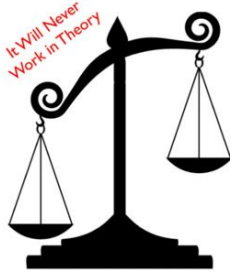
How to Recommend Tools for Finding and Fixing Software Errors

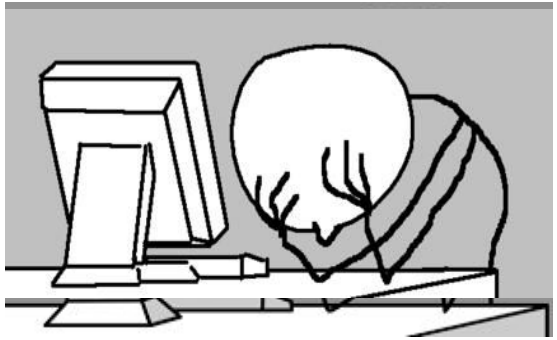
OR

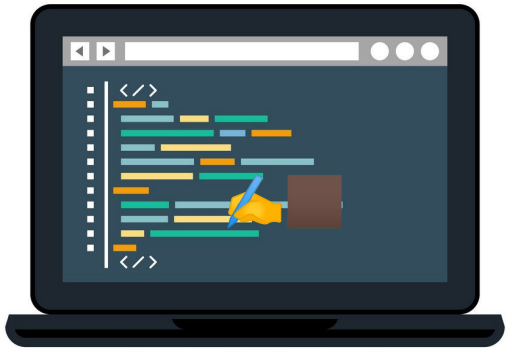
Overcoming the Unwise Monkeys of Tool Adoption

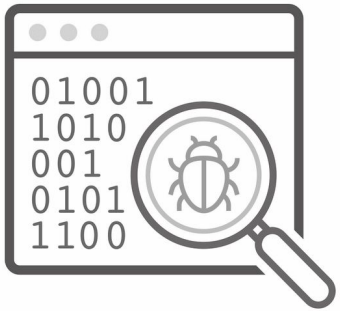
Chris Brown
Virginia Tech
It Will Never Work in Theory at Strange Loop

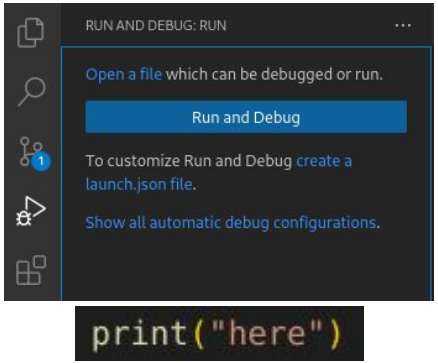


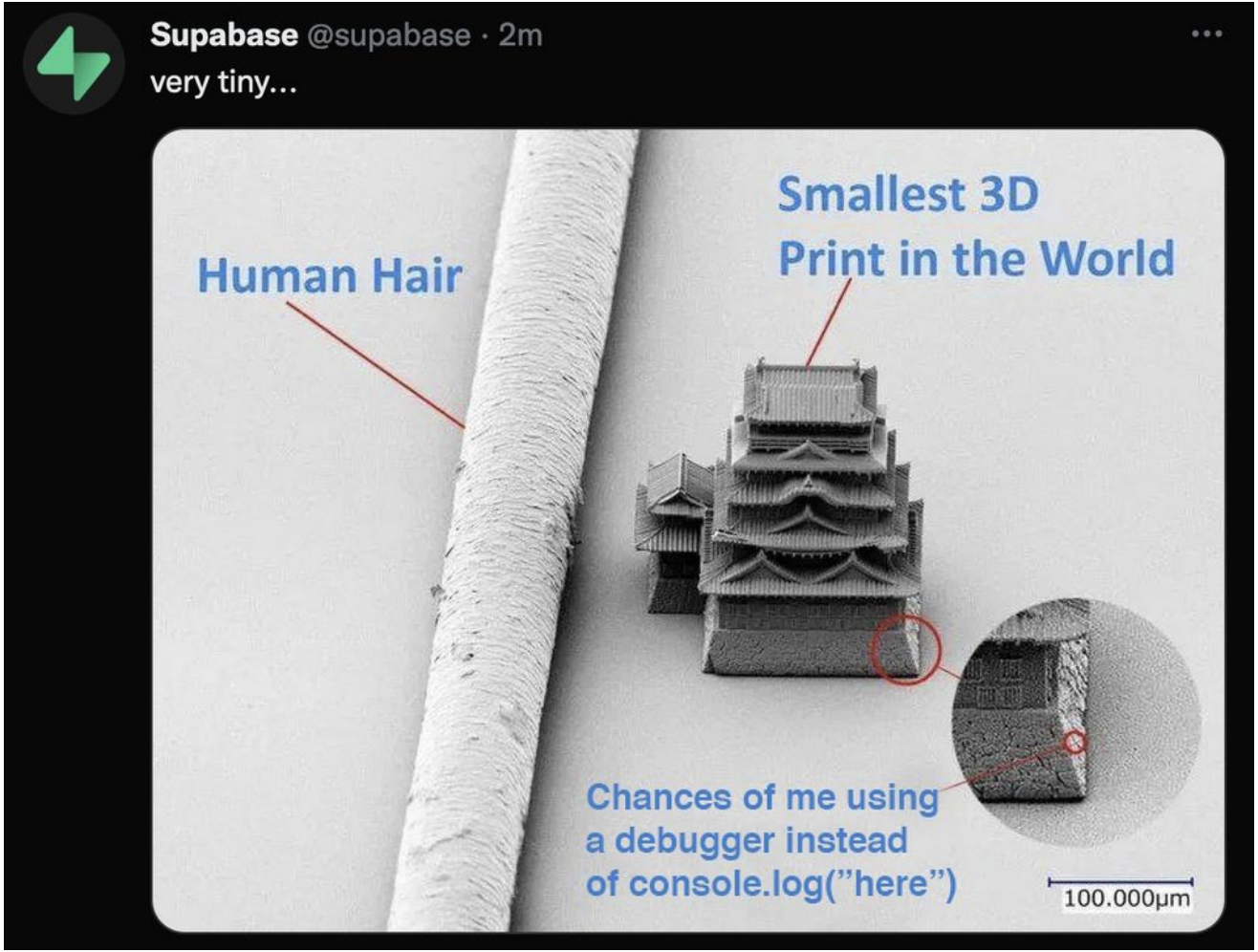
Finding and Fixing Errors










Supabase @supabase · 2m
very tiny...
Human Hair
Smallest 3D
Print in the World
Chances of me using
a debugger instead
of console.log("here")
100.000µm

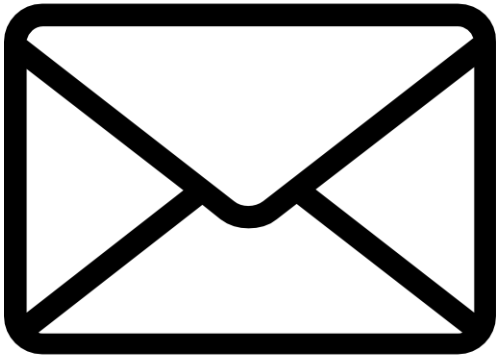
The Three *Unwise* Monkeys

Hear About
No Tools

See No Tools

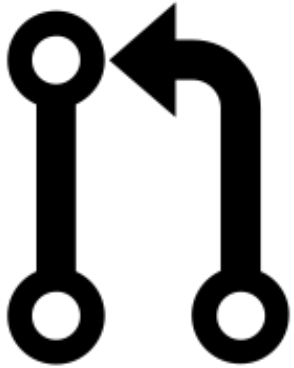
Speak About
No Tools

1



Automated Emails

2



Automated Pull Requests

Error Prone Static Analysis Tool #82

Open **cass-green** wants to merge 1 commit into `apache:master` from `cass-green:master`

Conversation 0 · Commits 1 · Checks 0 · Files changed 1

cass-green commented on Jan 31 • edited

Looks like you're not using any error-checking in your Java build. This pull requests adds a static analysis tool, Error Prone, created by Google to find common errors in Java code. For example, running `mvn compile` on the following code:

```
public boolean validate(String s) {
        return s == this.username;
}
```

would identify this error:

```
[ERROR] src/main/java/HelloWorld.java:[17,17] error: [StringEquality] String comparison
[ERROR]     (see https://errorprone.info/bugpattern/StringEquality)
```

```
88  88          </plugin>
    89  +       <plugin>
    90  +       <groupId>org.apache.maven.plugins</groupId>
    91  +       <artifactId>maven-compiler-plugin</artifactId>
    92  +       <version>3.5.1</version>
```

this pull request. Please feel free to add
ommendation useful.

[Brown, 2017]

Vest commented on Feb 7, 2019

This PR failed automatic checks, I think it should be closed.

gastaldi commented on Jan 29, 2019

Thanks for the contribution, but given the number of errors, I think it would cause more harm than good ;)

bendem commented on Jan 28, 2019 Contributor

This introduces a bunch of errors, can you check whether they are worth fixing or configure the plugin so as to ignore the false positives? https://travis-ci.org/fizzed/rocker/jobs/485416635
Also, you messed up the formatting of the pom.xml pretty bad.

yyuu commented on Jan 31 Owner

Umm. Your change itself looks good, but it seems CI's failing somehow 🤔 Will try digging it later....

https://travis-ci.org/yyuu/jetty-nosql-memcached/builds/486048677

gunnarmorling commented on Jan 31, 2019 Member

Hey; I appreciate your efforts, but this change removes quite a lot if important things from the POM file. Can you remove this, please? ErrorProne itself is something I could see us adding, but can you comment on the overhead in terms of build time it adds? Thanks!

jjlauer commented on Jan 28, 2019 Member

This is the 3rd or 4th bot now in the last few months submitting a PR to promote certain tools.

All checks have failed
1 errored check

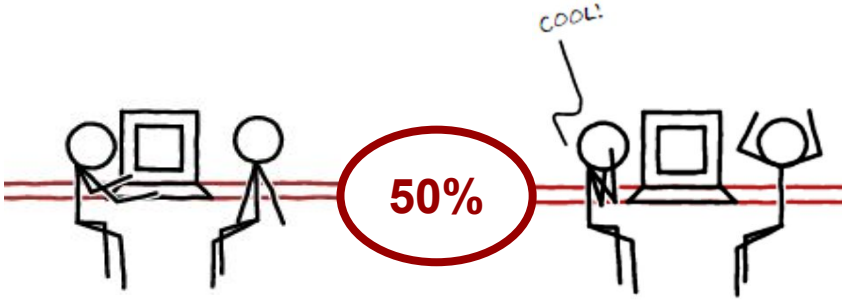
continuous-integration/travis-ci/pr — The Travis CI build

3

Peer interactions are effective!

Peer Interactions

[Murphy-Hill, 2011]



[Brown, 2019]

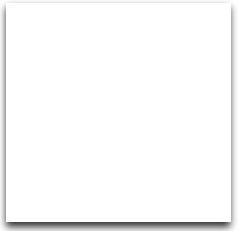
BUT…

They are also the most infrequent mode of tool discovery!

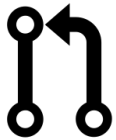
[Murphy-Hill, 2011]

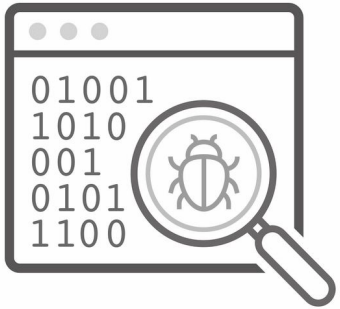
Opportunities for peer interactions are in decline:

- Remote Work 😷
- Global Software Engineering 🌎
- Physical Isolation 🧍
- Lack of scalability 📈
- Mandated tools 🏢
- Learning curve ⏱️
- Too many tools 🤯
- …



What Did We Learn?



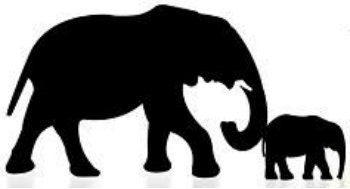


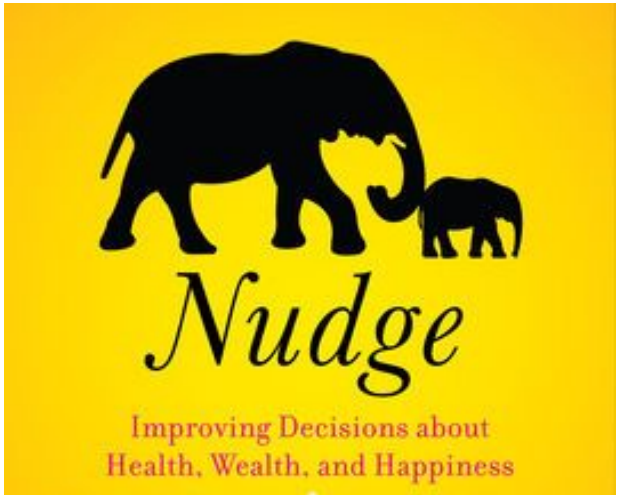
Hear About
No Tools

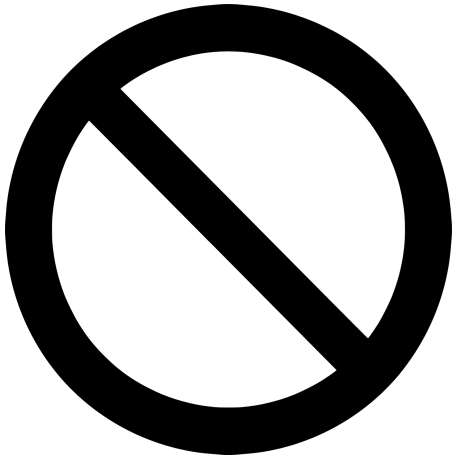
See No Tools

Speak About
No Tools

Solution: Nudge Theory







[Thaler and Sunstein, 2009]

Nudge Theory (cont.)

Choice Architecture: ***The framing and presentation of choices to decision-makers***

[Thaler, 2009]

Developer Recommendation Choice Architectures

No evidence for nudging after adjusting for publication bias

Maximilian Maier, František Bartoš, T. D. Stanley, +2, and Eric-Jan Wagenmakers Authors Info & Affiliations

July https://doi.org/10.1073/pnas.220030

Does this actually work in practice? #neverworkintheory

Actionable Recommendations

```
9  +  if status is True:
10 +    print 'passed'
```

cass-green now

Hi, the latest version of Python changes `prin` leading to a PEP 3105 warning here [1]. We re

`print('passed')`

This change will not impact the functionality o longer supporting Python 2 as of Jan. 1, 2020 project to Python 3. Thanks!

[1] https://www.python.org/dev/peps/pep-3105
[2] https://www.python.org/doc/sunset-python-

Reply…

```
9  +  if status is True:
10 +    print 'passed'
```

cass-green 33 seconds ago

Hi, the latest version of Python changes `print` to a built-in function instead of a statement, leading to a PEP 3105 warning here [1]. We recommend changing this line to

Suggested change

```
10 -    print 'passed'
10 +    print('passed')
```

Commit suggestion ▾ | Add suggestion to batch

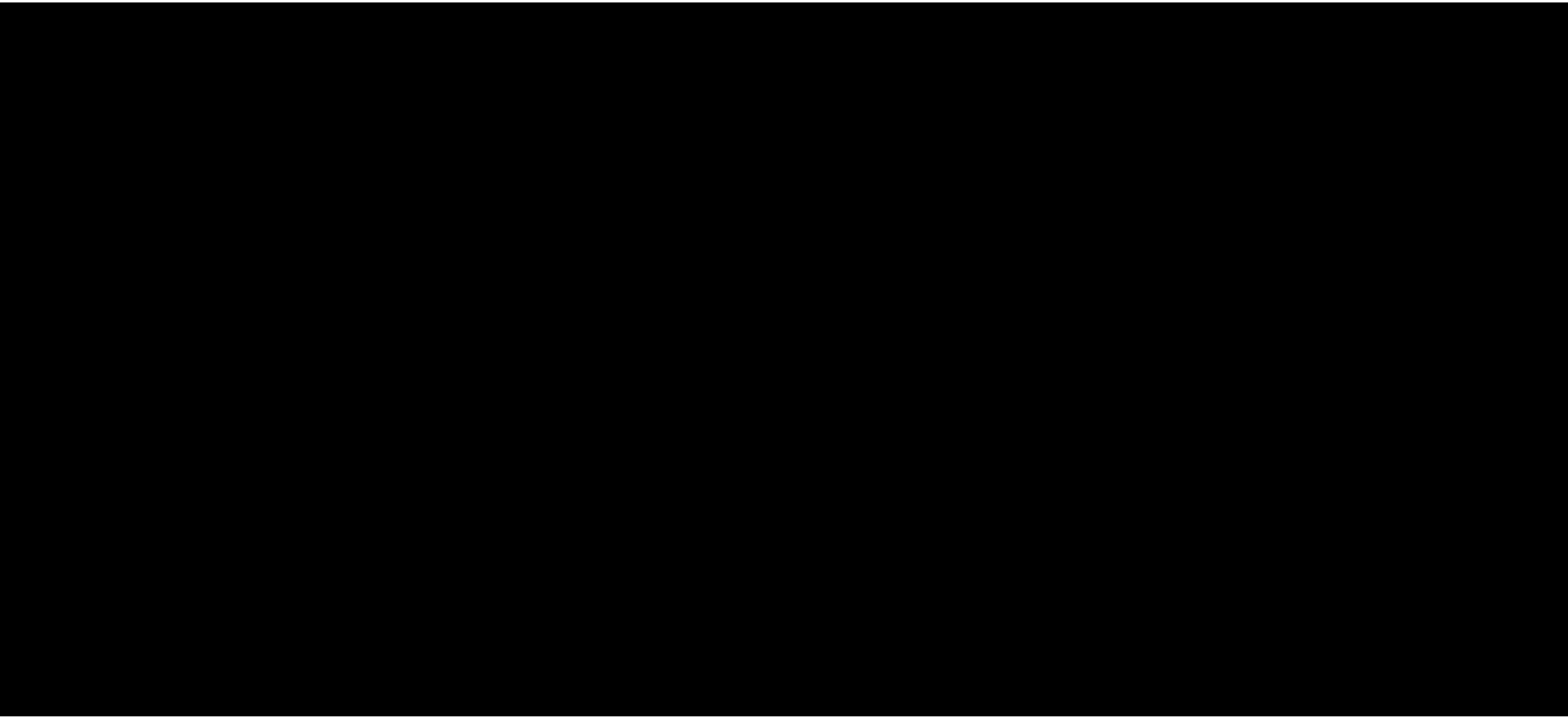
This change will not impact the functionality of your code. Additionally, Python is officially no longer supporting Python 2 as of Jan. 1, 2020 [2]. Please consider upgrading the code for your project to Python 3. Thanks!

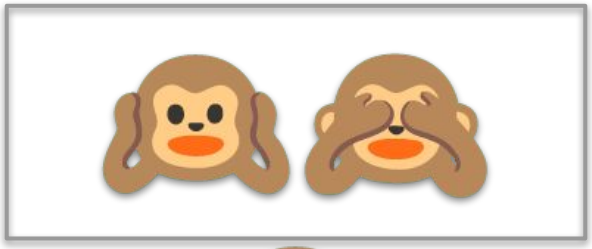
100%! ⭐

[1] https://www.python.org/dev/peps/pep-3105/
[2] https://www.python.org/doc/sunset-python-2/

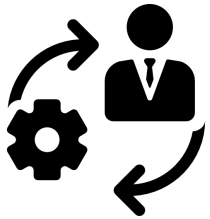
Reply…

[Brown, 2020]

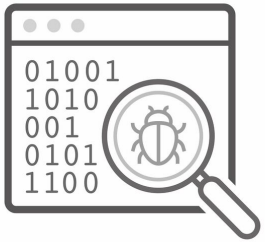




Researchers and Tool Bu[illegible]s:



Developers:



Thanks

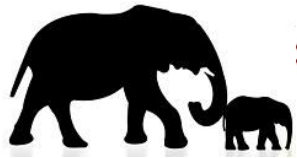
▸ *Would you like to get involved? Check out:*
https://se-participants.github.io!

D r Recommendation e Architectures

- **Actionability**
- **Feedback**
- **Locality**

Chris Brown

dcbrown@vt.edu
https://chbrown13.github.io
chbrown13
@d_chrisbrown2



EST. 2021 CODE WORLD NO BLANKET

VIRGINIA TECH

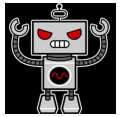
Backup Slides

Current/Future Work

Automated suggestions for other types of tools

- i.e. security, usability, etc. bugs

Other interdisciplinary concepts for recommendations

Dark Patterns 

Deceptive user interface designs to deceive users into adopting behaviors, usually beneficial to the designers.

Ex) hidden ads

Developer Recommendation Choice Architectures

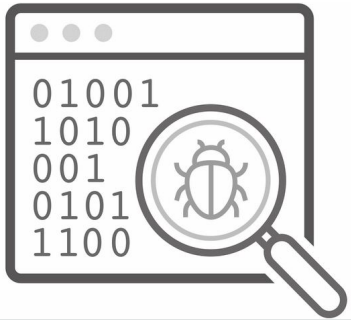
DRCA	Tool [16]	Definition
Actionability	Technology and decision aids	Introducing technology to aid decision makers in choice tasks
	Use defaults	The way decision makers initially encounter choice tasks
Feedback	Reduce number of alternatives	Limiting the number of choice options presented to decision makers
	Focus on satisficing	Helping users consider outcomes that lead to higher choice satisfaction
	Attribute parsimony and labeling	Limiting the number of characteristics presented with options
	Translate and rescale for better evaluability	Presenting attributes to increase impact and clarity
	Customized information	Personalization to account for individual differences between decision-makers
	Focus on experience	Considering the background and knowledge of decision-makers
Locality	Limited time windows	Providing time restrictions for users to make decisions
	Partitioning of options	Groups or categories of options or attributes
	Decision staging	Dividing decisions into multiple stages

Actionability

The ease with which users can act on recommendations

Default Rule
Automatic Enrollment
[Madrian, 2001]

Static Analysis
Splint (Secure Programming Lint)
[Evans, 2002]



Feedback

Information provided to users in recommendations to encourage adoption

Customized Information
Daily caloric intake
[Wisdom, 2010]

Compiler Error Messages
Argument structure
[Barik, 2018]

OpenJDK	`cannot find symbol`
	`symbol: variable varnam`
	`location: class Foo`

Jikes	`No field named "varnam" was found in type "Foo". However, there is an accessible field "varname" whose name closely matches the name "varnam".`

Locality: *Spatial*

The setting of recommendations to improve user behavior

Decision Staging
Healthy Convenience Lines
[Hanks, 2012]

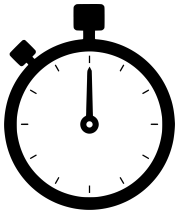
Flower
In situ navigation
[Smith, 2017]

SQLFileCache.java

createTables createProcedures executeSQLFile dropTables

```
26/** @param  fileName String path to SQL file*/
27public List<String> getQueries(String fileName)
28    throws Exception {
29  List<String> queries = cache.get(fileName);
```

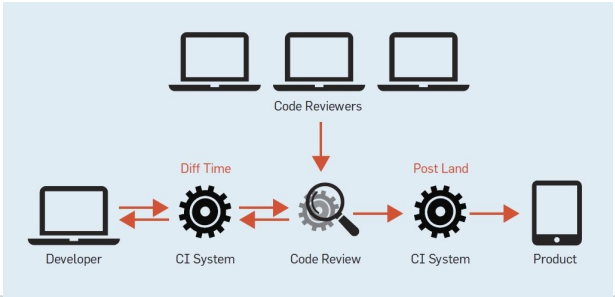
Locality: *Temporal*



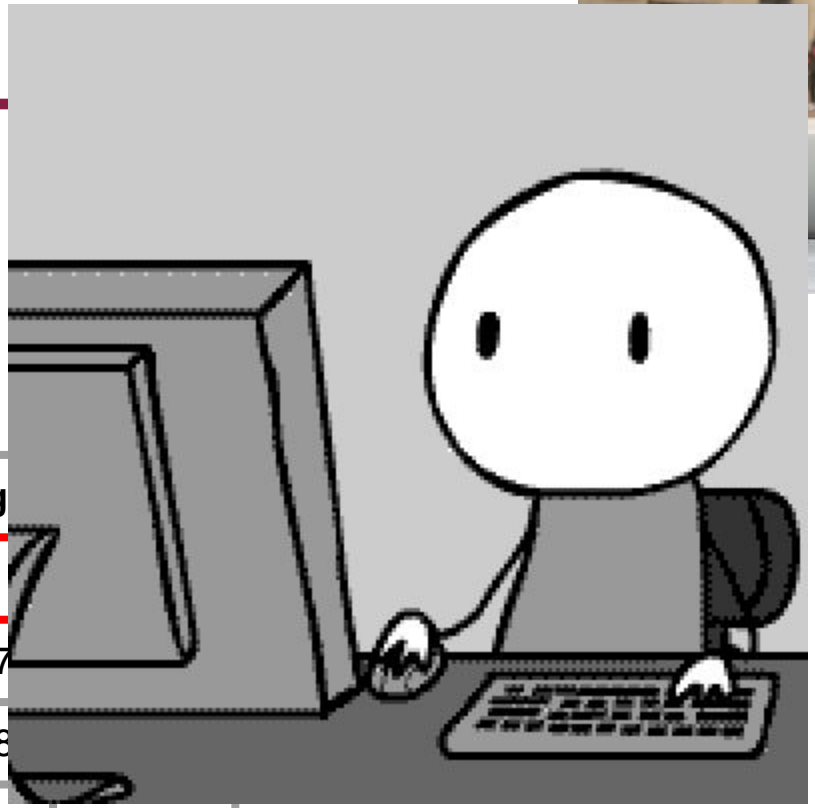
The setting of recommendations to improve user behavior

Time-limited windows
Present-biased farmers
[Duflo, 2011]

Scaling Static Analyses at Facebook
"diff time"
[Distefano, 2019]



Back-Up



	Averag	
Suggestions*		
Pull Requests	3.7	
Issues	2.8	
Email	2.36	2

Mhoye feedback

Strong opener has better opening

Break up decision-making into different slides (3)

- reading slides :/
- delivery

Not just what tools to recommend, how to get recommendations to land and stick

Get programmers to change code is hard, changing tools is harder

Tool finds to many errors

Get to email/PR sooner

* Can't see, can't hear, we have to talk about it!
  * Problem with talking hackernews vs colleagues

Need to break this cycle (find people at this vulnerable moment)

People will understand data but connect with stories

decisions about how we make decisions and improve processes (nod towards nudge theory, debate about whether it will work [suspect])

amount of data hasn't proved itself (made it look like it would be an easy thing to fix, but it's not...)

Feedback 2

How to find and fix errors before they find us

common set of failure modes (common theme) of tooling advocacy

All of these things work, people don't like them and don't want to engage with (do things that work but don't engage, work doesnt matter)

Nudge theory makes tool visible (show data to make it immediately something to react to)