

# Pycee 2.0: Enhanced Python Compiler Error Messages via Stack Overflow

## Sponsor

Christoph Treude  
School of Computer Science  
University of Adelaide  
Australia

## Project Abstract

Compilers tend to produce cryptic and uninformative error messages, leaving programmers confused and requiring them to spend precious time to resolve the underlying error. The goal of this project is to build the next version of Pycee, a plugin integrated with the popular Sublime Text IDE to provide enhanced compiler error messages for the Python programming language. Pycee automatically queries Stack Overflow to provide customised and summarised information within the IDE. See <https://cs.adelaide.edu.au/~christoph/esem19.pdf>.

## Project Description

To find help when confronted with compiler errors, programmers often take to online question-and-answer forums such as Stack Overflow to start discussion threads about the errors they encountered. We conjecture that information from Stack Overflow threads which discuss compiler errors can be automatically collected and repackaged to provide programmers with enhanced compiler error messages, thus saving programmers' time and energy. Based on a first prototype called Pycee (<https://github.com/emillieT/pycee>), the goal of this project is to (a) make Pycee available as a web application, and (b) improve Pycee.

(a) The first step of the project is to identify the best environment for deploying Pycee online. The main requirement is that Pycee should be really easy to access and demonstrate, without having to install anything (the current version requires an installation of SublimeText along with the Pycee plugin). For this purpose, it might make sense to develop a browser plugin which works with an existing online IDE such as Cloud9, or to fork an existing open-source Python web editor and deploy it on a server. In either case, the current code of Pycee needs to be adapted to the new environment.

(b) There are many ways in which Pycee can be improved, listed here in rough order of priority: (1) add systematic unit testing to the tool, (2) improve the accuracy of error messages by considering the offending code, (3) make the tool interactive, e.g., by allowing users to "scroll through" different interpretations of an error message and by allowing users to give feedback to an error message, (4) use natural language processing to transform the conversational style of Stack Overflow into text

more suitable for an error message, and (5) automatically fix the offending code if the tool is confident that it knows how to fix the error.

## Project Scope

The minimum requirement of the project is to make the current version of Pycee (the SublimeText plugin, <https://github.com/emillieT/pycee>) available in the form of a web application. Depending on the decisions taken in the first step of the project (see Project Description), this might involve forking an existing web-based Python editor, integrating Pycee into it, and deploying it on the web. Once this is achieved, there are a number of features that should be added, listed under Point (b) in the Project Description.

## Process Requirements

All development will happen in a public GitHub repository, with work items prioritised using GitHub issues. All code will be peer-reviewed by at least one other member of the team using the pull-based development model. Unit testing is required for the core functionality of the tool. Any agile-like process is fine, as long as it is properly explained and motivated.

## Environmental Constraints

Since Pycee 2.0 will be able to provide enhanced compiler error messages for the Python programming language, it would make sense for Pycee 2.0 to be written in Python (although this is not a hard constraint). As mentioned before, a key requirement is for Pycee 2.0 to be really easy to access and demonstrate, ideally by simply accessing a URL. Determining the best environment for this is the first step of the project.

## Project Restrictions

No specific restrictions. The first step of the project will be to determine the best possible environment for Pycee 2.0.

## Project License

GPL

## Level of Sponsor Involvement

Sponsor contact: [christoph.treude@adelaide.edu.au](mailto:christoph.treude@adelaide.edu.au)

I will be available on email throughout the project and will further set up monthly meetings (more if necessary). Most of the communication will likely happen through issues on a public GitHub repository.