

CONF-971099-1

BIOME: A Scientific Data Archive Search-and-Order System Using Browser-Aware, Dynamic Pages

19980219
066

Sarah V. Jennings
Research Associate, Transportation Center
University of Tennessee, USA
xqj@ornl.gov

Teresa G. Yow, Ph.D.
Systems Analyst, Computational Physics and Engineering Division
Oak Ridge National Laboratory, USA
tgy@ornl.gov

Vickie W. Ng
Systems Engineer
Hughes Information Technology Systems, USA
vw5@ornl.gov

RECEIVED
SEP 17 1997
OSTI

Introduction

The Oak Ridge National Laboratory's (ORNL) Distributed Active Archive Center (DAAC) is a data archive and distribution center for the National Air and Space Administration's (NASA) Earth Observing System Data and Information System (EOSDIS). Both the Earth Observing System (EOS) and EOSDIS are components of NASA's contribution to the U.S. Global Change Research Program through its Mission to Planet Earth Program. The ORNL DAAC provides access to data used in ecological and environmental research such as global change, global warming, and terrestrial ecology.

Web Pages Over Database

Because of its large and diverse data holdings, the challenge for the ORNL DAAC is to help users find data of interest from the hundreds of thousands of files available at the DAAC without overwhelming them. Therefore, the ORNL DAAC has developed the Biogeochemical Information Ordering Management Environment (BIOME), a customized search and order system for the World Wide Web (WWW). BIOME is a public system located at <http://www-eosdis.ornl.gov/BIOME/biome.html>.

Managing large amounts of data requires metadata, or data that describes the data, which is stored in a relational database management system. Several Sybase metadata databases form the heart of BIOME by managing to treat many different types of data in a consistent manner. Using metadata stored in a relational database management system allows for efficient searching of hundreds of thousands of records.

The data itself is stored on-line, off-line, and near-line. Small tabular datasets are stored on-line on spinning disk. CD-ROMs, tapes, and proprietary data are stored off-line. Larger datasets, i.e., satellite imagery, are stored near-line in a mass storage system. A browse capability allows users to preview near-line images by generating a thumbnail .GIF image of a larger imagery file. The location of the data is transparent in that the user does not need to know or care where the data is stored. With the exception of hard media (e.g., CD-ROMs) all data delivery is automated.

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED



MASTER

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, make any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Browser Aware, On-the-Fly HTML Pages and Graphs

The ORNL DAAC WWW site categorizes browsers based on their capabilities. Pages are created according to the ability of the user's browser to display them. High-end browsers can get pages with frames, tables, and Java applets in addition to the information available to character-based browsers.

Because many of our users are scientific researchers working in remote areas, we must balance their needs with those of users who have access to the latest technology. On-the-fly HTML page customization allows the ORNL DAAC WWW site to take advantage of the most innovative WWW features while maintaining backwards compatibility with older browsers and text-based browsers. In a one year period, 1152 unique browser/platform combinations accessed the ORNL DAAC site. Because of the browser-aware on-the-fly page creation capabilities of BIOME, the DAAC was able to respond to this challenge by presenting each combination with customized HTML pages.

The ORNL DAAC's WWW site is designed around include statements that pull in the appropriate "modules" for each browser. If a user's browser is capable of displaying a certain feature the section of the page that uses that particular feature is included. If not, that portion of the page is not included for display. Thus, the page is dynamically altered.

BIOME allows users to see a graph of selected data. Tabular data are parsed according to arbitrary classifications describing the configuration of the data. The GD1 library is then used to generate a plot of the data. The user's browser is sent a .GIF with the selected labeled columns plotted in color. This technique allows one graphic engine to display many different layouts of tabular data.

WWW-based Tools

As the complexity of the DAAC's data holdings has increased, the task of maintaining the databases has become increasingly difficult and time-consuming. Fortunately, custom WWW-based tools make the task of the database administrator less difficult.

For example, the DBA Maintenance Tool handles the ingest of new metadata by providing on-the-fly templates of database tables generated dynamically from Sybase's system tables. New data can be typed onto the templates, eliminating the need for manually constructing Sybase bulk copy files, a task that is tedious and error prone. In addition, the DBA Maintenance Tool easily handles updates to existing metadata, offering such options as global updates to the databases. Other options include automated bulk copies out of the database and the printing of the current structure for each table. The DBA Maintenance Tool also automatically generates a transaction log that provides a record of all DBA actions on the databases.

Conclusion

The ORNL DAAC provides WWW access to a large number of ecological and environmental datasets. The DAAC has accomplished this task by designing and offering a customized WWW search and order system that allows efficient and rapid data search and retrieval. By developing customized WWW tools to manage global ecological and environmental data, the ORNL DAAC has made an important contribution to NASA's Mission to Planet Earth Program.

Acknowledgments

Research sponsored by NASA under Interagency Agreement DOE No. 2013-F044-A1 under Lockheed Martin Energy Research Corp., contract DE-AC05-96OR22464 with the U.S. Department of Energy. "The submitted manuscript has been authored by a contractor of the U.S. Government under contract No. DE-AC05-96OR22464. Accordingly, the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes."

"The submitted manuscript has been authorized by a contractor of the U.S. Government under contract No. DE-AC05-96OR22464. Accordingly, the U.S. Government retains a non-exclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes."

M97009364



Report Number (14) CONF-97109 9--1

Publ. Date (11)

199708

Sponsor Code (18)

DOE/MA; NASA, XF

JC Category (19)

UC-900; UC-000, DOE/ER

DOE