CSSI Elements: Exosphere, A Researcher-Friendly Interface to the Cloud

A Data Management Plan created using dmptool

Creators: David LeBauer, First Name Surname, Chris Martin, First Name Surname

Affiliation: Non Partner Institution

Funder: National Science Foundation (NSF)

Template: NSF-CISE: Computer and Information Science and Engineering

ORCID iD: 0000-0001-7228-053X

Last modified: 01-22-2021

Copyright information:
The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customize it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal
CSSI Elements: Exosphere, A Researcher-Friendly Interface to the Cloud

Types of data

This project will neither collect samples nor generate scientific data. Data produced by the project will be related to the development, use, and distribution of software. The project will generate source code (text files), compiled software (binary files), and documentation (structured text markup files and compiled HTML). These may be accompanied by media files (images, diagrams, and videos). Anonymous user interviews and surveys will also be conducted, and this will generate response data.

Data and metadata standards

Documentation: user-facing and internal documentation will be written in a structured markup format such as Markdown.

File formats: all text files, using git for version control; GitLab for code hosting and continuous integration; Zenodo for archiving of two or more versioned releases per year that will including the addition of features proposed in this study.

Source code is written in a mix of programming languages, primarily Elm and JavaScript. Some code that is deployed to cloud instances will use Bash, Python, or Ansible (YAML). In any case, the project will use industry-standard formatting tools (e.g. elm-format, ESLint, Pylint) to ensure that source code complies with language-specific conventions.

We will provide metadata related to software and its use in standard formats designed for both humans and machines. Software metadata will be stored in the root directory of the repository - in JSON files following the Codemeta and Zenodo schemas to facilitate discovery and reuse. Furthermore, we will make relevant information available to contributors by maintaining a README.md, LICENSE, CONTRIBUTING.md, and related files in the root directory of the repository. A Dockerfile will describe dependencies and enable easy deployment. Continuous integration metrics will inform users of testing coverage and failures.

User survey response data will be stored in a standard format; this project's user survey may be integrated with the Jetstream Cloud user survey, in which case the data format will be determined by Jetstream personnel.

Policies for access, sharing, and privacy

All products of this work, and as much design and discussion activity as possible will be conducted in public forums including GitLab. All data and code generated within the scope of this project will be available in public repositories and databases. We will develop, test, and release code in GitLab. Gitlab repository contents and binaries will be automatically archived and assigned a DOI by Zenodo with each tagged release. Publications are budgeted to be open-access, and we will deposit publications and slides on pre-print servers and other public venues.

Anonymous user survey data will be published or shared only in aggregate form, and only after approval from the applicable institutional review boards (IRBs) to ensure the privacy of participants.

Policies for re-use, re-distribution, derivatives

We are committed to enhancing the value of research and furthering the advancement of public knowledge. The Exosphere software products are released under the permissive BSD 3-Clause License. All other resources developed during the course of the project will be made available to the scientific community with licenses that support further reuse, modification, and redistribution such as the BSD-3-Clause, CC-By, CC-0 Public Domain, or similar.

Plans for archiving and preservation
Code will be maintained on GitLab, with a mirror on GitHub, archives on Zenodo, and copies on no fewer than four personal computers. All code changes will be documented and stored using Git version control. Each release will be tagged in the git repository and archived on Zenodo.

Question not answered.

Roles and responsibilities

PI LeBauer will be responsible for ensuring adherence to this plan. LeBauer and Exosphere developers Julian Pistorius and Chris Martin have extensive experience with source code maintenance and software distribution and have been using the approach described in this plan for many years, including with Exosphere. If the PI leaves the project or institution, responsibilities will be transferred to collaborator Blake Joyce, PhD, Assistant Director of Research Computing at UA HPC center. All contributions to publications, software code, design, or feedback will be recognized appropriately through co-authorship and acknowledgements.