Data Curation and Distribution in Support of Cornell University's Upper Susquehanna Agricultural Ecology Program

Gail S. Steinhart*
Albert R. Mann Library
Cornell University
Ithaca, NY 14853
GSS1@cornell.edu

Brian J. Lowe
Albert R. Mann Library
Cornell University
Ithaca, NY 14853
BJL23@cornell.edu

Abstract

Effective documentation, curation, and provision of access to scientific data are essential to derive the full benefit of research data, both for participants in specific research projects and for the entire scientific community. Academic research libraries are positioned to be important partners in such endeavors, although success will depend in part on expanding and changing the customary roles of, and relationships between, researchers and libraries. Cornell University's Albert R. Mann Library is collaborating with the Upper Susquehanna Agricultural Ecology Program at Cornell to document and distribute the group's research data. In addition to collecting data and developing numeric and spatial models, the research group has access to approximately thirty years worth of observational data for their research sites, which are of significant value to environmental scientists. The approach includes identifying and using discipline-specific metadata standards in order to facilitate participation in discipline-specific data and metadata sharing initiatives, at the discretion of individual researchers. Training is provided for project collaborators in the use of existing metadata creation tools to create documentation for their datasets. "Pre-publication" data and metadata are stored in a database accessible only by project members, to facilitate early sharing and collaboration within the group. Complete, documented data sets and complete metadata records will then be deposited in Cornell's DSpace installation. As a test case, the historic data sets are being formatted and documented for deposit in DSpace. A public web portal provides information about the project and participants, as well as a future means of access to project datasets.