

# Bachelor Thesis Topic

## Adapting and Validating the Fragility Score

### Motivation and Background:

Web GUI tests rely on element locators to interact with application interfaces. The tendency of locators to break as applications evolve is a critical maintenance challenge in test automation [1, 2]. A recent study proposes the fragility score to predict the likelihood of a locator breaking [3]. However, the approach's predictions have only been validated on Selenium using a single small-scale web application (Addressbook with 27 tests), achieving a correlation coefficient of 0.64 with observed breakages.

The fragility score has never been adapted to or validated on modern testing frameworks like Cypress and Playwright, which employ fundamentally different locator strategies and ways of interacting with web elements. Understanding how fragility patterns translate across frameworks is essential, as Cypress and Playwright now dominate industry practice [4].

Simultaneously, Leon Adamietz's recent thesis introduced the Reproducible Locator Break Dataset (RLBD) [5], which contains 493 reproducible locator breaks extracted from real open-source repositories. This dataset provides an ideal opportunity to evaluate the fragility score in large-scale and modern web applications by associating higher fragility with the locator break in question.

This thesis has two connected goals: (1) adapting the Di Meglio fragility score methodology to Cypress and Playwright, and (2) systematically validating whether fragility scores can predict real locator breaks by measuring correlation against the RLBD reproduction dataset.

### Student Task and Responsibilities:

- Implement fragility score calculation for Cypress and Playwright, adapting Di Meglio's algorithm for each framework's selector syntax.
- Extract and score all 493 RLBD breaks (both old and new locators) using the adapted algorithms.
- Correlate fragility scores with observed breaks: compute point-biserial correlations and assess statistical significance via p-values.
- Analyze by framework (Cypress vs. Playwright) and locator type (CSS, ID, Text, TestID).

### Deliverables:

- Framework-specific fragility score implementation for Cypress and Playwright using the Di Meglio methodology, including unit tests and documentation.
- Empirical evaluation results for the RLBD dataset, using point-biserial correlation analysis with p-values, and statistical comparison against Di Meglio's Selenium baseline ( $r=0.64$ ).

### 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] Nass, Michel, Emil Alégroth, and Robert Feldt. "Why many challenges with GUI test automation (will) remain." Information and Software Technology 138 (2021): 106625.

[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] Di Meglio, Sergio, and Luigi Libero Lucio Starace. "Towards Predicting Fragility in End-to-End Web Tests." Proceedings of the 28th International Conference on Evaluation and Assessment in Software Engineering. ACM, 2024.

[4] ByBy. "Top 6 End-to-End Testing Frameworks for Web Apps." Retrieved April 1, 2026, from <https://byby.dev/e2e-testing-frameworks>

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

### **Contacts**

Thiago Santos de Moura, M.Sc. ([sq-teaching@rub.de](mailto:sq-teaching@rub.de))

Software Quality group, Faculty of Computer Science, Ruhr University of Bochum