4.6 Integrating Meteorological Tools and Data into Digital Libraries: A Strategy used in the Visual Geophysical Exploration Environment

Rajul Pandya[*1], Benedict Domenico[2], Mary Marlino[1], and Donald Murray[2]
(1) DLESE Program Center, Boulder, CO (2) Unidata Program Center, Boulder CO

1. INTRODUCTION

Geoscientists have access to a tremendous range of data, using a variety of technology across a number of sub-disciplines. Data distribution strategies range from web-based facilities that allow users to browse datasets on servers and download those of interest (e.g. IRIS Data Management Center), to real-time delivery that requires user's "subscribe" to certain datastreams whose products are delivered as soon as they are available (e.g. Unidata's IDD). More recently, client/server alternatives have been developed; these allow users running applications on remote servers to access data on the servers as if the datasets were on local disks (e.g. OPeNDAP or ADDE). None of these services, however, allow discovery of data – data users need to know a priori where to get the data.

THREDDS, THematic Real-time Earth Distributed Data Servers, (Unidata Program Center, 2003) builds on these systems by applying digital library technology to the discovery and use of Earth system data. THREDDS catalogs link datasets to URLS and metadata descriptions that can be included in digital libraries. THREDDS services negotiate the variety of data access protocols, and so provide more transparent access to distributed data. THREDDS also enables a semantic web to provide data users with the ability to search distributed data in the language of their own discipline. By providing a service layer between data users and providers, THREDDS can offer this functionality with little additional burden to data providers. Even so, THREDDS functionality is aimed primarily at researchers who have the domain knowledge to construct searches and evaluate results within the context of a scientific discipline.

The Digital Library for Earth System Education (DLESE; DLESE, 2001) has created a rich metadata framework and catalog to support the needs of geoscience educators and learners. Users search and browse among the 4000 or so community contributed and catalogued resources. The metadata framework supports discovery and evaluation of a number of resource types, including data.

*Corresponding author address: Rajul E. Pandya, DLESE Program Center/UCAR, P.O. Box 3000, Boulder, CO 80307. e-mail: pandya@ucar.edu

Most DLESE users, however, don't find anything but the simplest data sets useful. They don't have access to or knowledge of the kinds of tools that are required to meaningfully interact with the more complex data. Another barrier to educator or learner use of the data is the significant amount of domain knowledge necessary to be able to reliably connect appropriate data to learning goals. DLESE users have asked for data sets as part of integrated packages that included the tools to use the data and some guidance in the appropriate use of the data, not as a stand-alone library resource (Wright, M. J. and Sumner, T. R., Eds., 2003).

2. VGEE

The Visual Geophysical Exploration Environment addresses the needs of data users learners and educators by providing a way to connect data with curricula and tools to guide and facilitate meaningful and appropriate use of the data. The VGEE consists of four elements:

- ·An inquiry-based curriculum for guiding student exploration
- A learner-centered interface to a scientific visualization tool
- · A collection of concept models (interactive tools that help students understand fundamental scientific concepts)
- A suite of El Niño-related data sets adapted for student use

An example will illustrate the curriculums framework and how appropriate tools, data, and instructional resources are supported in that framework. In one topic, students begin by illustrating their own idea of typical ocean temperatures. Using a scientific visualization tool with a modified interface to support learning, students can compare their image to real-time temperature data. They can also construct three-dimensional, time-varying animations of the data and use these to explore relationships in the data. As an example, students in the past have constructed 3D animations that show the coincidence of persistant clouds and high SST, and their migration together across the Pacific. To understand these patterns, students use a collection of interactive online models. In this case, students can use an interactive model of positively buoyant air parcels to understand how hot, moist air rises to produce clouds. To help students connect this fundamental idea back to the data,

students can embed these models as probes in their visualization and see how these models represent the processes uncovered in real-time data.

The VGEE uses a number of technologies to support its integration. The scientific visualization tool is the Integrated Data Viewer (IDV; Unidata Program Center, 2003), which is installed, configured and updated over the internet. Data access is provided over the internet as well, using THREDDS middleware to provide access to real-time and archived distributed data, services that simplify the use of the data in educational settings (for example, by attaching intuitive names to variables and datasets), and services that support discovery of data. The anchore for both the IDV and the THREDDS is the the online curriculum, which is discoverable in digital libraries. This allows a student or educator to discover real-time data along with the tools and curricula they need to actually learn from the data.

By linking these three components, learners and educators always discover the data in contexts that ensure its usability and comprehensibility.

Figure 1: A path for data discovery by finding the VGEE in DLESE.

3. DISCOVERY IN DIGITAL LIBRARIES

The VGEE provides a model for discovering distributed data as well as the tools and curriculum that give data meaning and educational utility. Activities within the VGEE curriculum are catalogued as resources within DLESE. These activities, once discovered, include embedded links to the learner-centered visualization tool. The visualization tool itself is configured to link to THREDDS servers, which provide catalogs of distributed data relevant to the curriculum. This process is illustrated in Figure 1.

4. A GENERALIZABLE ARCHITECTURE

The VGEE/THREDDS/DLESE collaboration provides architecture to support other data-based curricula. The easy customization of the visualization tool means that new user interfaces can be built to accommodate different learners and goals. These interfaces could select appropriate interactive models from a library of such models. THREDDS services provide a similar library of data sets; by negotiating protocols and providing semantics for data sets, THREDDS gives educational developers transparent access to a variety of data types distributed over a number of sources. DLESE provides a forum to organize these development resources and create a community of developers.

5. ACKNOWLEDGMENTS

NSF NSF/CCLI #9972491 supports VGEE development. THREDDS is supported by NSF/DUE #0121623. Unidata's IDV is the basis for the VGEE visualization tool. Unidata is funded by NSF/ATM #9218790. DLESE is funded by NSF/EAR #0215640.

6. WEBSITE

http://www.dpc.ucar.edu/vgee

7. REFERENCES

DLESE: Digital Library for Earth System Education (DLESE). [Available online from http://www.dlese.org.]

Unidata Program Center: Integrated Data Viewer. [Available online from http://my.unidata.ucar.edu/content/software/metapps/index.html.]

Unidata Program Center: THematic Real-Time Environmental Data Distributed Services. [Available online from http://my.unidata.ucar.edu/content/software/thredds/index.html.]

Wright, M. J. and T. R. Sumner, Eds.: DLESE Developers' Workshop White Paper, Draft 4.0. [Available online from http://www.dlese.org/libdev/workshops/2003_Dev/Workshop_Report_Draft_4_dist.pdf.]