



Trolls: The Cost of Doing Nothing

Chris Priebe

CEO, Two Hat Security

A photograph of a young boy with short brown hair and glasses, wearing a light blue long-sleeved shirt. He is sitting at a desk, looking towards the camera while his hands are on a computer keyboard. The desk is cluttered with a large, white, boxy computer system. To the left of the keyboard is a large, white, printer-like device with a paper tray. The computer monitor is on the right side of the desk, displaying some text. The background is a plain wall.

“hello world”



hi game developers!

hi game developers!

hi game developers!

300M+

5card Flap15

Stuff

Blue

dude

Pen50gi

Michael56292

Kooldude233

hi game developers!

Babycrier

Tigerwolf





Deadly to @Lesdoggg 7m
 You just made a shitty game nobody liked. That's it. Nobody will care when you die.

Deadly to @Lesdoggg 7m
 I hope you enjoy your last moments alive on this earth. You did nothing worthwhile with your life.

Deadly to @Lesdoggg 8m
 If you have any kids, they're going to die too. I don't give a fuck. They'll grow up to be feminists anyway.

Deadly to @Lesdoggg 8m
 Your mutilated corpse will be on the front page of Jezebel tomorrow and there isn't jack shit you can do about it.

Deadly to @Lesdoggg 9m
 How's that for terrifying you stuck up cunt? I'm sick of you fucking feminist asshats.

Deadly to @Lesdoggg 9m
 I'm going to rape your filthy ass until you bleed, then choke you to death with your husband's tiny Asian penis

Deadly to @Lesdoggg 10m
 I've got a K-Bar and I'm coming to your house so I can shove it up your ugly feminist cunt.

Deadly to @Lesdoggg 11m
 Guess what bitch? I now know where you live. You and Frank live at _____

Deadly to @Lesdoggg 7m
 If I meet you for real I'd anal rape you with a huge chainsaw. Stop being the bitch you are.

Reply Retweet Favorite More

Jason Paulsen @jasonpaulsen
 Follow

@JulieDiCaro One of the Blackhawks player should beat you to death with their hockey stick like the WHORE you are. CUNT

Feb 10

@Lesdoggg SHES A FILTHY SLAVE AND A SCUMBAG WOMAN SHE SHOULDN'T BE ALLOWED TO BREED HER HORRIBLE KIND!!

Retweet Favorite

Leslie Jones @Lesdoggg
 Follow

Exposing

@Lesdoggg big lipped coon
 big lipped coon
 big lipped coon
 big lipped coon
 big lipped coon
 big lipped coon
 big lipped coon


7/18/16, 7:18 PM


Reply Retweet Favorite More


Lily Allen quits Twitter after trolls attack her over son's death


"My timeline is full of the most disgusting, sexist, misogynistic, racist shit." (84KB) ▼





 Lily Allen · 3h
[redacted] when I lay in a hospital bed with my deceased son stuck between my legs halfway out of my body for 10 hours.
57 replies · 11 retweets · 54 likes

 @lilyallen · 3h
[redacted] not to be a [redacted] but I very highly doubt it was 10 hours
14 replies · 1 retweet · 71 likes

 Lily Allen · 3h
[redacted] you are a [redacted] And it was. I was there, you weren't.
15 replies · 8 retweets · 106 likes

 [redacted] · 3h
[redacted] maybe if you didn't pump your body full of drugs you wouldn't have miscarried
19 replies · 13 retweets · 39 likes

 Lily Allen · 2h
[redacted] I didn't miscarry, I went into early labour and by son died from his chord wrapped round his neck.
62 replies · 10 retweets · 72 likes

 Dependable Cinema
Follow

@lilyallen The baby knew you were going to be such a horrible mother so it decided to go out on its own term

Twitter Inc

NYSE: TWTR - Feb 10, 4:33 PM EST

15.58 USD **↓0.83 (5.06%)**

1 day

5 day

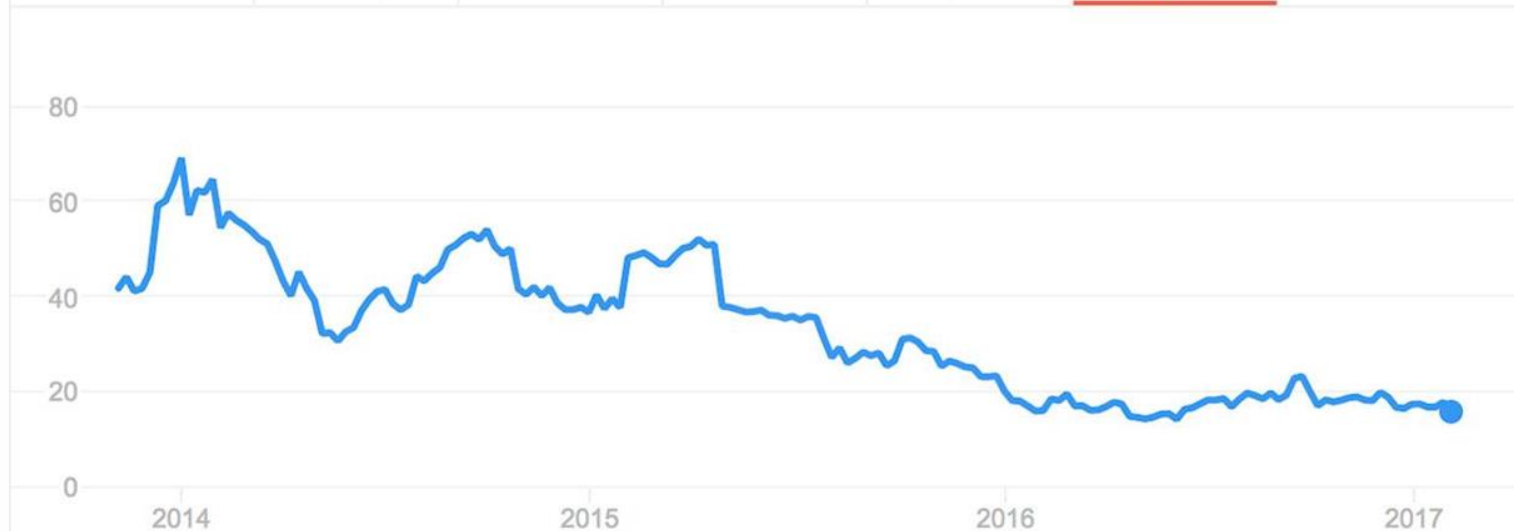
1 month

3 months

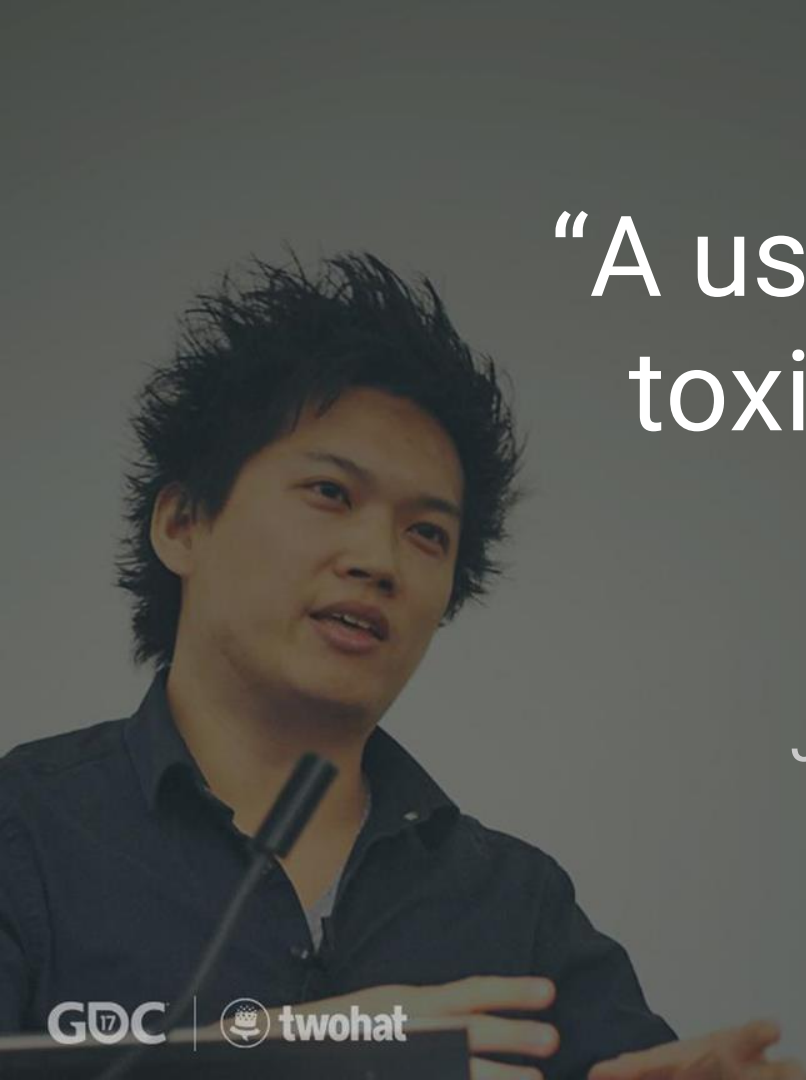
1 year

5 years

max





A photograph of Jeffrey Lin, a man with dark, spiky hair, wearing a dark blue button-down shirt. He is speaking into a black microphone and gesturing with his hands. The background is a solid dark grey.

“A user who experiences toxicity is **320%** more likely to quit.”

Jeffrey Lin, Riot Games, 2014

“I attended Jeffrey's talk because I'm a huge fan of Riot Games... We had the same complaints about toxicity in our game, so it clicked for me. It helped me recognize the potential risk of a toxic community and justified the monetary investment.”

Alexander Galasso

Monetization & Retention Director, Playtika



Toxicity

Mid 17th century Latin: toxicum 'poison,' Greek: toxikon (pharmakon) '(poison for) arrows'

The quality of being toxic or **poisonous**.*

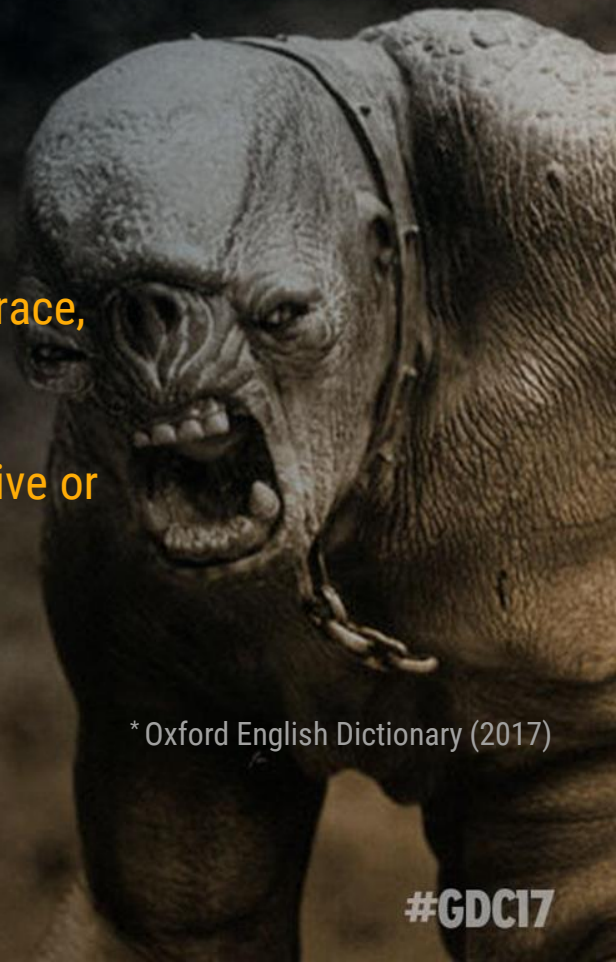
* Oxford English Dictionary (2017)

Troll

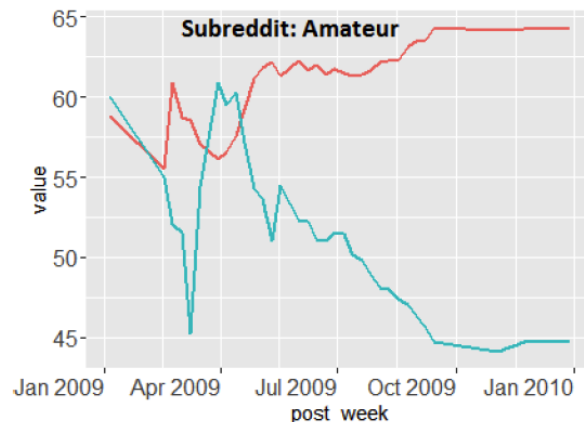
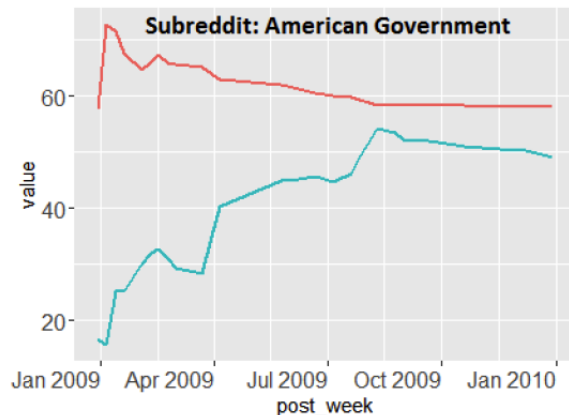
Orig. Old Norse, 1610s: 'not of the human race, evil spirit, monster'

A person who makes a deliberately offensive or provocative online post.*

* Oxford English Dictionary (2017)



“Small increases in toxicity *may* sometimes lead to growth.
But large increases in toxicity always lead to health decline.”



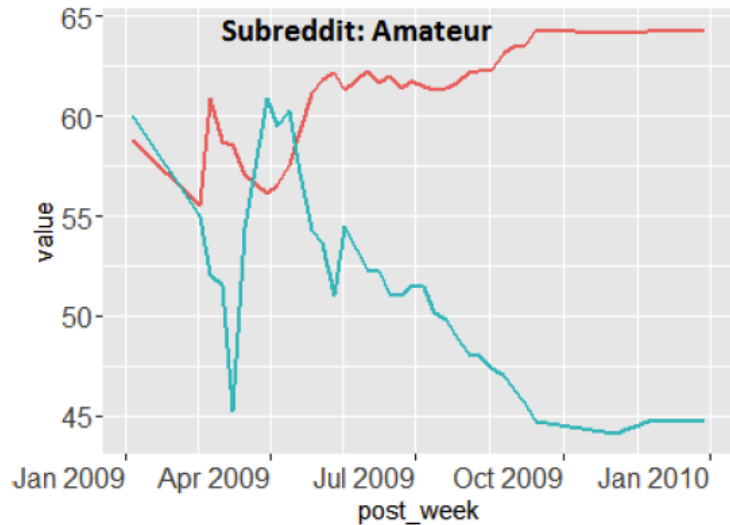
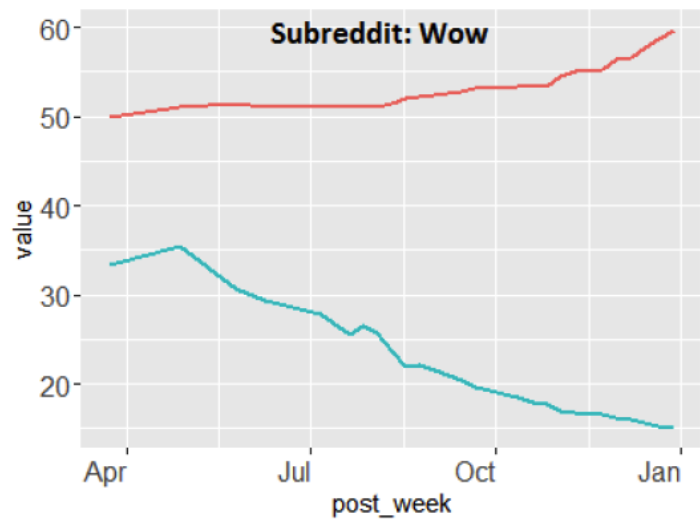
variable

— Healthscore Moving Average

— High-risk Post Percentage Moving Average

The Impact of Toxic Language on the Health of Reddit Communities

Shruthi Mohan, Simon Fraser University (2017)



variable

- Healthscore Moving Average
- High-risk Post Percentage Moving Average

The Impact of Toxic Language on the Health of Reddit Communities

Shruthi Mohan, Simon Fraser University (2017)

A silhouette of a warrior holding a spear, set against a blue abstract background with swirling patterns. The warrior is positioned in the lower-left foreground, facing right. The background is a vibrant blue with white and yellowish-green swirling lines, creating a sense of motion and depth.

“52% of MMORPG players reported that they had been cyber-victimized, and 35% said they had committed cyberbullying themselves”

Virtual Warfare: Cyber Bullying and Cyber-Victimization in MMOG Play
Ballard and Welch (2015)

The image features two cosplayers in elaborate fantasy costumes. On the left, a cosplayer with long blue hair and pointed ears wears a tan tunic and dark pants, holding a large, ornate golden staff. On the right, a cosplayer with long blonde hair wears a green and black outfit with large black wings and holds a long, dark staff with a green gem at the top. The background is dark with a starry, nebula-like pattern. The title text is centered over the image.

A TALE OF TWO COMMUNITIES





SIMILARITIES

TRIALS OF
SERATHIAN

COMPETITIVE MMOs

13+ TOU

61% M / 39% F



BASELINE

TRIALS OF
SERATHIAN

1 MILLION USERS



\$2.78 CAC*

\$13.51 LTV



5% CHURN

* Source: Fiksu Index, 2017



TARGETS

TRIALS OF
SERATHIAN

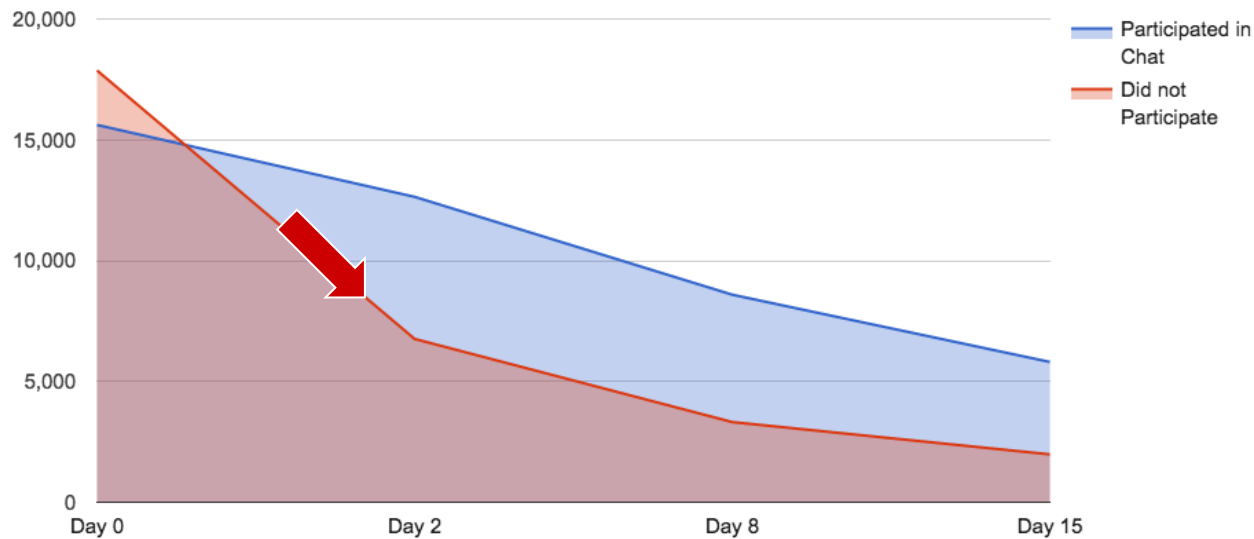
10%

Monthly Growth (MoM)

\$21M

Annual Revenue

Chat Participation & User Churn



	Day 0	Day 2	Day 8	Day 15	Total Users %
Participated in Chat	15,615	12,639	8,589	5,806	37.18%
Did not Participate	17,870	6,759	3,318	1,982	11.09%

Results: Users who engaged in chat

326%

more likely to keep playing
after the **first day**

335%

more likely to keep playing
after **2 weeks**



THE FORMULA

Loss rate with 10%
toxicity = 6.6%



Loss rate with 30%
toxicity = 9.8%

$$U' = U - (U * \text{Loss}) + \text{Acquired}$$



MODERATION

TRIALS OF
SERATHIAN



“We’ll incorporate moderation
from the start”



“We’ll let our users
deal with it themselves”



Q1

PATHS DIVERGE



218,460

Users Lost

10% toxicity

30% toxicity

324,380

Users Lost



-\$1,527,498

Acquisition Cost

-\$60,000

Moderation Cost

-\$1,821,956

Acquisition Cost



\$234,457

Difference



End of Year 2

Total Profit

\$47,177,580



The 7 Million Dollar Difference:

+\$7,392,722

Total Profit

\$39,784,858



User Attrition

-5,840,824

User Attrition

-8,672,738



POST MORTEM

Strategies That Didn't Work



Crowdsourced Everything

Muted



Oct 4 👍 8



I love this game, but it's unplayable. Trolls and toxic players are ruining this game. Reporting them does absolutely nothing, and the matchmaker still places you in the same games with them. This should absolutely never be able to happen.

If I report someone, why on earth would I ever want to be put in the same game as them? It's like reporting means nothing, and its just a fake button to make me feel better.

POST MORTEM

Strategies That Worked



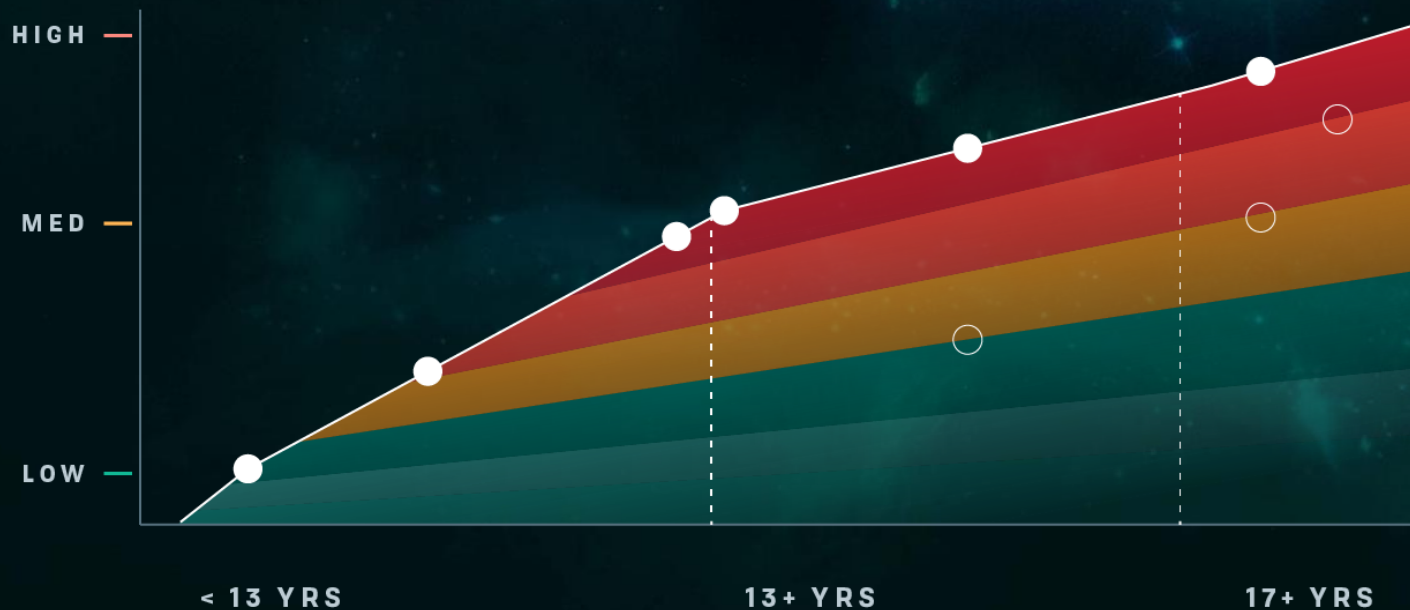
Knew the Resilience of Their Community

Divided Users into Quadrants

Used AI + Human Moderation

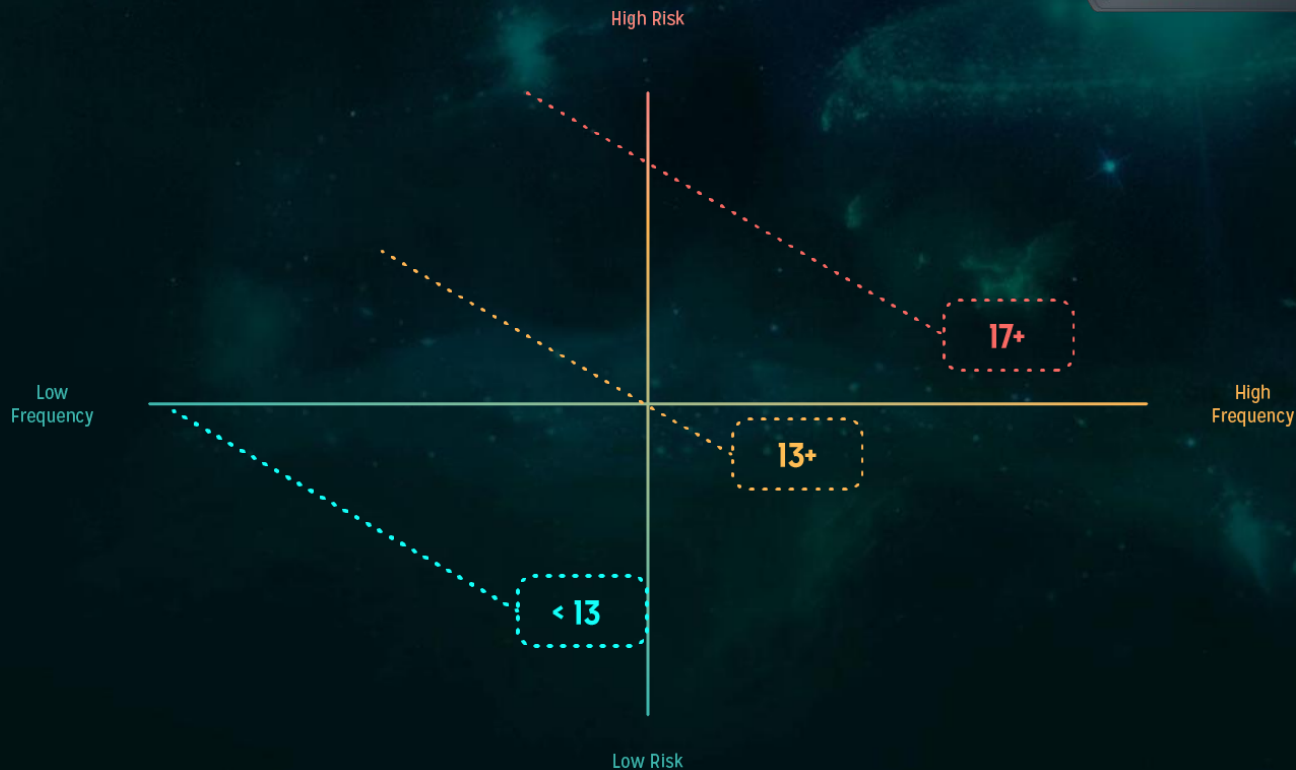
POST MORTEM | Strategies that Worked

Community Resilience



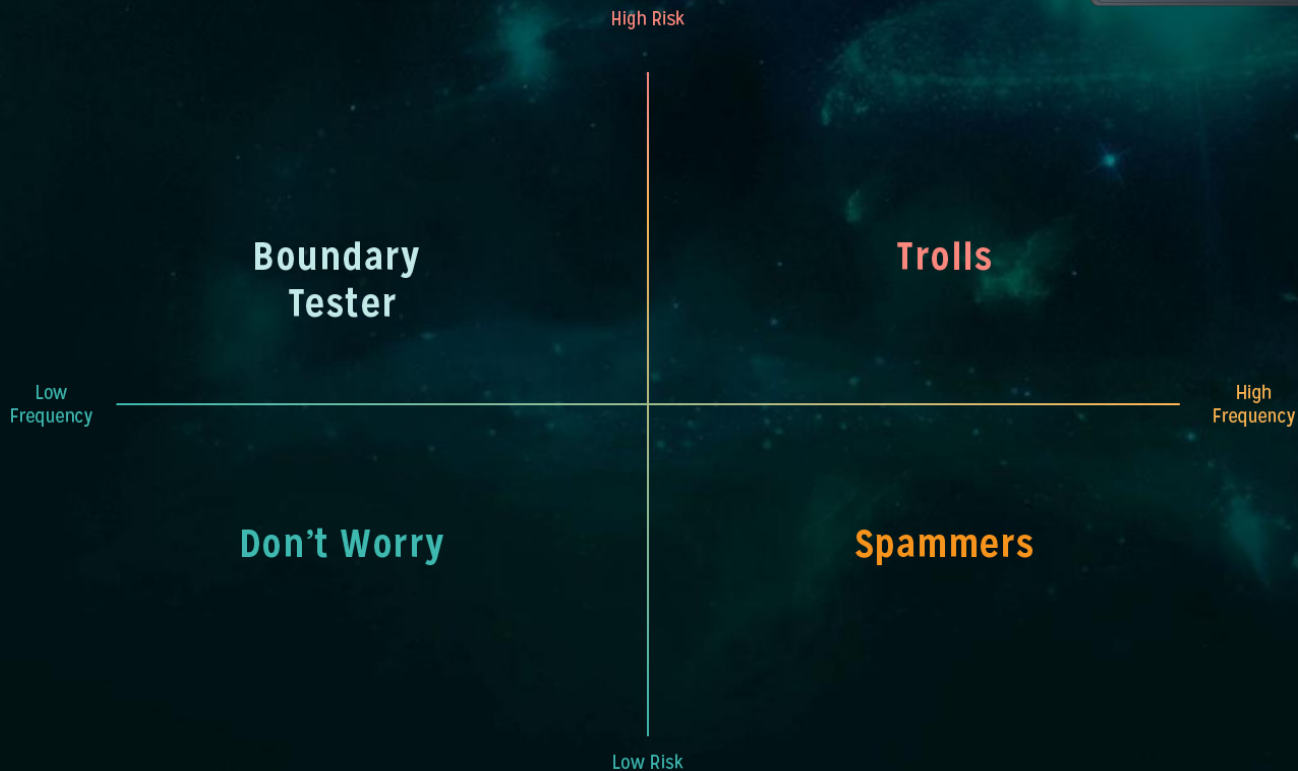
POST MORTEM | Strategies that Worked

Community Resilience



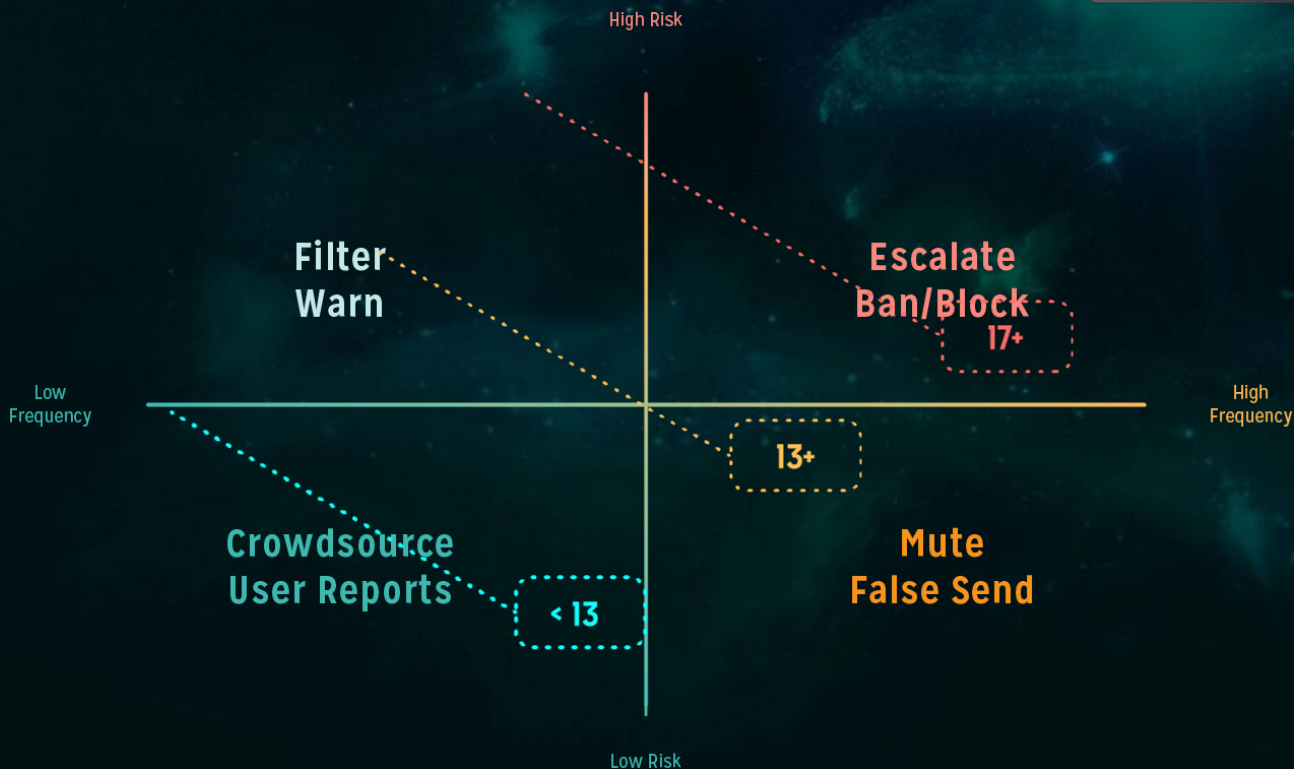
POST MORTEM | Strategies that Worked

Prioritized Risk Groups



POST MORTEM | Strategies that Worked

Defined Action Plans



Moderation + AI



Introducing

AutoMod



Moderation + AI



[Link to comment](#) · 2 months ago

With respect to the second-last paragraph, I think jerks sneaking around moderation is exactly what AutoMod is designed to overcome. Assuming it works as advertised, it would be able to detect if some new phrase is being used abusively, and block it or refer it to a human moderator. Basically, and hopefully, new forms of abuse would only work once.

^ | v · Reply · Share ›



12/14/16 4:18pm

Twitch, please share this magical technology with every FPS publisher.

↳ Reply

Your AutoMod setting:



Rule 1 **A Little Moderation**

Identity Language

Referring to race, religion, gender, orientation, disability, or similar

Some filtering

Sexually Explicit Language

Referring to sexual acts, sexual content, and body parts

No filtering

Aggressive Language

Hostility towards other people, often associated with bullying

No filtering

Profanity

Expletives, curse words, things you wouldn't say to grandma

No filtering

What are the Benefits of Doing Something?

Financial - Higher ROI

Brand - Protect Your Reputation

Community - Encourage Better Future Behavior



Thank you!

Chris Priebe

CEO, Two Hat Labs

a research division of Two Hat Security