NEWS

Long-term Bird Study Records Arctic Climate Change

PAGE 522

Alaska's summer of 2005 was the second warmest on record there, with a record retreat of arctic pack ice. As Alaskan temperatures gradually increase, artic birds, such as the black guillemots of Cooper Island, near Barrow, Alaska, are experiencing drastic habitat changes. Though these small black and white birds—the subjects of a long-term study of climate change—fared better this year than they have in the recent past (due to local cool conditions), they are nonetheless struggling to adapt as their artic island summer home becomes subarctic.

George Divoky, an ornithologist at the Institute of Arctic Biology, University of Alaska Fairbanks, discovered the Cooper Island colony of guillemots in the early 1970s and has spent every summer since 1975 there studying these birds. He presented his latest research during a 3 November talk in Washington, D.C.

Cooper Island is a small sand and gravel barrier island located about 40 kilometers east of Barrow, Alaska. For decades, it was a perfect home for the guillemots: It was close to the sea ice under which they catch arctic cod, and the birds could nest in debris left there in the 1950s during U.S. Navy ordnance testing.

Divoky was the first to note that the dates of the first egg laid in the colony and the median egg have been occurring about three days earlier per decade since the 1970s. He said this was happening due to climate change, and that it reflected the fact that the snow melt was also earlier, which was important because the birds could not enter their nests until then.

By examining records from Barrow, Divoky determined that the Cooper Island summer did not become long enough for guillemots to breed there until the 1970s, about the time when he discovered them. As the summer grew longer, the guillemot population continued to expand. But after 1990, the population declined by half, which corresponded with a shift in the arctic oscillation. Air pressure had been high over the North Pole, and winds had blown from the north in a constant northeast wind. When the North Pole pressure dropped around 1990, warm air started blowing, and the circulation of ice in the Arctic basin changed.

The area of winter ice has been decreasing by two percent per decade, with the summer ice decreasing by seven percent per decade, Divoky said. In addition, ice volume has decreased 40 percent in the last 50 years.

The sea ice has moved farther from shore, and guillemots now need to travel farther for their arctic cod or fish closer to the island for subarctic species, which have less energy content.

By 2000, Cooper Island was experiencing intense storms. In addition, horned puffins, a subarctic bird species, began to colonize the island and compete with guillemots for nesting space. By 2004, guillemot breeding success dropped to 12 percent from an average of 53 percent in 1975–1990.

Because the sea ice was closer to shore for much of the 2005 summer, breeding success climbed to 57 percent. Whether the guillemots will be so successful next year depends on the ice and their ability to adapt to the area's climate change.

—SARAH ZIELINSKI, Staff Writer

MEETINGS

Making Scientific Data Sets Easier to Find, Access, and Use

PAGES 522, 525

Over the years, a tremendous challenge has developed in the Earth and atmospheric sciences: that of managing and searching increasingly large volumes of data collected over relatively short time periods across multiple disciplines and from multiple platforms.

Significant resources have been spent often on activities that are program- or project-specific, which has resulted in many different standards, conventions, approaches, and tools for data documentation and management. However, it is still difficult to make data truly interoperable across the breadth and depth of the Earth and atmospheric sciences. Indeed, a key issue identified by scientists and technologists alike is the ability to use data (and its descriptive metadata), as described by different vocabularies, in a transparent, discipline-independent, and scientifically valid way. With the proliferation of vocabularies specific to individual projects, data repositories, domains, and even individual scientists, finding and using data has become a manual and often labor-intensive task.

The Marine Metadata Interoperability (MMI) project was created to help meet this challenge specifically for oceanographic data and to build upon previous program- or project-specific activities, such as the Global Change Master Directory (GCMD), the Semantic Web for Earth and Environmental Terminology, and the British Oceanographic Data Centre. MMI, now funded by the U.S. National Science Foundation (grant ATM-0447031), is led by an international steering committee (that includes this author) from a diverse set of government and academic and research institutions.

MMI seeks to identify best practices for making oceanographic data easy to distribute, advertise, reuse, and combine with other data sets. Its community-supported Web site, http://marinemetadata.org, provides the marine data management community with a comprehensive portal to general information, standards, ontologies, tools, usage guides or 'cookbooks,' and working examples. By simplifying the complex world of metadata into specific, simple guidance, MMI will allow scientists and data managers at all levels to apply good metadata practices from the start of a project.

Since the MMI project is a community-based effort, one of its primary tasks is to engage community members in creating solutions for metadata-related data management needs within oceanography. Toward this end, MMI sponsored a recent workshop, Advancing Domain Vocabularies.

A domain vocabulary is a resource that may be used to describe data in a standardized manner in order to facilitate interoperability with other related data sets. A vocabulary (usually consisting of a plain list of words, a dictionary, or even an Extensible Markup Language (XML) document) becomes an ontology when concepts are defined explicitly through the creation of machine-readable classes and subclasses, with relationships defined between them. Vocabularies and ontologies are becoming increasingly important in ensuring the effectiveness of data repositories, information portals, and digital libraries.

The workshop brought together 47 marine scientists, ontology specialists, and data management technicians, each with multiple areas of expertise in science and technology, from 40 different organizations. They collectively identified commonly-used vocabularies for describing oceanographic data sets for a select group of six domains: benthic habitat classifications, chlorophyll and pigments, conductivity-temperature-depth (CTD) parameters, Pacific Coast Ocean Observing System variables, sensors and instruments, and waves and currents.

Immediate and specific goals for the workshop included:

• Identifying from existing marine vocabularies recommended discovery terms (keywords or categories used to tag data sets so they can be found) and markup terms (which identify specific variables in a data set);
• Developing mappings of terms among the recommended vocabularies;
• Developing mappings of terms from other vocabularies to those in recommended vocabularies; and
• Demonstrating the value and utility of the above via a Web service application.

Long-term science goals that will be carried into subsequent efforts include being able to find science data without knowing the precise vocabulary used to label it, merging science data from different data sources, creating reference vocabularies that can be used for new data sets, and ultimately allowing more sophisticated and automated discovery and analyses.

To achieve these goals, the workshop organizers sought to connect data management professionals and ontologists with scientific domain experts in the shared pursuit of a common challenge. If the two sets of experts could unite on specific approaches, these and larger technical and scientific challenges might be addressed. Introductory sessions defined and demonstrated the value of vocabularies and their associated mappings (in finding terms, finding data, and using data) and provided final training in vocabulary and ontology concepts, the procedures of vocabulary mapping, and new applications and Web services that use the mappings. For this exercise, the most common mappings labeled each term as 'same as', 'broader than', or 'narrower than' another term. The properties of these relationships allowed them to be used to infer other mappings automatically. Such automated 'reasoning' illustrates in a simple way the strength of ontologies and associated semantic web concepts.

Workshop leaders sought a balance between providing technical detail and demonstrating progress. The detail was critical to achieving real technical results, but the goal of technical progress kept the meeting from getting hopelessly bogged down in details. Achieving this balance was particularly difficult because the different participant groups had different experiences and skills.

Participants in each of the six domain areas were charged with identifying the common metadata terms used in their domains in the context of existing vocabularies, which the team members primarily identified in advance of the workshop.

Each vocabulary was harmonized with the others prior to the workshop by translating them into the common Ontology Web Language (OWL) for subsequent mapping to other vocabularies. This translation was accomplished using a newly developed tool, Voc2Owl, which converts simple vocabularies (in ASCII format) to OWL. Relationships were then mapped between these OWL-formatted vocabularies using another new tool called Vocabulary Integration Environment (VINE). Both tools were developed for the workshop by Luis Bermudez, of MMI and the Monterey Bay Aquarium Research Institute, and are available from the MMI Web site.

Once the mappings were accomplished, they were then used for data discovery across several existing data systems on the Web already employing these vocabularies.

The success of the workshop can be measured by its productive outcomes, beginning with the 50 vocabularies harmonized to OWL that provided the input to the domain mappings. Over 800 mappings of terms in these harmonized vocabularies were generated directly by the participants, primarily in the domains of chlorophyll, CTD, and currents and waves. This resulted in 2200 automatically generated, or inferred, mappings as described above, for a total of over 3000 mappings. This was a significant accomplishment for a diverse group of participants new to the process and the tools.

As a result of the mappings, the participants were able to experiment with some live Web services, where they could use more specific terms to search more effectively through repositories (e.g., the broad GCMD keyword 'pigments' automatically related to specific pigments such as chlorophyll-*a*, chlorophyll-*b*, and beta-carotene). In addition, based on workshop feedback, the two essential software tools Voc2Owl and VINE have been updated and repackaged.

Although much work remains to be done to effectively find, access, and use scientific data, the results from this workshop are important for improved data discovery and the cost-effective use and interpretation of the discovered data. The environmental science community is still faced with a formidable array of disparate databases and portals across the Internet, as well as the complex infrastructure of ocean observing system repositories. However, the Advancing Domain Vocabularies workshop provided a firm foundation for future related activities, such as follow-on workshops and other activities that will connect the results of this workshop, such as the sensors ontology, to other community efforts.

The workshop, Advancing Domain Vocabularies, was held 9–11 August 2005 at the University Center for Atmospheric Research in Boulder, Colo. All workshop materials, proceedings, reports, activities planning, and the two software tools are available online at http://marinemetadata.org/workshop05/.

—DAWN WRIGHT, Oregon State University, Corvallis; E-mail: dawn@dusk.geo.orst.edu; STEPHANIE WATSON, Texas A&M University, College Station, E-mail: steph_watson@sbcglobal.net; JOHN GRAYBEAL and LUIS BERMUDEZ, Monterey Bay Aquarium Research Institute, Moss Landing, Calif.; E-mail: graybeal@mbari.org, bermudez@mbari.org

Workshop Highlights Progress in Solar-Heliospheric Physics

PAGES 525, 530

The Solar, Heliospheric, and Interplanetary Environment (SHINE) group is an affiliation of researchers dedicated to promoting an enhanced understanding of the processes by which magnetic fields, plasmas, and energetic particles are produced near the Sun and propagated through the interplanetary medium to Earth and other locations in the heliosphere. The group conducted its annual workshop in July to discuss recent developments in the study of solar variability and its impact on Earth's space environment. One hundred fifty-five scientists, including 27 students, participated in the plenary, working group, and poster sessions.

Student Day activities on 10 July consisted of tutorials given by experienced scientists: solar flares and particle acceleration (Robert Lin, University of California Berkeley), the origin of coronal mass ejections (CMEs) (Spiro Antiochos, Naval Research Laboratory, Washington, D.C.), connecting the Sun and heliosphere (Thomas Zurbuchen, University of Michigan, Ann Arbor), and acceleration and transport of solar energetic particles (SEPs) (Christina Cohen, California Institute of Technology, Pasadena). The tutorials were followed by student presentations on CMEs near the sun and in the interplanetary medium, solar wind, and SEPs.

In presentations designed to serve as an overview of issues that would be discussed later in the working group sessions, the plenary speakers addressed topics of interest to the entire SHINE community: the subsurface magnetic field structure and evolution (George Fisher, UC Berkeley), shocks and particle acceleration (Martin Lee, University of New Hampshire, Durham), particle acceleration near the Sun (Lin), and end-to-end modeling of CMEs and SEPs (Tamas Gombosi, University of Michigan, Ann Arbor). Funding agency representatives [Paul Bellaire, U.S. National Science Foundation (NSF); Madhulika Guhathakurta, NASA; David Byers, Air Force Office of Scientific Research] also made informative plenary talks.

Shine Campaign Events

Each year, the SHINE group focuses its attention on certain well-observed solar eruptive events, called 'campaign events,' that have significant heliospheric consequences. One session covered campaign events (12 May 1997, 1 May 1998, 21 April 2002, 24 August 2002) that were considered in the 2004 workshop. Information on the background corona and the CMEs propagating through it that was obtained from these events proved to be critical inputs to CME models (Richard Frazin, Univer-