ENHANCING ACCESS AND USE OF NASA SATELLITE DATA VIA TERRAFLY

William Teng
NASA Goddard Earth Sciences Data and Information Services Center
Distributed Active Archive Center (GES DISC DAAC)
Code 610.2, Greenbelt, MD 20771
Science Systems & Applications, Inc.
teng@disc.sci.gsfc.nasa.gov

Naphtali Rishe
Florida International University

Hualan Rui
NASA Goddard Earth Sciences Data and Information Services Center
Distributed Active Archive Center (GES DISC DAAC)
Code 610.2, Greenbelt, MD 20771
Science Systems & Applications, Inc.

ABSTRACT

NASA satellite data form a rich resource that is largely untapped by the applications user community, in part because of the complexity of using, and the cost of learning how to use, such data. These users are generally not interested in the data per se, but rather in one or more specific measurements (e.g., surface rain) from the data, which can then be seamlessly infused in their own environment (e.g., decision support systems). The Goddard Earth Sciences Data and Information Services Center Distributed Active Archive Center (GES DISC DAAC) has collaborated with the Florida International University's High Performance Database Research Center (FIU HPDRC) on an initial prototype effort, which has demonstrated the feasibility of making NASA satellite data more easily and seamlessly accessible, as a Web service, from the FIU's TerraFly environment. The latter is a Web-enabled system designed to aid in the visualization of spatial and remotely sensed data, by "flying" over the Earth's surface, via standard Web browsers. The other part of the collaboration is the GES DISC Giovanni, an online visualization and analysis system, which relieves the users of much of the data preparation work and provides a tool for easily and quickly obtaining information from the data, without having to download and handle large amounts of data. The prototype effects the seamless access of data by deep linking (i.e., geo-located) TerraFly and Giovanni, to enable a dynamic Web service, providing on-demand, near-real-time, satellite precipitation data. The Giovanni system is evolving towards a service-oriented architecture, thus making available to TerraFly users data not just from NASA but also potentially from many other sources, as well as making NASA satellite data potentially more widely accessible.

INTRODUCTION

NASA satellite data form a rich resource that is largely untapped by the applications user community, in part because of the complexity of using, and the cost of learning how to use, such data. So far, the approaches taken to ameliorate this situation have not sufficiently addressed the core needs of the applications community. The latter is generally not interested in the data per se, but rather in one or more specific measurements (e.g., surface rain) from the data, which can then be seamlessly infused in their own environment (e.g., decision support systems). The Goddard Earth Sciences Data and Information Services Center Distributed Active Archive Center (GES DISC DAAC) has collaborated with the Florida International University's High Performance Database Research Center (FIU HPDRC) on an initial prototype effort, which has demonstrated the feasibility of making NASA satellite data more easily and seamlessly accessible, as a Web service, from the FIU's TerraFly environment. The collaboration leverages TerraFly's capabilities and its large and diverse user community and the GES-DISC Interactive Online Visualization and Analysis Infrastructure (Giovanni), both mature systems (Rishe et al., 2004; Teng et al., 2005).

TERRAFLY

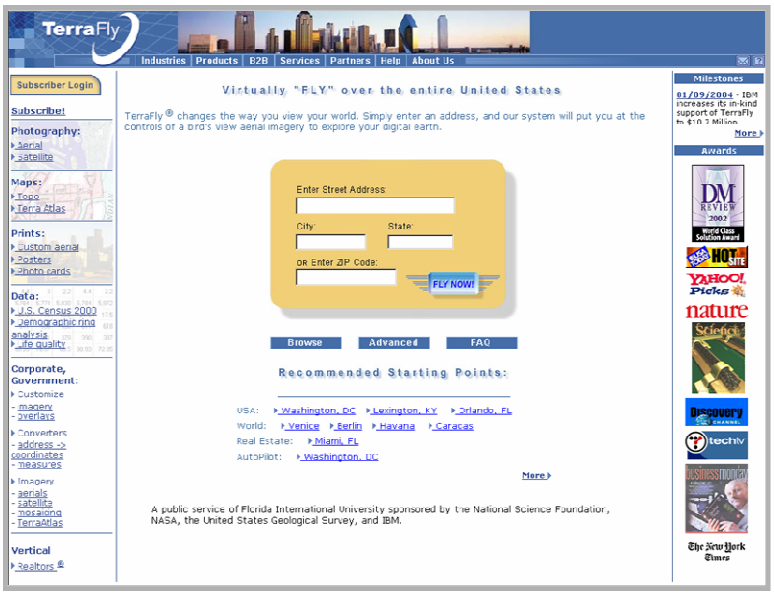
The state-of-the-art TerraFly environment, developed by FIU HPDRC in partnership with NASA, aims to deliver accurate geospatial data and imagery to science research and applications user communities. TerraFly is based on innovative semantic database technologies for data management and a rich set of tools for disseminating geospatial information via the Internet (Rishe et al., 2001; Rishe et al., 2003a,b,c; Chen et al., 2004). TerraFly enables users to "fly over" and manipulate the data retrieved from the database, which contains textual, remotely-sensed, and vector data (graphical maps), via applets using any standard Web browser (i.e., no specialized software is needed). TerraFly's easy-to-use, graphical user interface and portability make remotely sensed data readily available to the novice as well as expert user (Fig. 1). TerraFly is extremely popular with its diverse user communities (some 10,000 unique users per day) and is a NASA 2004 Spinoff (___, 2004). TerraFly offers a variety of data layers, such as U.S. Geological Survey's aerial photography, Landsat imagery, road overlays, biological field station locations, real-time stream monitoring stations, FEMA flood zone maps, Bailey's ecological regions, and PRISM mean annual precipitation data.

TerraFly's flexible Web dissemination of data is based on the "Best Available" theory, which is a pixel-by-pixel auto-mosaicking of imagery from a variety of sources to present a novice user with an imagery continuum composed of the best available pixels. The heuristic algorithm extracts the best pixels from different images by considering acquisition date, color palette, sharpness, and other factors. The novice user "flies" over the imagery without having to select a source of imagery. The advanced user, on the other hand, can select data sources, as well as define combinations of sources, spectral bands, and application of filters. Likewise for overlays (e.g., street names), different data sets of varying precision and comprehensiveness (e.g., NavTeq, Tiger) are merged and presented as a unified intuitive layer to the user.

TerraFly Features

TerraFly has many unique features and capabilities (summarized below), not found in other similar Web services, which are ideally suited to its science research and applications user communities.

- Overlaying of imagery and maps.
- Comprehensive imagery coverage of almost the entire U.S. at 1 m resolution and many urban areas at 30 cm resolution.
- Multiple levels of customizable overlays, both generic and user-specified.
- Large selection of imagery sources and automatic "best-available-pixel" mosaic.
- Topographic maps in addition to regular maps.
- Supports multiple browsers (e.g., Opera).
- Designed not just for general consumer use, but for a variety of science research and applications users (in Disaster Management, Homeland Security, Public Health, and Ecological Forecasting).
- Offers continuous flying mode via a Web applet-based technology (heuristic pre-fetching of tiles which enables the user to view a virtual stream of imagery as seen when "flying"), requiring no software installation by the user.
- Rich object search capability, especially the GeoQuery interface.
- Rich, user-customizable, PointData functionality showing data about the point clicked.
- Many modes of distribution of imagery and data.
- Multi-window synchronous visualization of various imagery sources (e.g., at different times, spectral combinations, instruments, resolutions).



http://terrafly.fiu.edu/

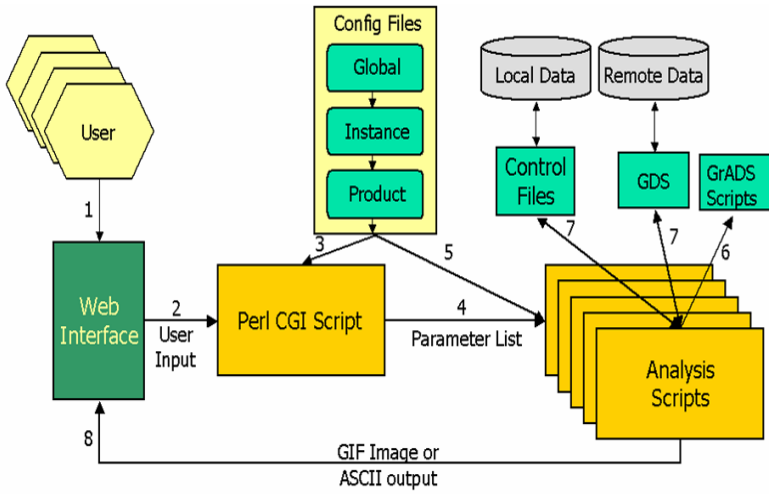
Figure 1. TerraFly environment, delivering geospatial data and imagery to science research and applications users.

GIOVANNI

Giovanni is the underlying infrastructure for a growing family of Web interfaces that allows users to analyze gridded data interactively online without having to download any data (Berrick et al., 2004). Through Giovanni, users can discover and explore GES DISC data using sophisticated analyses and visualizations. Giovanni-TRMM Online Visualization and Analysis System (TOVAS), the Giovanni instance used in the prototype with TerraFly, focuses on global precipitation data (Chiu et al., 2005). The principal design goal for Giovanni is to provide a quick and simple interactive means for science and applications data users to study various phenomena by trying various combinations of parameters measured by different instruments, arrive at a conclusion, and, as appropriate, generate graphics suitable for publications. Alternatively, Giovanni provides a means to ask relevant what-if questions and get back answers that could stimulate further investigations. All work can be done without having to download and preprocess large amounts of data. Current Giovanni instances support data from AIRS, Aura MLS, MODIS, TOMS & OMI, TRMM, and UARS HALOE, as well as agriculture- and ocean color-related project data.

Giovanni Data Flow

Giovanni consists of HTML templates, CGI scripts written in Perl and GrADS (Grid Analysis and Display System) languages. In addition, there is an image map Java applet with which a user can select a bounding box area to process. Access to data is via one or more GrADS-Data Servers (GDS) running on remote machines that have GrADS readable data. The data flow is illustrated in Figure 2.



http://giovanni.gsfc.nasa.gov

Figure 2. Giovanni data flow.

Giovanni Features

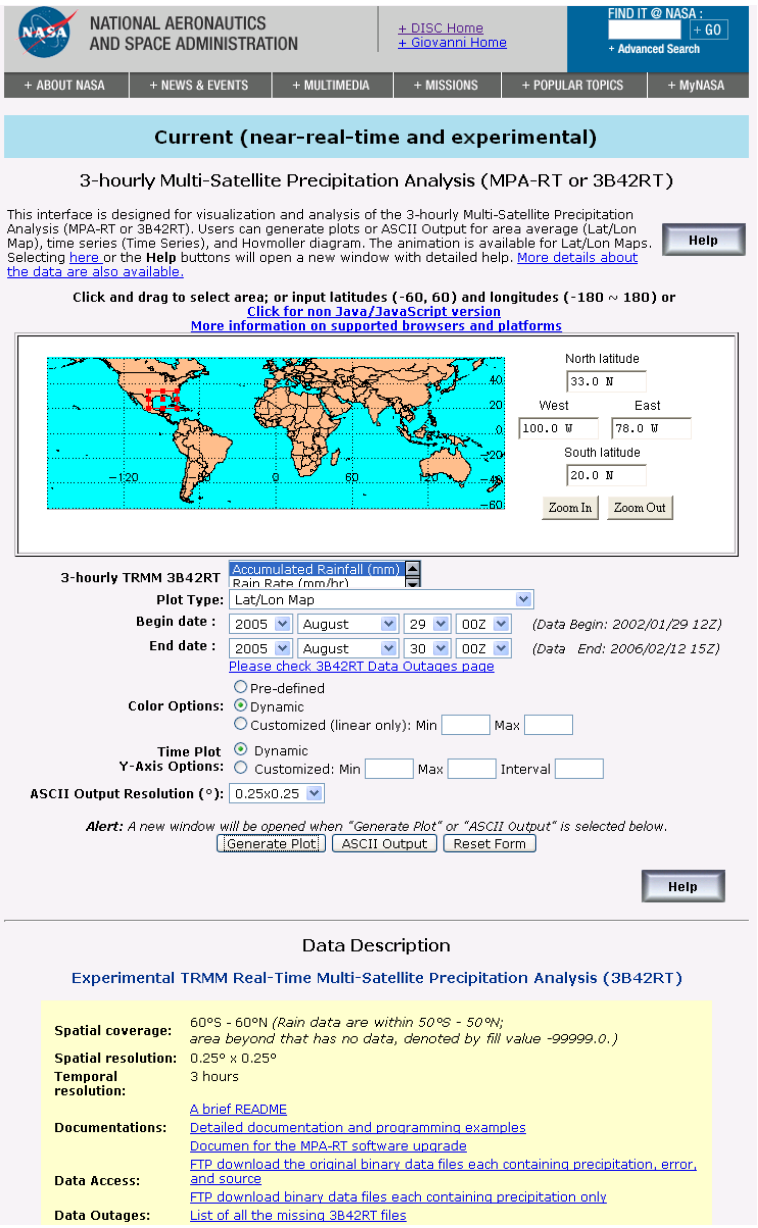
Current features of Giovanni include the following:

- Access to data from multiple remote sites as well as local sites.
- Server-side temporal and spatial subsetting.
- Server-side processing.
- Support for multiple data formats, including Hierarchical Data Format (HDF), HDF-EOS, network Common Data Form (netCDF), GRIdded Binary (GRIB), and binary.
- Support for multiple plot types, including area, time, Hovmoller, and image animation.
- Support for outputting data in ASCII format in multiple resolutions.
- Multi-parameter inter-comparisons.
- Easily configurable to support customized portals for measurements-based projects or disciplines.

Currently supported Giovanni output types include lat-lon map, time series, Hovmoller, animation, vertical profile, and ASCII. For parameter inter-comparisons, additional output types include overlay, scatter plot, correlation, and difference.

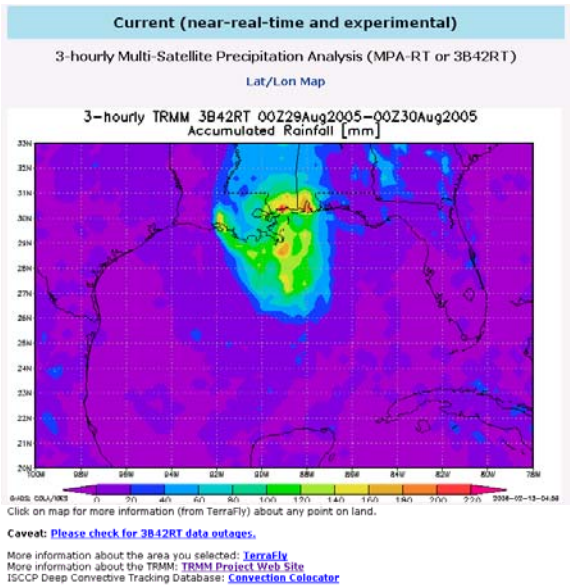
Giovanni User Interface

From the user's perspective, Giovanni is a simple Web application. A user can select either the Java or non-Java version. The resulting Web page allows the user to select a spatial area, a temporal range of the data, one or more parameters from the data set, and the output type (ASCII or one of several plot types). For plots, several color options are also available. Figure 3 shows the TOVAS user interface (with example user selections for Hurricane Katrina). Once the selections are made, the user has the option of generating a plot or outputting the results to an ASCII file that can be downloaded. Figure 4 shows an example area plot and part of its ASCII output for the selections shown in Figure 3. Links to the data are provided so the user can download the entire data set, if so desired.



http://disc2.nascom.nasa.gov/Giovanni/tovas/realtime.3B42RT.shtml

Figure 3. Part of Giovanni-TRMM Online Visualization and Analysis System (TOVAS), a simple and easy-to-use Web interface for access to global precipitation data. Example user selections are for Hurricane Katrina.



a

```
Selected parameter: 3-hourly TRMM 3B42RT Accumulated Rainfall
Selected area: lat=[20.0N,33.0N], lon=[100W,78W]
Selected time period: (00Z29Aug2005-00Z30Aug2005)
Undefined/Missing Value: -99999
Unit: (mm)
Latitude Longitude AccRain
 20.000 -100.000       0.0000
 20.000  -99.750       0.0000
 20.000  -99.500       0.0000
 20.000  -99.250       0.0000
 20.000  -99.000       0.0000
 20.000  -98.750       1.1700
 20.000  -98.500       3.3600
 20.000  -98.250       0.8700
 20.000  -98.000       9.0000
 20.000  -97.750       4.6500
 20.000  -97.500       0.5400
 20.000  -97.250       0.0000
 20.000  -97.000       0.0000
 20.000  -96.750       0.0000
 20.000  -96.500       0.0000
 20.000  -96.250       0.0000
 20.000  -96.000       0.0000
 20.000  -95.750       0.0000
 20.000  -95.500       0.0000
 20.000  -95.250       0.0000
 20.000  -95.000       0.0000
 20.000  -94.750       0.0000
 20.000  -94.500       0.0000
 20.000  -94.250       0.0000
 20.000  -94.000       0.0000
 20.000  -93.750       0.0000
 20.000  -93.500       0.0000
 20.000  -93.250       0.0000
 20.000  -93.000       0.0000
 20.000  -92.750       0.0000
 20.000  -92.500       0.0000
 20.000  -92.250       0.0000
 20.000  -92.000       0.0000
```

b

Figure 4. Example Giovanni-TOVAS (a) area plot and (b) ASCII output: rainfall accumulation from Hurricane Katrina between August 29 and 30, 2005, as it made landfall on the Gulf states.

INTEGRATING TERRAFLY WITH GIOVANNI

The FIU-GES DISC prototype effects the seamless access of data by deep linking (i.e., geo-located) TerraFly and Giovanni, to enable a dynamic Web service, providing on-demand, near-real-time, satellite precipitation data (with other measurements to be added). Users can click on any point within a Giovanni output plot and access, for an area the center of which is represented by the selected point, other data and information relevant to the clicked point. Conversely, users can access Giovanni precipitation data from several places within TerraFly.

Figure 5 shows an example scenario of a user seamlessly accessing data from TerraFly and Giovanni over Florida, during the occurrence of Hurricane Jeanne in 2004:

(a) Giovanni user displays a TRMM accumulated rainfall plot and clicks on the TerraFly deep link for more information about the displayed area, centered over the plot center, or clicks on any arbitrary point within the plot.

(b) User hyperlinks to a TerraFly PointData page, which shows a NASA mosaic of Landsat 5 imagery overlain with the names of roads, cities, NASA PRISM mean annual precipitation model data (shaded areas), and TRMM rainfall of the previous day.

(c) From the Environmental Data section of the PointData page, user submits a TerraFly GeoQuery, "100 nearest TRMM cells where the daily precipitation yesterday was between 148.98 and 160.41 mm," which, in this case, returns 14 found records,

(d) User clicks on a "details" button in the GeoQuery output, which hyperlinks to Giovanni and, in this case, displays a month-to-date TRMM plot.

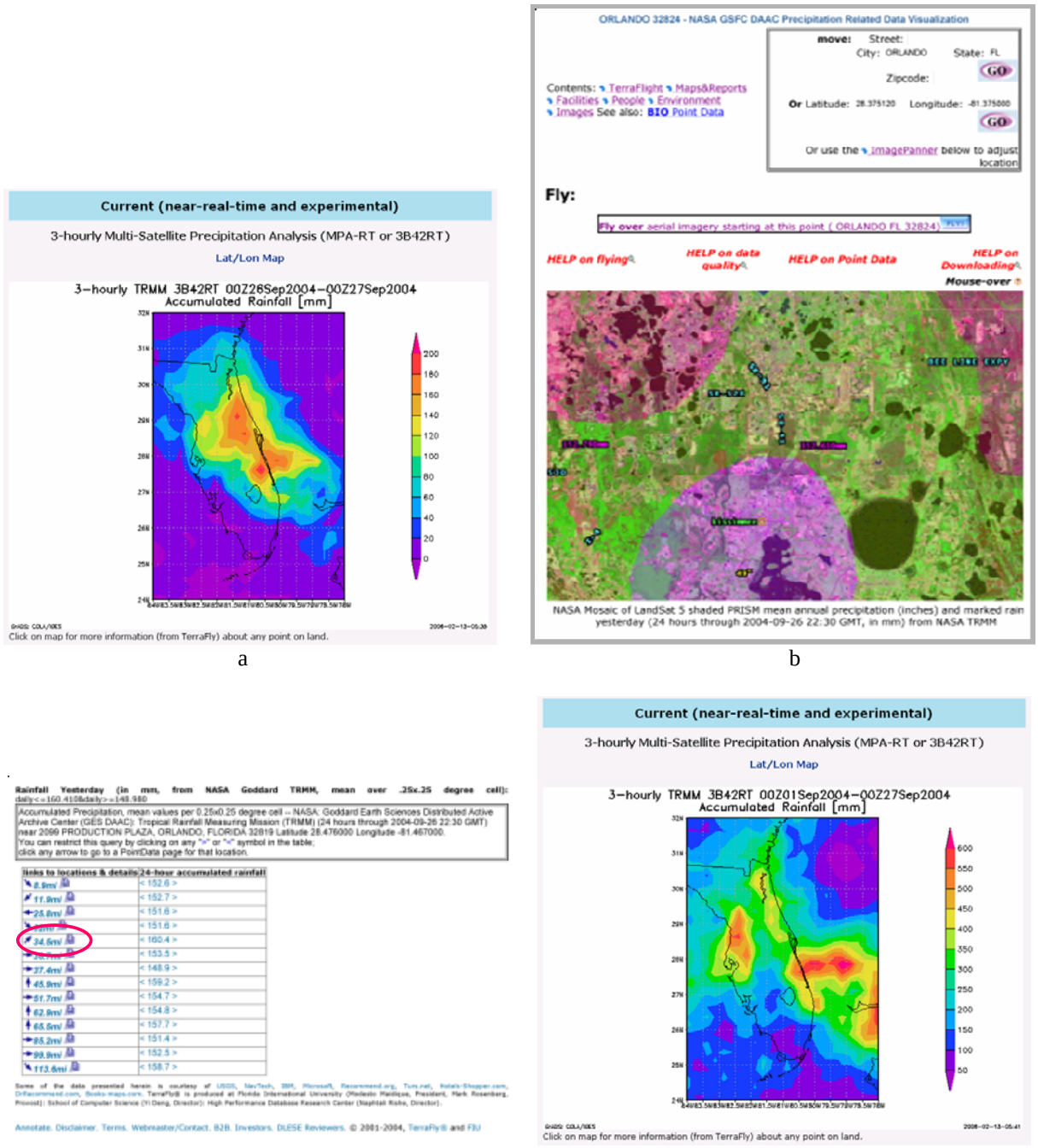


Figure 5. Example user scenario of seamless access of data from TerraFly and Giovanni over Florida, during the occurrence of Hurricane Jeanne in 2004.

Further development of this networked collaboration and integration of resources will enable the GES DISC to leverage the existing capabilities of TerraFly to better publicize and disseminate NASA satellite data to the diverse TerraFly user communities, including science, applications, education, as well as the general public. Conversely, the information available from Terrafly will be easily accessible to GES DISC users of TRMM and other NASA data, via Giovanni.

Future Plans

The Giovanni system is evolving towards a service-oriented architecture, which will make available to TerraFly users data not just from NASA but also potentially from many other sources, as well as making NASA satellite data potentially more widely accessible.

REFERENCES

Berrick, S., L. Pham, G. Leptoukh, Z. Liu, H. Rui, S. Shen, W. Teng, and T.Zhu (2004). Multi-sensor distributive on-line processing visualization and analysis system using Giovanni. In: *Proceeding from the IGARSS Conference*. September 15-24, 2004, Anchorage, AK.

Chen, S-C., X. Wang, N. Rishe, and M.A. Weiss (2004). A web-based spatial data access system using semantic R-trees. *Information Sciences Journal*, *167*(1-4):44-61.

Chiu, L., Z. Liu, H. Rui, and W. Teng (2005). Tropical Rainfall Measuring Mission (TRMM) data and access tools. In: *Earth System Science Remote Sensing*. J. Qu et al. (Eds.), Springer-Verlag.

Rishe, N., S-C. Chen, N. Prabhakar, M.A. Weiss, W. Sun, A. Selivonenko, and D. Davis-Chu (2001). Terrafly: a high-performance web-based digital library system for spatial data access. In: *Proceeding from the International Conference on Data Engineering*. Heidelberg, Germany.

Rishe, N., Y. Sun, A. Selivonenko (2003a). A rich-semantics GIS imagery composing framework. In: *The Ninth International Conference on Distributed Multimedia Systems.* Miami, FL, 144-147.

Rishe, N., W. Teng, H. Rui, S. Graham, and M. Gutierrez (2004). Web-based dissemination of TRMM data via TerraFly. In: *Proceeding from the AGU Fall Meeting.* December 13-17, 2004, San Francisco, CA.

Rishe, N., B. Worgsaroj, M. Chekmasov, A. Selivonenko, Y. Sun, S. Graham, and A. Mendoza (2003b). Autopilot scripting for the TerraFly GIS. In: *Proc. of the Ninth International Conference on Distributed Multimedia Systems.* Miami, FL, 98-100.

Rishe, N., J. Yuan, R. Athauda, S-C. Chen, X. Lu, X. Ma, A. Vaschillo, A. Shaposhnikov, and D. Vasilevsky (2003c). Semantic access: semantic interface for querying databases. In: *ACM SIGMOD Digital Symposium Collection DiSC01*.

Teng, W., H. Rui, N. Rishe, and R. Tetrault (2005). Enhancing access to NASA data via seamless integration into decision support systems: two examples. In: *Proceeding from the AGU Fall Meeting*. December 5-9, 2005, San Francisco, CA.

_____ (2004). A view from above without leaving the ground. *NASA Spinoff*. http://www.sti.nasa.gov/tto/Spinoff2004/er_6.html.

ACKNOWLEDGMENT

The work reported here is partly supported by NASA through REASoN CAN 02-OES-01. The authors acknowledge the contributions of members of the GES DISC REASoN project, the GES DISC Giovanni team, and the FIU TerraFly team.