APPENDIX A

# TASK FORCE CHARTER

<u>NSB-04-19</u>
February 5, 2004

## CHARTER FOR THE LONG-LIVED DATA COLLECTION TASK FORCE

Data collections, particularly digital data collections for research, have been increasing in number and size over the past couple of decades. They range from small single investigator collections to very large collections whose content is derived from instruments housed in large facilities. Over this same period, the National Science Foundation (NSF) obligations for support for both data collection and curation has been increasing.

The Foundation differs from agencies, such as NASA, NOAA, and the Department of Energy. Most frequently, their strategy is for the agency to own and manage the collections. As a consequence, they own and manage many fewer independent collections than the NSF supports. With ownership comes agency control for data format standards and access policies. The Foundation does not typically maintain data collections itself. It is individual researchers, consortia, and organizations that develop and maintain large facilities that manage the collections. This has resulted in a proliferation of data collections, large and small, across all disciplines. There is divergence in formats, access policies, and in quality.

It is timely to consider the policy ramifications of this rapid growth of data collections in the NSF supported community. This National Science Board task force will address the policy issues directly relevant to the NSF's style of data collection support. These policy issues and questions include:

- When, why, and for how long the NSF will fund data collections that are or appear to be very long-lived (decades)?
- Are there conditions under which it is appropriate for NSF to maintain a data collection intramurally, as most other agencies routinely do?
- What responsibility does NSF have to assure quality of the collections that it supports?
- What part, if any, should NSF play in creating and enforcing community technical standards, for example, for the use and form of metadata?
- How does NSF assure that data is accessible to a broad, diverse, and interdisciplinary community?
- Should the budget for collection curation be made more visible, or remain (more or less) integral to the many different research budgets?

- Is there a desired balance between expenditure on collection and curation?
- Since digital media are impermanent, migrations to new media are crucial if a collection is to persist. Under what conditions should NSF support such migration?
- What policies should guide relationships with national, international, and private agencies and organizations to cooperatively support data curation?

The objective of this National Science Board task force is to delineate the policy issues relevant to the National Science Foundation and its style and culture of supporting the collection and curation of research data. For those issues where guidance to the Foundation is appropriate, the task force should make recommendations for the National Science Board and the community to consider.