

Design strategies for building safer platforms

Kat Fukui

Product Designer @ GitHub

💖 **Kat Fukui**

Twitter: @katfukui

GitHub: @katmeister

Product Designer on the
Community & Safety team at GitHub

💥 Comics

👩 Full-process design

💜 Internet communities





GitHub

myoctocat.com/build-your-octocat



@annafil

@rebelagentm

@lindseyb

@spinecone

@gallexi

@benbalter

@katmeister

@dmleong

@antn

GitHub's Community & Safety team
builds systems that empower people to
grow inclusive and healthy communities
around their projects, while discouraging
behavior that is destructive or threatens
personal safety.

Building user safety into the foundation of technology we create is **everyone's responsibility.**

Any platform or feature can,
and will, **be abused.**

“How could this feature be used
harm someone?”

If users don't feel safe on a platform, **they will leave.**

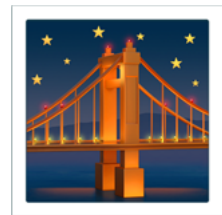
You don't need to be a designer to incorporate design strategies that get your team collaborating on safer, inclusive, more consensual features.



Understand with user stories



Build with safety principles



Bridge with acceptance criteria



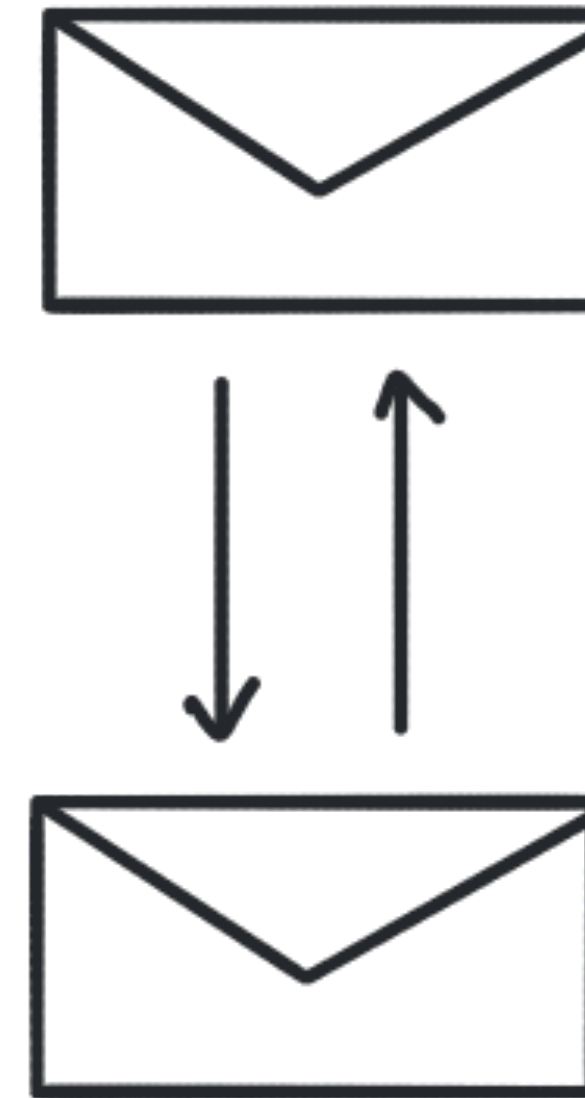
Scale with documentation

We'll learn about these design strategies through our lil friend, Blobbo!



Blobbo works on a B2B (blob to blob) platform.

They're working with their team to build a direct messaging feature.

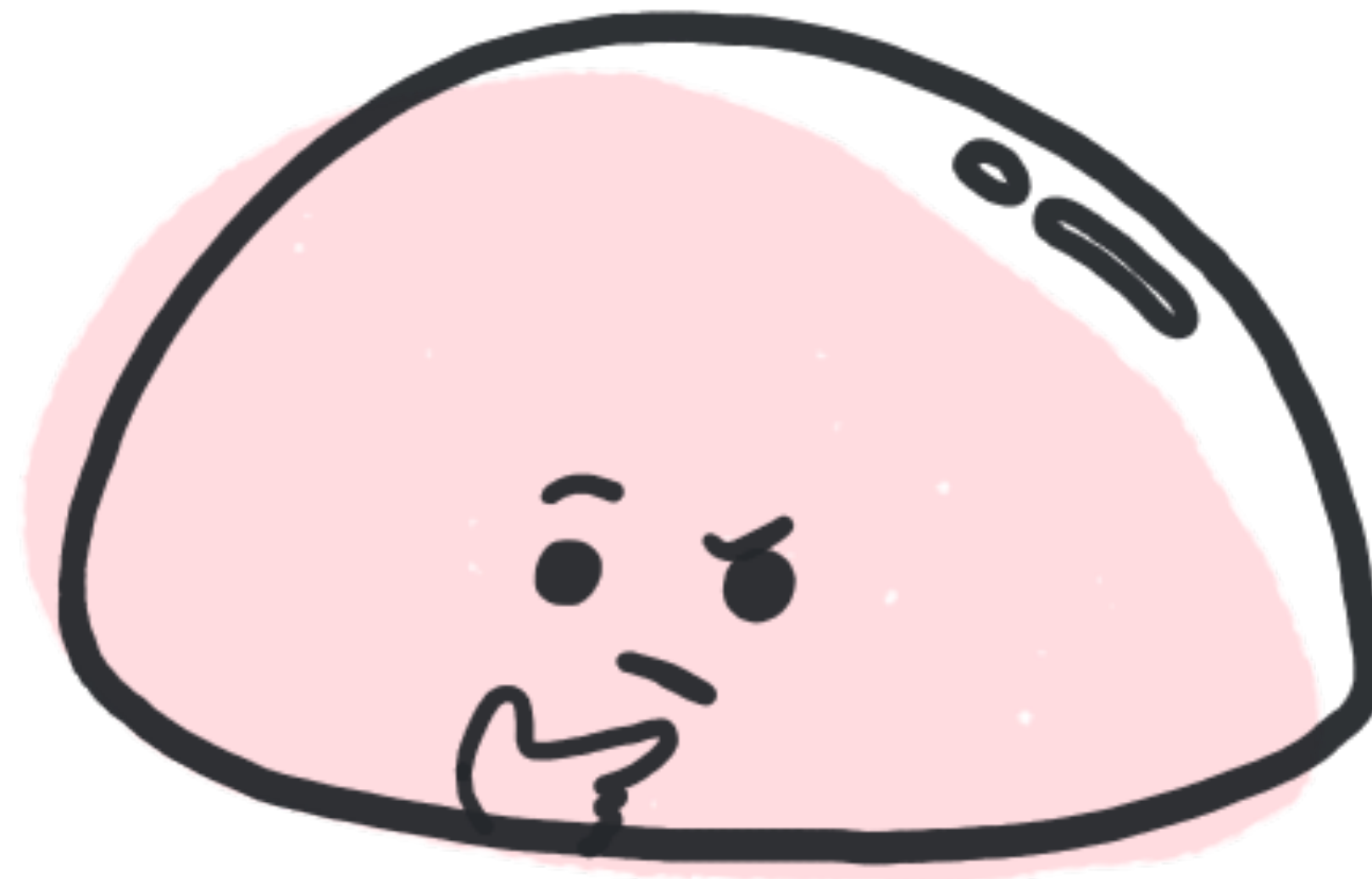


 **Understand with user stories**

“users want private spaces to chat”

“users want to stay on our platform without switching to other chat apps”

“users want to be more social with friends and strangers”



Create user stories of stress cases*
to better understand how your users
are feeling in scary situations—like
trying to escape abuse.

*From **Technically Wrong** by Sara Wachter-Boettcher

18% of respondents have personally experienced a negative interaction with another user in open source, but 50% have witnessed one between other people.

21% of people who experienced or witnessed a negative behavior said they stopped contributing to a project because of it, and 8% started working in private channels more often.

– ***Open Source Survey***

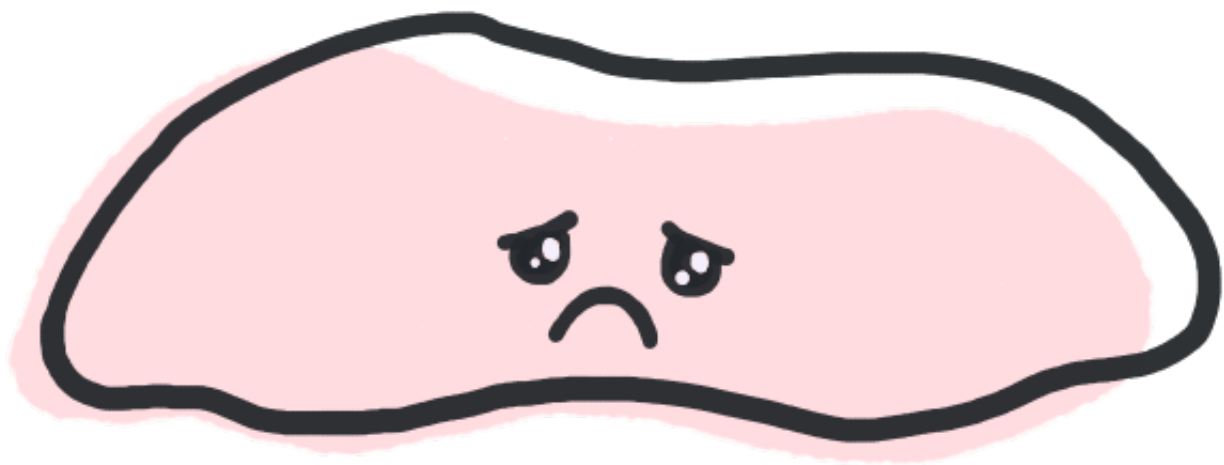
opensource-survey.org/2017

User stories are similar to user personas, but focuses more on their motivations and the context *outside the screen*—such as **mental health**.

What **problems** are they experiencing?

How are they **feeling**?

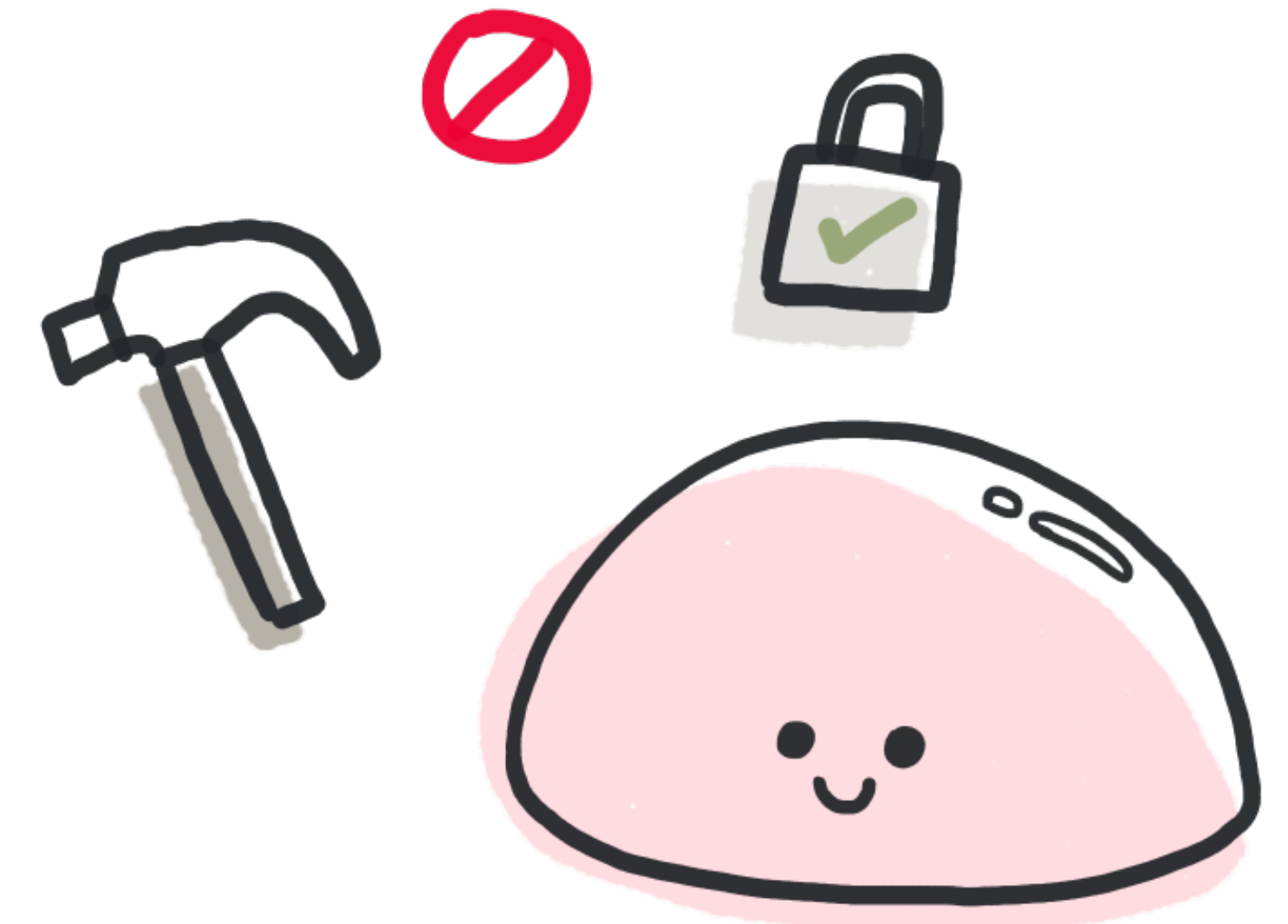
What does **success** look like?



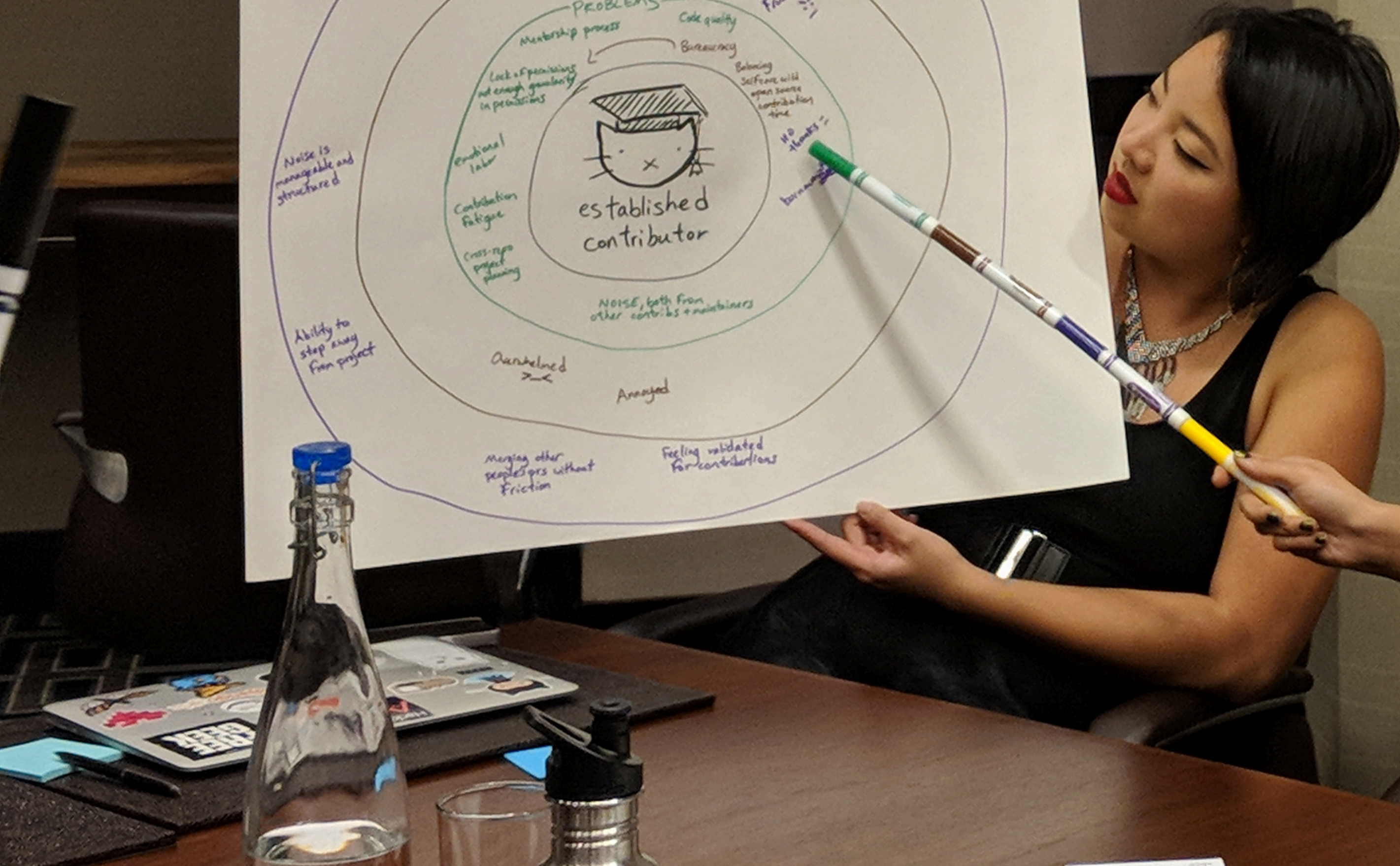
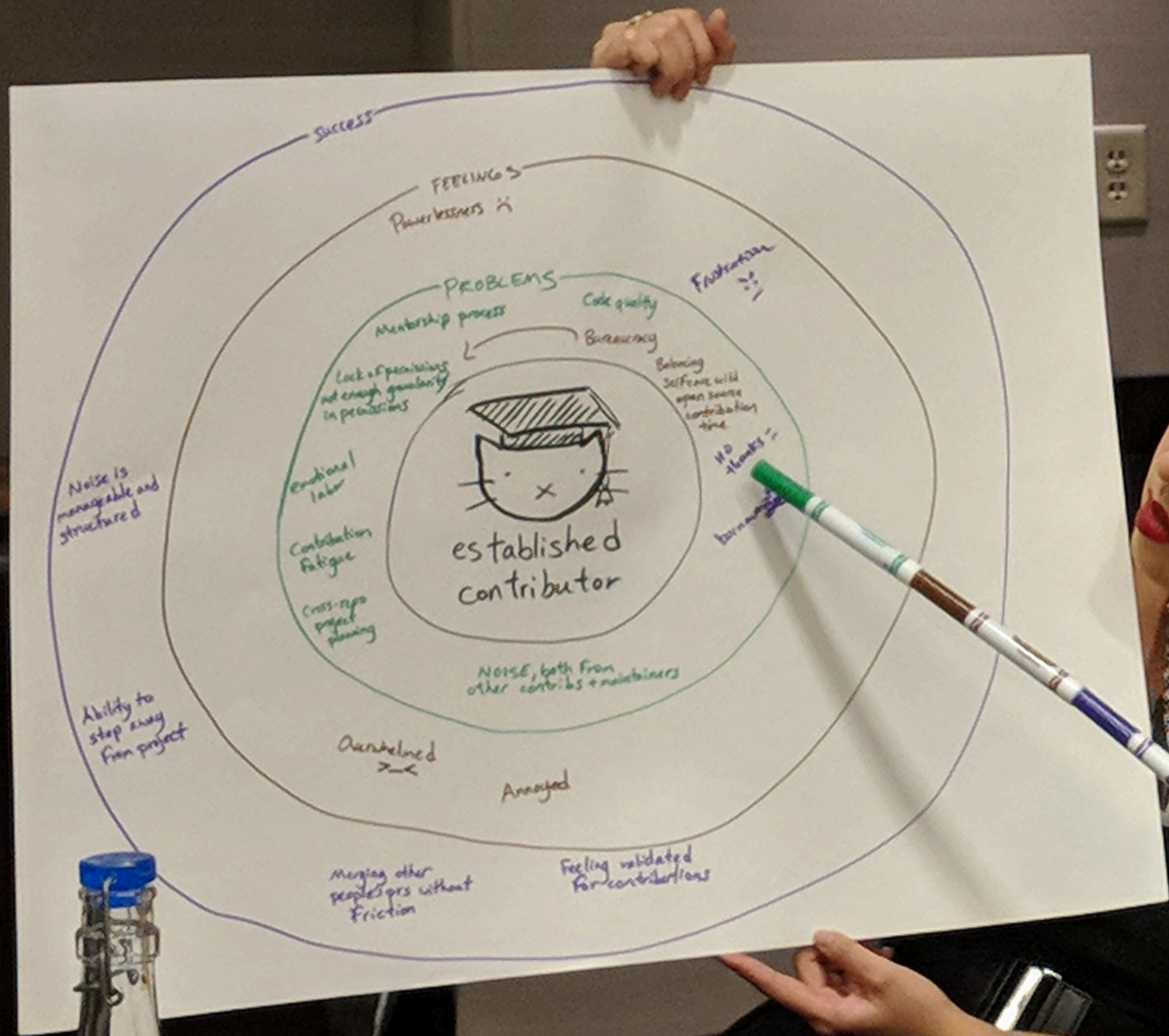
User is trying to escape harassing DM's from an abusive relationship



Abuser sends DMs from sock puppet accounts



✍️ ✨ Draw together!



User stories are great for aligning your team on the feature's vision, sharing specialized knowledge with other teams, and making quick decisions.

SEES HELP WANTED



DAY JOB FOR & NON-PROFITS



DAY JOB

FOR & NON-PROFITS

CREATES OWN PROJECT



CONTRIBUTOR → MAINTAINER
GAINS TRUST



AUTHOR

AUTHOR(S) WORK ON PROJECT & RELEASE

PROJECT GAINS POPULARITY

WORD SPREADS VIA MEETUPS, HN, G, etc.
MAINTENANCE WORK BEGINS.



MAINTAINER

- CODE RELEASES
CODE REVIEW
BUG FIXES
- COMMUNITY
NEW CONTRIBUTIONS
BUG REPORTS
MODERATION
- ISSUE TRIAGING
- DOCUMENTATION

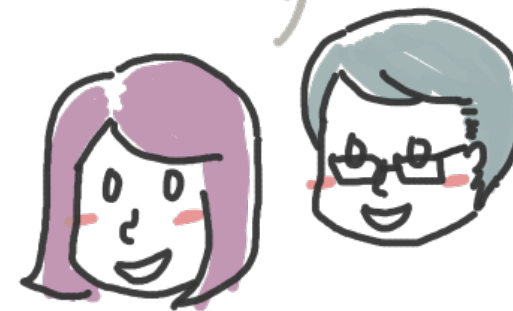


FIND HELP!!

- MORE CONTRIBUTORS
- CONVERT ACTIVE USERS TO MAINTAINERS

✓ IF SUCCESSFUL...

COMMUNITY IS HEALTHY!
- STRUCTURE PROJECT FOR SUCCESS



✗ IF UN-SUCCESSFUL...

OVERWHELMING TO MAINTAIN,
BURNOUT EMINENT 😞



MAINTAINER GIVES UP,
ABANDONS PROJECT.

PROJECT DIES or
COMMUNITY FORKS?

PROJECT HAND-OFF

- DON'T HAVE THE TIME
- COULDN'T FIND ENOUGH HELP

WORK ON OTHER STUFF

CHOICE A

CHOICE B







Build with safety principles

Define what **safety means specifically to your users** and create principles to guide you whenever making decisions.

They can be aspirational!

At GitHub, user safety means **ensuring everyone can participate** in communities and collaborate on code without risking their privacy or personal well being, regardless of their background and identity.

Starter principles:

-  Ask for consent
-  Encourage inclusive behavior, discourage destructive behavior
-  Minimize the impact of destructive content
-  Leave a papertrail

 **Ask for consent**

Consensual software is software that asks for the user's explicit permission to interact with them or their data.

Consensual software respects users' privacy and does not trick or coerce users into giving away permissions or data.

Consensual software asks for permission first, rather than begging for forgiveness later.

– ***Consensual Software***

consensualsoftware.com/what-is-consent

Asking for consent helps users **feel in control** of their experience on a platform.

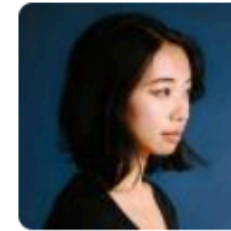
For actions that may **leak private information** (ex: their location) or **exploit a user's notifications**, ask for consent so users can opt-in to features or workflows.



throwawayblob wants
to message you.

Accept

Delete



feralkat invited you to collaborate

Accept invitation

Decline

 **Owners** of super-adventure will be able to see:

- Your public profile information
- **Certain activity** within this repository
- Country of request origin
- Your access level for this repository
- Your IP address

Is this user sending spam or malicious content?


[Block feralkat](#)



Consent-driven design also **helps users make more informed decisions** when interacting with other users, content, their personal data, etc.

  **katmeister** deleted a comment from **feralkat** on Apr 4



This comment was marked as disruptive content.


 Show comment

  **katmeister** locked as **off topic** and limited conversation to collaborators 14 days ago

  **katmeister** unlocked this conversation 14 days ago


atom / atom

 Watch 2,358


 Star 40,774

 Fork 7,559

 Code

 Issues

 Pull requests


 Insights

A user you've blocked has previously contributed to this repository.

[Collapse](#) ^

 evilscout

Visit your [blocked users](#) settings to turn off prior contributor warnings

 The hackable text editor <https://atom.io>

atom

editor

javascript

electron

windows

linux

macos



Atom

Free and open source text editor, brought to you by GitHub

San Francisco, CA <http://atom.io> atom@github.com

Verified



Repositories 271

People 301

Teams 1

We've verified that the organization **Atom** controls the domains:

- ✓ atom.io
- ✓ github.com

[Learn more about verified organizations](#)


Pinned repositories

atom

 The hackable text editor


JavaScript ★ 46.9k 🍴 10.3k

flight-manual.atom.io


 Documentation for Atom, generated by nanoc, hosted by GitHub Pages

CSS ★ 310 🍴 503

atom.io

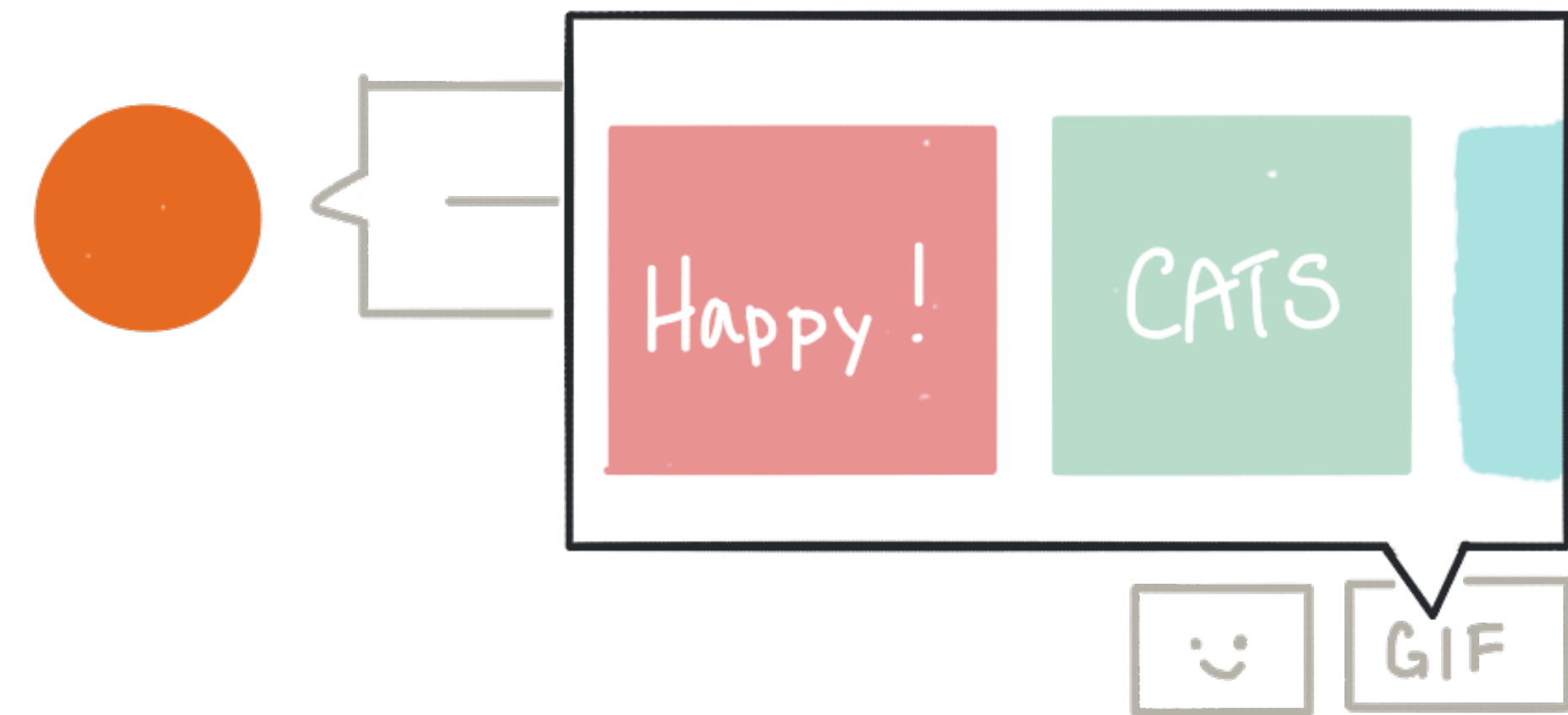
 A place for feedback on the atom.io website and package API

★ 63 🍴 42

 Encourage inclusive behavior,
discourage destructive behavior

Encourage behaviors that foster welcoming environments and helps participants of all backgrounds and identities to **be productive without risking personal safety.**

Ensure there's appropriate friction to discourage behavior that undermines productivity.



Encourage using trusted, "positive" GIFs



First-time contributor



Pick your reaction



fbdd

They're new to the community!



usagi commented 3 days ago

First-time contributor



Had some additons to documentation that might be helpful ✨

tetst #6

Edit New

Open LindseyB opened this issue 29 days ago · 3 comments



LindseyB commented 29 days ago • edited

Member - 😊 ...

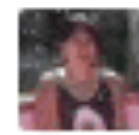
an edit



LindseyB commented 29 days ago

Member - 😊 ...

test



spinecone commented 29 days ago

- 😊 ...

COMMENTTTTTTTTTTTTTTTT



LindseyB commented 29 days ago

Member - 😊 ...

testing



Write Preview

AA B i “ < > @ 📌 ↶

Leave a comment

Assignees

No one—assign yours

Labels

None yet

Projects

None yet

Milestone

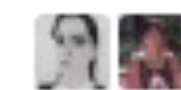
No milestone

Notifications

🔔 Unsubscrib

You're receiving notif because you authored thread.

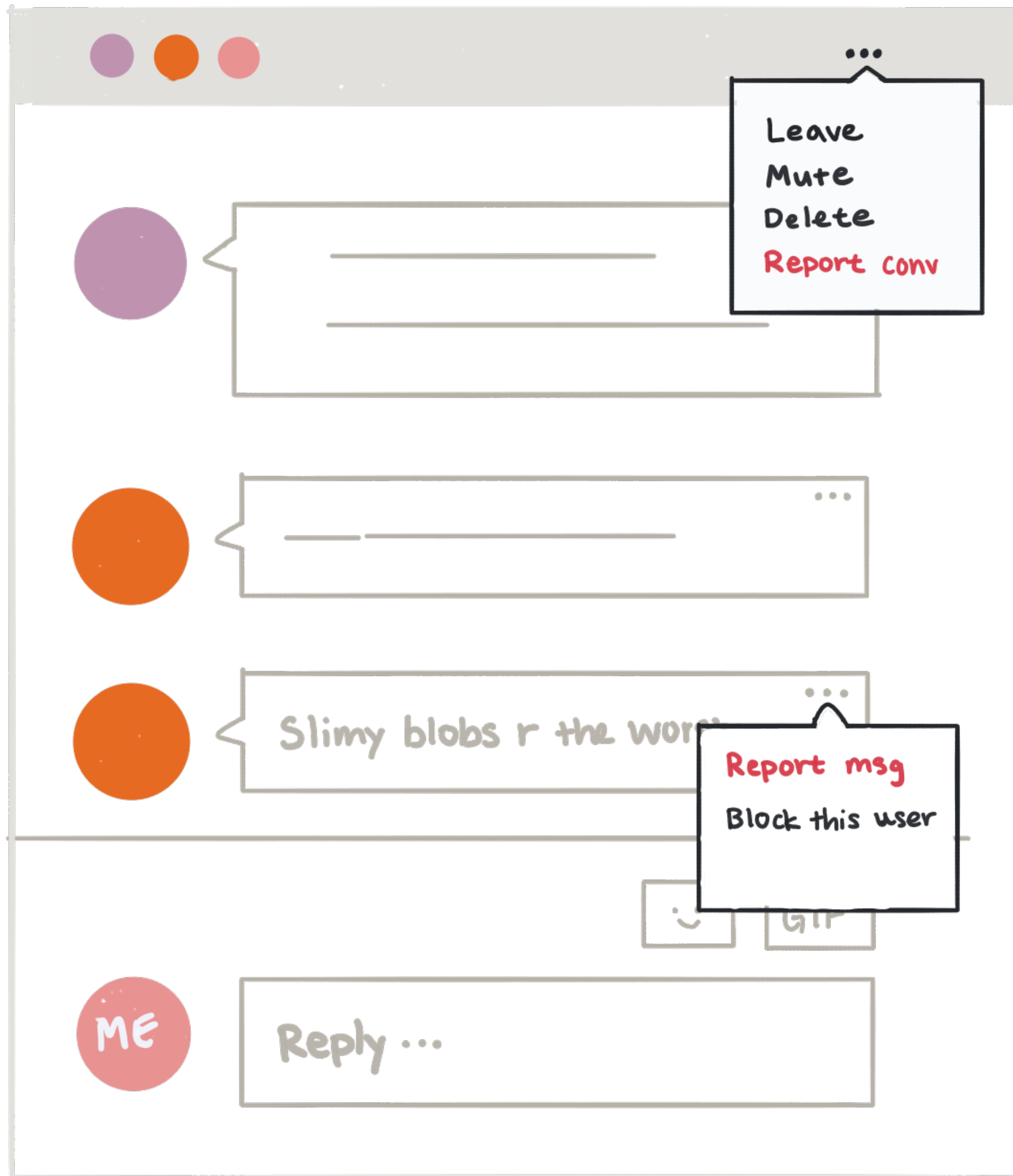
2 participants



! Minimize the impact of
destructive behavior or content

Because any feature can be abused, include mechanisms to deal with **unproductive content or users** in the worst case scenario.

Design a **tier of tools** ranging from least to most “nuclear” options.



Tools on the conversation level



Tools on the content and user level

tion to collaborators 14 days

B *i* “ ‹› ↻

- Copy link
- Quote reply
- Open new issue
- Edit
- Hide
- Delete
- Report abuse
- Block user

Temporary interaction limits

Temporarily restrict which users can interact with all of your organization's **public repositories** (comment, open issues, or create pull requests) for a **24-hour** period. This may be used to force a "cool-down" period during heated discussions.

Limit to existing users

Users that have recently created their account will be unable to interact with public repositories.

Enable

Limit to prior contributors

Users that have not previously committed to a repository's default branch will be unable to interact with the repository.

Enable

Limit to repository collaborators

 Enabled with 1 day remaining.

Users that are not collaborators will not be able to interact with the repository.

Disable





Leave a papertrail

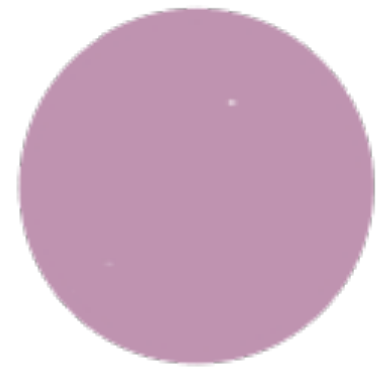
Add internal **audit logs**, especially to help your Support team when they investigate abuse reports.

Leave **timeline entries** in the UI to add context, especially in collaborative workflows.



edited

● **blahb** deleted their comment 5 min ago








  **katmeister** deleted a comment from **feralkat** on Apr 4

feralkat commented 26 days ago • edited by **katmeister** ▾



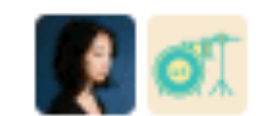
you're a horrible cat!

  **katmeister** locked as **off topic** and limited conversation to collaborators 13 days ago

 **Unsubscr**

You're receiving notific
because you authored

2 participants



 **Lock conversation**



Bridge with acceptance criteria

Acceptance criteria bridges the gap between design and engineering. By taking user stories and writing conditions for the features' functionality, the entire team understands what workflows will be built.

Blocking Users

Feature: Blocking Users

Description: Blocking is an affirmative step a user **or** organization takes against another user to express that other user's content.

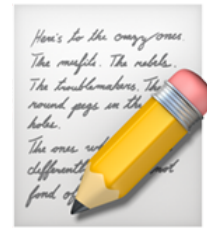
Background:

Given I am a user that wants to block another user
When I block that user

Scenario: Blocked user has starred my profile
Then their star is removed from my profile

Scenario: Issues

Then they cannot **open**, close, re-**open**, **or** assign **Issues**
And they cannot edit **or** delete comments
And they cannot mark it as a duplicate



Scale with documentation



Community & Safety design guidelines

Our work tends to touch many parts of the GitHub platform where a community exists. The goal of these guidelines is to help cross-functions build in areas we've thought critically about—user blocking, reporting, content moderation, and comments.

Getting started

- [Our design principles](#)
- [Research](#)
 - [User stories](#)
 - [Community moderation](#)

Styleguides


- [Kebab menu](#)
- [Reason dropdown](#)



Community & Safety design principles

Building safety into the foundation of our product development processes is every team's responsibility. Anything we build must be a trusting experience for our users. We have a moral obligation to ensure we are scaling our platform to be **inclusive** and welcoming. GitHub influences collaboration, behavior, and safety across online communities and we should lead by example. This document outlines the principles that Community & Safety follows when making decisions, and provides a guide for product and engineering teams to incorporate into their own processes.

- [Definitions](#)
- [Principles](#)
 - [Encourage inclusive behavior, discourage destructive behavior](#)
 - [Enforce baseline platform expectations and empower community governance](#)
 - [Any feature can be abused](#)
 - [Minimize the impact of destructive content/users](#)
 - [Ask for consent](#)
 - [Context is !\[\]\(e40bb48ad1470e3a14017c64c5673877_img.jpg\)](#)
 - [Leave a papertrail](#)
 - [Accessibility](#)
- [Development guidelines](#)
- [Internal resources and stakeholders](#)
- [Recommended reading](#)

 **katmeister** add a clarifying sentence

5f777b8 on Oct 5

1 contributor

65 lines (47 sloc) | 3.92 KB

Raw

Blame

History



Kebab menu

The kebab menu provides more secondary options to a user in a dropdown. This guide aims to help you implement a kebab menu consistently and responsibly. 🍡

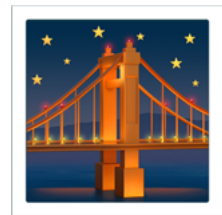
- [A brief history](#)
- [Should I use the kebab menu?](#)
- [Implementing the kebab menu](#)
 - [Group by type](#)
 - [Order from least to most destructive](#)
 - [Visual indicators](#)



Understand with user stories



Build with safety principles



Bridge with acceptance criteria



Scale with documentation

Let's continue to prioritize humans in the technology we create, and work towards **standardizing an open framework** for the industry to build safer platforms.



Thank you!

Twitter: @katfukui

GitHub: @katmeister

Email: kat@github.com