Ledley, TS, A Prakash, C Manduca, and S Fox, (2008), Recommendations for Making Geoscience Data Accessible and Usable in Education, Eos, Vol. 89, No. 32, (DOI: 10.1029/2008EO2003). To view the published open abstract, go to http://dx.doi.org and enter the DOI.
Direct link for AGU members: http://www.agu.org/journals/eo/eo0832/2008EO320003.pdf#anchor

Recommendations for Making Geoscience Data Accessible and Usable in Education

Eos, Vol. 89, No. 32, 5 August 2008

Geoscience and planetary science are awash in data. Over the past decade, our field has transformed into one that acquires and uses extensive remote-sensing data. Enabling state-of-the-art models of complex systems, these data focus our understanding of the future impact of our actions by examining the present and past.

Our scientific and technical communities have been pushed by society to make scientific data accessible and usable by the educational community. Much geoscience data, however, are difficult for educators and students to find and use. They are generally described by metadata (i.e., information about the data) that are usually narrowly focused and cryptic, which challenges educators and researchers to discern whether a dataset is relevant to their needs and to access and use the data effectively.

Data-related research initiatives (Earth Observatory, EarthScope, Global Land Cover Facility, Linked Environments for Atmospheric Discovery, NASA Earth Observations, The Geosciences Network, and the Thematic Realtime Environmental Distributed Data Services (THREDDS) project) as well as the Distributed Active Archive Centers and the Global Change Master Directory (GCMD) have sought to include metadata for educational use. However, these metadata focus on supporting discovery or access once a specific data requirement has been determined. What is needed is metadata that outline the data's range of pedagogical uses to support the design of curriculum materials and student research activities.

Overview

To enable easy access and use of geoscience data, we propose that the educational and scientific communities collaboratively implement 1) the educationally relevant review criteria for data-rich Web sites, which we refer to as "data sites"; and 2) the educationally relevant metadata for datasets, both of which are described below. In order for the implementation of these recommendations to effectively support the educational community's use of geoscience data, it is critical that enacting these recommendations be viewed as a collaborative effort between the scientific and educational communities.

It is important that students and teachers access and use geoscience data. Three key reports—The National Science Education Standards [*National Research Council*, 1996; 2000], the Benchmarks for Science Literacy (Project 2061) [*AAAS*, 1993], and the Revolution in Earth and Space Science Education report [*Barstow and Geary*, 2002]—emphasize that students should learn science through inquiry and should understand the concepts and processes that shape our natural world. One way to involve students in scientific inquiry is to provide the data and analysis tools they need to explore concepts and answer questions.

The advantages of using data in science education are highlighted in the Using Data in the Undergraduate Science Classrooms report [*Manduca and Mogk*, 2002]. This report was the outcome of a workshop sponsored by the National Science Digital Library with funding from the U.S. National Science Foundation (NSF) on the topic in April 2002. These advantages include 1) preparing students to address real-world problems; 2) preparing them to critically evaluate the validity of data evidence and their conclusions; 3) teaching quantitative skills, technical methods, and scientific concepts; and 4) building scientific, technical, quantitative, and communication skills [*Manduca and Mogk*, 2002, p2].

In response to these reports, the digital library community is exploring ways to make geoscience

data more accessible and usable in the educational community. Increasingly, datasets are being made available by the scientific and technical communities to broader audiences through data-rich Web sites (data sites) that provide interfaces to, and supplementary information about, the data. Data sites can serve educator needs to access geoscience data for their students; however, the data sites need educationally-relevant metadata to make them more discoverable and usable. Additionally, digital libraries are looking to ensure the quality of their collections. The educationally relevant metadata described here, and the scientific metadata, can enable a data site to meet the review criteria proposed below, thus ensuring quality.

The Data Access Working Group (DAWG), composed of individuals from the digital library, data providers, research, and education communities, was first assembled in 2001 to identify the hurdles of using data as a learning tool and to outline strategies for overcoming these hurdles. The DAWG focused on two critical areas: (1) Educationally relevant review criteria for data sites, including those sites that are not part of a finished educational product, that would prevent their inclusion in collections of educational resources that use scientific data; and (2) educationally relevant metadata standards for data sites in digital libraries, to facilitate the sites' discovery, access, and use for student learning.

Educationally Relevant Review Criteria for Data Sites

These criteria are appropriate for the broad range of data-related resources that support teaching about the Earth system. The data-related resources include data, data access portals, data visualization tools, models, model data and visualizations, data processing sites, and data location sites. The review criteria can be applied to data sites which are diverse in the technologies they use, the nature of their data, and their approaches to annotation and description. The review criteria acknowledge this diversity by focusing on how effectively each site as a whole meets key educational needs, rather than prescribing particular technical approaches or data organization schemes.

Recommended Review Criteria for Data Sites

The recommended review criteria are listed below and can also be found at the Using Data at the Classroom portal (http://serc.carleton.edu/usingdata/site_criteria.html). The term 'curriculum developer', used in the criteria, refers to anyone developing educational materials using the data.

1. Curriculum developers can find and use appropriate data easily.
    - The level of knowledge for use is clear.
    - Information is provided on the relevance of data to societal, educational, or scientific problems.
    - The data site points to examples of educational use of data.
    - Data are archived for reliable, long-term access.
    - The interface to, and the organization of, the data is user-friendly.
    - Descriptive information (e.g., metadata) is audience-appropriate.

2. Curriculum developers can ascertain data quality and its impact on their conclusions.
    - Information is provided about overall data collection, quality, reduction, and limitations. Data sites include sources of error and limitations of the collection process as well as inaccuracies and uncertainties from models and from particular choices of representations.
    - Information about the accuracy of individual data sets, points, and analyses is provided.

3. Data are provided in ways that facilitate their manipulation through a variety of tools, including:
    - Providing the data in common formats that are accessible via desktop tools (spreadsheets, geographic information systems, image analysis programs);
    - Providing tools that are integrated into, or downloadable from, the site; and
    - Exposing the data via protocols to make them accessible through third-party tools. These tools should be easily acquired, easy to use, and reliable.

4. The data site supports, through these tools, data manipulation to answer questions by:
    - Using data contained within the site or combined with data from other sites;
    - Generating appropriate visualizations; and
    - Comparing students' own data to the data on the site.

Educationally Relevant Metadata Standards for Datasets

To implement the recommended review criteria for data sites, data require educationally relevant metadata. However, existing science education metadata such as the Alexandria Digital Library/Digital Library for Earth System Education/NASA (ADL/DLESE/NASA: ADN) metadata focuses on completed educational materials, not on the educational uses of datasets which are not themselves educational activities. We recommend a set of metadata for datasets that: targets educators; extracts the educationally relevant portions of information from the Federal Geographic Data Committee (FGDC) and ADN metadata; and provides insight into possible uses of the data.

The Science Education Resource Center (SERC) at Carleton College, with input from DAWG and the broader science education community, has created the template for a "DataSheet" and for educationally relevant metadata for datasets, and has begun building a collection of DataSheets (http://serc.carleton.edu/usingdata/browse_sheets.html). A DataSheet concisely describes a scientific dataset for educators interested in teaching a specific scientific concept or analysis technique with the dataset it describes. A DataSheet highlights the connections between the dataset it describes and science topics; it explicates how to acquire the data and how to include (when available) links to classroom activities that use the data. A DataSheet also provides the information needed for a dataset to meet the educationally relevant review criteria for datasets, mentioned above.

Recommended Educationally Relevant Metadata for Datasets

The following DataSheet components are recommended as educationally relevant metadata for datasets (full descriptions of these fields are available at http://serc.carleton.edu/usingdata/about_datasheets.html, and an online template for contributing a DataSheet is available at http://serc.carleton.edu/usingdata/contribute_datasheet.html).

A DataSheet has the obvious components such as author, title, Web addresses to the data, the organization archiving the data, and a graphic representation of the data. Other key elements of a DataSheet are:

1 a description of the data, how they are presented, and their geospatial and temporal extent;

2 the use and relevance of the data that describes the data's importance;

3 how the data might be used in teaching, including scientific topics and data-use skills;

4 how to explore the data, including the nature and presentation of the data; how to access, visualize, and manipulate the data; acronym and jargon definitions; and analysis tools to manipulate and analyze the data;

5 how the data was collected, its limitations, and sources of error; and

6 scientific and educational references and resources that use the data.

It is vital that the scientific and educational communities collaborate to develop DataSheets. The AccessData Workshops (http://serc.carleton.edu/usingdata/accessdata) present a model of an effective collaborative mechanism for creating both DataSheets and educational modules using the data [*Ledley et al*., 2005] from research projects.

The DataSheet template enables scientists and educators to jointly develop educationally relevant metadata for the benefit of both communities. One pathway for creating a DataSheet is first to review

the Web site (http://serc.carleton.edu/usingdata/about_datasheets.html) which describes and exemplifies the fields in a DataSheet, and then to complete and submit the online DataSheet template at http://serc.carleton.edu/usingdata/contribute_datasheet.html. We will work with submitters to identify the most effective way of incorporating their DataSheet(s) into the collection of DataSheets (http://serc.carleton.edu/usingdata/browse_sheets.html). Once developed, DataSheets can be cataloged in digital libraries (e.g. Digital Library for Earth System Education (DLESE), National Science Digital Library (NSDL) using the ADN scheme, making the data discoverable by the broader educational community. Widespread use of DataSheets will enhance science learning by facilitating the broad use of the data they describe.

Web Sites

Earth Observatory, http://earthobservatory.nasa.gov/.

EarthScope, http://www.earthscope.org/home, http://www.earthscope.org/data.

Global Change Master Directory, http://gcmd.nasa.gov/index.html.

Global Land Cover Facility, http://glcf.umiacs.umd.edu/index.shtml.

Linked Environments for Atmospheric Discovery - LEAD, https://portal.leadproject.org/gridsphere/gridsphere

NASA Earth Observations - NEO, http://neo.sci.gsfc.nasa.gov/Search.html.

The Geosciences Network -GEON, http://www.geongrid.org/, http://www.geongrid.org/resources/data.

Thematic Real-time Environmental Distributed Data Services (THREDDS): http://www.unidata.ucar.edu/projects/THREDDS/,

Acknowledgments

DAWG collectively developed these recommendations. The contributing members of the DAWG can be found at

http://serc.carleton.edu/files/usingdata/dawg/dawg_members_contributing_to_r_1192558649.pdf. This

effort was supported by NSF grants EAR-0305058, EAR-0620360, and GEO-0304762. Any opinions, findings, conclusions or recommendations expressed in this article are those of the author(s) and do not necessarily reflect the views of NSF.

References

AAAS (1993), Benchmarks for Science Literacy, 418 pp., Oxford University Press, New York, Oxford.

Barstow, D., and E. Geary (Eds.) (2002), Blueprint for Change: Report from the National Conference on the Revolution in Earth and Space Science Education, 100 pp., TERC, Snowmass, Colorado, http://www.earthscienceedrevolution.org/acknowledge/acknowledge01.cfm.

Ledley, T. S., L. Dahlman, B. Domenico, and M. R. Taber (2005), Facilitating the Effective Use of Earth Science Data in Education through Digital Libraries: Bridging the Gap between Scientists and Educators, IEEE Technical Committee on Digital Libraries Bulletin, 2(1), , http://www.ieee-tcdl.org/Bulletin/v2n1/ledley/ledley.html.

Manduca, C. A., and D. Mogk (2002), Using Data in Undergraduate Science Classroom: Report from an Interdisciplinary Workshop, edited, Carleton College, http://dlesecommunity.carleton.edu/research_education/usingdata/.

National Research Council (1996), National Science Education Standards, 262 pp., National Academy Press, Washington D.C., http://www.nap.edu/books/0309053269/html/.

National Research Council (2000), Inquiry and the National Science Education Standards, 203 pp., National Academy Press, Washington DC, http://www.nap.edu/books/0309064767/html/.

Author Information

Tamara Shapiro Ledley, TERC, Cambridge, Mass.; E-mail: tamara_ledley@terc.edu; Anupma Prakash, Geophysical Institute, University of

Alaska Fairbanks, Fairbanks; and Cathryn A. Manduca and Sean Fox, Science Education Resource Center (SERC), Carleton College, Northfield, Minnesota