



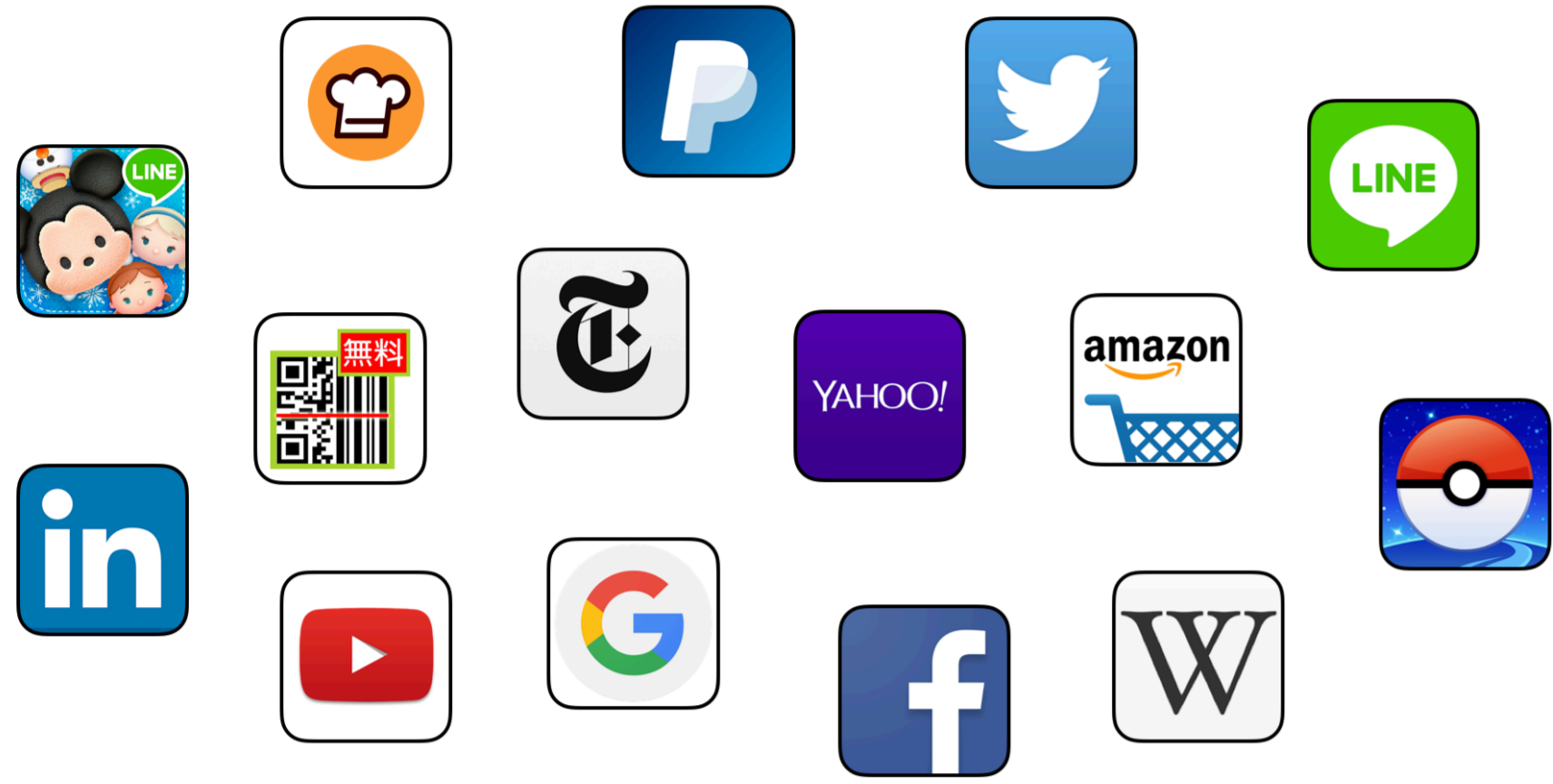
Automatically Translating Bug Reports into Test Cases for Mobile Apps

Mattia Fazzini*, Martin Prammer*, Marcelo d'Amorim[^], Alessandro Orso*

*Georgia Institute of Technology, [^]Federal University of Pernambuco



Mobile Applications



Mobile App Bug Reports

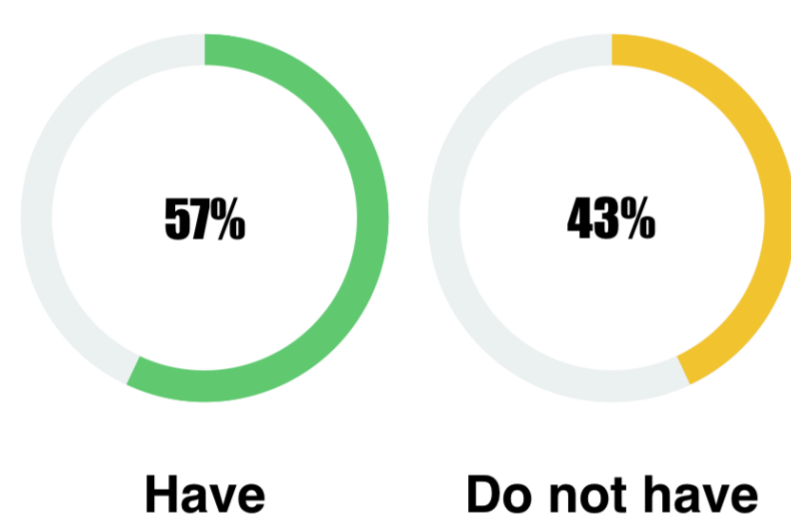
One bug report per day

[Bhattacharya et al., CSMR 2013]

Important aspects of a bug report

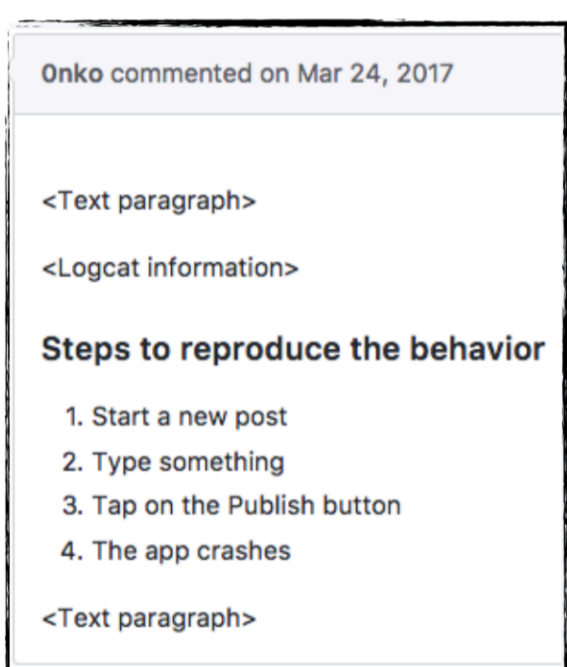
- Observed behavior
- Expected behavior
- Steps to reproduce
- Stack trace

How many bug reports have steps to reproduce?



Motivating Example

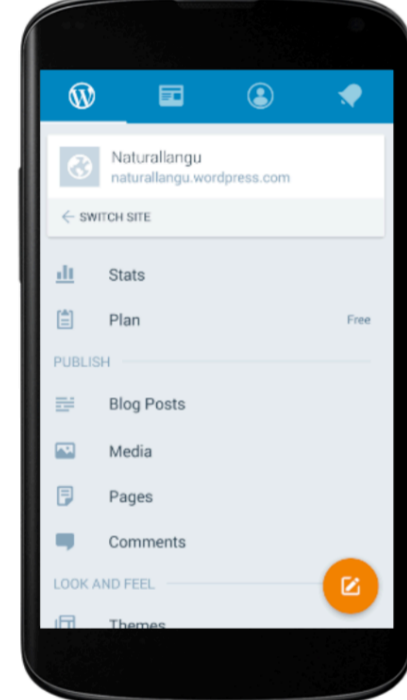
Bug Report



Developer

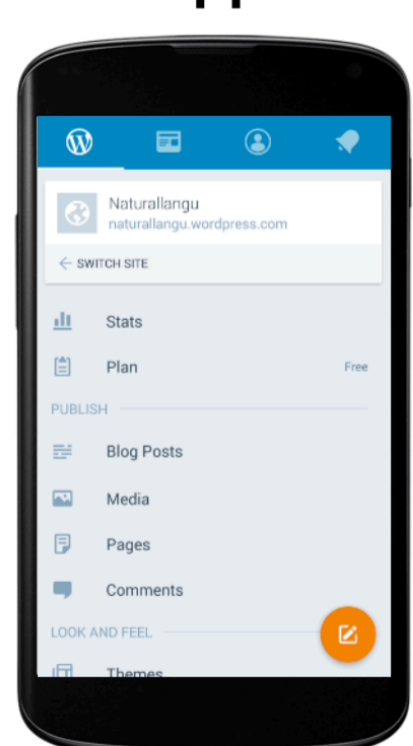


WORDPRESS App



Motivating Example

WORDPRESS App



Test Case

```
public void test() {
    //start a new post
    onView(withId(R.id.fab.button))
        .perform(click());

    //type something
    onView(withId(R.id.post_title))
        .perform(typeText("GhbcVK91"));

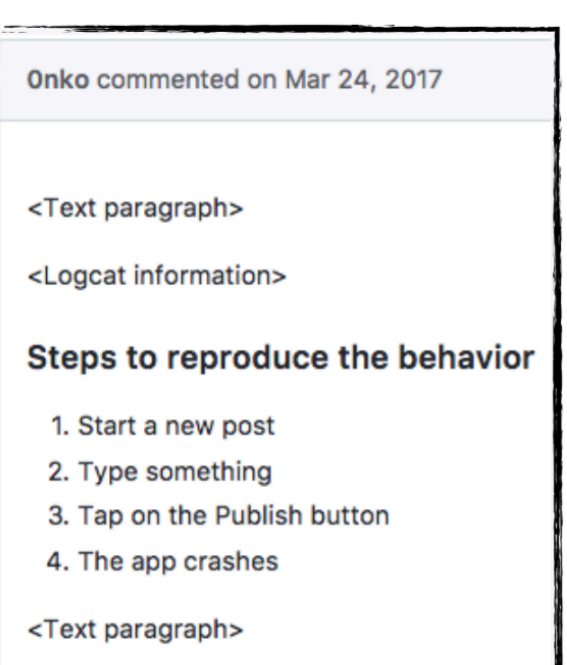
    //tap on the publish button
    onView(withId(R.id.menu_save_post))
        .perform(click());
}
```

Developer



Motivating Example

Bug Report



Test Case

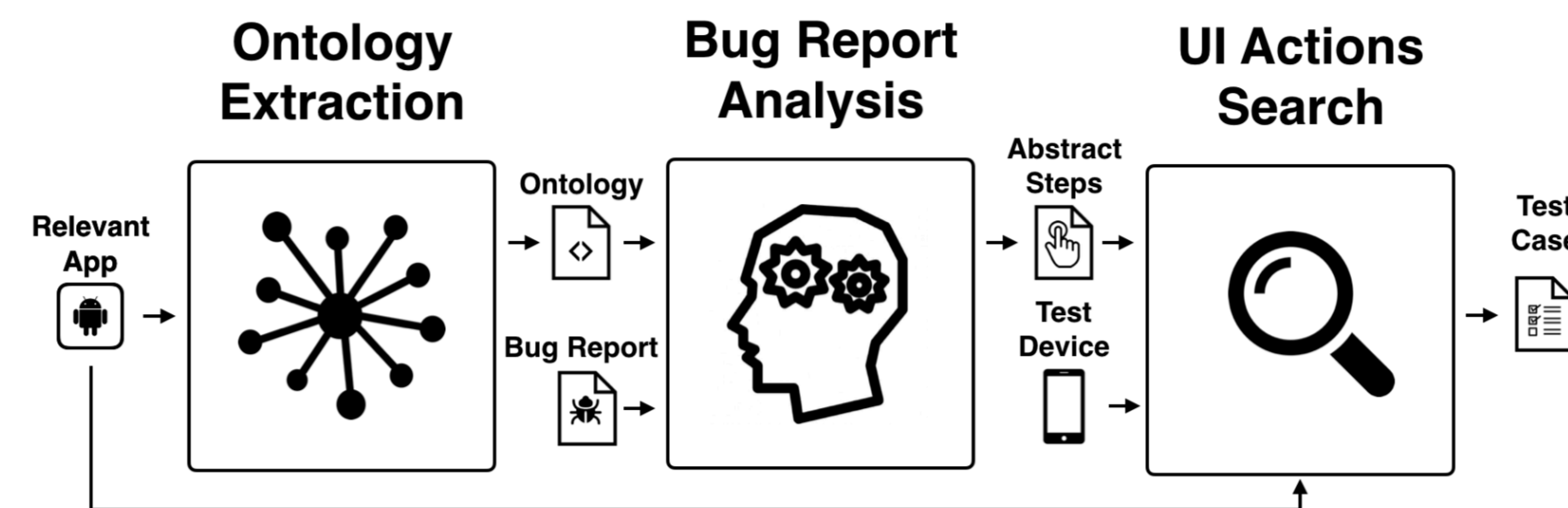
```
public void test() {
    //start a new post
    onView(withId(R.id.fab.button))
        .perform(click());

    //type something
    onView(withId(R.id.post_title))
        .perform(typeText("GhbcVK91"));

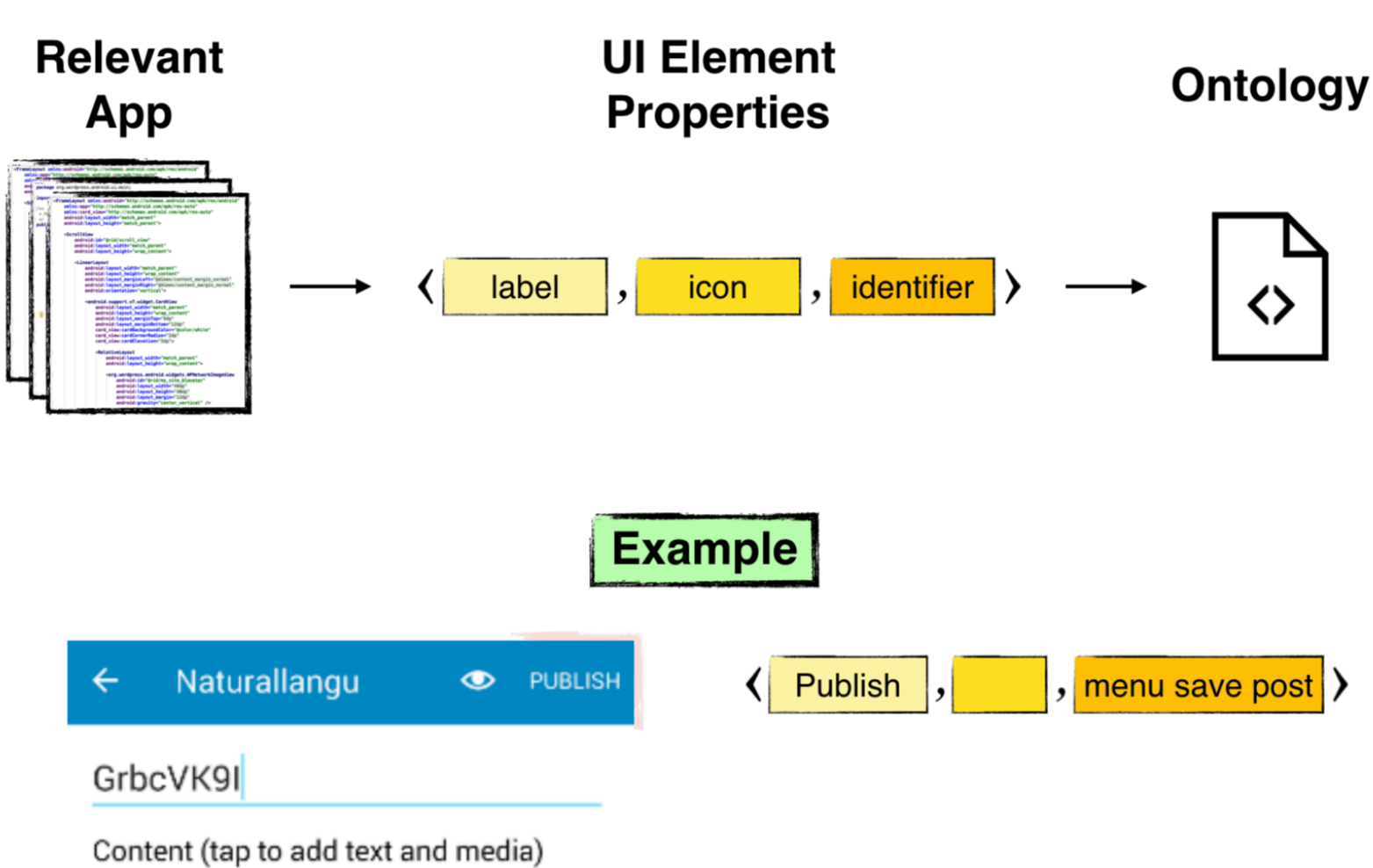
    //tap on the publish button
    onView(withId(R.id.menu_save_post))
        .perform(click());
}
```

Yakusu

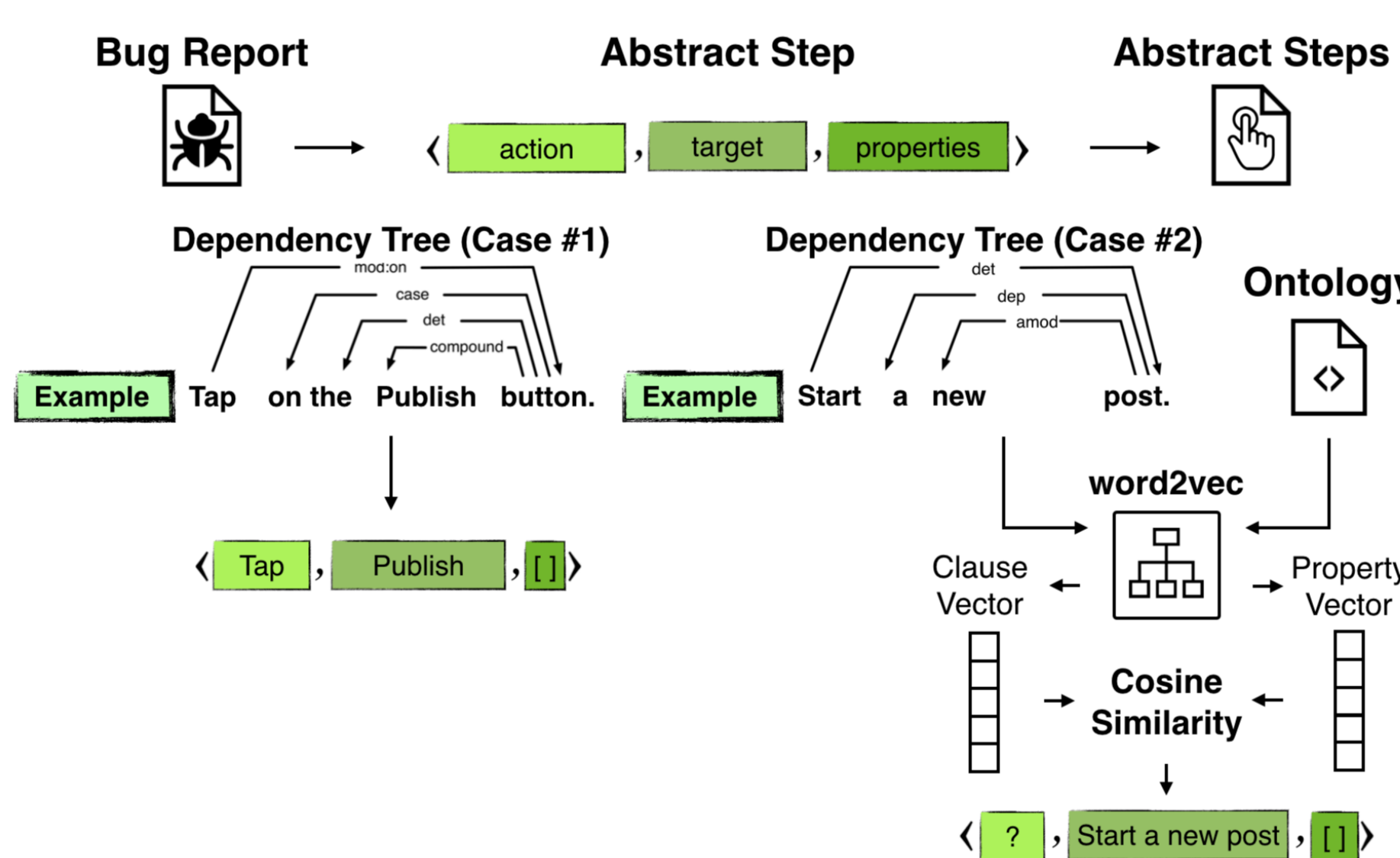
Technique Overview



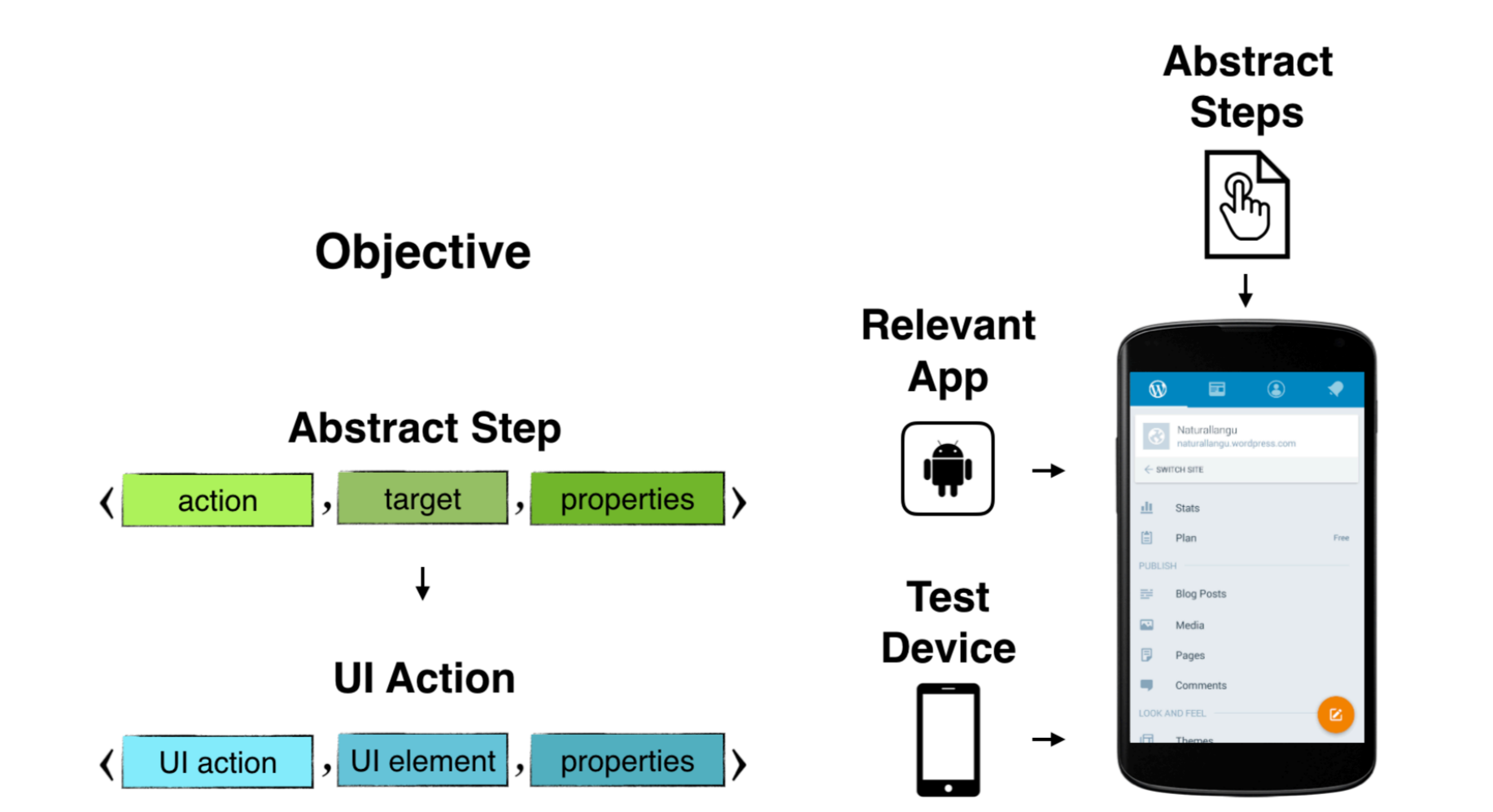
Ontology Extraction



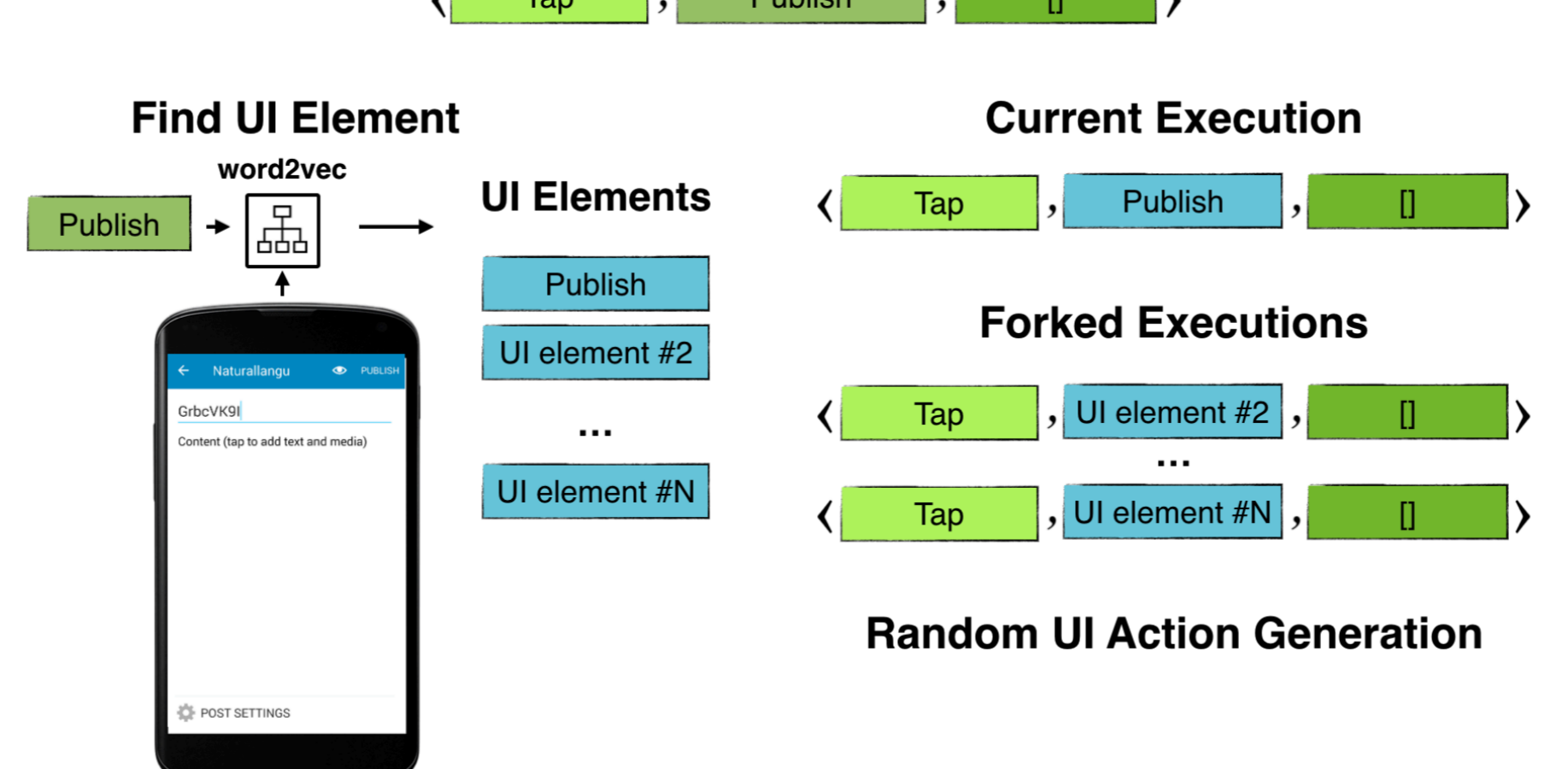
Bug Report Analysis



UI Actions Search



Abstract Step



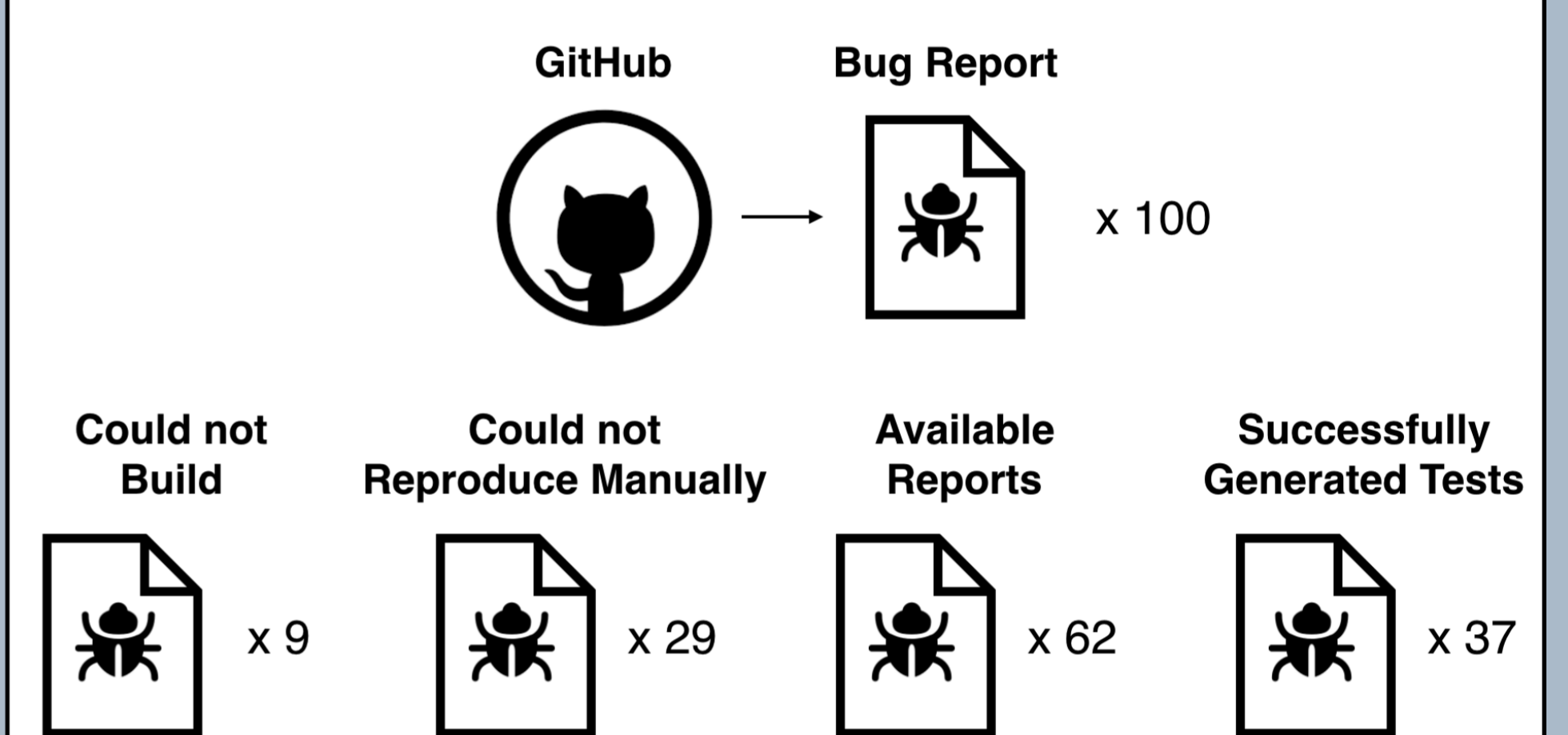
Empirical Evaluation

Research Questions:

RQ1: Can YAKUSU translate bug reports written in natural language into executable test cases?

RQ2: What is the cost of running YAKUSU?

Benchmarks and Setup



Research Question #1

ID	Name	Actions		Steps		Search				Tests	
		A _o	A _i	AS	AS _g	S _g	S _p	H	R	TC _s	TC _a
01	TACHYOMI	4	1	-	4	10	1	1	-	5	-
03	TWIDERE	2	-	2	-	12	1	-	-	2	-

9 bug reports had at least one implicit action

7 bug reports with at least one generic abstract step

19 bug reports reproduced exploring only one execution

6 bug reports reproduced exploring more than one execution

4 test cases with additional actions

Research Question #2

ID	Name	Cost		
		T _{oe}	T _{ora}	T _{uis}
01	TACHYOMI	54s	18s	1m12s
03	TWIDERE	2m34s	15s	1m07s
05	SIGNAL	1m23s	17s	6m57s
08	REDREADER	25s	19s	1m25s
99	FLASHCARDS	42s	15s	5m08s

UI Actions search phase takes the longest on average

Average analysis time is 5m00s and median time is 2m58s

Longest analysis time is 29m41s

Future Work

- Extend ontology
- Learn macro-step associated with the app
- Interpret non-crashing failures from bug reports
- Generate test cases from written/spoken test specifications
- Identify duplicate bug reports