**Data Management and Sharing Plan**

Transparency and accountability are recognized as important means to ideological and communal bridge building that aims to strengthen social cohesion and political confidence among the public. The sharing of source materials and procedural methodology is important for the continued development of sound performance measurement and program evaluation research among organizers, as well as support of open scientific inquiry by permitting the reproduction of executive education and advanced training programs, making it possible for the testing of existing approaches, as well as the advancement of new approaches. Such information is also pertinent for policymakers, as they often rely upon assessors for their expertise in informing and improving programmatic services and activities to establish a meaningful and systematic link between the social value and financial cost of a public program. Such efforts also facilitate the education and training of assessors and new developers/administrators and enable the exploration of projects and programs not yet envisioned by initial researchers.

The proposed work will produce a number of different data artifacts. The dissemination of these artifacts, to the maximum extent possible, is essential to the goal of broad impact. The research team have carefully formulated distinct data management policies according to artifact types and sensitivities. It is anticipated that the proposal will result in the following three categories of data:

**Data Stewards**

The PI and Co-PIs will be designated as the stewards for all data resulting from that task, assuming responsibility for the management of the data and determining an appropriate classification (public, conditionally sharable, or private) for particular data artifacts.

**Data-Handling Policies**

The research team will make data artifacts available to other researchers as well as the general public to the greatest possible extent, as consistent with privacy considerations. This approach will be to identify which datasets are to be made as public (P), conditionally releasable (S), confidential (C), or educational (E). Specification of the associated policies for each below.

Policy (P), (S), (C), or (E) will be applied to category (1) data as deemed appropriate by the associated data steward. Policy (P) will be applied to category (2) data and policy (E) will be applied to category (3) data by default.

The data-handling policies are as follows:

* Policy (P): Public data. Public datasets will be those suitable for posting online, e.g., data derived from public sources, or the outputs of experiments (e.g., data, source code) that themselves do not involve any privacy-sensitive data. The policy will be to retain these data for five years from the date of publication of any paper relying on the data. The data will also be retained for a longer period of time if possible, giving explicit priority to the goal of ensuring long-term scientific reproducibility. Public data will be made available via a project website or a public cloud. Larger data sets that cannot be disseminated by either such means will be stored locally, and instructions will be published for interested researchers and others to obtain access to the data. Releasing all source code resulting from the proposal as open-source software under suitable nonrestrictive licenses will be adhered to, as well as making use of repositories that support this practice.
* Policy (S): Conditionally releasable data. Some data artifacts produced by the work will carry either temporary sharing limitations (e.g., individually requested moratorium on the release of personal data) or permanent ones. If so, such data will be retained for the same duration of time as specified in policy (P). These data will not be made public but stored locally with appropriate access-control mechanisms to restrict both external and internal access or in a cloud with protections that are suitable to the sensitivity of the data. Should researchers or others submit appropriate requests for data access, the request will be confirmed as appropriate and a practicable minimal-release strategy will be determined, specifically exploring time-limited and sanitized data-sharing approaches, as well as whether data should be released directly or through a query interface. Data will be released as expeditiously as possible, consistent with resource and policy constraints.
* Policy (C): Confidential data. A data steward may deem some data temporarily or permanently unsuitable for release outside the institution. In the proposed work, for example, public data sets that are deemed sensitive (e.g., data published on the Internet that was obtained via a security breach may be obtaoned). At the time of data collection, the steward will determine whether it is appropriate to erase the data. (For example, overly sensitive data not employed in research may be summarily deleted.) Otherwise, the data will be preserved according to Policy (S), but with no access granted outside the institution of the data steward.
* Policy (E): Educational data. The data produced in curriculum development in the context of this project will be handled under Policy (P). These data will be made publicly accessible on the website of the data steward or in an appropriately locatable and accessible public archive.

**Data Storage and Lifetime**

The volume of data produced in this proposal will be small enough to permit handling within the existing data storage facilities of respective universities. At a minimum, data will be stored for the duration of the project. If it is anticipated that the storing of most data for a considerably more extended period of time, the research team will store for as long as is practical both data required to reproduce published experiments and data of public value. All data will also be stored in suitable standard formats and will confirm that university facilities include access controls and encryption as suitable for the handling of specific data artifacts.

**Vulnerability Disclosures**

This research project will involve evaluating potential products or systems in a way that surfaces security vulnerabilities. The research team will adhere broadly to community-standard responsible disclosure practices. Specifically, the research team will adhere to the following protocol in disclosing a vulnerability:

* The research team will identify primary stakeholders, entities developing or managing the affected systems or data, as well as secondary stakeholders, those potentially harmed by the vulnerability, e.g., users of the impacted system or subjects of the relevant data. The research team will work as advocates for secondary stakeholders throughout the disclosure process.
* The research team will then notify primary stakeholders of the vulnerability and provide tangible evidence so that they can confirm and assess its scope. The research team will seek to make this disclosure as expeditiously as possible and advise primary stakeholders on technical remediation strategies, as appropriate.
* By default, the research team will make a public disclosure that specifically identifies the vulnerability, modifying this approach if it may bring about harm to secondary stakeholders and working with primary stakeholders to determine the appropriate level of detail to disclose about the vulnerability. Upon discovery of the vulnerability, the investigators will set a target date for public disclosure, which will be XXX days from private disclosure of the vulnerability.
* The evaluators will circulate drafts of the public disclosure to primary stakeholders, soliciting their feedback and working with them to ensure that details are correct and amending the disclosure as appropriate, taking into account any harm that may affect primary and secondary stakeholders as a result of disclosure.
* The investigators will publish the disclosure, including both technical detail and explanations accessible to secondary stakeholders, as warranted by the vulnerability.