Secondary Use of Data

Project Info.

File No: 17625
Project Title: Evaluation of Student-Written Test Suites
Principal Investigator: Dr. Michael Miljanovic (Faculty of Science (2700))
Start Date: 2024/01/01
End Date: 2024/04/30
Keywords: NSE- Information Technology

Project Team Info.

Principal Investigator

Prefix: Dr.
Last Name: Miljanovic
First Name: Michael
Affiliation: Faculty of Science (2700)
Position: Assistant Professor
Email: michael.miljanovic@ontariotechu.ca
Phone1:
Phone2:
Fax:
Primary Address:
Institution: #University of Ontario Institute of Technology (UOIT)
Country: #Canada
Comments:

Other Project Team Members

Prefix	Last Name	First Name	Affiliation	Role In Project	Email
Dr.	Bradbury	Jeremy	Faculty of Science (2700)	Collaborator	jeremy.bradbury@ontariotechu.ca

Ms.	Showler	Amanda	Faculty of Science (2700)	Student Lead/Post-Doctoral Lead	amanda.showler@ontariotechu.net

Common Questions

1. 1: Research Team

#	Question	Answer
1.1	As the PI, have you completed the TCPSII training tutorial online or an approved equivalent? This REB application will not be approved without it.	I have completed an approved equivalent. Details are provided in my response to question 1.2 and I have attached a copy of the certificate to this application.
1.2	If you have completed an approved equivalent, provide details below. Attach the completion certificate to this application (use the attachments tab).	
1.3	As the PI, do you have the required professional expertise and qualifications for this research?	I do not have the required professional expertise and qualification, but a member of my research team does (details are in my response to question 1.4 below).
1.4	Which member(s) of your team has the necessary qualifications to conduct this research (if not the PI)? What additional steps are necessary to ensure that you or your research team will have the necessary qualifications?	My undergraduate honours thesis supervisor, Dr. Michael Miljanovic, has the necessary qualifications to conduct this research.
1.5	Applicable to all PIs, student/post-doctoral PIs, and Co-PIs listed.: As researchers, are there any interpersonal relationships (family, close friendships, colleagues, etc.), financial partnerships, other economic interests (i.e. spin-off companies in which researchers have stakes or private contract research outside of the academic realm) or any other incentives that may compromise the integrity or respect for the core principles of this policy?	No.

1.6	If you answered yes to question 1.5, describe how you plan to minimize this conflict of interest. NOTE: While it may not be possible to eliminate all conflicts of interest, researchers are expected to identify, minimize, or otherwise manage their individual conflicts in a manner that is satisfactory to the REB (TCPSII 7.1).	
1.7	Are all members of the research team listed in the 'Project Team Info' tab?	Yes, I have added all project members in the 'Project Team Info' tab with the correct role in project.

2. 2: Purpose & Background

#	Question	Answer
2.1	What type of research is your project?	Undergraduate Capstone Research Project

2.2	In 500-600 words, situate the proposed research in the scholarly literature (including some references) and provide a rationale for the study to justify the study's purpose. Describe the project, the overarching research issues/problem and specific research question(s), and, if applicable, hypothesis. Describe the anticipated contribution of the research. Assume the REB reviewers are knowledgeable about research but not familiar with the discipline. IMPORTANT: SAVE THIS APPLICATION OFTEN.	Software testing is an essential skill in the software industry and yet, it is reported that CS graduates' testing ability falls short of industry expectations [1]. This research project aims to evaluate student-written test suites to gain insights into their overall quality and completeness. Through this evaluation, we will identify common strengths and areas of improvement in categories such as test suite completeness and accuracy. The contributions of this research will provide instructors with insights into the testing ability of students throughout a software quality assurance course. This will be valuable for enhancing software testing courses at the undergraduate level and future research into bridging the gap between student learning objectives, test development ability, and industry expectations.In the research project, we propose to evaluate student-written test suites by using GitHub repositories contributed to by students for a course project. They are written by senior undergraduate students from two cohorts of the course CSCI 3060U Software Quality Assurance from Winter 2023 and Winter 2024. The GitHub repository is a database of changes the students made to their project (called a version control history), a student-written test suite in .txt files and the source code. The instructional team, including the TAs, are given access to the private repositories for the course project evaluation so the data we propose to collect is already available to them. The course project has 6 phases where students work in groups of up to 4. The students are given a specific list of requirements to implement for the project, so although their implementations will differ, the software features/behaviour will be the same. The two cohorts' project

		instructions will differ on software theme (bidding system vs. Steam game store system) but they will have identical project structure and milestones. This course project was designed specifically for this course by Dr. Bradbury. No changes will be made to the course project for the research project.We are requesting the student submissions since there is a lack of available datasets of student-written test suites where the students did not have access to the source code prior [2]. In the project phase 1, the students write test cases from the instructor's requirements before they start writing their software implementation. In addition, this course project varies from other projects since it has a unique component where in phase 4 the students exchange part of their code base (frontend code written in C++). This means that they will understand the intent of the software from the project requirements specified by the instructor that they already implemented. However, they will be unfamiliar with the software implementation from the other group. For the remainder of the project, the students are asked to integrate this new software implementation with their test suite and the remaining part of their code base (backend code written in Java or Python). This is unique to an undergraduate course, yet it is common in industry that a programmer will be given a new code base where they understand the intent and are expected to learn the implementation. We want to study this dataset because it is a unique opportunity to study student behaviors when developing the test suite and when they are editing the test suite for integration with a new code base after the source code swap.The following are example research questions:- What are the most

		common successes in student-written test suites? (Identify the strengths)- What are the most common errors in student-written test suites? (Identify the areas of improvement)- How do students respond to feedback on their test suites?References:[1] Gina R. Bai, Justin Smith, and Kathryn T. Stolee. How students unit test: Perceptions, practices, and pitfalls. In Proceedings of the 26th ACM Conference on Innovation and Technology in Computer Science Education V. 1, ITiCSE '21, page 248–254, New York, NY, USA, 2021. Association for Computing Machinery[2] Jeffrey C. Carver and Nicholas A. Kraft. Evaluating the testing ability of senior-level computer science students. In 2011 24th IEEE-CS Conference on Software Engineering Education and Training (CSEET), pages 169–178, 2011

3. 3: Funding Information

#	Question	Answer
3.1	Have you or will you apply for any funding related to this research?	No, this project is unfunded.
3.2	Will you be using any other source of funding for this project (start-up funds, professional development funds, personal funds, etc.)?	No.
3.3	If you answered yes to question 3.2, describe the sources of internal or personal funding that will be used for this research.	

4. 4: Partnerships and agreements screening questions

#	Question	Answer
4.1	Does your research proposal involve the exchange of any proprietary/potentially commercial items (tangible or intangible), personal information, confidential information, materials, human resources or funds to/from Ontario Tech?	No
4.2	If YES, name the institution.	

4.3	Are you signing an external agreement with an institution governing the use of data?	No
4.4	If YES, who are you signing an external agreement with?	

5. 5: Original Research & External Approvals

#	Question	Answer
5.1	Describe in lay language the research questions and hypothesis of the original research.	What are the most common successes in student-written test suites?What are the most common errors in student-written test suites?How do the student-written test suites change over the duration of the project?There is no hypothesis yet as this is an exploratory study.
5.2	Evaluate and comment on the degree of expectations the individuals who originally provided the research data had regarding their data being kept confidential and unused for other purposes.	The undergraduate students who originally provided the research data would not expect the data would be used for other purposes. The students submitted the research data for a course project.
5.3	Have you attached (using the attachments tab) a copy of the consent form which was originally used for the initial collection of data?	No, - I have provided an explanation in my response to question 4.4 below.
5.4	If you answered no to question 4.3, please provide an explanation for not attaching the original consent form.	The original purpose of the data collection was for a course project and did not require consent.
5.5	Does this application require external approvals or permission?	Yes, I have provided details in my response to question 4.6 below and have also attached the approval message/permission letter to this application (using the attachments tab).
5.6	If you answered yes to question 4.5 above, from which organizations will/have you obtained permission?	We obtained permission from the Dean of Science, Greg Crawford, since the requested anonymized data is from two cohorts of CSCI 3060U (a course offered by the Faculty of Science). We attached the consent from the Dean to collect the secondary use data under Supportive Documentation (filename: Dean Approval - Secondary Data Collection for Research Study.pdf).

6. 6: Research Participants

#	Question	Answer

6.1	Approximately how many adults (18 years of age and older) contributed to the data set you intend to reuse?	180
6.2	Approximately how many minors contributed to the data set you intend to reuse?	0
6.3	Who were the participants in the data set? What were the demographic details (age, gender, etc.)?	The participants in the data set are senior undergraduate students at Ontario Tech University who were enrolled in CSCI 3060U course in the Winter 2023 and Winter 2024 cohorts.
6.4	Did the original research involve a vulnerable population?	No
6.5	Did the original research involve Indigenous Peoples in Canada (i.e. First Nations, Inuit, Métis)?	No
6.6	Do you have any exclusion criteria for the data set?	No.
6.7	If you answered yes to question 5.6, describe the exclusion criteria.	
6.8	Indicate how you will obtain free and informed consent from the research participants, if applicable. Attach a copy (using the attachments tab) of the consent form and/or information sheet that will be given to the research participants.	For the previous cohort from Winter 2023, the permission to obtain the course project data secondary use was obtained from the Dean of the Faculty of Science (Greg Crawford). For the upcoming cohort in Winter 2024, there is a consent form that will be given to research participants to opt-in to the data collection. The consent form is a Google form that the student will use to register their group for the course project. We have attached the consent form under Supportive Documentation (Consent Form.pdf).

7. 7: Risk & Benefit Assessment

#	Question	Answer
7.1	Check any possible risks to the participants. Provide details in question 6.3 below.	Not applicable
7.2	Are there any possible risks to participants greater than they encounter every day?	No.
7.3	If you checked any risk involved in question 6.1 OR answered yes to question 6.2, explain in detail the risk to the participants.	

7.4	Describe how any risks described in 9.3 will be managed. If needed, include the availability of medical expertise or clinical expertise, etc.	
7.5	Is there any deception involved?	No.
7.6	If you answered yes to question 6.5 above, describe the nature of the deception, why it must be used, and the procedures that will be used to protect the participants. Attach the second consent form required for full disclosure on deception (use the attachments tab).	
7.7	Is there potential for participants to feel coerced into participating in this research?	Yes, details are in my response to question 6.8 below.
7.8	If you answered yes to question 6.7 above, explain how you manage any potential coercion to participate and how this will be managed.	For the upcoming cohort in Winter 2024, the Principal Investigator for this research, Dr. Michael Miljanovic, is the professor of CSCI 3060U so there is a possibility that students would feel obligated to allow their course project data to be used. This has been managed by the course project data collection occurring in the lab sections. Dr. Miljanovic will not be aware of who is participating in the study since the opt-out process and data collection will be completed by the Student Lead (Amanda Showler). In addition, the labs where the course projects are collected and graded are managed by the teaching assistants.
7.9	Describe any risks that may occur to the community and/or environment and how they will be managed.	N/A
7.10	Discuss any direct benefits to the participants from their involvement in the project. Comment on the potential benefits to the community and/or environment that would justify involvement of participants in this study.	There are no direct benefits to the participants from their involvement in the project. A potential benefit to the community is the enhancement of future software quality course offerings based on findings. In addition, we intend to bring this research to a conference and pursue a publication to disseminate our findings into the CS education community.

8. 8: Data Storage and Confidentiality

#	Question	Answer
8.1	How will you ensure the anonymity of the research participants? If anonymity is not be guaranteed, explain how the research participants will be informed of that fact.	Student source code and test suite .txt files will be anonymized using the following procedure:- Source code will be reformatted using a standard style- File names will be reviewed and identifiable information will be removed- Source code comments will be reviewed and identifiable information will be removed- Test suite .txt files (e.g., input, expected output) will be reviewed and identifiable information will be removed- Student names will be anonymized
8.2	Will anyone have access to the data who is not listed as an investigator on this application (e.g. research assistant, analyst, etc.)?	No.
8.3	If you answered yes to 7.2, provide a list of people who will have access. NOTE: For all personnel listed below, they must sign a confidentiality agreement. Attach a copy of the confidentiality agreement template to this application (use the attachments tab). Signed copies of the confidentiality agreement are not required for submission.	
8.4	Describe in detail where the data will be housed securely (e.g. encrypted USB key, cloud storage (such as Google Drive), encrypted files, etc.).	The data will be housed on a secure Ontario Tech University server.
8.5	What steps will you take if there is a security breach of participant data (e.g. loss of USB key or a hack of your cloud storage solution)?	If there is a security breach of participant data, we would immediately report it to the Office of the University Secretary and General Counsel and the Office of Research Services. We would follow the procedure described in the Privacy Breach Protocol from the Information and Privacy Commissioner of Ontario.
8.6	What is the anticipated data storage end date (if the data is being kept indefinitely, please state so)?	August 31, 2024

8.7	Describe the procedures that will be used to ensure the confidentiality of data both during the conduct of the research and in the release of its findings.	No identifiable information from the students will be stored. The release of findings may include the course project requirements for reproducibility and the student data will not be released. The data will be described in the findings as student-written programs and test suites from a senior undergraduate course project.

9. 9: Other Details

#	Question	Answer
9.1	This section is for you to inform the REB of anything you feel is pertinent to the review of your application that you were unable to convey elsewhere in this application.	

10. 10: PI Certification

#	Question	Answer
10.1	I certify the information provided in this application is complete and accurate. I have complied with the TCPSII and UOIT's policies and procedures governing the protection of human participants in research.	I understand and agree.
10.2	I will report any adverse or unanticipated events (unanticipated negative consequences or results affecting participants) to the UOIT Ethics & Compliance Officer as soon as possible.	I understand and agree.
10.3	Any additions or changes in the approved research protocol will be submitted to UOIT's REB for approval prior to implementing them.	I understand and agree.
10.4	If my research remains active beyond the expiry date assigned, I will renew annually in accordance with the Tri-Council Policy Statement.	I understand and agree.
10.5	I will complete and submit a Completion Form to the Research Ethics Coordinator at UOIT once the research has completed.	I understand and agree.

10.6	I take full responsibility in ensuring that all other researchers (including the student/post-doctoral PI, if applicable) in this research follow the protocol as outlined in the application.	I understand and agree.

Attachments

Doc / Agreement	Version Date	File Name	Description
Consent Letter	2023/11/23	Consent Form.pdf	N/A
Consent Letter	2023/12/05	Consent Form v2.pdf	N/A
Recruitment Materials	2024/01/30	Flyer_Revised2.png	Flyer revised 2
Response to Clarifications	2023/11/23	17625 Clarification Letter Response - Miljanovic - 23 Nov.docx	N/A
Supporting Documentation	2023/09/21	Dean Approval - Secondary Data Collection for Research Study.pdf	N/A
TCPS2 Certificate	2013/07/03	miljanovic_michael_tcps2_core_certificate.pdf	N/A
TCPS2 Certificate	2023/09/18	showler_amanda_tcps2_core_certificate.pdf	N/A