

New Collaborative Relationships: The Role of Academic Libraries in the Digital Data Universe

Suzanne Carbotte, Program Director, Marine Geoscience Data System, Lamont-Doherty Earth Observatory, Columbia University

Comments regarding challenges facing the long-term preservation and management of digital data;

How do we ensure the engagement of data producers at various stages in the data life cycle. To develop comprehensive digital data resources of maximum use to the research and education community requires the active involvement and engagement of individual scientists who are essential data producers throughout the data life cycle. A diverse range of individual scientists may be involved at various stages throughout the data lifecycle as field data are processed, reprocessed for new applications, integrated into data syntheses, and higher level derived data products are developed. It must be easy (transparent?) for scientists to document and contribute their data products, scientists need to have their contributions adequately protected and acknowledged, and new rewards for contribution are needed.

Inadequate enforcement of data policies. New governance structures for enforcement of data policies are needed. In some realms of scientific research, existing NSF data policies have been difficult to enforce, partly because appropriate digital data repositories have not existed, but also because mechanisms to document compliance are not in place. Another aspect of the problem is that the scientist may be the only one who knows of the existence of data and compliance of the individual scientist with a data policy must be based on trust and commitment to data preservation as part of the scientific process.

Inadequate incentives for scientist participation in data preservation. New incentives for scientists to contribute to data collections go hand in hand with the need to fully engage data producers in the data preservation process and the need for new structures for enforcing data policies. We need to change how we reward and credit scientists for data contribution. Contribution to databases needs to be part of the publication process, and we need a new system of professional recognition that acknowledges the value of data production and contribution to data collections. New partnerships with the academic journals will be needed to develop policies for publication, which include linking publications to digital data resources.

How do we ensure the long-term security of digital data collections in an uncertain funding climate? New scenarios and partnerships to ensure long-term funding are essential for both the development and security of digital data collections. To adequately manage and preserve the complex heterogeneous data that are produced in an increasingly multidisciplinary research environment requires data managers with a high level of expertise in both the domain sciences as well as information technology. Such people are difficult to find and keep in an uncertain and short term funding climate.