

Public Access and Use of Electronically Archived Data: Ethical Considerations

Ecologists today are increasingly expected to electronically archive their data sets and make them publicly accessible. For example, NSF now requires that LTER sites have major data sets documented and available on-line within a reasonable time of collection. This is generally interpreted as 3–5 years, a lag designed to allow those who collected data a reasonable period in which to publish results. The expectation that data will be archived and shared may well spread to all NSF-sponsored research. Indeed, NSF's Grant General Conditions state that NSF "expects investigators to share with other researchers, at no more than incremental cost and within a reasonable time, the data, samples, physical collections and other supporting materials created or gathered in the course of the work."

Although some ecologists have made their data available following publication, making data available before they are published, or possibly even before they are analyzed, makes many ecologists uneasy and raises

some ethical issues. Although fears of being scooped or contradicted by one's own data may be unwarranted, such fears could promote hiding or poor documentation of data, or hasty and premature publication, therefore undermining progress in the field. In addition, unpublished data, particularly unanalyzed data, often contain data entry and other errors, errors that are usually discovered during analysis by those who collected them. Public access to such data could result in erroneous conclusions by others if those who collected the data were not consulted.

We believe that making more data accessible to more ecologists has the potential for greatly increasing communication, collaboration, and synthesis within the ecological community. However, for the latter to occur, it is imperative that the ecological community develop, and individual members abide by, a set of ethical guidelines that promotes data sharing, collaboration, and scientific advances, while at the same time acknowledging and respecting those individuals whose time, energy, and intellectual efforts produced the data. Currently, no ethical code of conduct exists within the ecological community governing the access and use of archived data. Thus, we strongly encourage ESA to develop such a code, by which all members of the ecological community would be expected to abide. Ideally such a code would be developed collaboratively with the ecological societies of other countries, since the ecological community is global, and electronically archived data can be accessed from anywhere in the world.

Currently, ESA has a general code of ethics <<http://esa.sdsc.edu/codeofethics.htm>>. This code lists several principles of conduct that are relevant to the issues we have raised concerning the electronic archiving of data. These include:

- Researchers will not submit for publication any manuscript containing data they are not authorized to use.

• ESA assumes that the principal investigator(s) of a research project retain the right to control use of resulting unpublished data, unless otherwise specified by contract or explicit agreement.

We believe ESA needs to clarify whether these principles are also intended to apply to the use of electronically archived data. Specifically, we believe that ESA needs to address the issue of electronically archived data in its code of ethics and/or develop a code specifically addressing this topic. In order to encourage the development of such a code (whether a supplement to the current code of ethics or a separate code), we offer the following working draft that ESA could use as a starting point for such an effort.

A Proposed Code of Ethics for Access and Utilization of Archived Data

Science is a community enterprise that prospers most when there is open dissemination of ideas and data. Thus, The Ecological Society of America supports the archiving and public accessibility of ecological data and urges all ecologists to make their data accessible in this way. It is expected that raw and summary data will be made available promptly following any publication of the data. In addition, all summary and raw data, published or not, normally should be made publicly accessible no more than 5 years following their collection or determination. Should further delay be necessary, reasons for this delay must be publicly provided, along with an expectation of the release date. It is expected that those who gathered the data will document the archived data sets sufficiently to permit their use and interpretation by others. Furthermore, all individuals who access and use data gathered by others are ethically obliged to respect and acknowledge the time, energy, and intellectual effort expended by those who produced the data. Any indi-

vidual who analyzes or presents archived data, or any portion of the data, in any form or fashion, is required to have contacted those who produced the data and to have thoroughly discussed with them issues of data quality, interpretation, acknowledgment, recognition, and authorship before any dissemination of ideas, results, or conclusions based on the data occurs.

If such a code of conduct were to exist, then individuals and research teams archiving data sets could post a copy of the code to their archive web site reminding users of their ethical obligations. Authors of manuscripts submitted to ESA journals, and authors of posters or oral presentations at ESA-sponsored events, could be required to indicate whether any archived data sets other than those of the author(s) were used in the papers or presentations. If so, the authors could be required to provide the names of those individuals whose data sets were used and to indicate how they had resolved issues of data interpretation, acknowledgment, recognition, and authorship. This could be accomplished via a standardized form provided all authors by the editors or ESA. Should questions remain, editors and meeting organizers could then contact the individual(s) whose archived data sets were being used to obtain additional information and perspective. Although ESA could require such information only for its own journals and meetings, it would be expected that ESA members would abide by the ethical code of conduct no matter where or in what form any dissemination of ideas, results, or conclusions based on another's archived data occurs.

We believe that such a code of ethics would provide the assurances that both data gatherers and data users need to make data sharing a win-win situation. By providing leadership in developing and implementing such a code, ESA could help the ecological community adapt to, and benefit from, the newly emerging culture of open data sharing.

*Mark A. Davis
Department of Biology
Macalester College
Saint Paul, MN 55105
E-mail: davis@macalester.edu*

*David Tilman, Sarah E. Hobbie, and
Clarence L. Lehman
Department of Ecology,
Evolution, and Behavior
University of Minnesota
Saint Paul, MN 55108*

*Peter B. Reich
Department of Forest Resources
University of Minnesota
Saint Paul, MN 55108*

*Jean M. H. Knops
School of Biological Sciences
University of Nebraska
Lincoln, NE 68588*

*Shahid Naeem
Department of Zoology
University of Washington
Seattle, WA 98195*

*Mark E. Ritchie
Ecology Center and
Department of Fisheries and Wildlife
Utah State University
Logan, UT 84322*

*David A. Wedin
School of Natural Resource Sciences
University of Nebraska
Lincoln, NE 68583*