

How do we keep devs safe online?

Community Safety at Scale

DENAE FORD ROBINSON
It Never Work In Theory: Live!
April 2022

 @denaefordrobin
<https://denaeford.me>



Jeff Atwood ✅
@codinghorror

...

“being online did not make most people more aggressive or hostile; rather, it allowed a small number of aggressive people to attack a much larger set of victims”



theatlantic.com

Why the Past 10 Years of American Life Have Been Uniquely Stupid
It's not just a phase.

Latest

The Atlantic

First, the dart guns of social media give more power to trolls and provocateurs while silencing good citizens. Research by the political scientists Alexander Bor and Michael Bang Petersen [found](#) that a small subset of people on social-media platforms are highly concerned with gaining status and are willing to use aggression to do so. They admit that in their online discussions they often curse, make fun of their opponents, and get blocked by other users or reported for inappropriate comments. Across eight studies, Bor and Petersen found that being online did not make most people more aggressive or hostile; rather, it allowed a small number of aggressive people to attack a much larger set of victims. Even a small number of jerks were able to dominate discussion forums, Bor and Petersen found, because nonjerks are easily turned off from online discussions of politics. Additional research finds that women and Black people are harassed disproportionately, so the digital public square is less welcoming to their voices.

(Haidt, 2022)



Jeff Atwood ✅
@codinghorror

...

“being online did not make most people more aggressive or hostile; rather, it allowed a small number of aggressive people to attack a much larger set of victims”

Latest

The Atlantic



theatlantic.com

Why the Past 10 Years of American Life Have Been Uniquely Stupid
It's not just a phase.

use aggression to do so. They admit that in their online discussions they often curse, make fun of their opponents, and get blocked by other users or reported for inappropriate comments. Across eight studies, Bor and Petersen found that being online did not make most people more aggressive or hostile; rather, it allowed a small number of aggressive people to attack a much larger set of victims. Even a small number of jerks were able to dominate discussion forums, Bor and Petersen found, because nonjerks are easily turned off from online discussions of politics. Additional research finds that women and Black people are harassed disproportionately, so the digital public square is less welcoming to their voices.

(Duggan, 2017)
(Blackwell, 2017)



Jeff Atwood ✅
@codinghorror

Little Window

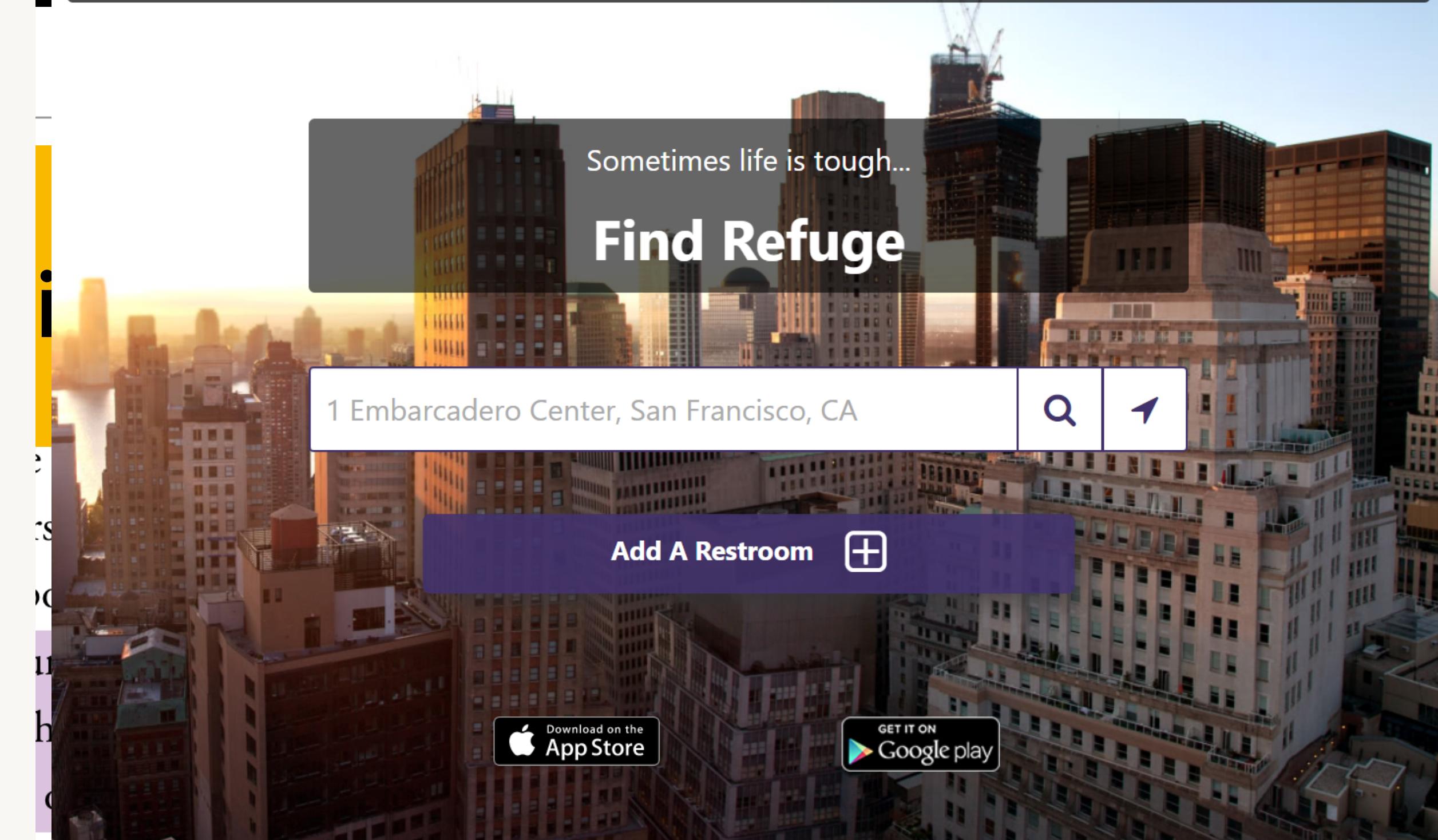
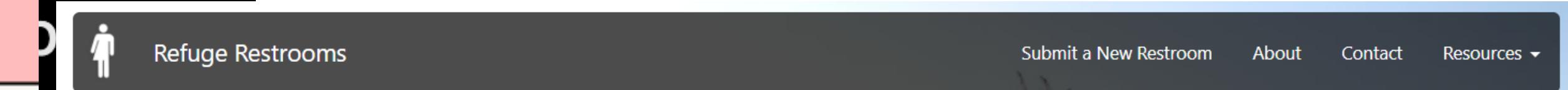


Hello 🌻, I am Little Window - a chat bot to help you get the information you need.



Type here...

Submit



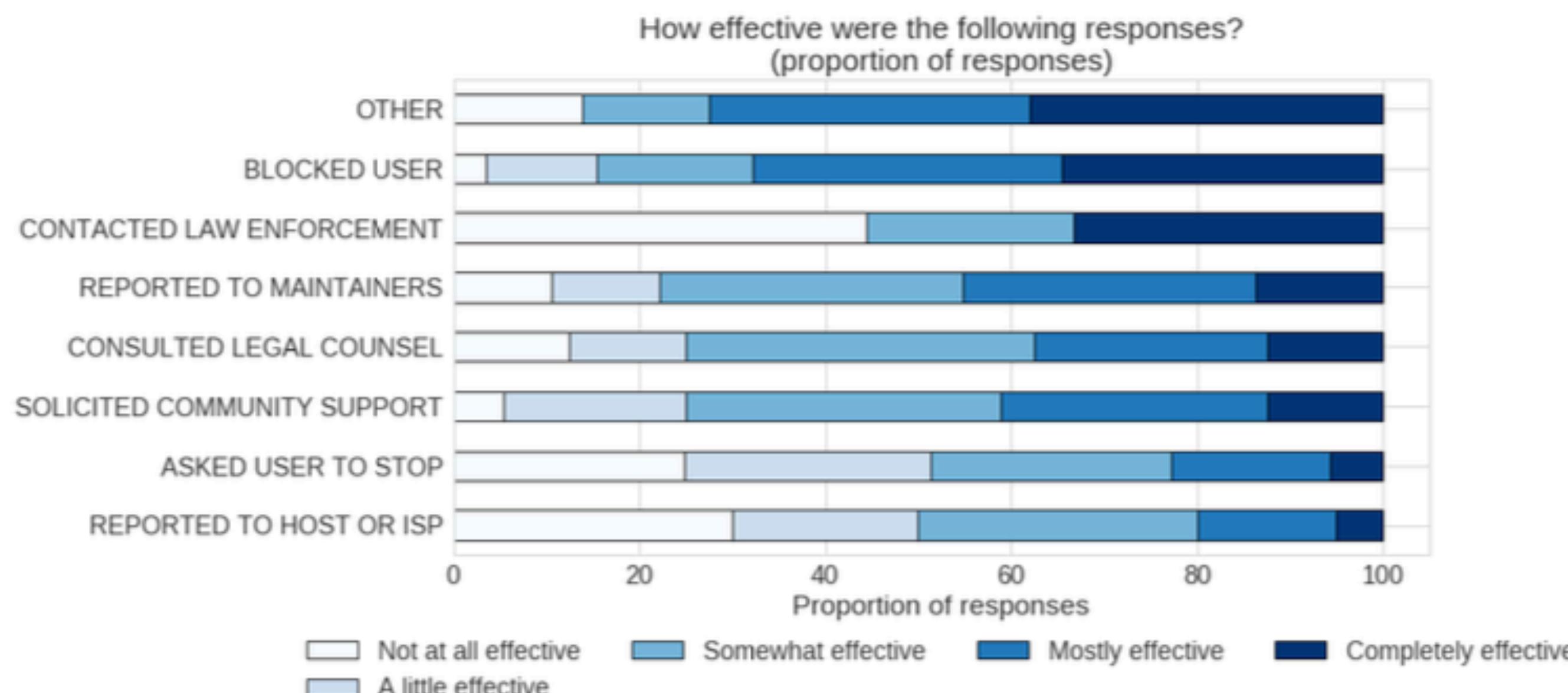
ums, Bor and Petersen found, because nonjerks are easily turned off from line discussions of politics. Additional research finds that women and Black people are harassed disproportionately, so the digital public square is less welcoming to their voices.

Safety Matters + This is Why

- Psychological Safety → Quality of Work → Quality of Product
- How do we protect contributors from inappropriate behavior?
- How do we keep people safe and feeling protected?

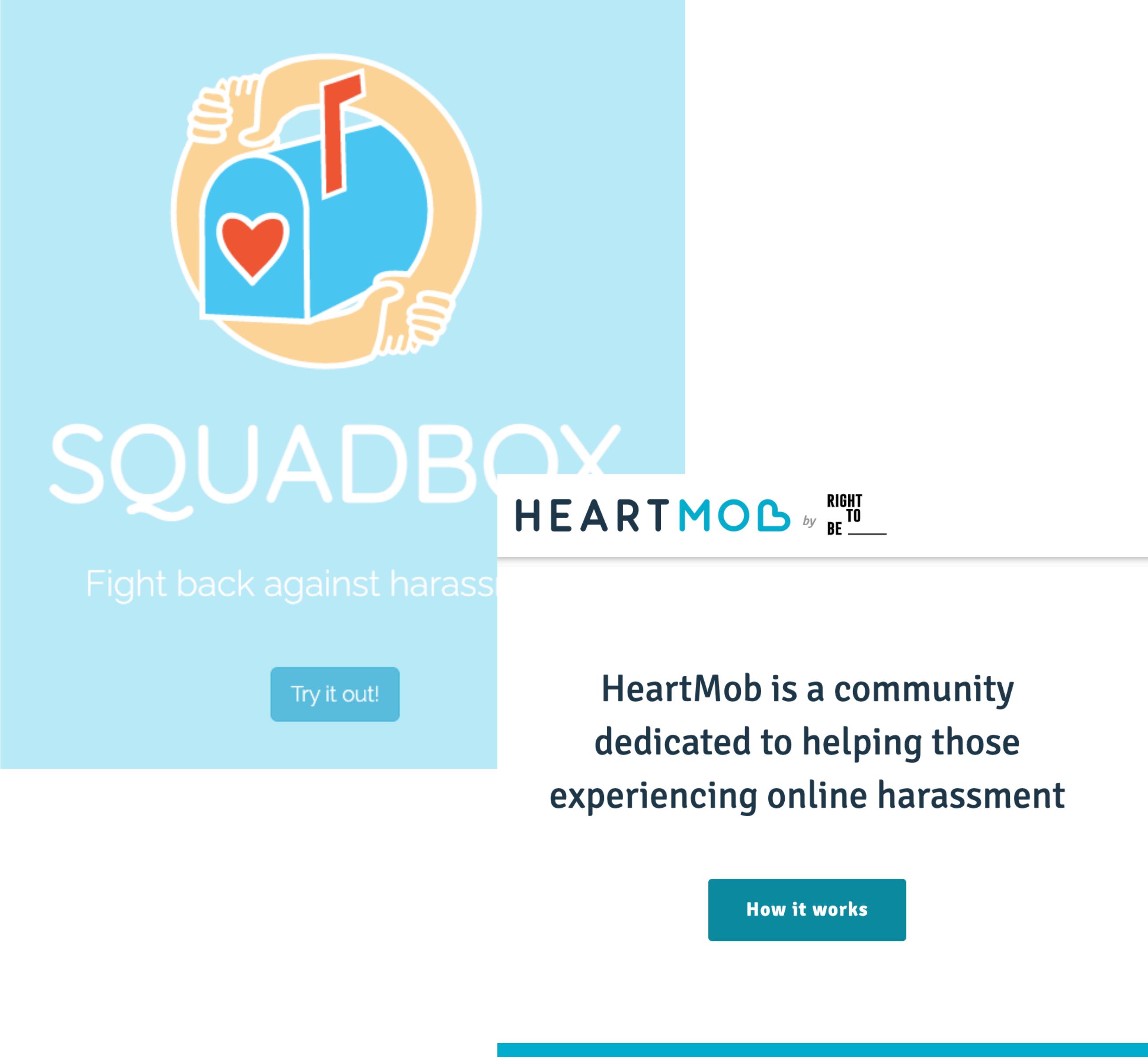
62%

(Global Wellness Survey, 2016)



(Geiger, 2017)

Empirical Research on Safety



The screenshot shows the homepage of the Squadbox website. At the top is a large logo featuring two hands holding a blue shield with a red heart in the center. Below the logo, the word "SQUADBOX" is written in large white letters. To the right of "SQUADBOX" is the "HEARTMOB" logo, which includes the text "by RIGHT TO BE". Below the main title, there is a sub-headline "Fight back against harassment". A blue button labeled "Try it out!" is located on the left side. On the right side, there is a text block: "HeartMob is a community dedicated to helping those experiencing online harassment". Below this text is another blue button labeled "How it works".

- Amy Zhang et al.'s work on promoting safety through Squadbox (squadbox.org)
- iHeartMob.org - a community dedicated to offering support to people experiencing online harassment
- What does this look like for developers as they engage in communities?

What has already been successful

- Temporary interaction limits
- Reporting mechanisms for maintainers
- Advanced blocking features
- Reporting features

The screenshot shows a modal window titled "Temporary interaction limits". It explains that users can temporarily restrict interactions for a 24-hour period. Three options are listed: "Limit to existing users", "Limit to prior contributors", and "Limit to repository". The "Limit to repository" option is checked and highlighted in green, with the status "Enabled with AI". Below this, a "Report content" section is shown, which states: "This content violates @octo-org's Code of Conduct and should be submitted to the maintainers of this repository for review." A "Choose a reason for reporting this content" dropdown is present, along with links to learn more about requesting moderation and reporting abuse. Finally, two buttons are at the bottom: "Report to repository admins" and "Report abuse to GitHub Support".

Temporary interaction limits

Temporarily restrict which users can interact with your repository (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 the repository.

Limit to prior contributors
Users that have not previously committed to the repository's master branch will be unable to interact with the repository.

Limit to repository
Enabled with AI

Report content

This content violates @octo-org's Code of Conduct and should be submitted to the maintainers of this repository for review.

Contact GitHub Support about this user's spammy or abusive behavior.

Choose a reason for reporting this content

Choose a reason

Learn more about requesting that maintainers moderate content.

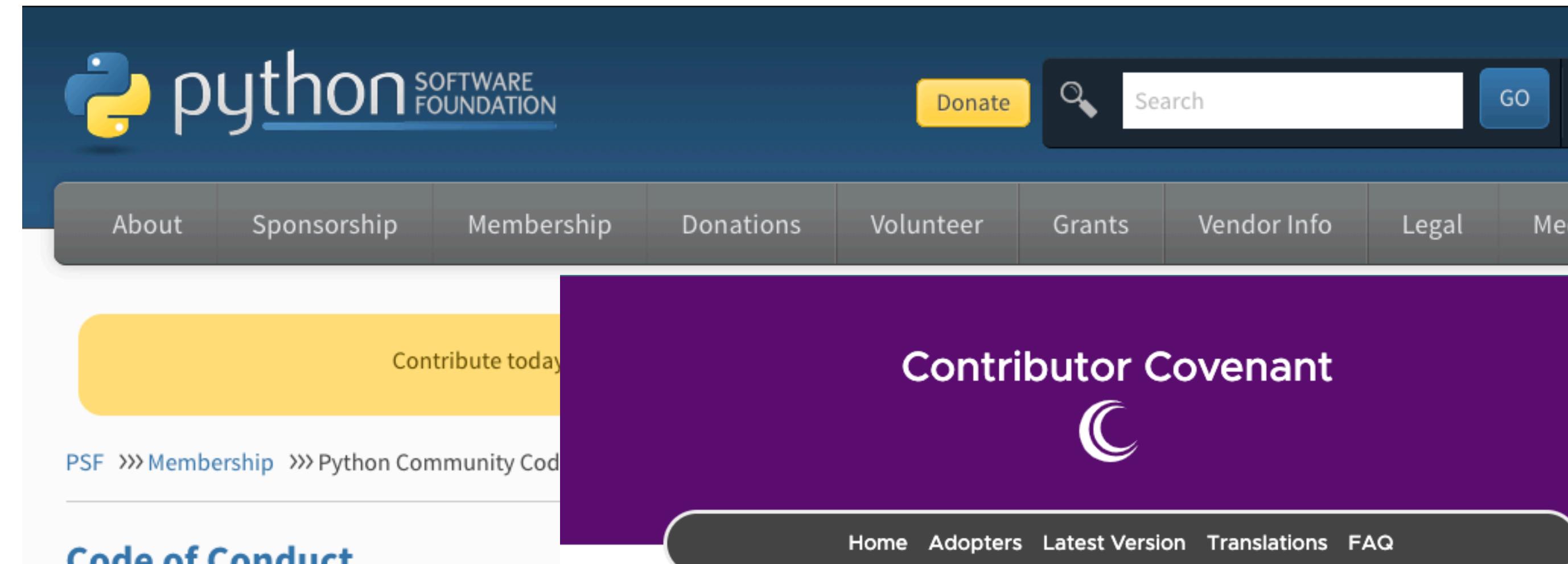
Learn more about reporting abuse to GitHub Support

Report to repository admins

Report abuse to GitHub Support

Code of Conduct Research

- Coraline Ada Ehmke's Contributor Covenant: <https://www.contributor-covenant.org/>
- Research on 52K repos->3800 code of conduct conversations. After removing noise + machine commits...
 - 32% of were centered around the creation
 - 16% of on moderating behavior

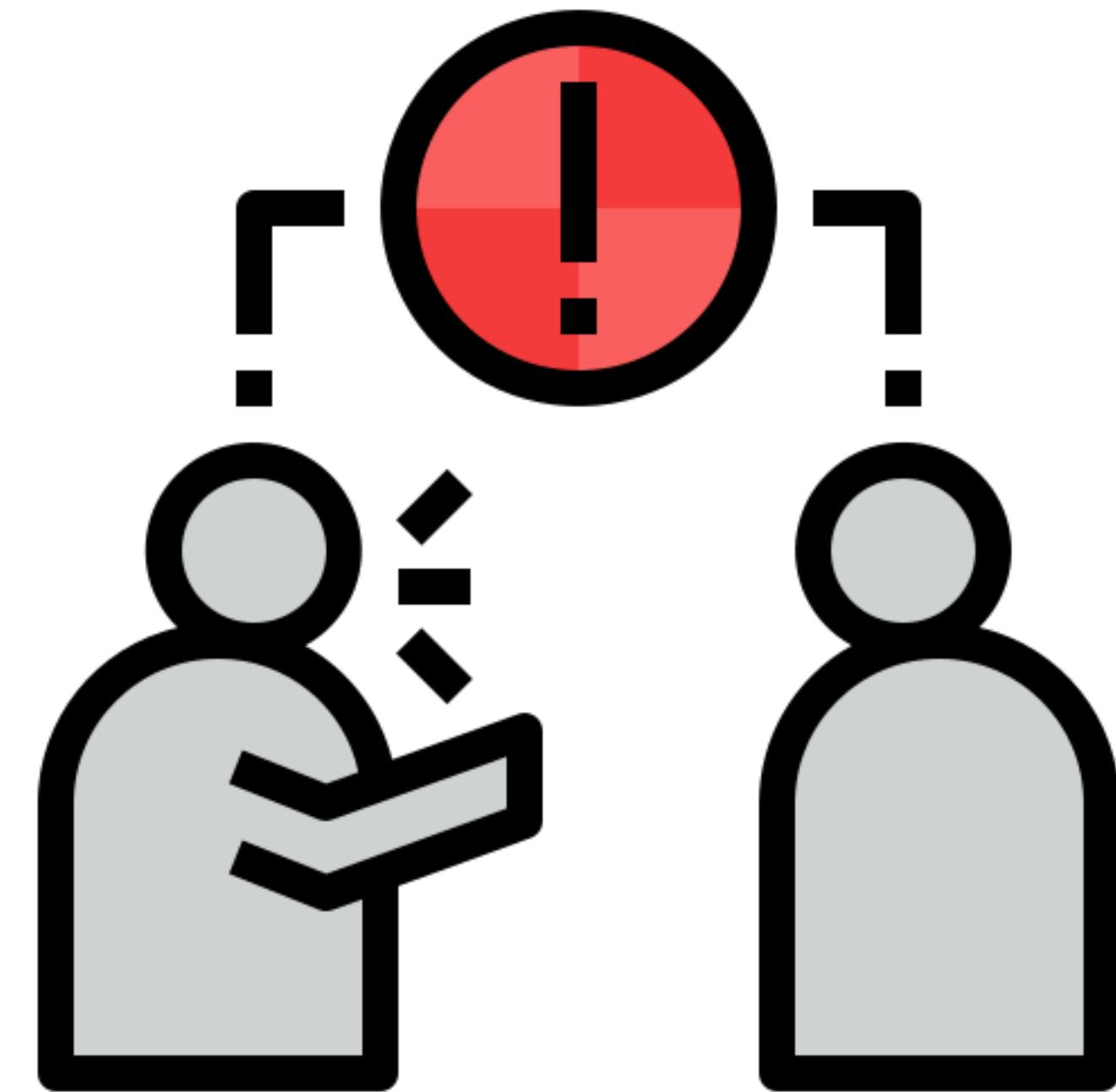


Opportunities to Explore

Encouraging Safe Communities Implicitly and Explicitly



Promote desired behaviors by highlighting



Sanctions to discourage undesired behaviors

We need human+automated approaches to check how a code of conduct can be enforced.

**What can you do as a
developer today?**

Community Safety 101



- If you care about it, its measured
- If it's important, there is a person accountable for it
- If there's accountability for it, progress can now be reasonably tracked

Where the action is happening

And where you can join forces

THE LINUX FOUNDATION PROJECTS

CHAOS About Community Events Software Metrics DEI Event Badging Participate

Psychological Safety

Question: To what extent do community members feel safe within a community, including adding contributions, influencing change, bringing their authentic selves, and generally participating in the project?

Description

A key component that is integral to the goal of community inclusion, is psychological safety. Psychological safety can be defined in several different stages, according to the [Center for Creative Leadership](#):

- Stage 1 – Inclusion Safety: Inclusion safety satisfies the basic human need to connect and belong. In this stage, you feel safe to be yourself and are accepted for who you are, including your unique attributes and defining characteristics.
- Stage 2 – Learner Safety: Learner safety satisfies the need to learn and grow. In this stage, you feel safe to exchange in the learning process, by asking questions, giving and receiving feedback, experimenting, and making mistakes.
- Stage 3 – Contributor Safety: Contributor safety satisfies the need to make a difference. You feel safe to use your skills and abilities to make a meaningful contribution.
- Stage 4 – Challenger Safety: Challenger safety satisfies the need to make things better. You feel safe to speak up and challenge the status quo when you think there's an opportunity to change or improve.

*Looking for people interested in testing out interventions at scale

<https://chaoss.community/>

What your community must have today

1. See if your repositories have codes of conduct
2. Propose your project adopts the contributor covenant
3. Be transparent about the safety procedures in place



SIGHT Community Framework

Safety
Inclusion
Growth
Honor
Trust

Thank you

What your community must have today

1. See if your repositories have codes of conduct
2. Propose your project adopts the contributor covenant
3. Be transparent about the safety procedures in place



SIGHT Community Framework

Safety
Inclusion
Growth
Honor
Trust



@denaefordrobin

<https://denaeford.me>