



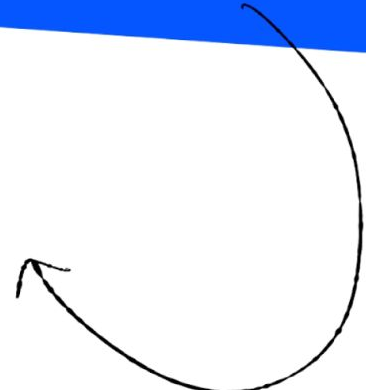
How Atlassian delivers actionable insights that empower teams to improve productivity



Rebecca Fitzhugh
Senior Principal Engineer



Maurits Evers
Principal Data Scientist



Hi, we're from Atlassian



Jira



Confluence



Trello

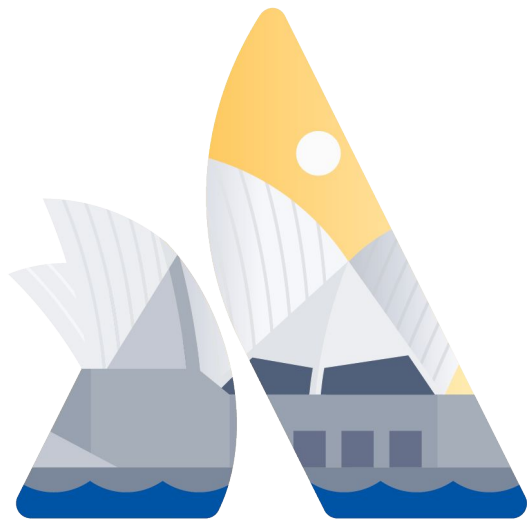


Bitbucket

Unleash the potential of every team



We're an Australian company



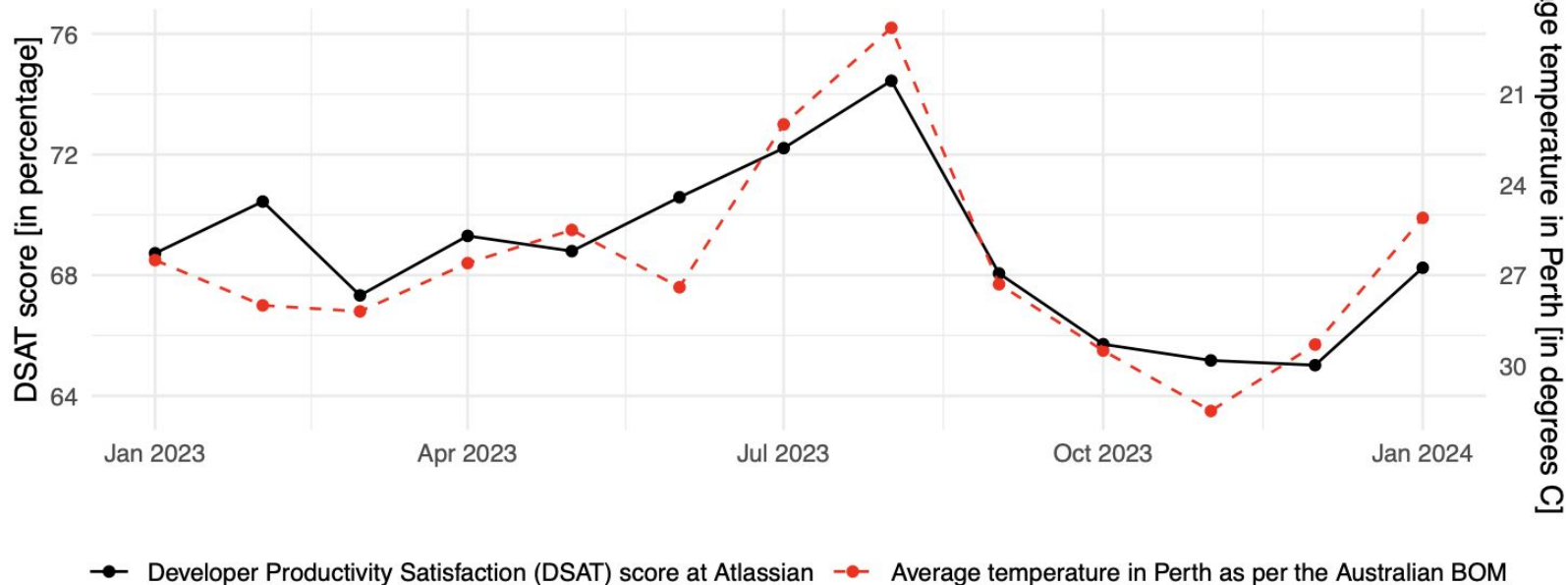
**Australian weather
drives Atlassian
developer productivity**

Developer Productivity Satisfaction at Atlassian

inversely correlates with

Average outside temperature in Perth, Australia

$r \sim -0.88$, $p < 0.01$





Spurious correlation

Two quantities exhibit similar trends & patterns



Causation

Careful when inferring causation from correlation



Spurious correlation

Two quantities exhibit similar trends & patterns



Causation

Careful when inferring causation from correlation

Understanding Dev Prod is hard!

Agenda

Dev Prod @ Atlassian

A short overview

Vital Signs

Measuring Dev Prod

Our 5R framework

Improving Dev Prod

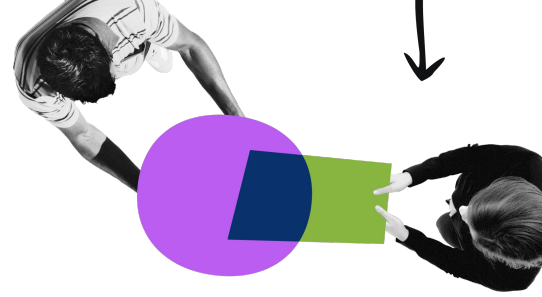
Project Sonic

One of our success stories



Developer Productivity @ Atlassian

**We believe happy engineers
are productive ones.**



Our Approach



**Awesome
Tools**



**Amazing
Engineering
Culture**



**Empowered
Teams**

Data driven

We empower teams to better understand their engineering health, make data-driven decisions, and drive better business outcomes.





Vital Signs

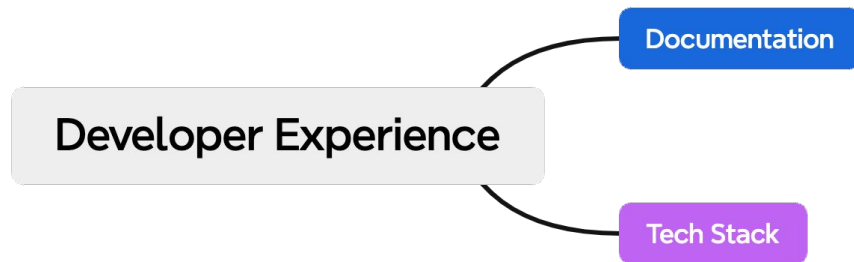
What is a Vital Sign?



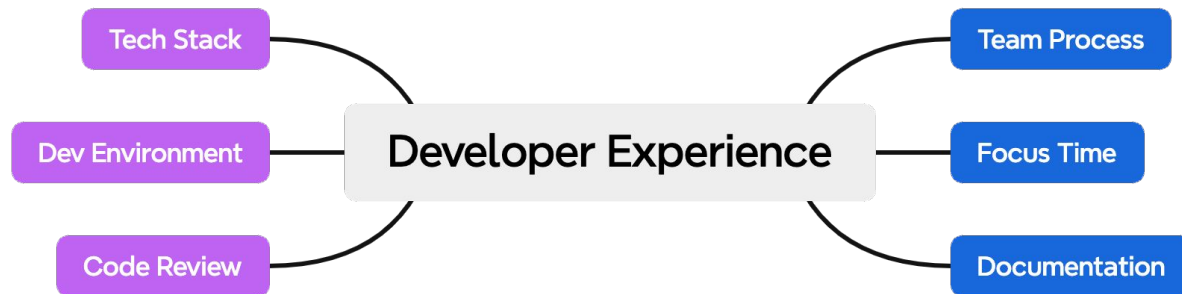
Vital signs represent our diagnostic toolkit.

They focus on inputs and consider the realities of software engineering to help us understand friction in the development lifecycle.

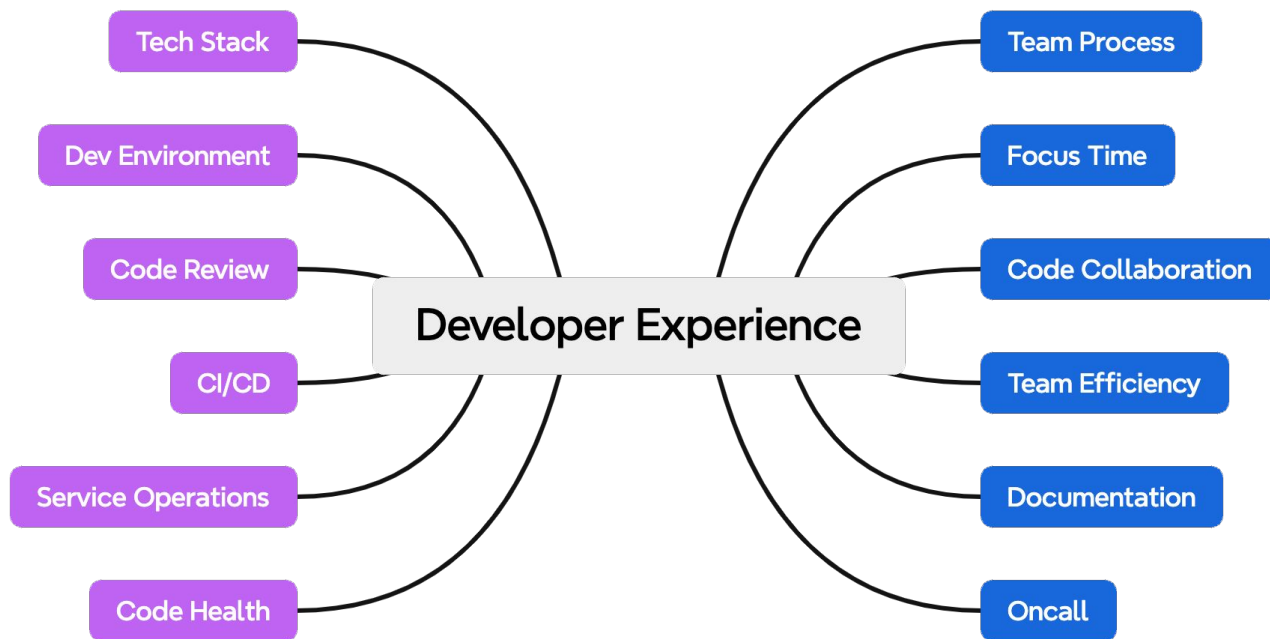
Vital Signs: Dimensions



Vital Signs: Dimensions



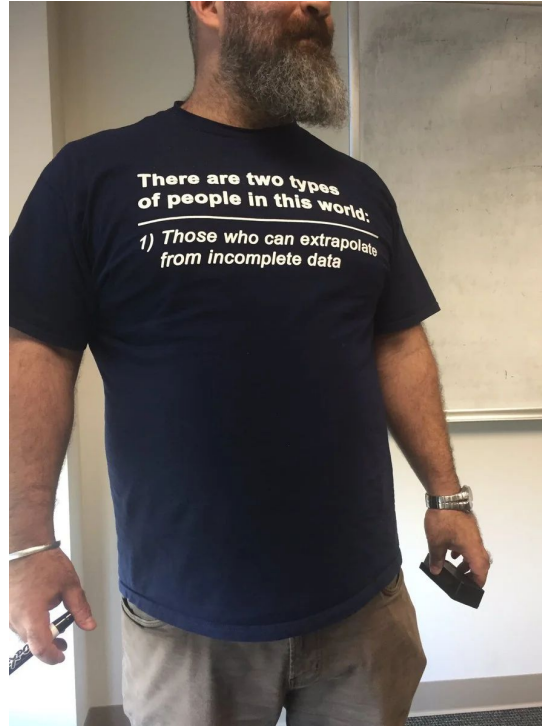
Vital Signs: Dimensions





Dev Prod *is* a measurement problem!

Your friendly resident data
scientist



Vital Signs: Design



Quantitative signals

Measurement comprised of data provided by *systems* or software



Qualitative signals

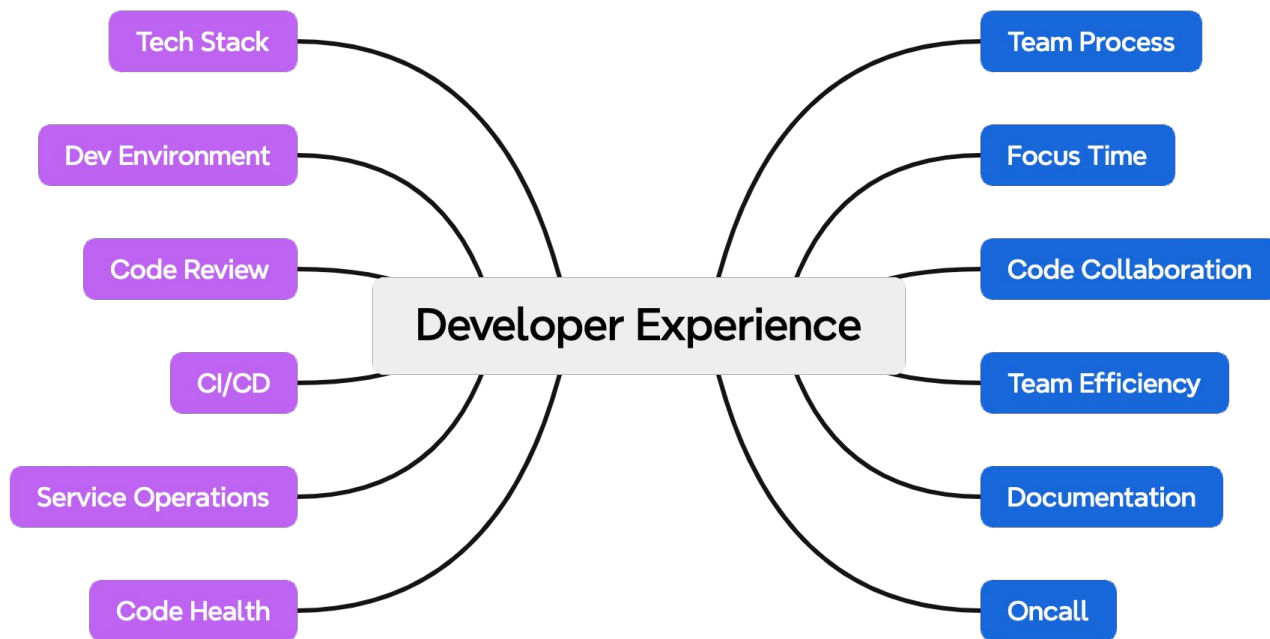
Measurement comprised of data provided by *humans*

“The alternative to good design is always bad design. There is no such thing as no design.”

Vital Signs: Example

	System Data	Human Data
Code Review	PR cycle time - the duration from opening a pull request (PR) by a PR author to deployment into a production environment (or to a build endpoint). PR pickup time - the duration from opening a pull request to the first review action. PR review time - the duration from the first review action to its final approval. PR cycles - how many times a PR was updated or reworked during the review time. PR complexity - defined by number of files changed and number of comments in review; higher complexity is associated with longer pre-merge time. PR review comment length - signal potentially correlated with PR complexity or cycle time.	• The amount of time I have to wait when I request a code review is acceptable. • Our code review process results in us releasing high-quality code

Vital Signs: Dimensions



All measurements are noisy



Stochastic uncertainty

Things change even if there was no real change to productivity



Epistemic uncertainty

Measurement is not accurate or is a proxy of what we really want

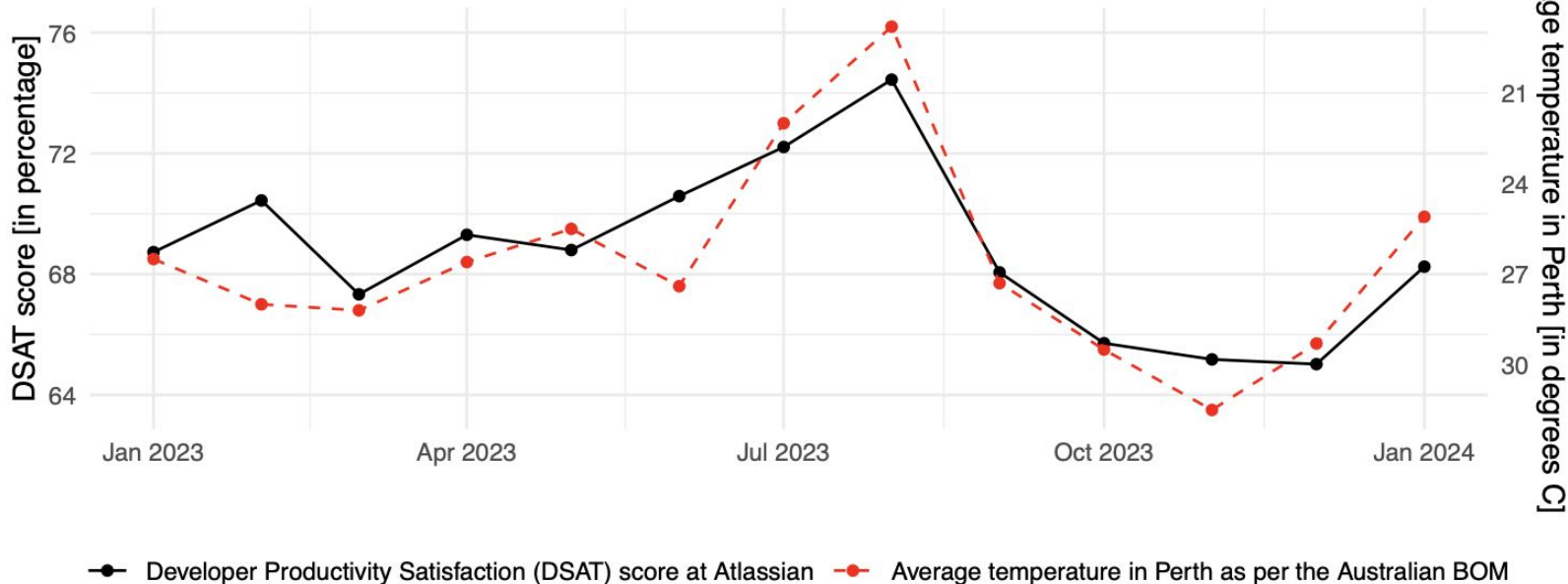
Remember the weather in Perth?

Developer Productivity Satisfaction at Atlassian

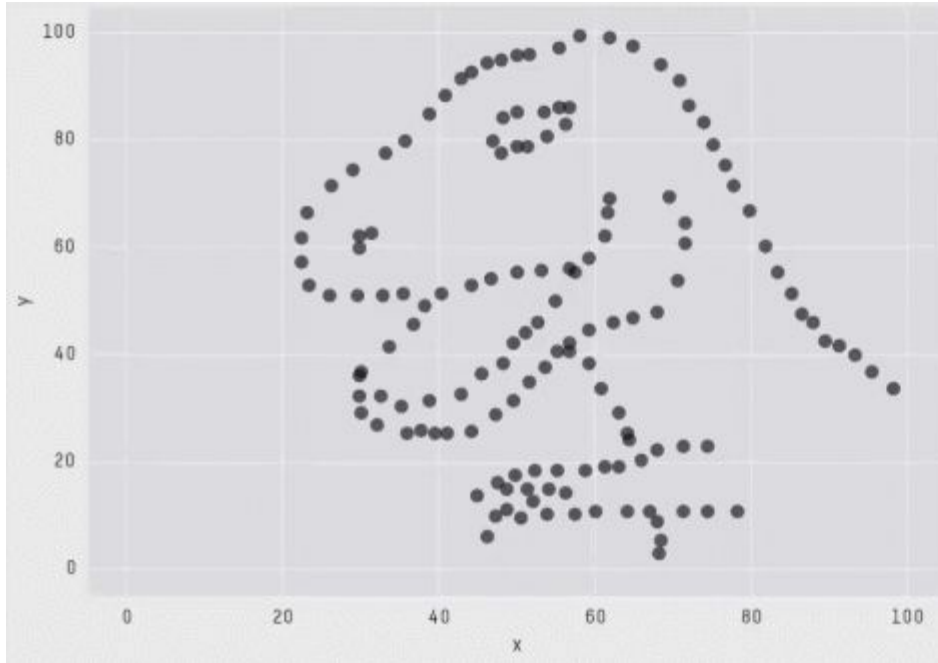
inversely correlates with

Average outside temperature in Perth, Australia

$r \sim -0.88$, $p < 0.01$

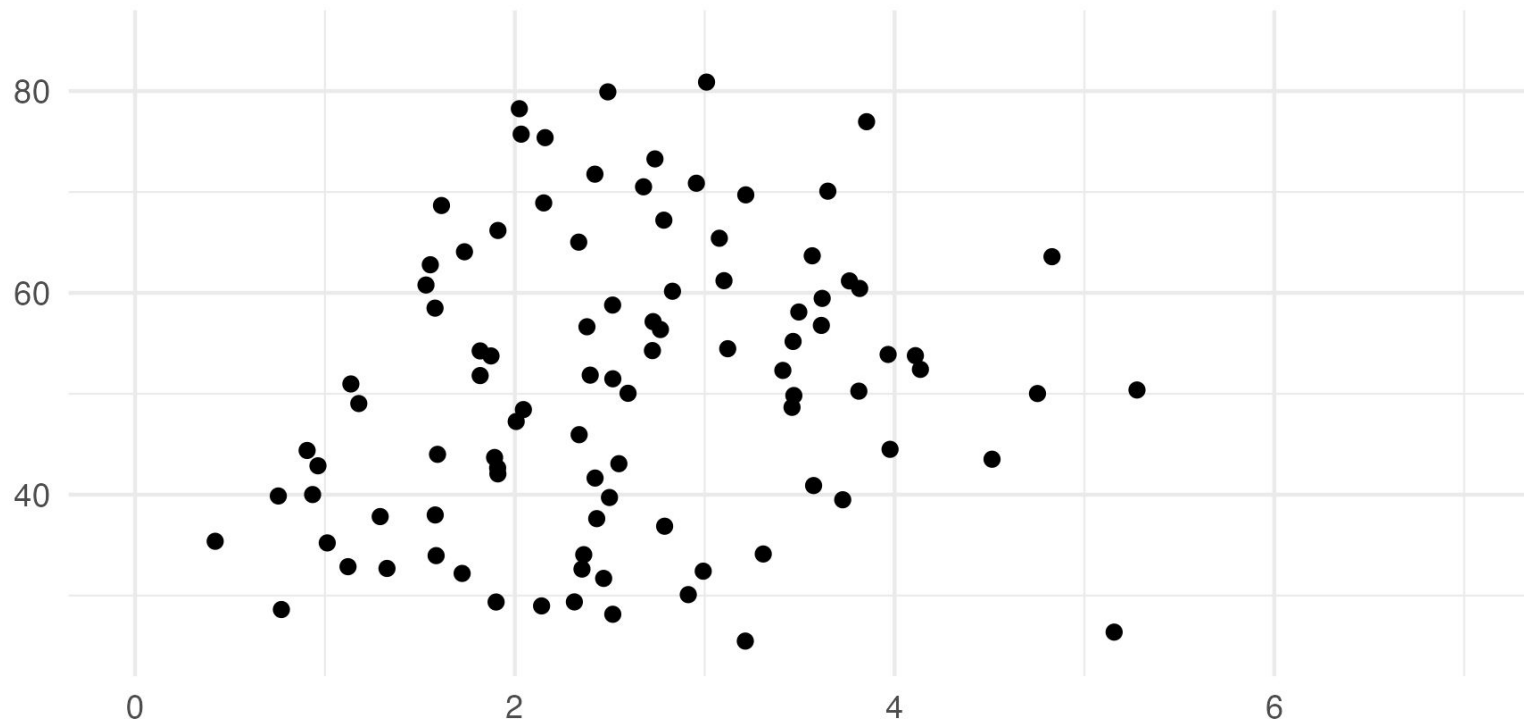


Behold: The Datasaurus

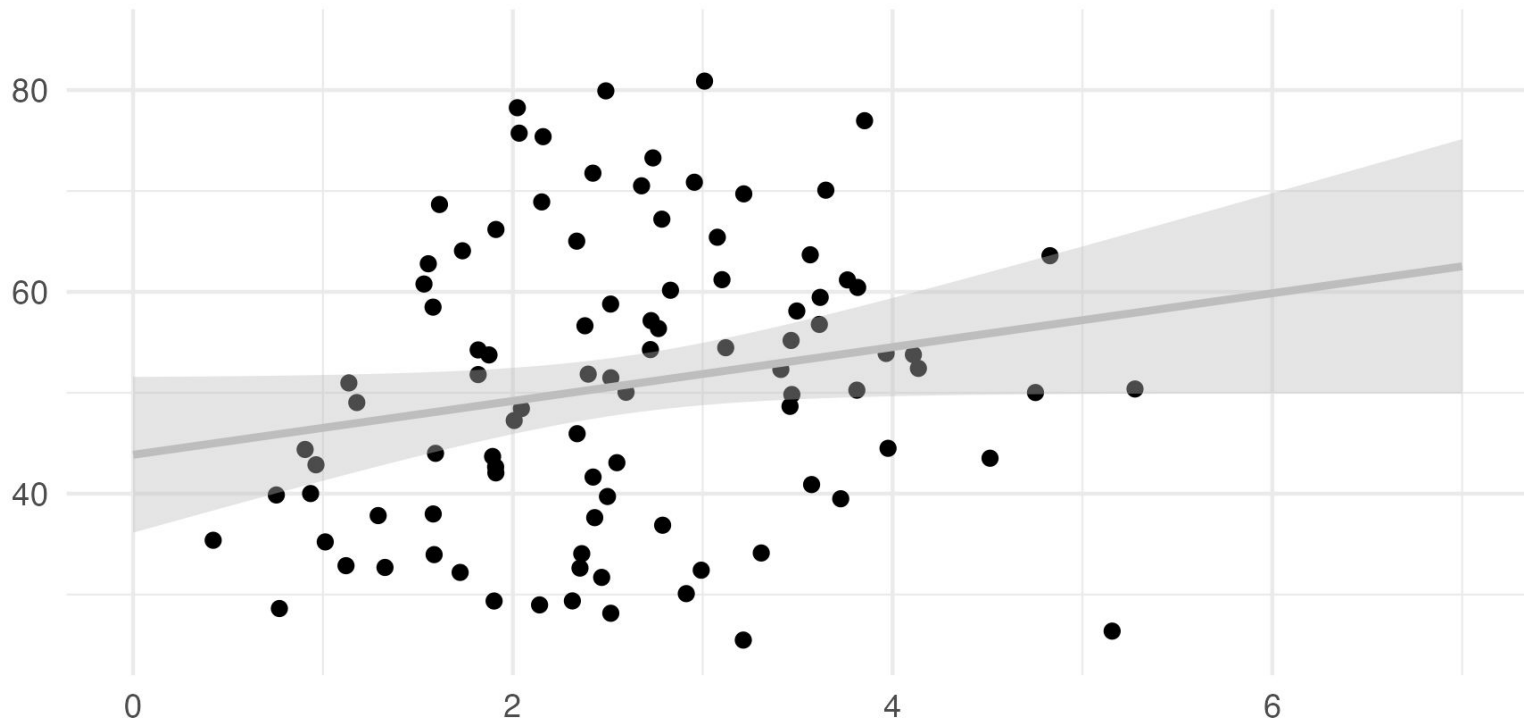


X Mean: 54.2659224
Y Mean: 47.8313999
X SD : 16.7649829
Y SD : 26.9342120
Corr. : -0.0642526

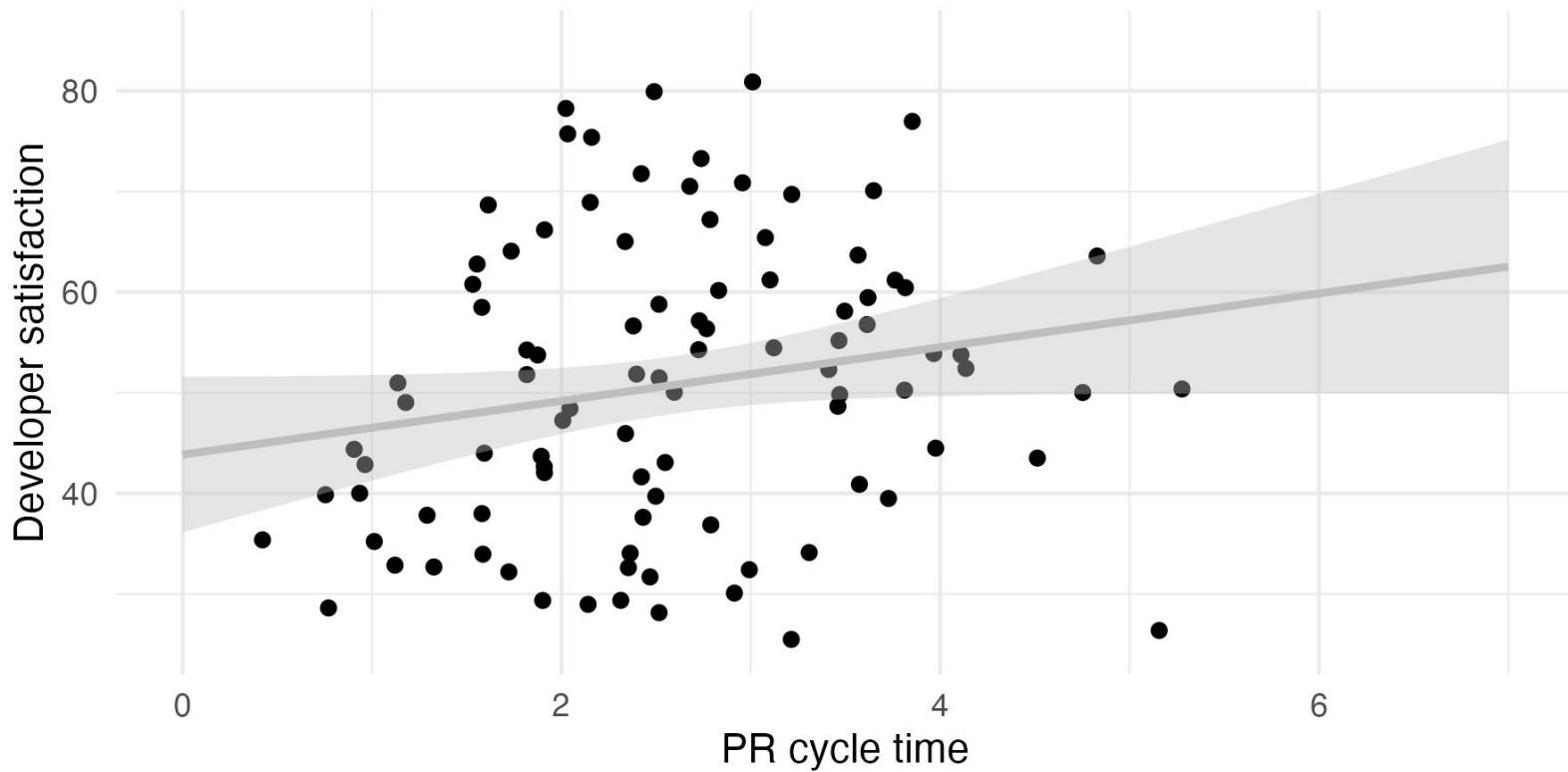
Trendspotting



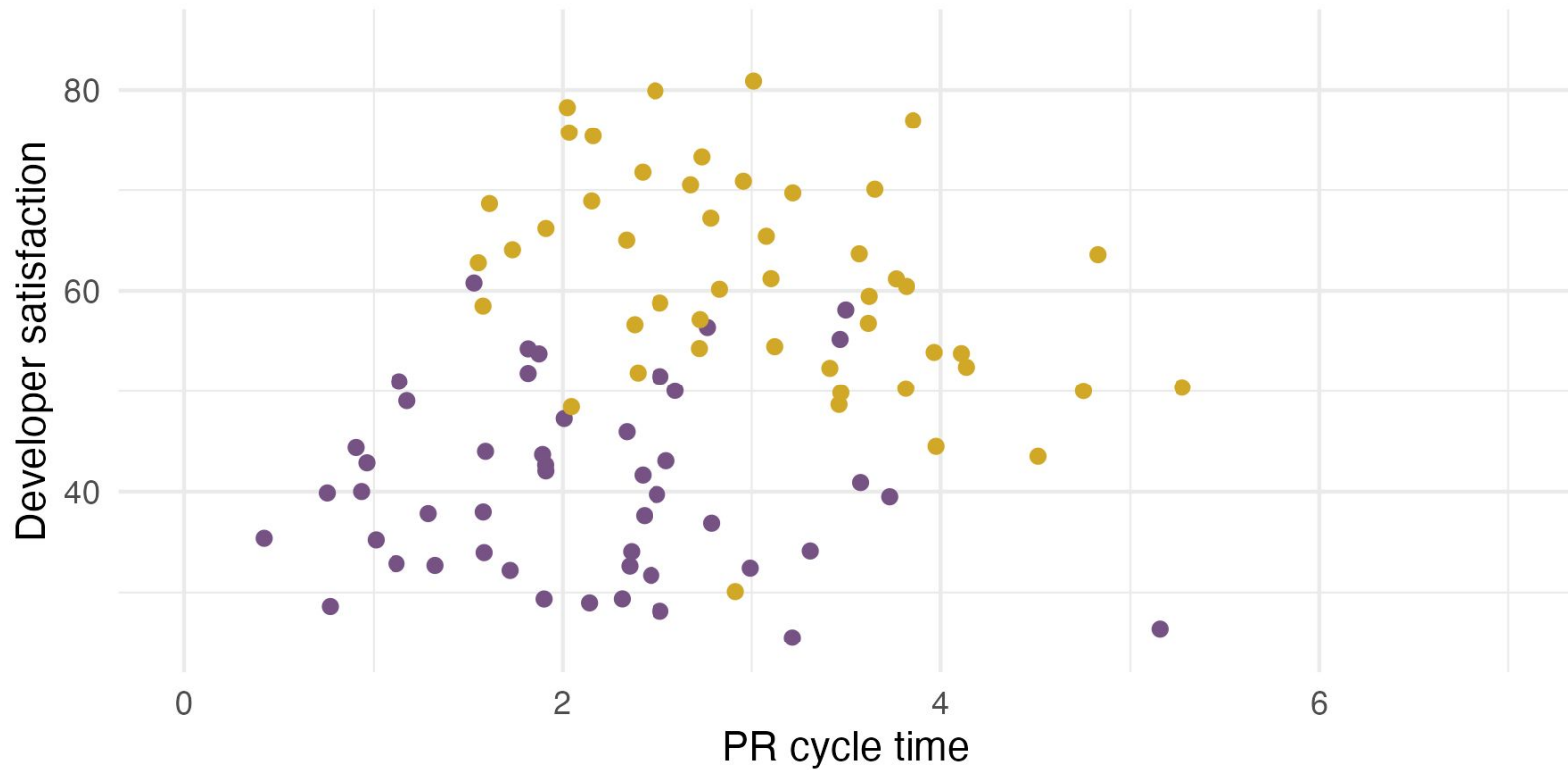
Trendspotting



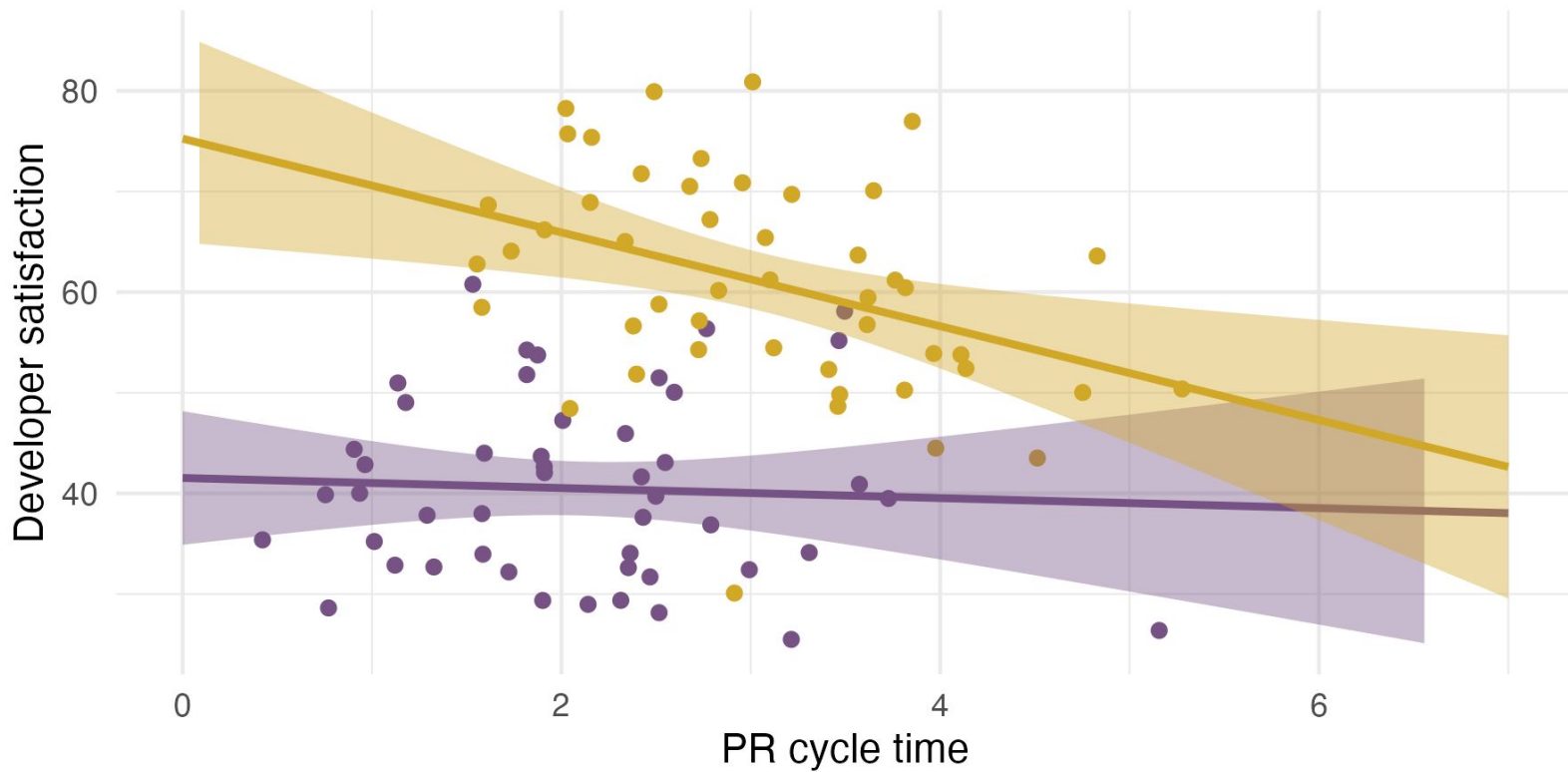
Trendspotting



Trendspotting



Trendspotting



From vital signs to metrics to insights



(Minimal) vital sign/metric design

- Should be meaningful & relevant
- Can be measured/estimated (baselines, predictions)
- Can be changed by the business/team



From vital signs/metrics to insights

- Model the data-generating process
- Test the model

“All models are wrong, but some are useful” (George Box)

From vital signs to metrics to insights



(Minimal) vital sign/metric design

- Should be meaningful & relevant
- Can be measured/estimated (baselines, predictions)
- Can be changed by the business/team



From vital signs/metrics to insights

- Model the data-generating process
- Test the model

"Before we can improve any system we must understand how the inputs affect the outputs of the system."
(Donald Wheeler)

03

5R framework

We empower teams by providing

the right **insight** 💡



We empower teams by providing

the right **insight** 

to the right **person** 



We empower teams by providing

the right **insight** 

to the right **person** 

at the right **time** 



We empower teams by providing

the right **insight** 

to the right **person** 

at the right **time** 

in the right **place** 



We empower teams by providing

the right **insight** 

to the right **person** 

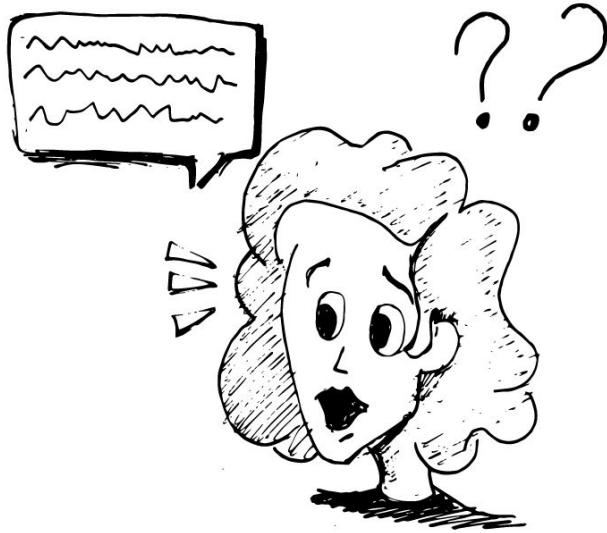
at the right **time** 

in the right **place** 

to do the **right thing** 

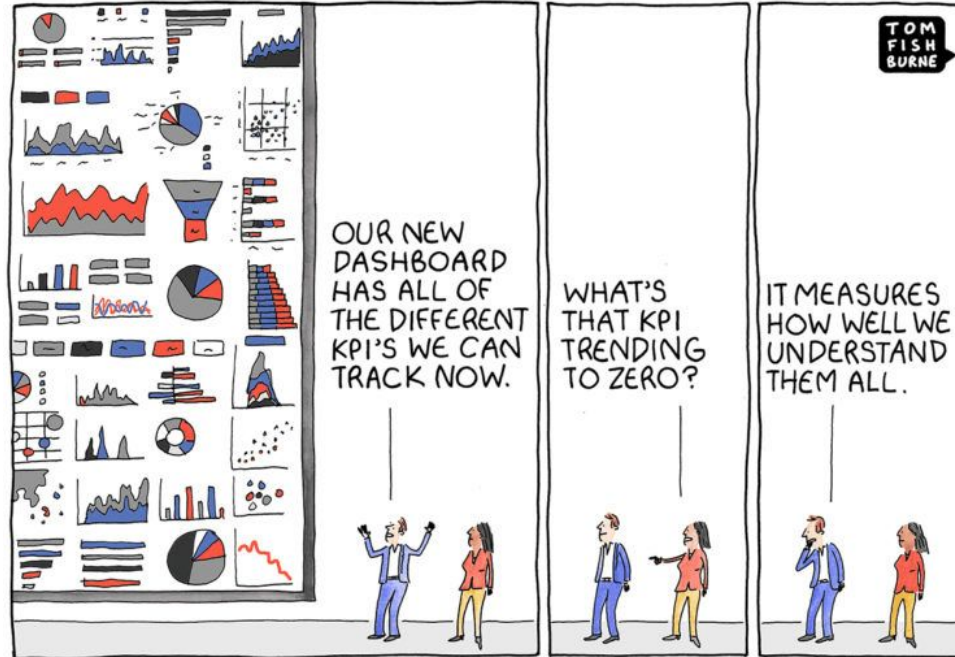


the right **insight** 💡

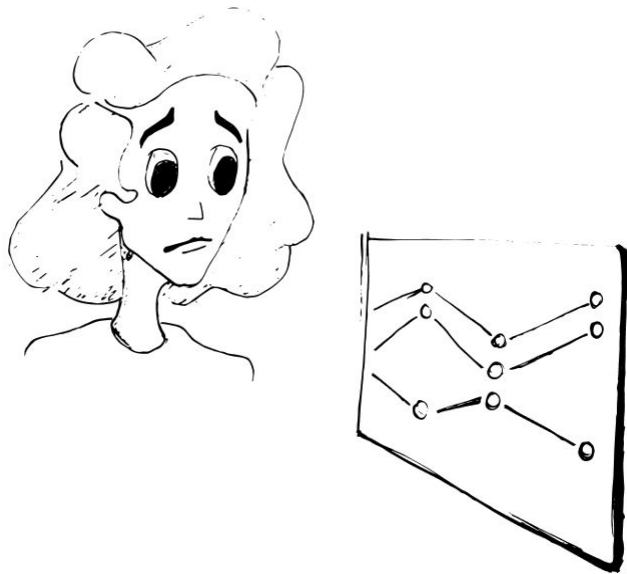


Our insights help teams
interpret data about their
engineering health and
recommended practices to
help improve their
effectiveness.

the right **insight** 💡

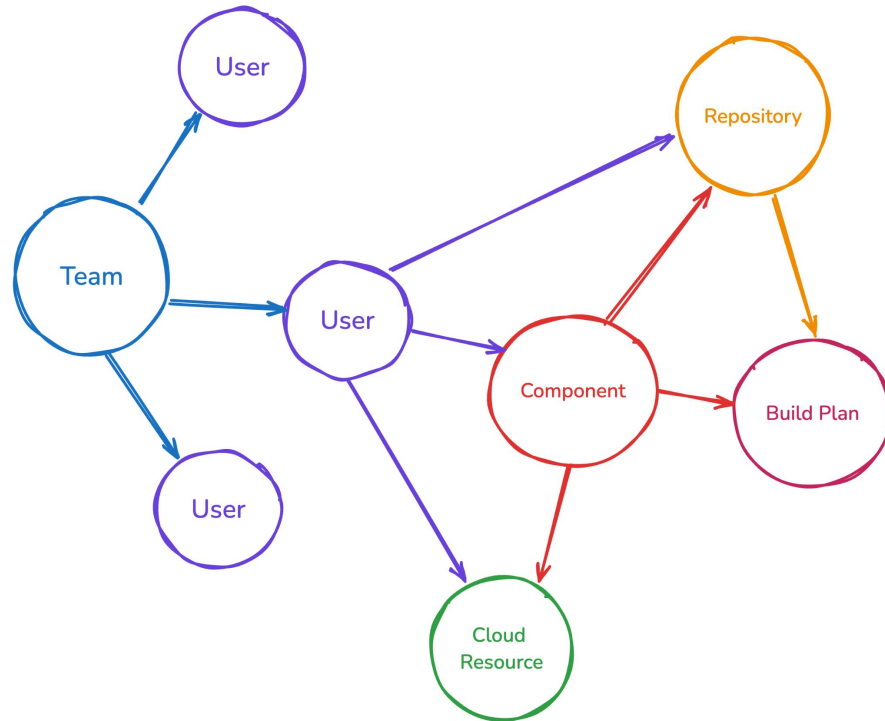


to the right **person** 



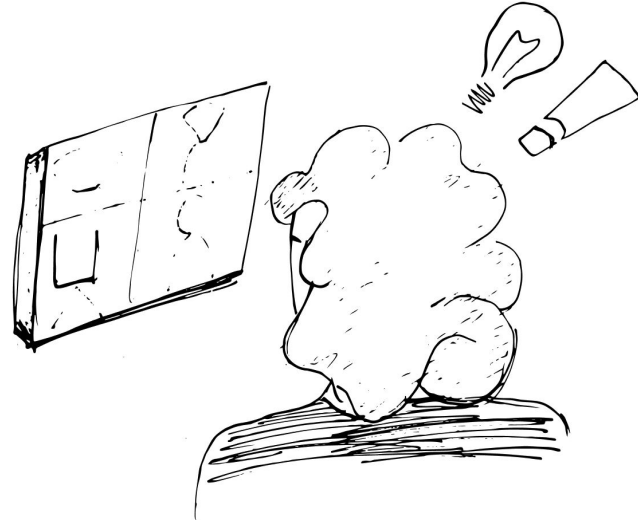
We leverage *context* to
intelligently *target* teams
through accurate attribution
and cohorting.

to the right **person** 

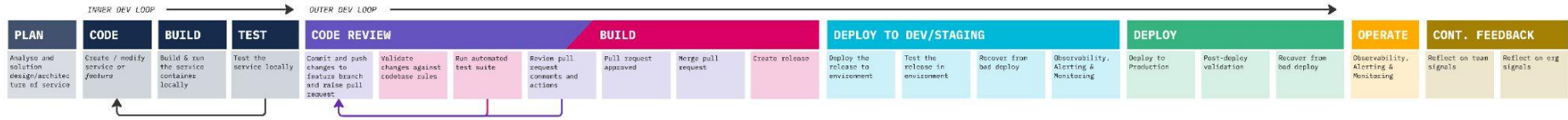


We *intervene* and bring
attention to issues as *early* as
possible.

at the right **time** 🕒

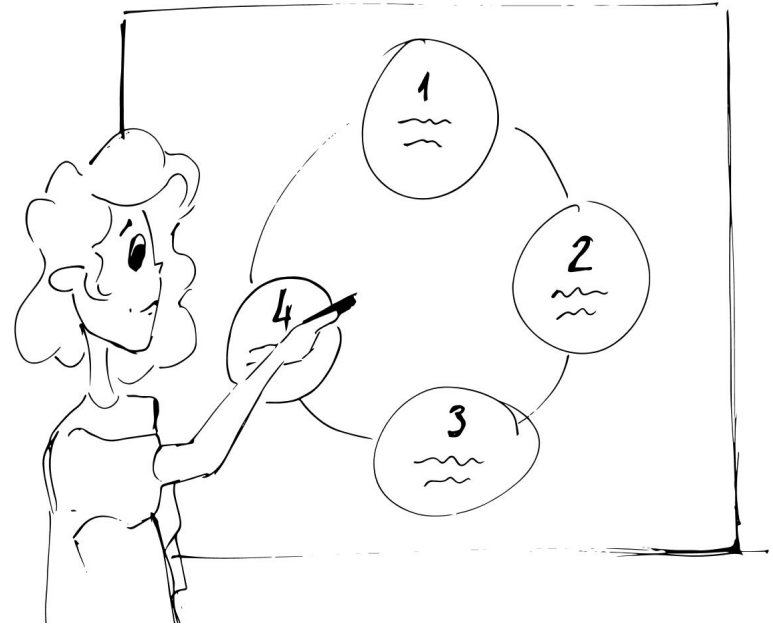


at the right **time** 

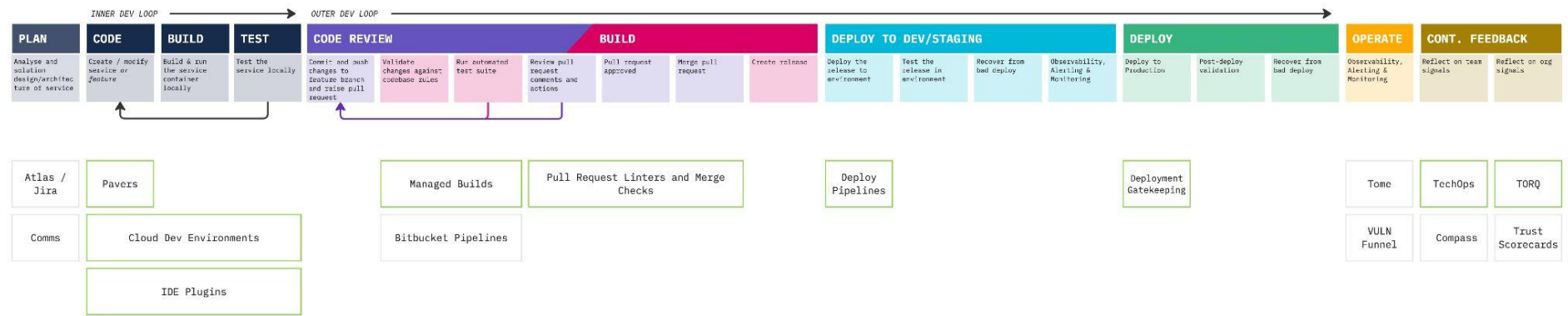


We *intervene* and bring attention to issues through as *non-disruptively* as possible.

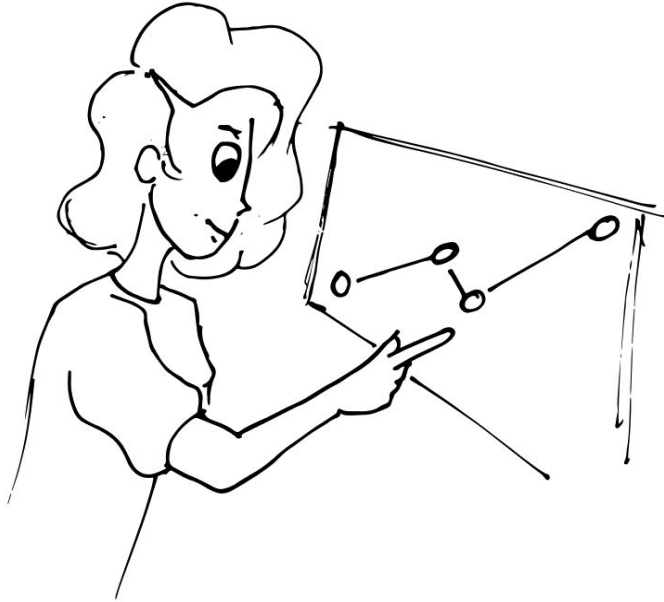
in the right **place** 



in the right **place**



to do the right thing 💪



We recommend the use of practices that are *correlated* with *positive* outcomes; teams adopt and see improvement over time.

to do the right thing 💪

Waiting Time for Review MEDIAN

Median of total time that the PR author is waiting for a reviewer's action. Includes pick up time from PR open to first review action.

🚨 **4.05** HOURS



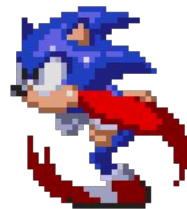
🔔 Recommendations

< 2 / 2 >

Use Slack bot to assign and nudge PR reviewers. [More info](#) 🔗

04

Project Sonic



Existing friction



Long feedback loops

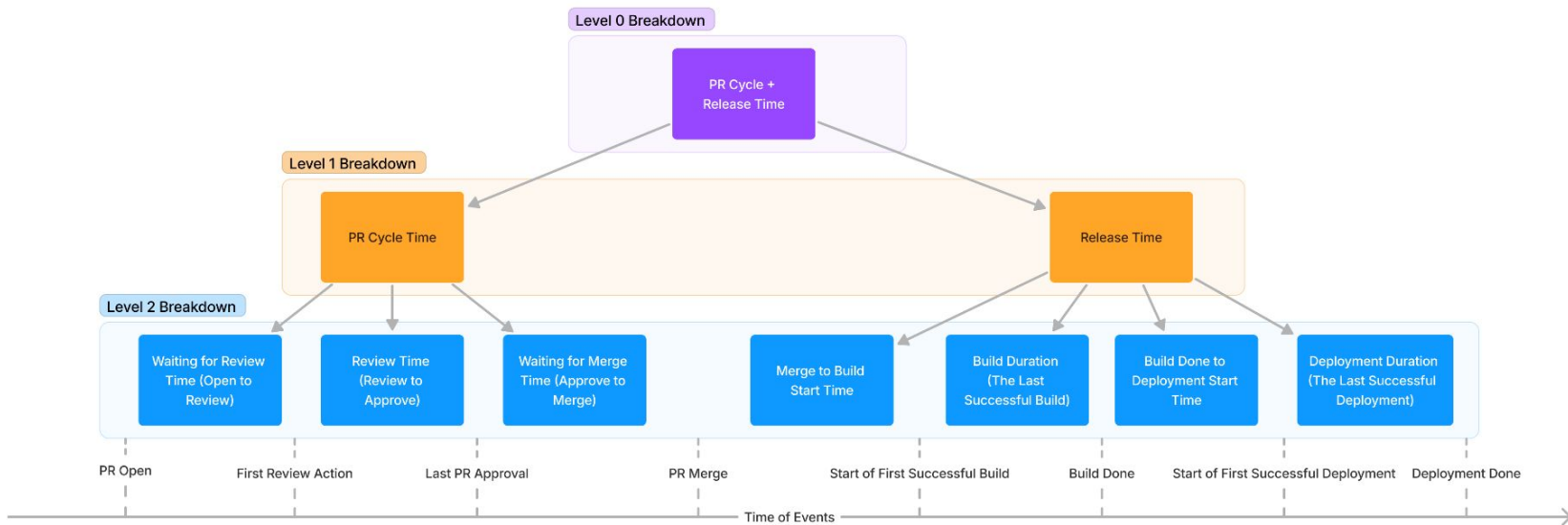
Delays between actions and feedback slow development creating friction in our ability to deliver customer value as fast as possible.



Poor visibility

Rudimentary developer metrics with little insight into bottlenecks in the software development lifecycle.

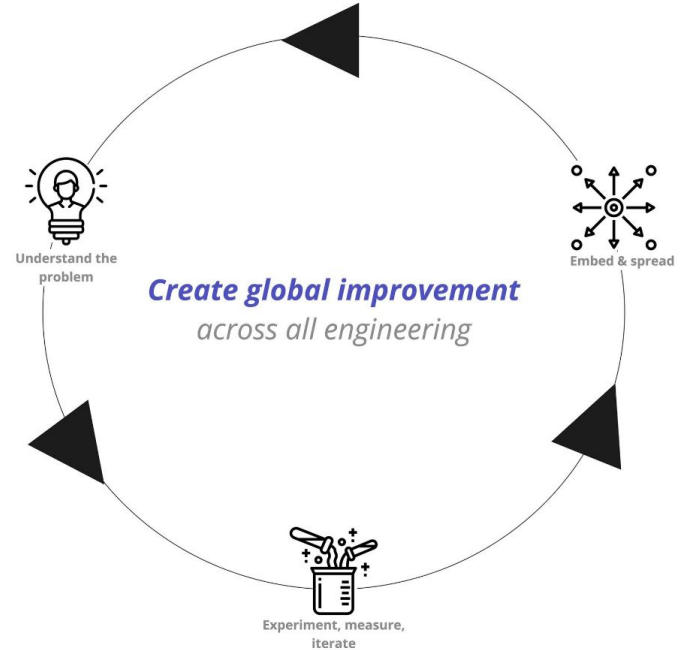
Develop vital signs for PR cycles



Identify PR practices



Identify patterns & opportunities



Better code review experience

The screenshot shows the Bitbucket web interface. On the left is a sidebar with navigation links: Source, Commits, Branches, Pull requests (selected), Pipelines, Deployments, Jira issues, Downloads, Snyk, Snyk security, and Repository settings. The main area displays a pull request for a repository named 'forge-pull-request-...'. A modal titled 'Giving meaningful peer feedbacks' is open in the center. It provides instructions on how to create a new, top-level comment and lists available labels: BLOCKING, question, issue, and todo. Each label has a description and a corresponding action. The modal also includes a 'Learn more' link and a 'Close' button. On the right side of the interface, there are sections for 'Request changes', 'Approve', '4 of 4 checks passed', '2 reviewers', '0 tasks', '0 Jira issues', 'Reports NEW', 'Pull Request Validation', 'Pull request best practices', 'Pull Request Guidelines', 'Review Checklist', and 'Conventional Comments'.

Bitbucket Your work Pull requests Repositories Projects People More [Create](#)

Search

JS forge-pull-request-... sashifqbal/s

Source
Commits
Branches
Pull requests
Pipelines
Deployments
Jira issues
Downloads
Snyk
Snyk security
Repository settings

NOISSUE: A
sashifqbal/s

Me

Describe

0 attachments

Security

Activity

SUNDAY

TUESDAY

Giving meaningful peer feedbacks

When creating a new, top-level comment (i.e., not a response to an existing comment) in a pull request, use the following format:

Format

```
1 <label>: <comment>
```

See below for available labels

label		Action
BLOCKING		
question:	question is should be used if you have a potential concern but are not quite sure if it's a problem. Asking the author for clarification or investigation can lead to a quick resolution. <pre>1 question: can we safely assume that this will always be defined? 2 3 If the auth call doesn't succeed, this may explode.</pre>	Address first Questions may block code review and should be prioritised
issue:	issue highlights a specific, functional problem with the code under review. This can be user-facing or behind the scenes. If you are not sure if a problem exists or not, consider leaving a question. <pre>1 issue: this endpoint doesn't follow the existing API conventions. 2 3 Should be camelCase not snake_case.</pre>	Address before merge Issues may require more time to resolve, and should be prioritised.
todo:	todo is a small, trivial, but necessary change. Distinguishing 'todo' from 'issue' helps direct the reader's attention to comments requiring more involvement. DO <pre>1 todo: update the readme to reflect this change to the setup flow</pre>	Address before merge

[Learn more](#) [Close](#)

11 + * Note: These are picked from here: <https://developer.atlassian.com/platform/spinnaker/default-pipelines/soaking/#st>

Michael Tokar 2024-09-17

Request changes **Approve** ...

4 of 4 checks passed

2 reviewers

0 tasks

0 Jira issues

Reports **NEW**

Pull Request Validation

Pull request best practices

For quick reference, these resources mirror Atlassian's [pull request best practices](#)

Pull Request Guidelines [Learn more](#)

Review Checklist [Learn more](#)

Conventional Comments [Learn more](#)

For support, reach out on Slack at [#practices-pullrequests](#)

What we achieved



Teams empowered with data insights

Quantitative signals used as diagnostics for teams to detect and understand their developer productivity hotspots.

Reduced cycle time by 51%

Indicator for impact of improving a team's software development lifecycle processes (e.g., code reviews)

Improved developer sentiment

The code review category in the DSAT survey is our highest sentiment rating, sitting at >80% *favorability*



Thank you!