ANALYSIS AND MODELING SYSTEMS (SAMS)

Charles Vermillion/P.I.
Univ. of Chicago
Chicago, IL

Paul Chan/Co.I.
John Hill/Co.I.
Robert Jaske/Co.I.
Gilbert Rochon/Co.I.
Fran Stetina/Co.I.
STATUS REPORT

SPATIAL ANALYSIS AND MODELING SYSTEM (SAMS)

P.I. - Charles Vermillion, GSFC, Code 970.1
Co.I. - Fran Stetina, GSFC, Code 970.1
- Paul Chan, Science Systems & Applications, Inc.
- John Hill, Houston Advanced Research Center

The objective of this project is to develop a uniform environmental data gathering and distribution system to support a) emergency management for environmental disasters, and b) the calibration and validation of remotely sensed data. Initial activities will be to select a data test site and to demonstrate multi-discipline applications using simulated or satellite data in a non real-time mode.

The investigators have arranged collaboration with the CERL Laboratory of the Corps of Engineer (COE) in integrating its Geographic Information System (GIS), the Geographic Resources Analysis Support System (GRASS), into SAMS. CERL will also provide hydrological models (HEC-I and II) to be integrated into SAMS. The collaboration with the Federal Emergency Management Agency (FEMA) will allow this project to tap into FEMA's resources in GIS and emergency management tools.

We choose to use rainfall and flooding as the testbeds for the SAMS concept because of the abundance of data and the availability of models. We will integrate into SAMS (in a UNIX, GRASS and X-Windows environment) capability to display and process GOES data and analyze GOES generated rain-rate maps. GRASS is compatible with the majority of the data to be input to the selected hydrologic model (i.e. topography, land-use, soils, rainfall, stream gauge, etc.). The choice of the data test site has been narrowed down to West Virginia or Galveston, Tx, based on the availability of test data.

The expected results in six months are:

a) Design concept of SAMS,
b) Demonstration of a pilot module, and
c) Project implementation plan based on the pilot demonstration.
PRINCIPAL INVESTIGATORS PRESENTATIONS

INTRODUCTION TO CADET CENTER FOR ADVANCED DATA EVALUATION TECHNOLOGY

Cathy Schulbach/P.I.
NASA Ames Research Center
Moffet Field, CA

C. Jorgensen/Co.I.
INTRODUCTION TO CADET
Center for Advanced Data Evaluation Technology

Cathy Schulbach /c_schulbach
7/25/91

OUTLINE
- INTRODUCTION
- GOAL
- CADET OVERVIEW
- OBJECTIVES
- APPROACH
- KEY CADET FOCUS
- FUTURE MILESTONES

TECHNOLOGY PROBLEM
approx. 1% of EOS sensor data can be transmitted

HUMAN PROBLEM
DATA COLLECTION
DATA UNDERSTANDING
IMPLICATIONS

- The "BRUTE FORCE" method of trying to collect, save, and analyze "everything" is intractable and may not produce the best science.
- Smarter and more automated approaches are required.

GOAL

- DEVELOP AND DEMONSTRATE IMPROVED METHODS OF HANDLING LARGE SCIENTIFIC DATA FLOWS RESULTING FROM CURRENT AND PLANNED NASA MISSIONS.
  - Real-time
  - On-board

OBJECTIVES

- Integrate existing work in neural networks, photonics, parallel processing, and dependable systems.
- Develop, test, and evaluate new concepts for model-capturing and novelty detection mechanisms (e.g. neural net learning).
- Implement advanced data analysis technology using advanced processors (e.g. optical processors, IWARP, ES-KIT).
**ELEMENTS OF THE MODELLING PROCESS**

- Developing a model from data
  - numerical analysis
  - Bayesian estimation
  - neural network learning
- Applying a model
  - expectations vs. observations—prediction
  - novelty detection—Δ change
- Developing algorithms
  - mapping to computer architectures
  - optimizing performance

**FUTURE MILESTONES**

**FY92**
- Use the Failure Environment Analysis Tool (FEAT) to model a sensor system.
- Demonstrate the use of sensor overlays to display earth science data.
- Demonstrate the ability of a neural network front end to automatically capture underlying regularities in a real-world sample of spectrographic sensor data.

**FY93**
- Develop parallel code implementing a novelty filter to redirect a data-recording device to dynamically respond to changes in measured, earth science phenomena.

**FY94**
- Validate the above concepts on the Ames Advanced Data Systems and Software Test Facility.
APPLIED INFORMATION SYSTEMS
RESEARCH PROGRAM WORKSHOP

PROGRAM PRESENTATIONS
FLIGHT PROJECTS OFFICE
INFORMATION SYSTEMS TESTBED
(FIST)

Patricia Liggett
Jet Propulsion Laboratory
Pasadena, CA
(818) 357-4619
Flight Projects Office
Information Systems Testbed (FIST)

P. Lippard

---

To perform technology evaluation and prototyping of information systems to support SFOC and JPL flight projects in order to reduce risk in the development of operational data systems for such projects.

---

FIST reduces risk by providing support for:

- Requirements Clarification
- Design Validation
- Design Evolution
Criteria

- Applied research and development using commercially available products and systems.
- Results of prototypes and evaluations are reported in the FIST Quarterly Report and presented during demonstrations.
- FIST guidelines are provided by SFOC compatibility requirements.

Quarterly Reports

- FIST Quarterly and SFOC Prototype Interim Reports are available through:
  FPSO Library
  Mail Stop T-1607
  Jet Propulsion Laboratory
  4800 Oak Grove Drive
  Pasadena, California 91109

SYSTEM BUILDING BLOCKS

- Open Systems Architecture
- Server/Client Model
- Iterative Design and Development
- Commercial and Public Products

OPEN SYSTEMS ARCHITECTURE

- Limited Constraints
- Well Defined Interfaces
- Data Format Agreement
- Soft Serve versus Hard Freeze
SERVER/CLIENT MODEL

- Limited and Well Defined Interdependence
- Loosely Coupled
- Services Provided to Requesting Client
- Role Reversal

ITERATIVE DESIGN AND DEVELOPMENT

- Provides Users with Early View of System
- Adaptive to Changing Technology and Requirements

Technology Transfer

Feed Forward

Prototype - Pilot - Project

Feed Backward

May 10, 1991