RUHR UNIVERSITÄT BOCHUM RUB

Bachelor Thesis Topic

Extending the Reproducible Locator Break Dataset

Motivation and Background:
Web GUI testing is critical for ensuring the reliability of modern web applications, yet GUI tests often break as applications evolve, even when the underlying functionality remains intact [1, 2]. While mitigation strategies exist, they still need to be properly evaluated on a large-scale dataset of locator breaks. A recent thesis by Leon Adamietz [3] created the Reproducible Locator Break Dataset (RLBD) by analyzing over 300 open-source repositories from the E2EGit dataset [4], identifying 10,872 potential locator breaks across Playwright and Cypress frameworks. Leon focused his efforts on the top 25 projects with the most possible locator breaks, yet only four projects (and 493 locator breaks) could be configured to reproduce them, due to environmental complexity, dependency challenges, and project-specific configurations.
The thesis aims to revisit the non-working cases from this top 25 and also explore projects from the top 50. The idea is to extend reproducibility from 493 to approximately 1,000+ locator breaks (roughly doubling the dataset). This would become a truly publishable research artifact and resource suitable for large-scale studies on locator break mitigation, test repair, and GUI testing robustness [5].

Student Task and Responsibilities:
- Analyze the 21 projects from Leon's top 25 that were not executed. Diagnose root causes and create targeted Docker/Makefile solutions for each project, when possible.
- Explore the top 50 (projects 26-50). Create, when possible, Docker-based reproduction files following Leon's standardized approach.
- Incrementally extend the dataset using Leon's create-reproduction dataset script, adding reproducible breaks to the SQLite database and testing each configuration until the target of 1,000+ breaks is reached.
- Document all reproduction blockers encountered, categorize by type (dependency conflicts, environment-specific issues, framework incompatibilities, configuration errors), and record solutions for each project.

Deliverables:
- An extended Reproducible Locator Break Dataset (RLBD v2) with approximately 1,000+ reproducible breaks from the top 50 projects, stored in Leon's SQLite schema with all required metadata.
- Categorized documentation of all reproducibility challenges encountered (environment setup, dependency conflicts, framework issues, configuration errors) and solutions implemented for each project.
- Ideal outcome: A dataset paper ready for publication describing the methodology, dataset scale/structure, empirical findings, and research applications.

Pre-Requisites: (Programming Languages, OS, Skills, Papers, etc)
- Experience with Python, JavaScript, and TypeScript.
- Familiarity with Playwright and Cypress frameworks, and a solid understanding of web technologies (DOM, CSS selectors, HTML attributes).
- Knowledge of containerization (Docker and Docker Compose) is essential to reproducing the pipeline extension.

[1] Leotta, Maurizio, et al. "Challenges of end-to-end testing with Selenium WebDriver and how to face them: A survey." 2023 IEEE Conference on Software Testing, Verification and Validation (ICST). IEEE, 2023.

[2] Hammoudi, Mouna, Gregg Rothermel, and Paolo Tonella. "Why do Record/Replay Tests of Web Applications Break?" 2016 IEEE International Conference on Software Testing, Verification and Validation (ICST). IEEE, 2016.

[3] Adamietz, Leon. "Identification and Reproduction of Locator Breaks in Evolving Web Applications." Bachelor's Thesis, Ruhr University Bochum, 2026.

[4] Di Meglio, Sergio, et al. "E2EGit: A Dataset of End-to-End Web Tests in Open Source Projects." 2025 IEEE/ACM 22nd International Conference on Mining Software Repositories (MSR). IEEE, 2025.

[5] Di Meglio, Sergio, et al. "Towards predicting fragility in end-to-end web tests." *Proceedings of the 28th International Conference on Evaluation and Assessment in Software Engineering*. 2024.

Contacts

Thiago Santos de Moura, M.Sc. (sq-teaching@rub.de)
Software Quality group, Faculty of Computer Science, Ruhr University of Bochum