

NSSDC OMNIWeb: The First Space Physics WWW-Based Data Browsing and Retrieval System

The Third International WWW Conference '95: Technology, Tools, and Applications
Fraunhofer Institute for Computer Graphics

April 10 - 14, 1995

G. Jason Mathews
Syed S. Towheed

National Space Science Data Center
National Aeronautics and Space Administration
Goddard Space Flight Center
Greenbelt, Maryland 20771, U.S.A.

Abstract

The National Space Science Data Center (NSSDC), located at NASA's Goddard Space Flight Center, provides access to a wide variety of data from NASA spaceflight missions. Traditionally, the NSSDC has made data available in a variety of hard media and via on-line systems. However, the WWW provides an exciting new enhancement by allowing users to examine and browse the data before retrieval. This paper presents OMNIWeb, a WWW-based system that allows users to produce plots and retrieve data. The browsing and retrieving capability was designed to aid researchers in identifying trends and to obtain the data through an uninterrupted process.

Keywords: Data Analysis, Data Retrieval, Scientific Data Server, Client/Server, User Interface Design, CDF, IDL.

1. Introduction

Users of science data are often in the difficult position of having to perform a suite of tasks to obtain data. These tasks can be broadly defined in three categories:

1. Finding the source for the data.
2. Obtaining descriptions of the data.
3. Obtaining the data.

Once the source for the data is located, users are still faced with the task of obtaining sufficient information about the data (metadata) to determine if the data will be adequate for their needs. Often, the metadata reside on some on-line information server (NASA Master Directory, NASA Master Catalog) that provides high level descriptions of the data; however, rarely is the information server coupled with the data server (NDADS, NODIS). Even after the user has found the data archive and obtained the metadata, there is still the additional task of determining which portion of the data is most likely to contain the phenomena or feature of interest to the user. With textual browsing, the user still cannot visualize the features of interest by just looking at a listing of numbers and, therefore, is most likely to request the entire data set.

OMNIWeb is a prototype data system that is the next logical development of World Wide Web (WWW) applications in the NASA environment where the focus is to provide more sophisticated visualizations of data in addition to on-line data retrieval systems. OMNIWeb is designed to exploit the hypertext feature of the WWW and offer a single "one-stop-shop" interface to the metadata and the data as well as an additional data browse capability to NSSDC's near-Earth heliosphere data (OMNI). The goal is to empower the user by providing a cohesive, consistent, and transparent interface that eases the process of metadata and data discovery and thus gives the researcher more time to examine the actual data itself.

2. Methodology and Interface Design

We applied the well-tested axioms of interface design, namely, user control, transparency, flexibility, and learnability [COX93]. All components of the OMNIWeb system adhere to these axioms and provide the user with the most user-friendly interface possible under the current constraints of the Hypertext Mark-up Language (HTML), and WWW server and client software. In an effort to apply these axioms, we took a **functional** approach to systems design and envisioned **what** tasks the user would be interested in performing and **how** those tasks ought to be performed. It is important to note that we, as developers, looked at the entire project from the users' perspective and not from the developers' point of view. Therefore, we modeled a "session," that is, a typical user interaction with the system starting from the moment users arrive

at the "home page" (Figure 1) to the time the data have been downloaded to their computer system. Defining a "session" helped to identify four functions that the user would want to perform with OMNIWeb.



Figure 1. OMNIWeb "Home Page."

While building OMNIWeb, particularly the server-side software, we were eager to ensure that each component is as generic as possible. The same components should be reusable by other WWW-based data systems in the future. (See section 5.0, Beyond OMNIWeb.)

The interface to OMNIWeb is composed of the "home page" with links to the OMNIWeb Browser, the OMNIWeb Retriever, a feedback form, and an extensive help file. The goal in the interface design was to utilize the visual feature of WWW to create an attractive and consistent interface. Banners and logos distinguish the pages and provide a visual clue as to where the user is and who is responsible for providing the service. Graphic bullets are used to highlight specific links and draw the user's attention. We also wanted to empower the user with control elements so that the entire system is easily navigable. A "control panel" was created from graphic buttons that appear near the top and bottom of most pages. The "control panel" lets the user jump to other important areas in the system.

2.1 OMNIWeb Browser

The OMNIWeb Browser is a forms-based interface that allows the user to input a start and stop date and select up to three parameters from the available data. The OMNIWeb Browser passes the selected parameters to a Common Gateway Interface (CGI) script called the OMNIWeb Plotter, which in cooperation with graphical data analysis software creates a GIF image that is transmitted back to the user's client.

2.2 OMNIWeb Retriever

The OMNIWeb Retriever is almost identical to the OMNIWeb Browser except that it allows the user to obtain OMNI data. The process of data retrieval begins as users identify themselves by providing an E-mail address. The user then has to make a selection from two data delivery options; that is, the data are displayed to the user's WWW client or copied to NSSDC's FTP site. The user also has a choice of data formats for FTP retrieval. The rest of the OMNIWeb Retriever is almost identical to the OMNIWeb Browser, namely, the way in which the user selects a start and stop date and the required parameters from the data. By default, the most common parameters are automatically selected; however, all of the available parameters may also be selected.

2.3 OMNIWeb Feedback

To collect ideas and user opinion on how OMNIWeb could be further improved, there is a forms-based feedback mechanism available from the OMNIWeb Browser, the OMNIWeb Retriever, and the "home page." The inclusion of the feedback form is based on the philosophy that the best designer of any computer system is the user.

2.4 OMNIWeb Help

The OMNIWeb Help is designed to be a user's guide as well as a context-sensitive help file. When a user selects the "Help" button or any of the hyperlinks from the data parameters, the link takes the user to the most appropriate section of the help document. The location of the section is determined by which page the user is on and which feature the user is trying to exercise. For example, the "Help" button on the OMNIWeb Browser will jump to the Browser section of the help document.

2.5 Error Handling

To make OMNIWeb as user-friendly as possible, extensive error trapping was included. The most likely errors are caused by the user's not having enough information on how to use OMNIWeb. Whenever an error is encountered, OMNIWeb displays an error message with a link to the

appropriate section in the help file. This method ensures that the user understands the nature of the error and gets instant help on how to use OMNIWeb.

OMNIWeb was designed to treat an error not as a user problem but either as a failure to present the user with clear instructions or a system problem. A typical error might result from invalid input values or missing fields. Other more serious errors might result from server problems such as running out of disk space.

3. Server-End Components

OMNIWeb is constructed of CGI shell scripts and HTML generators, fill-out HTML forms, data listing tools, data subsetting and conversion tools, a graphical data analysis engine, an E-mail handler, and a cron clean-up script. The system diagram in **Figure 2** shows all components of the entire system and how they interact. The user has access to the system by interacting with the client layer, which represents the forms and HTML pages that maintain the state of the system. Components in the client layer control the scripts in the CGI layer, by passing various input parameters from submitted forms via the Hyper Text Transfer Protocol (HTTP) server.

All components of the system below the server layer are hidden from the user. The user moves from one page to another as a seamless transition, since it is not apparent whether the document is a static page (e.g., OMNIWeb Browser) or one of the dynamically created HTML pages, such as the output from the OMNIWeb Plotter. The system layer is composed of the low-level tools that access the data. The physical data files are represented on the data layer, from which the output is generated and displayed at the user's WWW client. The diagram shows the entire system from the user to the developer with the multiple layers that tie them together.

3.1 OMNIWeb Plotter

OMNIWeb provides the user with a capability to query the data set for specific parameters, such as plasma temperature and ion density, over a time period from the OMNIWeb Browser and view a time series plot of the selected parameters. The plot is created by the OMNIWeb Plotter with the graphics engine and is displayed as an HTML document. An example of a typical plot is shown in **Figure 3**, which covers the time period of January 1, 1973, to January 31, 1973. The figure shows an X-Y Plot with time on the x-axis and the user-selected parameters on the y-axis. A different panel is generated for each parameter with the appropriate scaling of the y-axis. The gaps in the plots, such as for day 9, indicate no data were available for that time period.

3.1.1 IDL

Plots are generated automatically in the form of GIF files created by a graphical data analysis engine. The engine is a product from Research Systems Inc. called the Interactive Data Language (IDL). IDL is primarily an interactive data analysis and visualization tool that offers an interpreted programming language and extensive library for plotting and analyzing data on many computing

environments. However, IDL has the ability to compile modules fast, run in batch mode, perform all graphics output in memory, and write the results to a GIF file, which makes IDL ideal as a graphics engine for WWW-based data systems.

Jason Mathews
 Syed Towheed
 National Space Science Data Center
 NASA Goddard Space Flight Center
 December 10, 1994

NSSDC OMNIWeb System Diagram

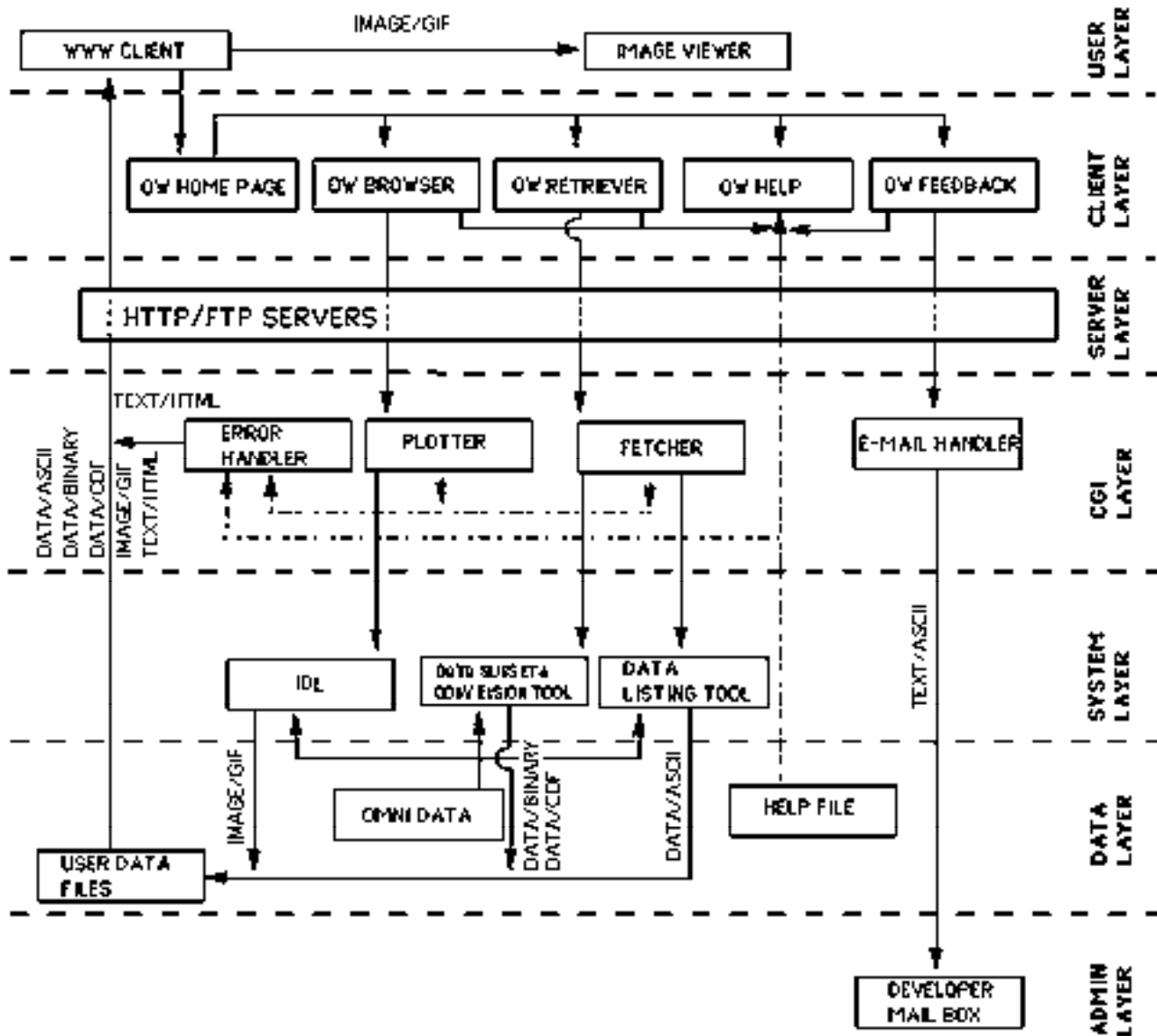


Figure 2. OMNIWeb System Diagram.

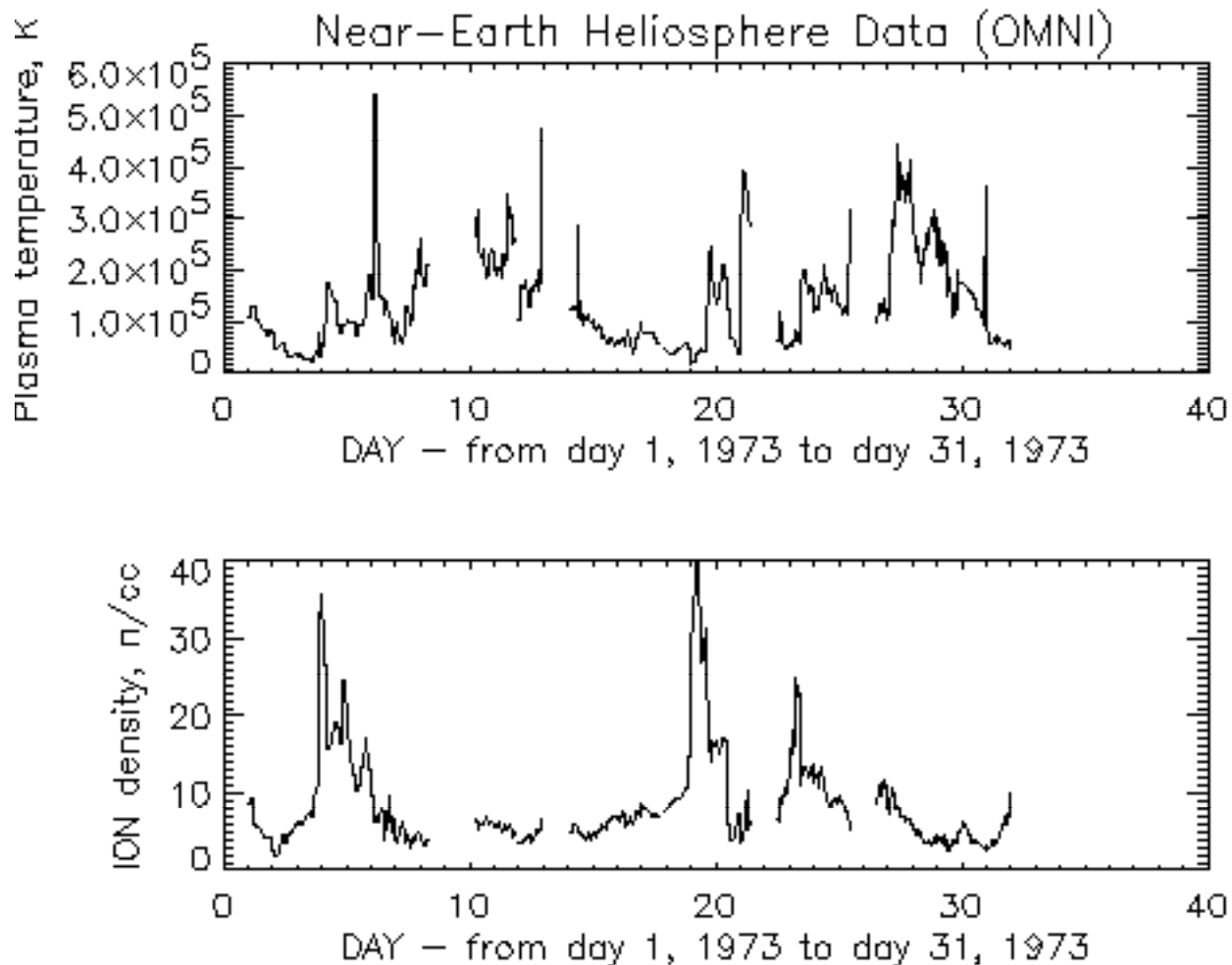


Figure 3. Sample Plot from OMNIWeb.

IDL is called from the OMNIWeb Plotter with the posted input of the user's completed form, which is formatted with a postquery-like filter and set in the appropriate shell environment variables as parameters. An IDL calling sequence is built up from these parameters, and then IDL is executed with the output written to a GIF file. For practicality, OMNIWeb Plotter limits the number of panels to four because beyond four panels the scaling of the y-axis becomes too small to distinguish the labels or much of the data.

Custom plots and images can be created by writing routines that call the appropriate IDL graphics, mathematical, and statistical library functions. Some special handling of the data is required for the OMNI data set because it spans a period of over 30 years with hourly measurements. A requested data period with less than two days of data is plotted with a labeling of the x-axis by hour; a period less than two years labels time by day number, and anything greater than two years uses a month/year label.

3.2 OMNIWeb Fetcher

The user-selected parameters from the OMNIWeb Retriever are passed to the OMNIWeb Fetcher. The OMNIWeb Fetcher is a shell script that produces a subset of the data in either ASCII or binary format with the Data Listing Tool and Data Subset and Conversion Tool, respectively.

3.2.1 Data Listing Tool

Requesting the data as ASCII will generate a report with the Data Listing Tool that can be displayed directly to the WWW client as TEXT/HTML through the OMNIWeb Fetcher script or from an FTP connection. An example of an ASCII listing with the default 12 variables for a period of 24 hours is shown in **Figure 4**. Each variable is formatted with a format specification defined for that variable. Within the OMNI metadata is a FORMAT attribute with a value indicating how an application should format the value of the variable. For instance, the year variable has a FORMAT of "I2", which can be used to print an integer value with two digits. This attribute information allows generic tools such as the Data Listing Tool to work with different data sets and produce the desired results.

Listing for OMNI data from 73001 to 73001											
YR	DAY	HR	B	F	THETA B	PHI B	T	N	V	PHI V	THETA V
73	1	0	0.0	0.0	0.0	0.0	111415	8.5	511	0.0	0.0
73	1	1	4.3	2.8	-6.0	11.0	111415	8.5	511	0.0	0.0
73	1	2	4.8	4.3	-14.0	283.0	111415	8.5	511	0.0	0.0
73	1	3	5.7	4.2	-14.0	295.0	129077	9.4	510	0.0	0.0
73	1	4	6.2	4.2	-42.0	288.0	129077	9.4	510	0.0	0.0
73	1	5	5.5	3.7	53.0	330.0	129077	9.4	510	0.0	0.0
73	1	6	5.5	4.7	38.0	332.0	129898	5.9	534	0.0	0.0
73	1	7	4.6	3.5	3.0	313.0	129898	5.9	534	0.0	0.0
73	1	8	4.0	2.7	19.0	346.0	129898	5.9	534	0.0	0.0
73	1	9	4.0	3.2	-10.0	4.0	106285	5.5	523	0.0	0.0
73	1	10	4.0	3.0	6.0	342.0	106285	5.5	523	0.0	0.0
73	1	11	4.2	3.0	-12.0	339.0	106285	5.5	523	0.0	0.0
73	1	12	0.0	0.0	0.0	0.0	89247	5.2	520	0.0	0.0
73	1	13	0.0	0.0	0.0	0.0	89247	5.2	520	0.0	0.0
73	1	14	4.2	4.0	53.0	33.0	89247	5.2	520	0.0	0.0
73	1	15	4.0	2.1	38.0	331.0	85643	4.4	515	0.0	0.0
73	1	16	4.3	2.9	31.0	330.0	85643	4.4	515	0.0	0.0
73	1	17	3.5	2.6	10.0	321.0	85643	4.4	515	0.0	0.0
73	1	18	4.0	3.7	20.0	0.0	73684	4.2	494	0.0	0.0
73	1	19	4.6	4.5	11.0	10.0	73684	4.2	494	0.0	0.0
73	1	20	4.6	4.6	0.0	12.0	73684	4.2	494	0.0	0.0
73	1	21	4.7	4.7	-1.0	8.0	81888	4.2	495	0.0	0.0
73	1	22	0.0	0.0	0.0	0.0	81888	4.2	495	0.0	0.0
73	1	23	0.0	0.0	0.0	0.0	81888	4.2	495	0.0	0.0

Figure 4. Sample ASCII Data Listing from OMNIWeb.

3.2.2 Data Subset and Conversion Tool

The Data Subset and Conversion Tool performs the task of producing and converting the data from its native format to the user-selected binary format. The native format of the data is a machine-independent format known as the Common Data Format (CDF)[CDF94 and GOU94].

CDF is not just a data format but a data interface for applications that allows transparent access to self-describing data. The OMNI data are stored in CDF because CDF is unique in the sense that the data is portable across disparate platforms and it allows conversion into the native integer and floating point representation from any of the following machines:

- DEC Alpha/OSF1
- DEC Alpha/OpenVMS (D_FLOAT and G_FLOAT)
- DECstation
- HP 9000 series
- IBM PC
- IBM RS-6000 series
- Macintosh
- NeXT
- SGI Personal Iris, Power Series, and Indigo
- Sun (SunOS and SOLARIS)
- VAX

CDFs may be copied to any of the supported computers and read by any CDF application available for that computer, but the raw OMNI data are readable only on the specific machine because of different memory word representations. For example, the double-precision floating point representation on the supported computers are either IEEE 754 floating-point, Digital Equipment Corporation's (DEC) D_FLOAT floating-point, or DEC's G_FLOAT floating-point. There are also two different byte orderings for the computers that use IEEE 754, which is either big-endian or little-endian. The CDF is self-describing in that it contains all the metadata that describe the structure of the data set, while the user will need to keep track of which variables are a part of the raw binary file, what data types they are, what is the record length, and for what machine. OMNIWeb supports all of these data formats, and users decide how the data should be formatted to suit their needs.

3.3 E-Mail Handler

The E-mail handler is based on the `post_query.c` code provided with the NCSA `httpd 1.1` package released to the public domain. It composes the input given from the feedback form into an E-mail message, sends the message to the OMNIWeb developers, and responds with HTML output for an acknowledgment of acceptance or rejection.

3.4 Cron Cleaner

The files created in the user data area can accumulate quickly, so there is a need to clean up old data files. A daily cron process will delete all files older than two days and remove empty directories as a result of deleting the old files. The user has two days to retrieve data files, which are located in a directory created with the user name entered from the OMNIWeb Retriever.

4. Advanced Features for OMNIWeb

As users try out the different interfaces and approaches to data and information access via the WWW, the developers and users will quickly determine what is popular and what is not. Developers will in turn adopt those interfaces that are best for their own systems. Likewise, OMNIWeb will grow as feedback from users suggest features and capabilities that would be helpful.

In the near future there are plans for improving OMNIWeb. OMNIWeb Browser currently offers a selection from a list of variables and the time range (start/stop dates) to create a plot. We hope to extend this interface with an "options" feature that saves the selections on the OMNIWeb Browser as hidden fields on a new dynamically created form with advanced graphics options to customize the output to a much greater degree. Submitting this new form will generate the plot however requested. The initial browsing interface uses default graphics options that should satisfy most users' requests, and this new form will provide more control. Some advanced features might include other plot types such as histograms, plot and background colors, image size, plot symbols, logarithmic scaling of the axes, and the ability to spawn an external image viewer rather than generate an HTML document with in-line GIF images.

5. Beyond OMNIWeb

The developers intend to apply the OMNIWeb-type interface to other data sets available at the NSSDC that are of special interest to users. Creation of a common set of HTML forms, CGI scripts, data reporting tools, and data browsing tools that can be used for other data with little or no modification is the ultimate goal for this project.

The OMNI data set has only 0-dimensional data, such as one temperature value per record. Other data sets that have multidimensional data, such as vectors or images, would allow the graphic tools to produce 2-D or 3-D plots and images as appropriate. The OMNIWeb Browser interface is not limited to space physics data nor scientific data in general but is applicable to other disciplines and other data sets that are suitable for visualization.

OMNIWeb and systems like it have the potential of radically shifting the paradigm with which researchers currently acquire data. Not only does it offer instant access to data but has the added feature of supporting interactive data analysis. The WWW has allowed the folding of network-based data retrieval and analysis seamlessly into the core of research activity.

6. Acknowledgments

The authors are indebted to Dr. Joseph H. King, NSSDC, for providing the OMNI data set and to Mr. Nathan L. James for providing unlimited access to the NSSDC On-line Data and Information System (NODIS). The IMP 8 magnetic field and plasma data were provided by Drs. Ron Lepping, GSFC, and Al Lazarus, MIT.

7. Authors

Jason Mathews is a computer engineer at the National Space Science Data Center Interoperable Systems Office, Code 633.2, NASA/Goddard Space Flight Center. He has earned a B.S. in computer science from Columbia University and an M.S. in computer science from the George Washington University. The focus of his research at GSFC is the development of portable and reusable software tools. In addition to creating OMNIWeb, he develops advanced scientific data analysis systems, data management tools, and electronic information management systems.

Syed S. Towheed is a systems programmer at the National Space Science Data Center, NASA/Goddard Space Flight Center. He is the coordinator of NSSDC's WWW development activities and has played a central role in applying WWW technology to NSSDC's mission of providing access to its data holdings. Mr. Towheed was one of the early proponents of using WWW in the NASA environment and was among the first to develop a unique look-and-feel for WWW services. Over the last year Mr. Towheed has developed several systems, including the

NSSDC CD-ROM Catalog, NSSDC Planetary Sciences, NASA Space Science Education, NSSDC Solar Physics, NSSDC Life Sciences, NSSDC Shoemaker-Levy 9 system, and NSSDC OMNIWeb system.

8. References

- [CDF94] *CDF User's Guide, Version 2.4*, NSSDC/WDC-A-R&S 94-01, February 1994.
- [COX93] Cox, K., and Walker, D., *User Interface Design*, Prentice Hall, Singapore, 1993.
- [GOU94] Goucher, G. W., and Mathews, G. J., *A Comprehensive Look at CDF*, NSSDC/WDC-A-R&S 94-07, August 1994.
- [HOR94] Horowitz, R., King, J. H., *NSSDC Data Listing*, NSSDC/WDC-A-R&S 94-09, October 1994.
- [IDL94] *IDL Scientific Data Formats, Version 3.6*, Research Systems Incorporated, Boulder, Colorado, April 1994.
- [KIN77] King, J. H., *Interplanetary Medium Data Book*, NSSDC/WDC-A-R&S 77-04, September 1977.
- [KIN94] King, J. H., and Papitashvili, N. E., *Interplanetary Medium Data Book Supplement 5, 1988-1993*, NSSDC/WDC-A-R&S 94-08, September 1994.
- [LEE94] Berners-Lee, T., and Connolly, D., *HyperText Markup Language Specifications - 2.0*, CERN, October 1994.

9. Related URLs

- [OMNIWeb System] <http://nssdc.gsfc.nasa.gov/omniweb/ow.html>
- [NSSDC Home Page] http://nssdc.gsfc.nasa.gov/nssdc/nssdc_home.html
- [CDF Home Page] http://nssdc.gsfc.nasa.gov/cdf/cdf_home.html
- [IDL Home Page] http://sslaboratory.colorado.edu:2222/projects/IDL/idl_ssl_home.html