1-26-2021

# Collaborative Research: Indian Education in Computing: a Montana Story

Kristin Searle
*Utah State University*, kristin.searle@usu.edu

# Data Management Plan

## 1 Public Access

To disseminate the research output to the broad scientific community, we will maintain a public website that will provide links to our algorithms, software, data, and publications. Through this website, the general public can easily access our latest research results and educational materials developed from the proposed project.

**Course Materials.** The website will hold material for (or links to) courses related to the proposed work, including data science and STEM Education courses. Assignments developed integrating the research into the courses will be specifically highlighted on the website.

**Data.** Data supporting the conclusions from published articles will be made available under an Attribution 4.0 Creative Commons License (`https://creativecommons.org/licenses/by/4.0/`) in Zenodo (`https://zenodo.org`), a data repository hosted by CERNa memory institution for High Energy Physics, known for its pioneering work in Open Access. Zenodo is funded by CERN, the European Commission via OpenAire projects, and the Alfred P. Sloan Foundation. All open content in Zenodo is available for the public to view and download, free of charge, via an online interface. Data depositors may opt to embargo their content for a limited period of time before it is made public, or they may opt to restrict their data, approving access only to requesters who meet certain requirements. Zenodo ascribes to the FAIR Guiding Principles for scientific data management and stewardship, aiming to ensure that data is findable, accessible, interoperable, and reusable (`http://about.zenodo.org/principles/`).

Zenodo is fully run on open source products, with technical, security, and preservation features that follow community best practices. Data files and metadata are backed up nightly and replicated into multiple copies. All data files are stored with a MD5 checksum, and files are regularly checked against their checksums to ensure that file content remains constant. Data retention is guaranteed for the lifetime of the repository. In case of closure of the repository, best efforts will be made to integrate all content into suitable alternative institutional and/or subject based repositories (see `http://about.zenodo.org/infrastructure/` and `http://about.zenodo.org/policies/`).

**Source Code.** Any source code developed under this project will eventually be released under a BSD-style license, and a link will be provided on the project website mentioned above.

**Survey Data.** Qualtrics online survey system is a web-based service that allows users to create a survey, collect and store data securely, analyze responses, and present results using charts and graphs. MSU has a site license for Qualtrics, and the service has been vetted and authorized by University IT and Office of Legal Counsel.

**Manuscripts.** Papers and manuscripts documenting results will be submitted to conferences and journals for dissemination to the community.

## 2 Internal Data Management among the Team

**Git** Developmental work—including collaborations on papers—will be maintained in private git repositories hosted on GitHub and accessible to members of the research team.

**Box: Storage, Backup, and Security.** Box cloud storage system enables collaboration between teams within MSU and across institutions by allowing users to edit, download, and share files. Box provides password-protected links, file-locking, and access statistics. Box functions as a cloud backup system by syncing content between local machines and the cloud. Box employs multiple data centers with strict security policies, and provides redundancy, uninterruptible power and backup systems, and fire/flood detection and prevention. Box's security and encryption are suitable for most personally-identifiable information, with the exception of data which must meet HIPAA data security requirements. Any questions related to the storage of personally identifiable information and applicable security standards may be directed to MSUs Legal Counsel.

# 3 IRB

For the evaluation of the grant goals, Mark Quinn (MSU IRB chair) has provided a Notification of 45 CFR 690.118; see the Supplementary Documents. The grant goals have two types of human subjects: students in grades 4 through 8 and teachers:

- *Students:* For RQ3, both survey data (attitudinal surveys and formative assessments) and observational data will be collected.
- *Teachers:* For RQ2, we will collect data through content knowledge surveys, classroom observations, and recorded interviews.

Survey data will either be on paper, or collected via the online tool Qualtrics. Data not collected in surveys will be stored in a Box folder, with access limited to those who are directly working on the data analysis.

All personnel involved with developing and administering such surveys will complete CITI certification in the *Human Subjects Research* and *Responsible Conduct of Research (RCR)* modules.