UNIVERSITY OF CALIFORNIA AT BERKELEY

BERKELEY • DAVIS • IRVINE • LOS ANGELES • MERCED • RIVERSIDE • SAN DIEGO   SAN FRANCISCO • SANTA BARBARA • SANTA CRUZ

CONSENT TO PARTICIPATE IN RESEARCH

Understanding programmer behavior by collecting program edit histories

Introduction

My name is Justin Lubin. I am a graduate student at the University of California, Berkeley in the Department of Computer Science working with my faculty advisor, Professor Sarah E. Chasins. I am inviting you to participate in a research study that I am running because you indicated on a survey that I sent out that you write code in a statically-typed functional programming language.

Purpose

The purpose of this research study is to better understand how statically-typed functional programmers write code by analyzing how their code evolves over time.

Procedures

If you agree to take part in this study, you will be asked to do the following until you decide to discontinue participation in the study (with a maximum participation duration of two years):

1. Perform a one-time installation of an editor plugin that will **log every edit and compilation that you make to a programming project of your choice**. The installation process will prompt you for some basic demographic information, such as the number of years of experience you have.
2. In the unlikely event that any identifiable or sensitive private information (such as an email or password) makes it into your source code, redact it using a provided automated tool.
3. Respond to a prompt from the plugin (at most once per day), which will ask you to a) re-affirm your consent to this study, b) confirm that you have redacted any identifiable or sensitive private information, and c) provide consent to upload your recent edit data to a secure server.
4. Write code as usual for your programming project.

The logging plugin will only activate for programming projects that you explicitly enable it for.

Installation should take no longer than 30 minutes. Using the automated redaction tool should take about 1 minute each time. Responding to the plugin's prompts should take about 5–10 seconds each time.

Data Collected

Throughout the course of the study, we will collect the following four types of information from you:

1. Your initial screening survey response (kept private)
2. Your signed consent form (kept private)
3. Your logged programming edit history (to be publicly published)
4. Basic demographic information such as years of experience (to be publicly published)

Although we will not publicly link any of this information with you, we will **publicly publish your logged programming edit history and basic demographic information** at the conclusion of this study, so it is important that you are **willing to have the code you are working on be made open source** and to use the automated redaction tool to redact any identifiable or sensitive private information.

Risks

The biggest risk you face in participating in this study is a **breach of confidentiality**. For example, if you mistakenly type a password into your code then immediately delete it after realizing you should not have typed it, the logging software will store both the entering and the deletion of the password.

For this reason, we will provide you with a tool to automatically redact information of your choice. The logged data will leave your computer only after you have confirmed that you have used the automated redaction tool if necessary. **Any redacted information will *not* be transmitted to us or stored in any way**; once you redact the information locally, it is completely overwritten.

If at any point you wish to retroactively apply a redaction to uploaded data, please email Justin Lubin at justinlubin@berkeley.edu, **but do not include the information to be redacted in the email**. We will ensure that we redact the relevant information in a secure way that respects your confidentiality. We will give you at least one month of notice before publicly publishing any of your data.

Benefits

There will be no direct benefit to you from participating in this study. However, it is hoped that the information gained from this study will help authors of tools for statically-typed functional programming languages align their tools to programmers' needs and existing behaviors.

Confidentiality

Your study data will be handled as confidentially as possible. When results of this study are published or presented, individual names and other personally identifiable information will not be used. Identifying and sensitive information will be removed from the data by your use of the automated redaction tool, but **your (redacted) code history will be publicly accessible, which the public may be able to associate with other places you have posted your code online, such as your GitHub profile**. We may use the information that we collect from you for future research studies or distribute it to other investigators for future research studies without additional informed consent from you.

Compensation

You will not be paid for taking part in this study.

Rights

Participation in research is completely voluntary. You may discontinue participation in the study at any time and/or request to view or delete all or some of the data collected from your participation by emailing Justin Lubin at justinlubin@berkeley.edu. Alternatively, to discontinue participation in the study, you may simply uninstall the logging plugin from your editor.

Questions

If you have any questions or concerns about this study, please contact Justin Lubin via email at justinlubin@berkeley.edu. If you have any questions or concerns about your rights and treatment as a research subject, you may contact the office of UC Berkeley's Committee for the Protection of Human Subjects at 510-642-7461 or subjects@berkeley.edu.