

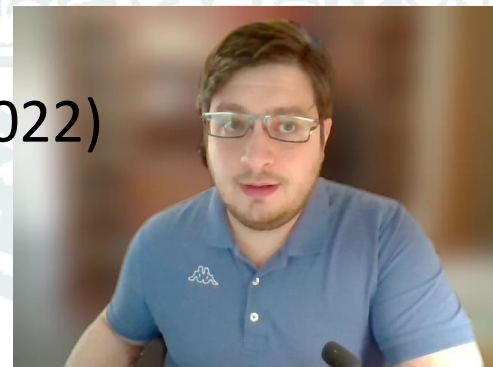
# ReCover: a Curated Dataset for Regression Testing Research

F. Altiero, A. Corazza, S. Di Martino, A. Peron, and **L. L. L. Starace**

*Università degli Studi di Napoli Federico II, Naples, Italy*

[luigiliberolucio.starace@unina.it](mailto:luigiliberolucio.starace@unina.it)

19th International Conference on Mining Software Repositories (MSR 2022)



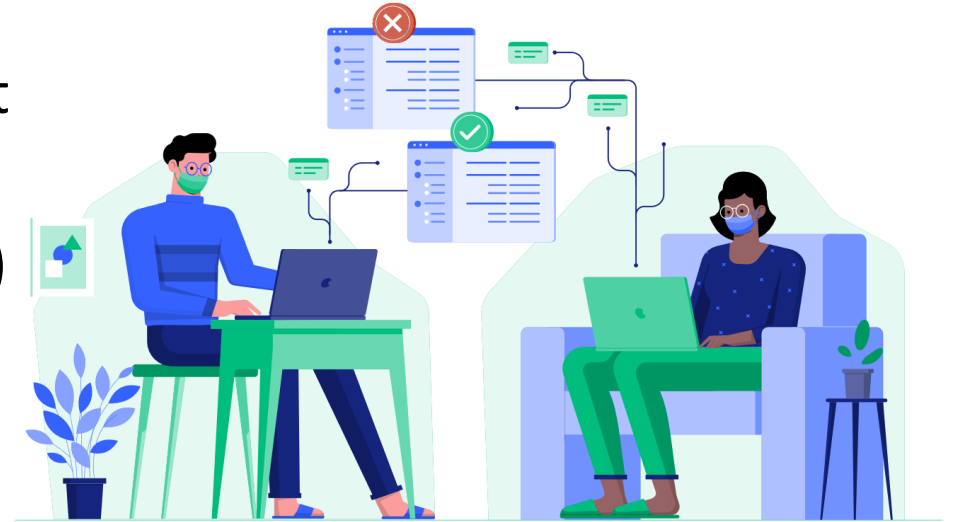
# Regression Testing

- Gain confidence that **maintenance changes** in the codebase do not introduce **faults** in previously-validated functionalities.
- The **Retest-all** approach is often unfeasible due to time/resource constraints.
- More refined strategies are needed!
- Hot and longstanding topic for the SE community, with a number of techniques being proposed



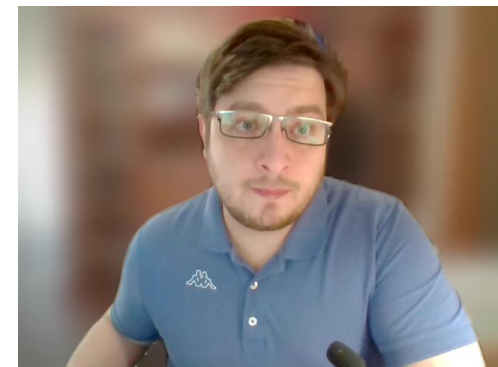
# Datasets for Regression Testing Research

- Collections of relevant software project **version pairs**.
  - The older version works fine (all tests **pass**).
  - A fault has been introduced in the more recent version (one or more tests **fail**).
- Datasets with **real faults** (e.g.: RTPTorrent<sup>1</sup>) have been proposed and used in the literature.

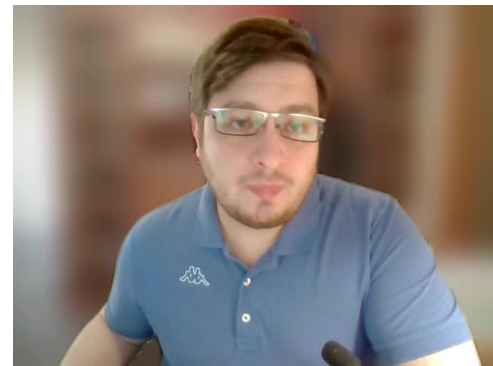


- Why yet another dataset for regression testing research?

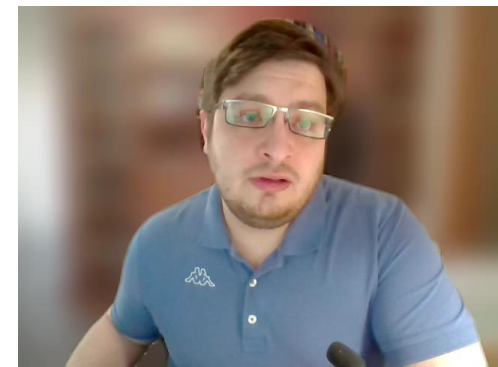
[1] Mattis, Toni, et al. "RTPTorrent: An open-source dataset for evaluating regression test prioritization." *Proceedings of the 17th International Conference on Mining Software Repositories*. 2020.



# Limitations of Existing Datasets



# Limitations of Existing Datasets

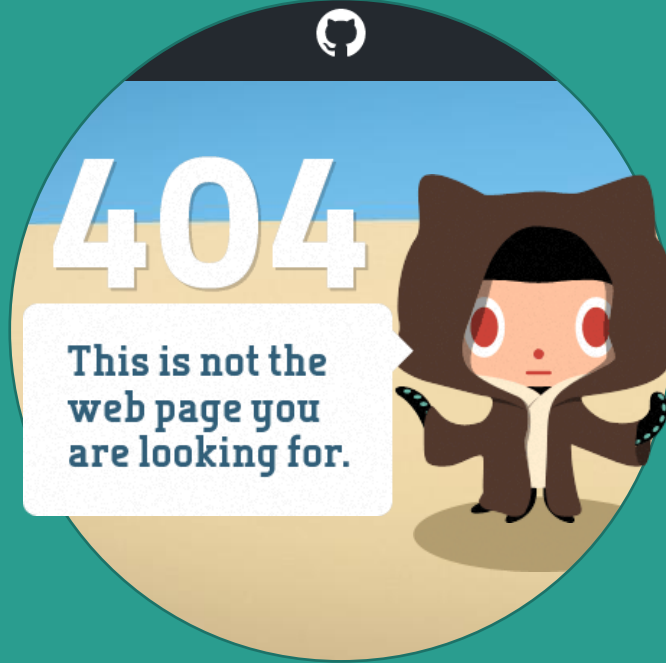


# Limitations of Existing Datasets



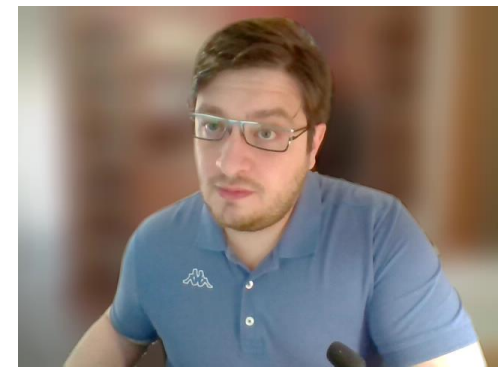
Coverage Reports not directly available

The illustration shows a person in a blue shirt sitting at a desk with a laptop, looking at a large screen displaying various data visualizations including a bar chart and a pie chart. The scene is set within a circular frame on a teal background.



References often out of date

The illustration features a large '404' error code in the center. To the right is a cartoon cat character wearing a brown hoodie. A speech bubble from the cat says, 'This is not the web page you are looking for.' The scene is set within a circular frame on a teal background.



# Limitations of Existing Datasets

An illustration of a person sitting at a desk with a laptop, looking at a large screen displaying various data charts and graphs. The scene is set in a circular frame with a teal background.

Coverage Reports not directly available

An illustration of a 404 error page. The number '404' is prominently displayed in the top left. A small GitHub logo is at the top center. A cartoon cat character in a brown hoodie stands on the right, pointing to a speech bubble that says 'This is not the web page you are looking for.' The scene is set in a circular frame with a teal background.

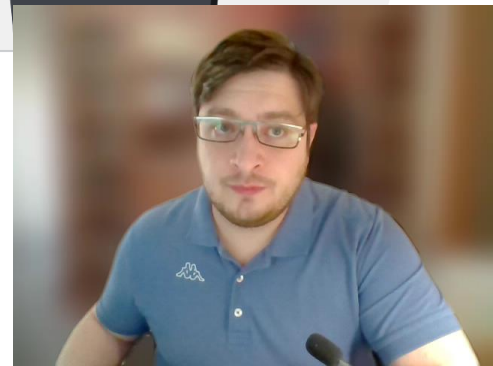
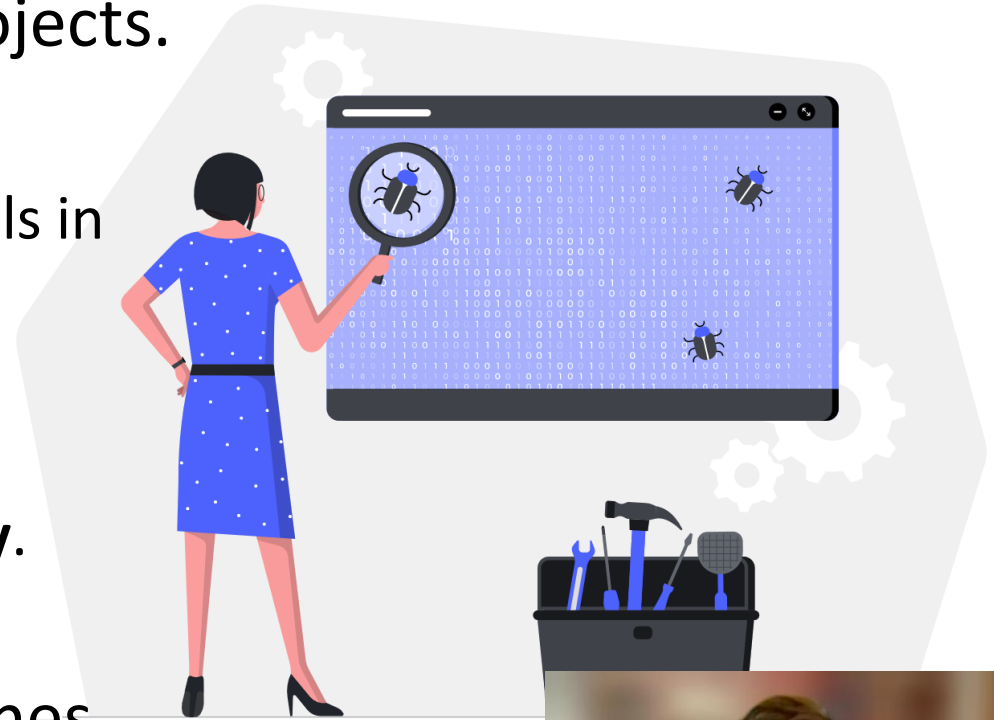
References often out of date

An illustration of a person in a blue cap and overalls looking at a computer monitor. The monitor displays a sad face with 'X' eyes and a speech bubble that says 'BUILD ERROR!'. The scene is set in a circular frame with a teal background.

Build not always straightforward

# The Proposed ReCover Dataset

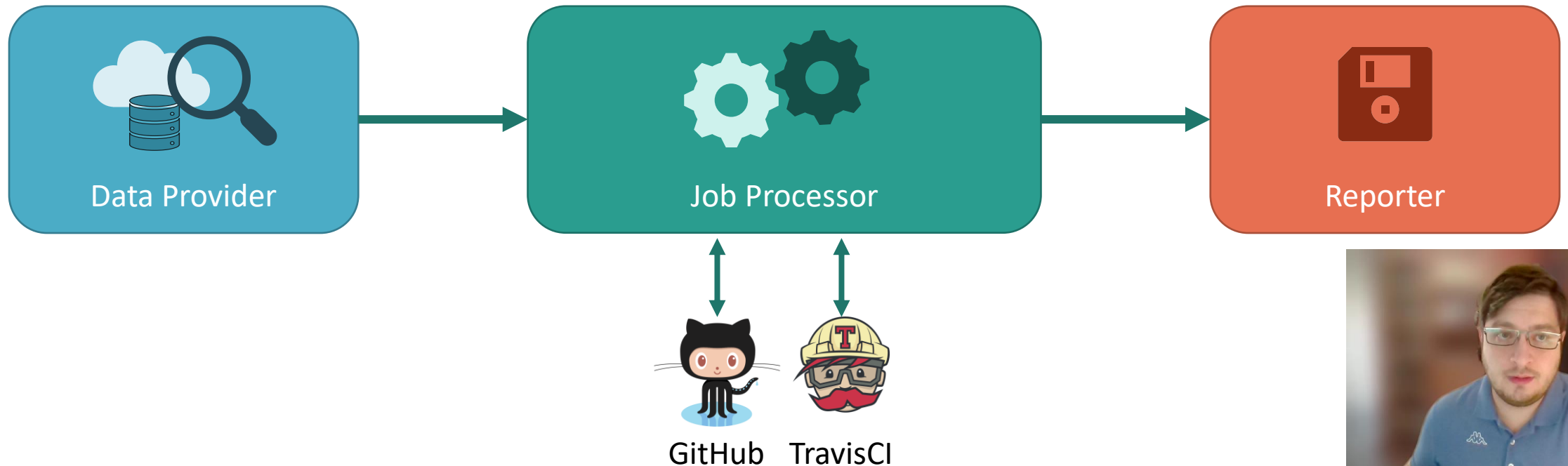
- **114** version pairs from **28** Java (Maven) projects.
- Each version pair includes:
  - At least one **real fault** (i.e.: at least one test fails in the most recent version).
  - Full, **buildable** source code of both versions.
  - Fine grained test execution/coverage reports.
  - A **Docker** container for immediate **replicability**.
- Collected projects vary in size and number of test cases, ranging from small to large ones.





# The ReCover Mining Tool

- ReCover has been obtained by automatically mining TravisCI/GitHub with a tool we designed.
- The tool can be easily **extended** to support additional data sources.



# ReCover: a Curated Dataset for Regression Testing Research



<https://doi.org/10.5281/zenodo.5913165>

- Real faults
- Full source code
- Fine grained test coverage reports
- Docker env. for hassle-free replication

Any questions? Drop us an email!

[luigiliberolucio.starace@unina.it](mailto:luigiliberolucio.starace@unina.it)

