SAMPLE Data Management framework for a software development project

Purpose

* Provide continuity of effort through the period of project performance.
* Provide for reliable archiving of materials (versions, configuration, datasets) used in the production of publications and reports.
* Provide for reliable archiving of materials as records of their state at milestones such as project phases, pivotal tests, and personnel changes.
* Provide for access to materials by personnel responsible for project oversight.

Applicability

* [Use this section to make it clear to all personnel when, and to what work, the practices apply.]

Practices

* Identifying working materials
  + Under direction of the PI[[1]](#footnote-0), personnel assigned to the project shall determine materials that are necessary for continuity of the project, including source code, information necessary to build from source code, training and test datasets, etc.
* Safekeeping of working materials
  + All materials necessary for project continuity shall be kept in repositories residing in network-accessible storage with backup.
  + Means of version and configuration control, and definition of checkpoints at which repositories are to be updated, shall be established by the project PI in consultation with the project team.
  + Repositories shall be accessible at all times to the project PI.
  + Materials may be incorporated by reference. This is strongly encouraged for readily accessible software libraries, industry standard data sets, and other materials that are readily identified and readily retrieved or reproduced.
* Archiving at project milestones
  + Materials shall be archived on an event basis as identified below. The Project PI may specify additional conditions for archiving.
  + Archiving may be by reference to versioning and configuration provided by a version control system.
  + Materials may be incorporated by reference. This is strongly encouraged for readily accessible software libraries, industry standard data sets, and other materials that are readily identified and readily retrieved or reproduced. In the interest of reproducibility, incorporation by reference should take versioning of outside resources into account.

|  |  |
| --- | --- |
| **Event** | **Contents** |
| Publication or conference presentation | All materials necessary to recreate results, per journal/conference standards and professional norms |
| Formal test or demonstration | Materials necessary to reproduce the project team’s contributions to the test or demonstration. |
| In advance of substantial change in project personnel contributing to production or revision of source code | Current versions of working materials |
| Completion of a project phase, coincident with mid and final reports | Current versions of working materials |

Implementation

* Roles and responsibilities
  + The investigator responsible for the project (PI, co-PI, project supervisor, as applicable) is responsible for ensuring that practices are followed by project personnel. In a multi-investigator project, substantive departures from practices shall be reported to the Lehigh PI.
* Tools and project-specific implementation
  + Tools such as archiving and versioning controls may be selected per the nature of each project and the practices of the laboratory.
  + For each project, the PI shall create a record of how required practices are implemented and will ensure project personnel adherence to those practices.

References

* DOE policy on data management [https://www.energy.gov/datamanagement/doe-policy-digital-research-data-management](https://www.energy.gov/datamanagement/doe-policy-digital-research-data-management#Principles)

1. See Roles and Responsibility for details for multi-investigator projects [↑](#footnote-ref-0)