



A vision for end-to-end data services to foster international partnerships through data sharing

M. Ramamurthy and T. Yoksan

University Corporation for Atmospheric Research, Unidata, Boulder, United States (mohan@ucar.edu)

Increasingly, the conduct of science requires scientific partnerships and sharing of knowledge, information, and other assets. This is particularly true in our field where the highly-coupled Earth system and its many linkages have heightened the importance of collaborations across geographic, disciplinary, and organizational boundaries. The climate system, for example, is far too complex a puzzle to be unraveled by individual investigators or nations. As articulated in the NSF Strategic Plan: FY 2006-2011, "...discovery increasingly requires expertise of individuals from different disciplines, with diverse perspectives, and often from different nations, working together to accommodate the extraordinary complexity of today's science and engineering challenges." The Nobel Prize winning IPCC assessments are a prime example of such an effort.

Earth science education is also uniquely suited to drawing connections between the dynamic Earth system and societal issues. Events like the 2004 Indian Ocean tsunami and Hurricane Katrina provide ample evidence of this relevance, as they underscore the importance of timely and interdisciplinary integration and synthesis of data. Our success in addressing such complex problems and advancing geosciences depends on the availability of a state-of-the-art and robust cyberinfrastructure, transparent and timely access to high-quality data from diverse sources, and requisite tools to integrate and use the data effectively, toward creating new knowledge.

To that end, Unidata's vision calls for providing comprehensive, well-integrated, and end-to-end data services for the geosciences. These include an array of functions for collecting, finding, and accessing data; data management tools for generating, cataloging, and exchanging metadata; and submitting or publishing, sharing, analyzing, visualizing, and integrating data. When this vision is realized, users — no matter where they are, how they are connected to the Internet, or what computing device they use — will be able to find and access a plethora of geosciences data, experience how all of the aforementioned services work together, and use our tools and services both productively and creatively in their research, education, and other activities.

Permit me to elucidate on what that vision really means for you by drawing a simple analogy. Most of you are familiar with Amazon and eBay e-commerce sites and content sharing sites like You Tube and Flickr. On the eBay marketplace, people can sell practically anything at any time and buyers can share their experience of purchasing a product or the reputation of a seller. Likewise, at Amazon, thousands of merchants sell their goods and millions of customers not only buy those goods, but provide a review or opinion of the products they buy and share their experiences with the purchase. Similarly, You Tube and Flickr are sites tailored to video- and photo-sharing, respectively, where users can upload their own content and share them with millions of other users, including family and friends. What all these sites have enabled is a sense of a virtual community in which users can search and browse products or content, comment and rate those products from anywhere, at any time, and via any Internet-enabled device like an iPhone, laptop, or a desktop computer. In essence, these enterprises have fundamentally altered people's buying modes and behavior toward purchases. I believe that similar approaches, appropriately tailored to meet the needs of the scientific community, can be adopted to provide and share geosciences data in the future. For example, future case-study data access systems, in addition to providing datasets and tools, will provide services that allow users to provide commentaries of a weather event, say a hurricane, as well as provide feedback on the quality, usefulness and interpretation of the datasets through integrated blogs, forums and Wikis, upload and share products they derive, ancillary materials that users might have gathered (such as photos and videos from the storm), and publications and curricular materials they develop, all through a single data portal. In essence, such case study collections will be "living" or dynamic, allowing users

to be also contributors as they add value to and grow existing case study collections.

At Unidata, our goal is to provide a portfolio of integrated data services toward realizing the vision presented here so that the geosciences community can continue to address societally relevant problems such as weather prediction, atmospheric and oceanic variability, climate change, and the water cycle, and advance scientific discovery.