Using the SEAS Code Repository For Research

Overview

This page discusses how to apply the SEAS Code Repository at http://code.seas.harvard.edu to research projects. It assumes that you are familiar with the service, so please review the docs:

- Getting Started with code.seas
- Advanced Features of code.seas

While intuitively obvious, it’s important to remember that the most precious data of the research process is the files, datasets, and information that results from pure intellectual work. This is opposed to data that is derived from this work (such as outputs of data analysis, simulations, etc.) which could in time be reproduced. It’s this sort of data, such as:

- source code and related files
- paper drafts
- working notes
- documentation
- figures
- scripts
- unique data input sets

that the SEAS Code Repository aims to host. The underlying goal is to make it easy for the researcher to capture, safeguard, track, and archive this type of data centrally, and to remove that difficulty from the research process.

Approaches

Tracking and Managing Individual Work

At it’s simplest, the SEAS Code Repository can be used by individuals to manage their development work, including not only code, but papers, datasets (text-based) and similar intellectual products. In this capacity, code.seas.harvard.edu provides the following benefits:

- Centralized storage of critical data (i.e. code, paper drafts, figures, etc.)
- Access from anywhere via the web
- Automatic backup
- Revision control
- Basic wiki

Collaboration Among Research Project Members

Core Researcher Collaborations

In addition to the benefits to individual researchers, the SEAS Code Repository can make it easier for researchers to collaborate on a project. By creating Teams, project management can be shared among multiple researchers, and development work (or paper writing work) distributed without causing major headaches, and with built in tools for resolving conflicts (in the data).

External Researcher Contributions

In addition, external collaborations with researchers who aren’t core project members is made possible by using merge requests. Interested, non-core collaborators can clone a codebase, make changes, and then submit those changes for review by core project members. This allows external researchers to contribute to projects – even making significant changes to the codebase – without touching the main codebase directly.

Publishing

Lastly, a researcher or collaboration can easily publish the fruits of their work using the SEAS Code Repository by exposing certain repositories to the public. With a single click an internal project can be made available readonly to the rest of the world using a number of protocols, or as a downloadable “tarball” of files.