

Data Management Plan

Human Subject Data

This project will collect human subjects data consisting of background demographics (including age, gender, ethnicity, field of study, and prior computing experience), career information (including undergraduate and/or graduate education, publication histories, and industrial experience), and extent and kind of participation in open and web-enabled science projects and communities. This data will be used to determine what impact the training provided by this project has on participants' careers and productivity.

The demographic data will be collected from questionnaires and interviews administered by the PI and student/postdoctoral researchers associated with this project, and will be entered into an electronic database. Since this data will be from human subjects, approval for human subjects research will be obtained through the University of California - Berkeley Review Board. To ensure confidentiality, each subject will be assigned an arbitrary code, which will be associated with all data relating to that subject. One file that contains the correspondence between subject names and codes will be kept in an encrypted, password-controlled file accessible only to the PI and assessment researcher. The de-identified electronic data will be preserved on DVDs and external hard drives.

If requested, access to the de-identified data will be provided by contacting the principal investigator. Data will in principle be available for access and sharing as soon as is reasonably possible, normally not longer than one year after publication of the data. The data will be preserved for at least three years beyond the award period, as required by NSF guidelines.

We do not anticipate that significant intellectual property issues involved with this data will arise. However, in the event that discoveries or inventions are made in direct connection with these data, access to the data will be granted upon request once appropriate invention disclosures and/or provisional patent filings are made. The data acquired and preserved in the context of this proposal will be further governed by the University of California - Berkeley's policies pertaining to intellectual property, record retention, and data management.

Educational Materials

All of the educational materials produced by and for this project will be made freely available under the Creative Commons – Attribution (CC-BY) license. Authorship and contribution will be tracked to ensure that contributors are credited for their work. All materials will be archived in publicly accessible version control repositories, and made available for download from the Software Carpentry web site and university web sites in widely-used formats, including (but not limited to) HTML pages, PDF documents, and slide decks compatible with common presentation tools. Finally, we shall provide metadata such as prerequisite descriptions, lesson plans, and exercises in standard e-learning formats such as SCORM.