Data Management Plan

Dissemination of data, training in statistical tools relevant to these data, and broader public access to compiled materials are a major focus of the proposed project. We will present findings at professional scientific meetings across several disciplines (i.e., primatology, ecology, evolutionary biology, paleoanthropology, and biological anthropology). We will also formally announce the database in relevant journals (e.g., *American Journal of Physical Anthropology, International Journal of Primatology, Evolutionary Anthropology, Journal of Human Evolution, Ecology*, and *Paleobiology*) and promote the availability of the data through online listservs that reach thousands of subscribers daily, such as ECOLOG. Data access will be made available through a web-based portal based at Arizona State University (ASU) *and* through popular biological data repositories (e.g., Dryad and The Knowledge Network for Biocomplexity).

Project databases will be stored on individual computers for daily workflow, backed up on a cloud-based server (e.g., Dropbox), and housed on servers at the University of Massachusetts, Amherst, School of Human Evolution and Social Change at ASU, and Institute for Social Science Research at ASU, enabling us to interface the collected data with Google Earth for both scientific analysis and public outreach. A web-based data portal will include features that allow users to visualize data geographically, perform advanced search queries, and download data of interest.

The web-based data portal will be modeled after the Hadar Geoinformatics Project database, where initial descriptive and geospatial information on fossil collections can be found. The Hadar Geoinformatics Project database was developed previously by Campisano et al. (2008) and is an internally funded project at Arizona State University that currently incorporates all paleoanthropologically relevant datasets from the lower Awash regions of Ethiopia into a single, geospatial database. We plan to use this geoinformatics framework to guide our mapping and database protocols in the proposed project to enhance our data gathering and sharing capabilities. The Hadar Geoinformatics Project is one of the major contributing projects to the NSF-funded Paleocore Project working to develop a data standard for paleontology and paleoanthropology. Paleocore is based on an existing bioinformatics data standard called Darwin Core. Data standards facilitate the integration of data collected from disparate sources, which is a focus of the current project as we are collecting and standardizing data for subsequent research and public access. Furthermore, data standards provide unambiguous definitions for database terms that are critical for ensuring the long-term stability and usefulness of the database for future researchers.

Our plans will allow researchers to access the databases through the web-based portal. We will also develop a public outreach portion on the web that is simple and easy to use and can inform the public about extant primate biodiversity, primate biodiversity of the past, and primate evolution.