Data Management Plan

# Types of data

This research will produce a set of software for scalable software tuning, a suite of

performance data, and some documents and publications.

# Standards for data and metadata format

All data will be stored in different formats, appropriate for the type of data being stored. For example, appropriate data formats for papers include Word files, latex files, and PDF files. As another example, appropriate data formats for repository data include Attribute Relation File Format (.arff) or Comma-Separated Values (.csv). The research templates developed as part of this proposal will dictate what the best practice is for data formats, but those who wish to share data will ultimately decide what format to store their data in.

# Policies for access and sharing

The data and software will be shared with public through repository. Repository data will be freely available to the world, and licensed under a Creative Commons Attribution 4.0 International license (<https://creativecommons.org/licenses/by/4.0/>). This data will be made freely available in our repository, accessible through the internet.

# Policies and provisions for re-use, re-distribution, and derivatives

All published and open source data may be re-used, re-distributed, and derived as long as the original data is not misrepresented and the materials do not violate the acceptable use of the IP holders. Others may change or alter the open source data for personal purposes provided that these changes are made clear in any publications, re-distributions, or derivations.

# Plans for archiving data and other research products

Repository data will be hosted on the institutional hardware; papers will be published at major conferences and journals in SE, compiler, and data analytics.