Context Matters

Event Detection

Expression
GGWP Behavioral Models Detect Events

Baseline Behavior and Detecting Anomalies

Delivery Channel
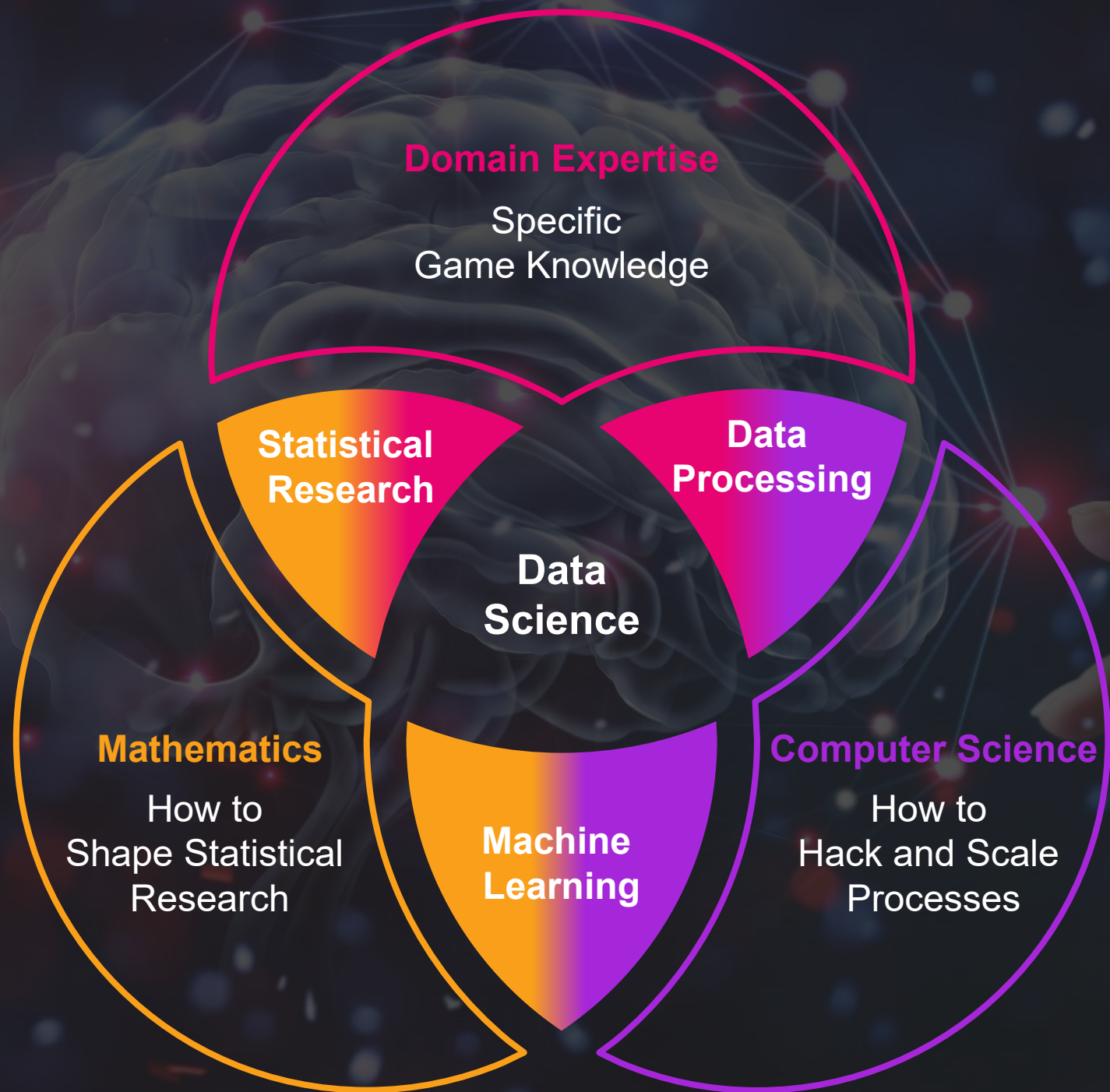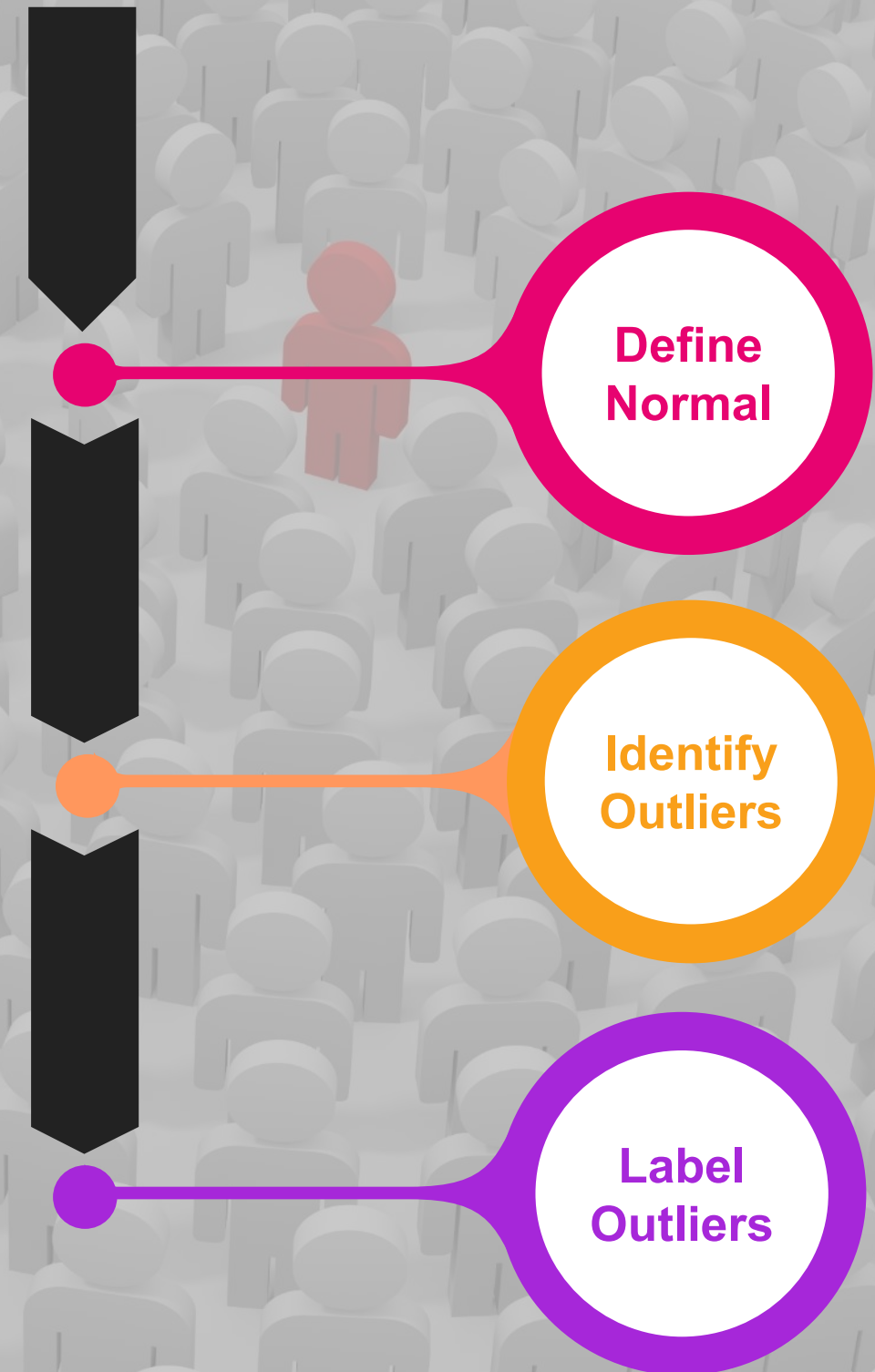
Root Cause
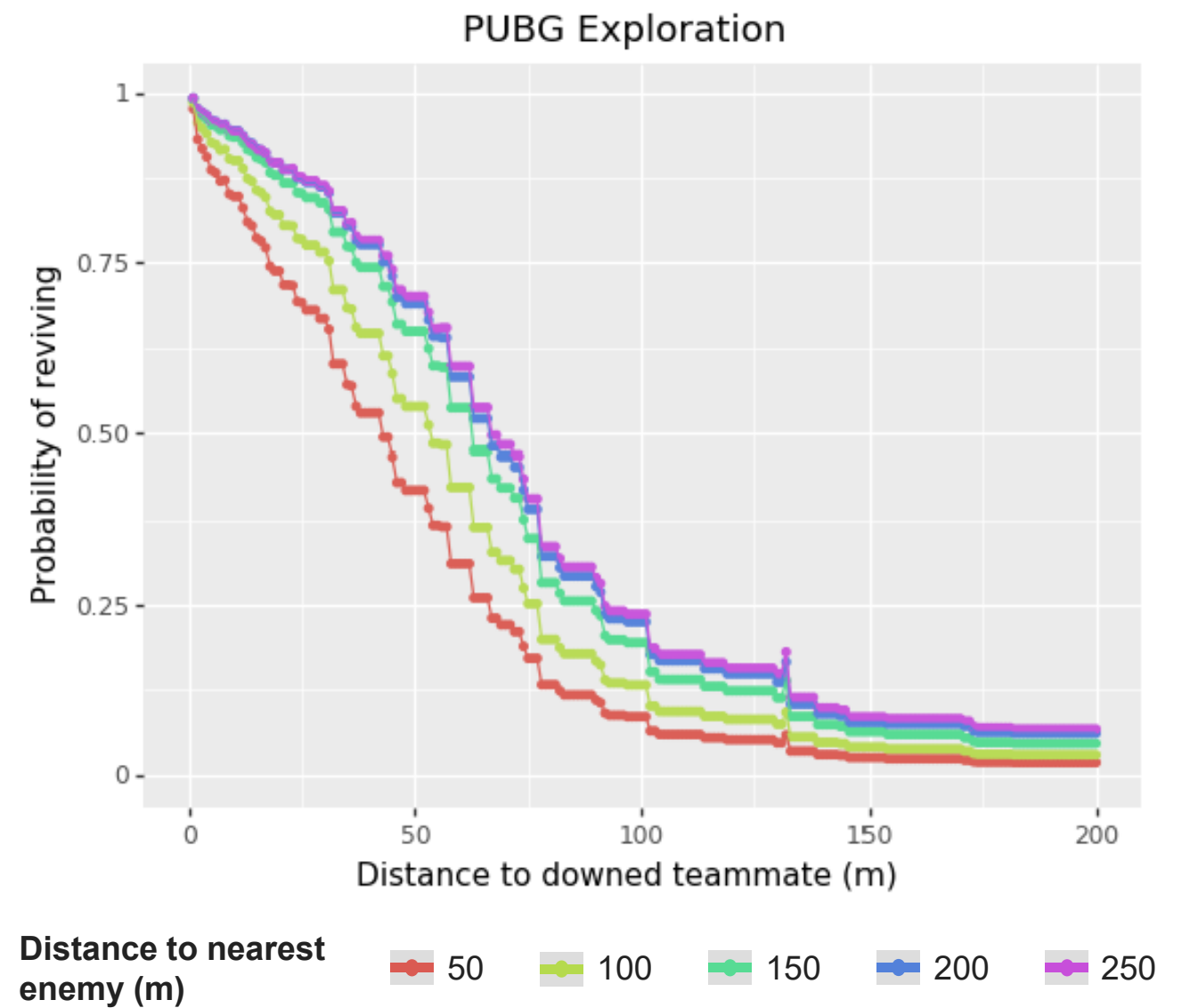
Impact

Baseline Models – Ex. PubG Revive Probability

Event Detection - Ex. PubG Reviving a teammate or not?

Legend (map):
- ● Player location
- ● Downed teammate
- ● Enemies

Probability of reviving teammate
- 90%
- 70%
- 50%
- 30%
- 10%

Map distances: 17m 52m 66m 89m 138m

Map labels: Monte Nuevo, Pecado, Ladrillera, Chumacera, Los Le[...], Valle del Mar

Decision making becomes complex when there are multiple factors involved

- Distance to downed teammate / Distance to nearest enemy
- Health
- Time available
- Number of previous revives
- Teammates positions
- Circle position

Probability of reviving teammate

High? →

Revive? → Incident

1. The scale of the background map is distorted to facilitate the visualization

What Led to the Event?

Expression
GGWP Behavioral Models detect events

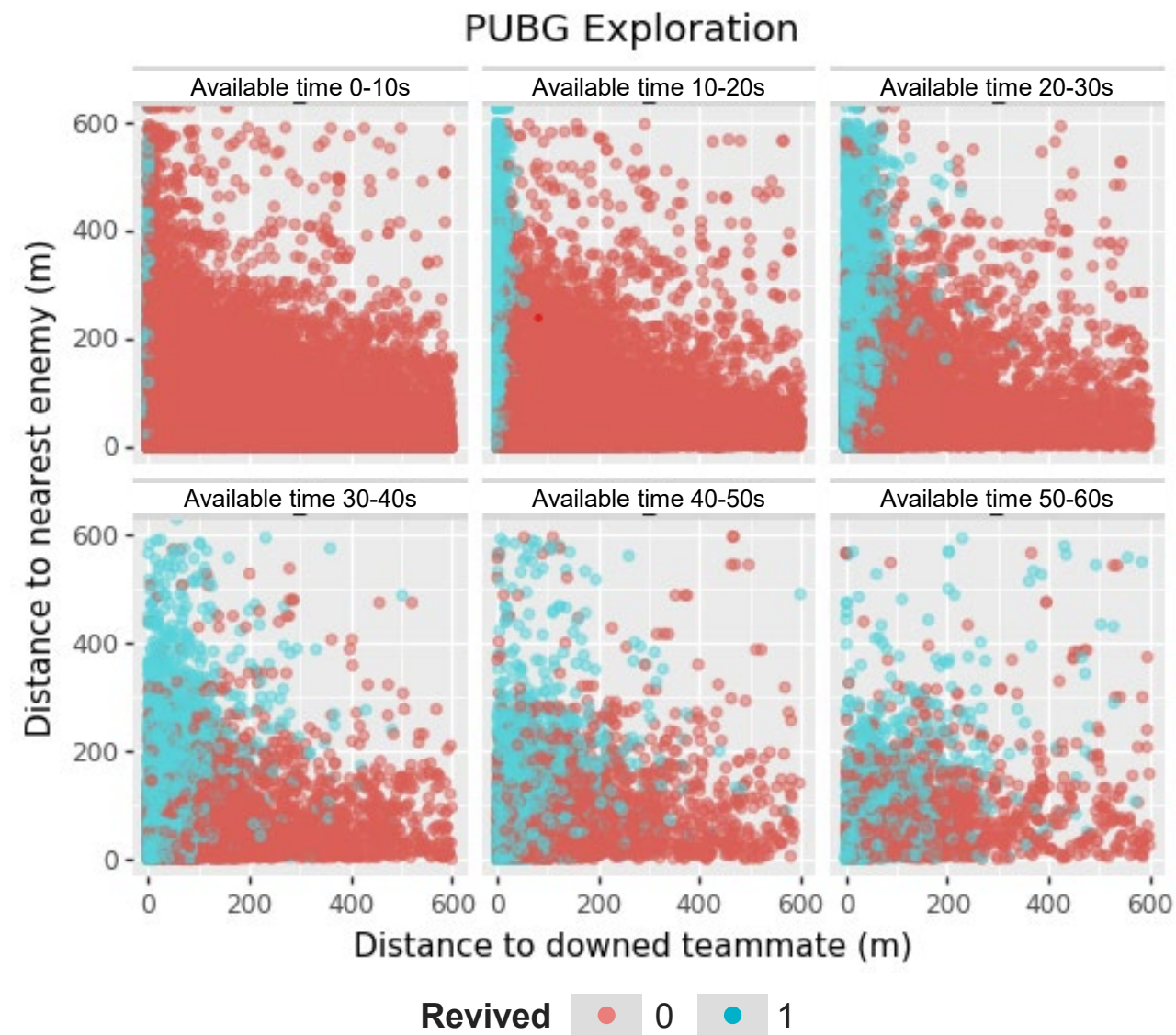Baseline Behavior and Detecting Anomalies

Delivery Channel

Root Cause
GGWP Severity Modifiers adjust for context of events

Types of Players

Expectations

Initiation and Triggering
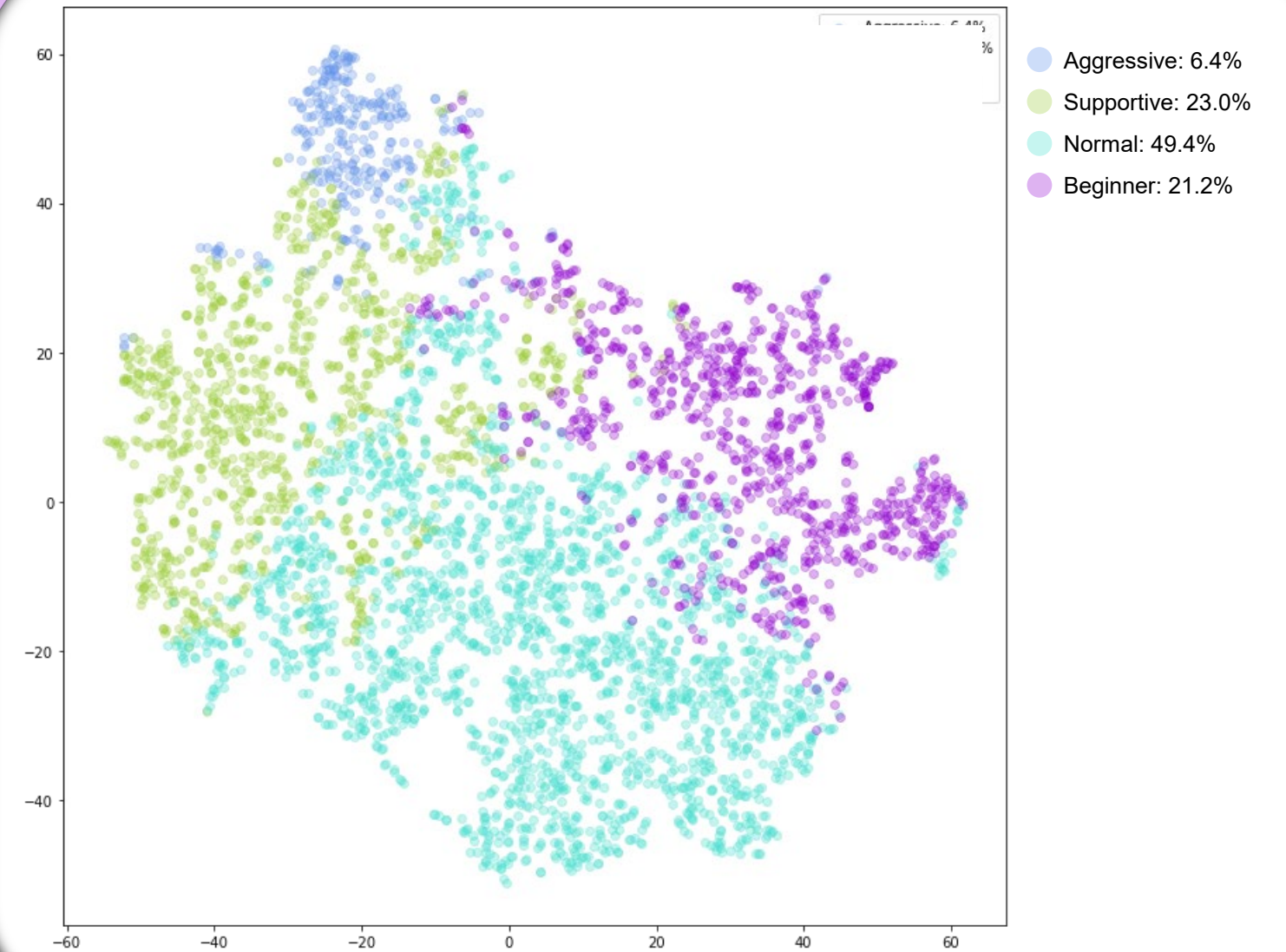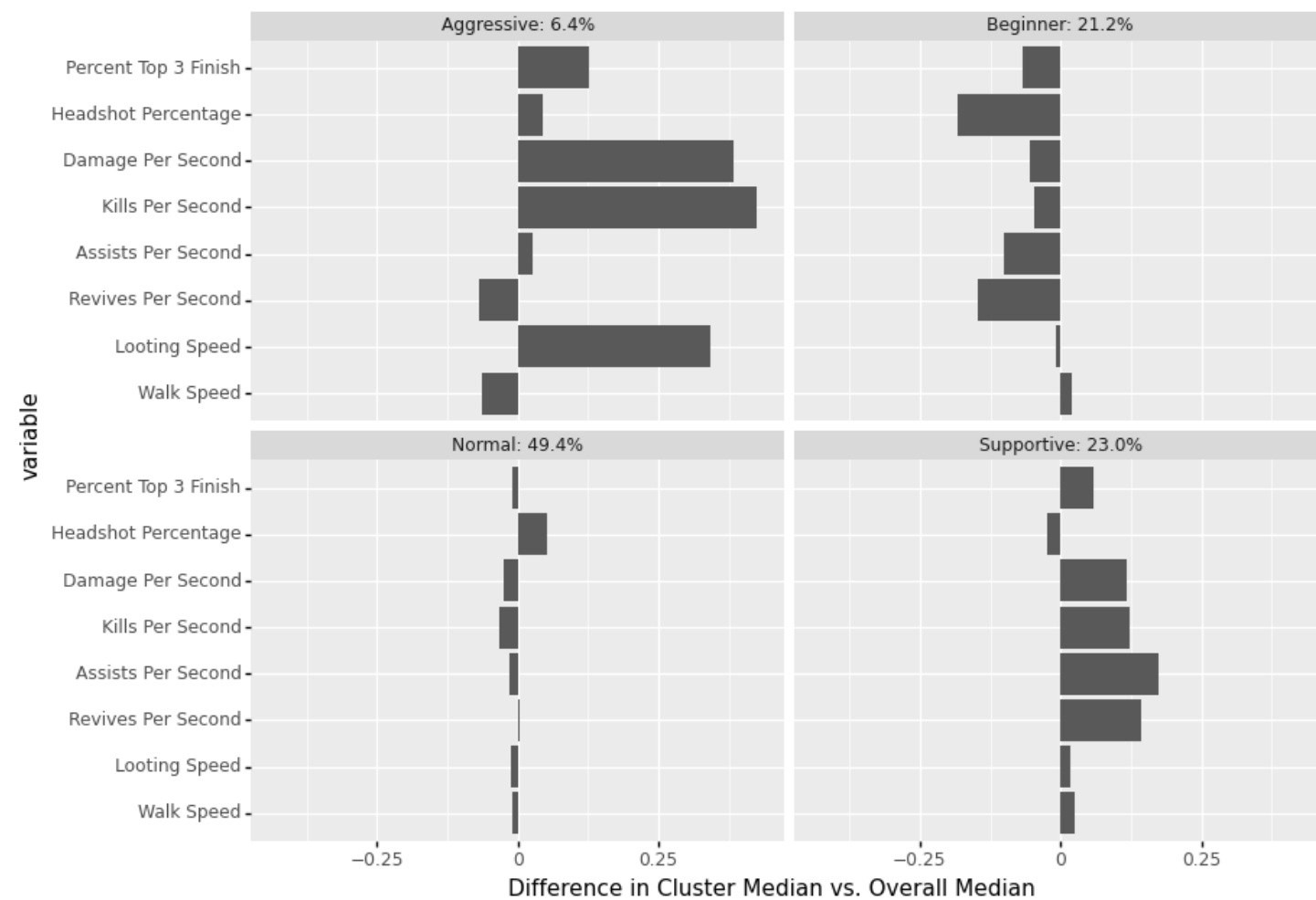
Social Factors

Impact

Types of Players – Ex. PubG Personas

Expectations (Skill Level) – Ex. Valorant Friendly Fire by Rank

	Breach Aftershock	Brimstone Incendiary	Brimstone OrbitalStrike	Killijoy Nanoswarm	Phoenix Blaze	Phoenix HotHands	Raze PaintShells	Raze Showstopper	Sova HuntersFury	Sova ShockBolt	Viper SnakeBite
Iron	0.0136	0.0271	0.0118	0.0157	0.0225	0.0294	0.0314	0.0067	0.0111	0.0169	0.0079
Bronze	0.0069	0.0158	0.0077	0.0077	0.0093	0.0163	0.0198	0.0038	0.0106	0.0153	0.0112
Silver	0.0051	0.0115	0.0084	0.0081	0.0067	0.0152	0.0156	0.0058	0.0103	0.0125	0.0087
Gold	0.0028	0.0101	0.0076	0.0081	0.0048	0.0122	0.0178	0.006	0.0071	0.0111	0.0071
Platinum+	0.0052	0.006	0.0053	0.0049	0.003	0.0128	0.0175	0.0054	0.0089	0.0103	0.0064

Triggering – Ex. Offenders That May Not Have Initiated Toxicity
Dataset: League Tribunal Case Chatlogs

19.3%
Percent of total cases where the reported player does not initiate toxicity

Flagged by GGWP Toxicity Models

Reviewed Case Examples

[Ally] (00:00:21) Ezreal: ask myself sometimes

[Ally] (00:00:24) Ezreal: who needs a map

[Ally] (00:00:25) Ezreal: gg

[Ally] (00:00:32) Ezreal: sounds gay

[Ally] (00:00:33) Ezreal: im in

[Ally] (00:01:11) Ezreal: crank?

[Ally] (00:01:13) Ezreal: da fuk u doing ?

[Offender] (00:01:21) Blitzcrank: im going top

[Offender] (00:01:21) Blitzcrank: told u

[Offender] (00:01:22) Blitzcrank: dumbass

[Ally] (00:01:15) Xin Zhao: voli help plz

[Enemy] (00:01:21) Elise: first time, don't flame if i screw :p

[Enemy] (00:01:29) Pantheon: lol ok

[Enemy] (00:02:09) Annie: omg

[Ally] (00:02:11) Xin Zhao: idiot

[Ally] (00:02:14) Fizz:

[Enemy] (00:02:15) Annie: how u jungle -.-"

[Ally] (00:02:22) Fizz: fu

[Enemy] (00:02:22) Annie: smite noob

[Enemy] (00:02:32) Elise: look what you are doing

[Enemy] (00:02:38) Annie: dude i made basic atks!

[Ally] (00:02:39) Xin Zhao: no ganks mid

[Ally] (00:02:40) Xin Zhao: fking idiot

[Ally] (00:02:44) Xin Zhao: then i troll

[Enemy] (00:02:45) Annie: dont blame if u cant smite! Noob

[Offender] (00:02:46) Kassadin: dont cry

Triggers – Valorant Friendly Fires Detected by Score Difference

Team Losing

<=-8	-7	-6	-5
0.121	0.08	0.065	0.064

Close Match

-4	-3	-2	-1	0	1	2	3	4
0.047	0.051	0.038	0.048	0.056	0.049	0.035	0.045	0.045

Team Winning

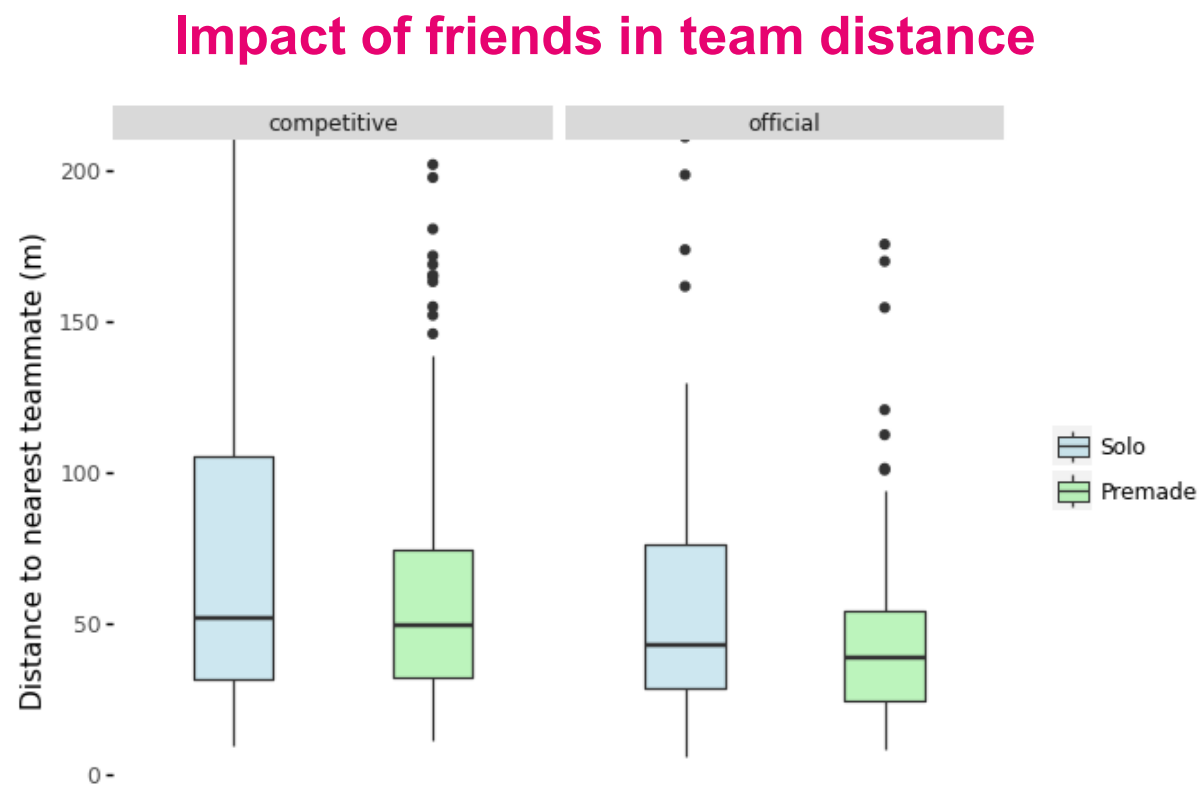| 5	6	7	>=8
0.059	0.052	0.065	0.08

Social Factors –
Ex. PubG Distance in Premade vs Solo Queue

Players with friends on their squad tend to walk closer together

Impact of friends in team distance

Match type	Playing with Friends	Avg, distance to nearest teammate (m)
Competitive	Yes (Premade)	76
	No (Solo)	112
Official (Casual)	Yes (Premade)	67
	No (Solo)	144

How are the victims impacted?

Expression
GGWP Behavioral Models detect events

Delivery Channel

Root Cause
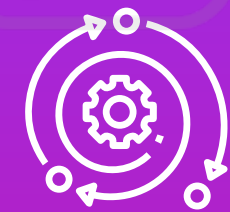GGWP Severity Modifiers adjust for context of events

Impact

Baseline Behavior and Detecting Anomalies

Types of Players

Expectations
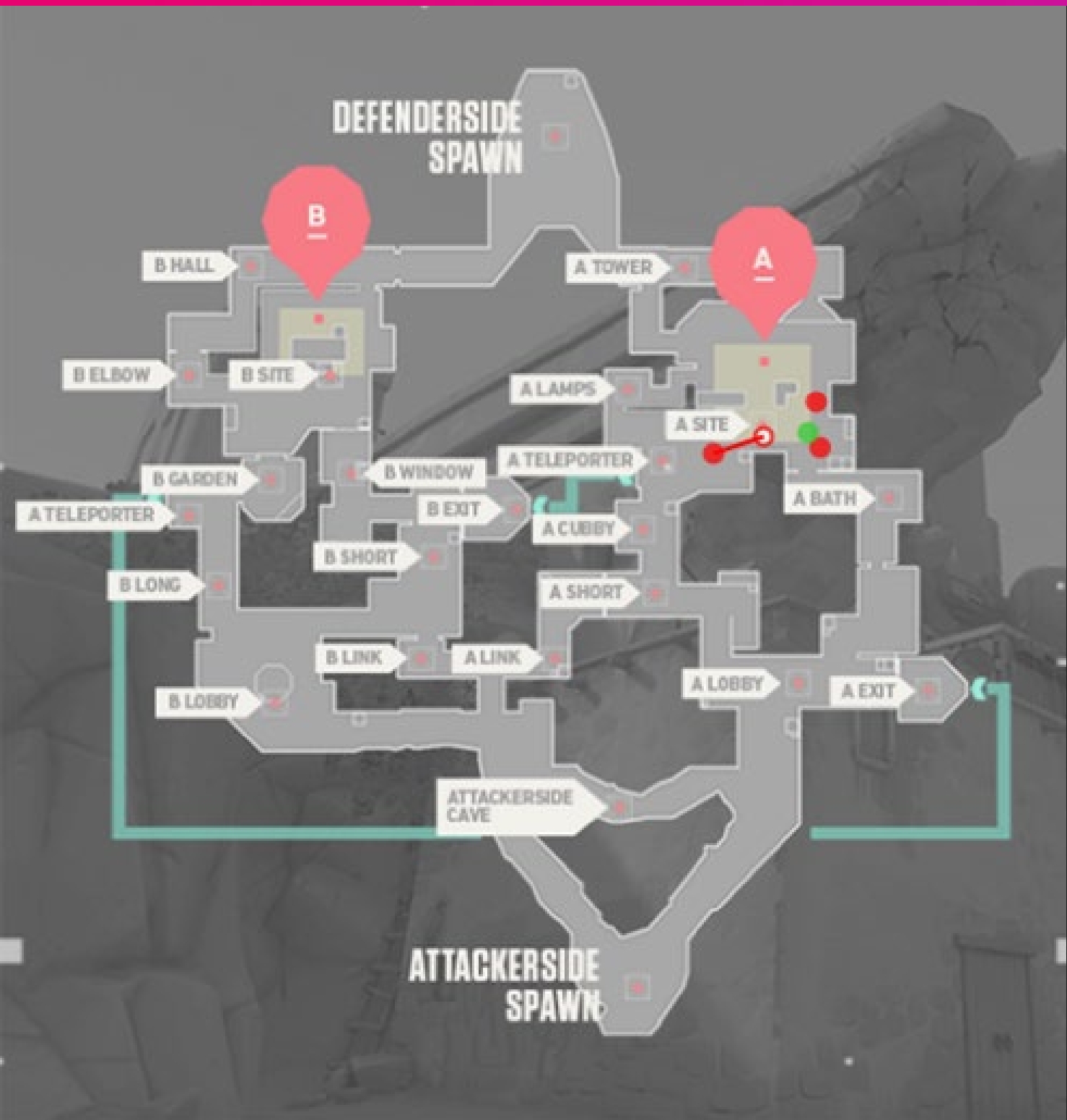
Initiation and Triggering

Social Factors

Game State
Win % Rating

Game Situation – Ex. Valorant Friendly Fire
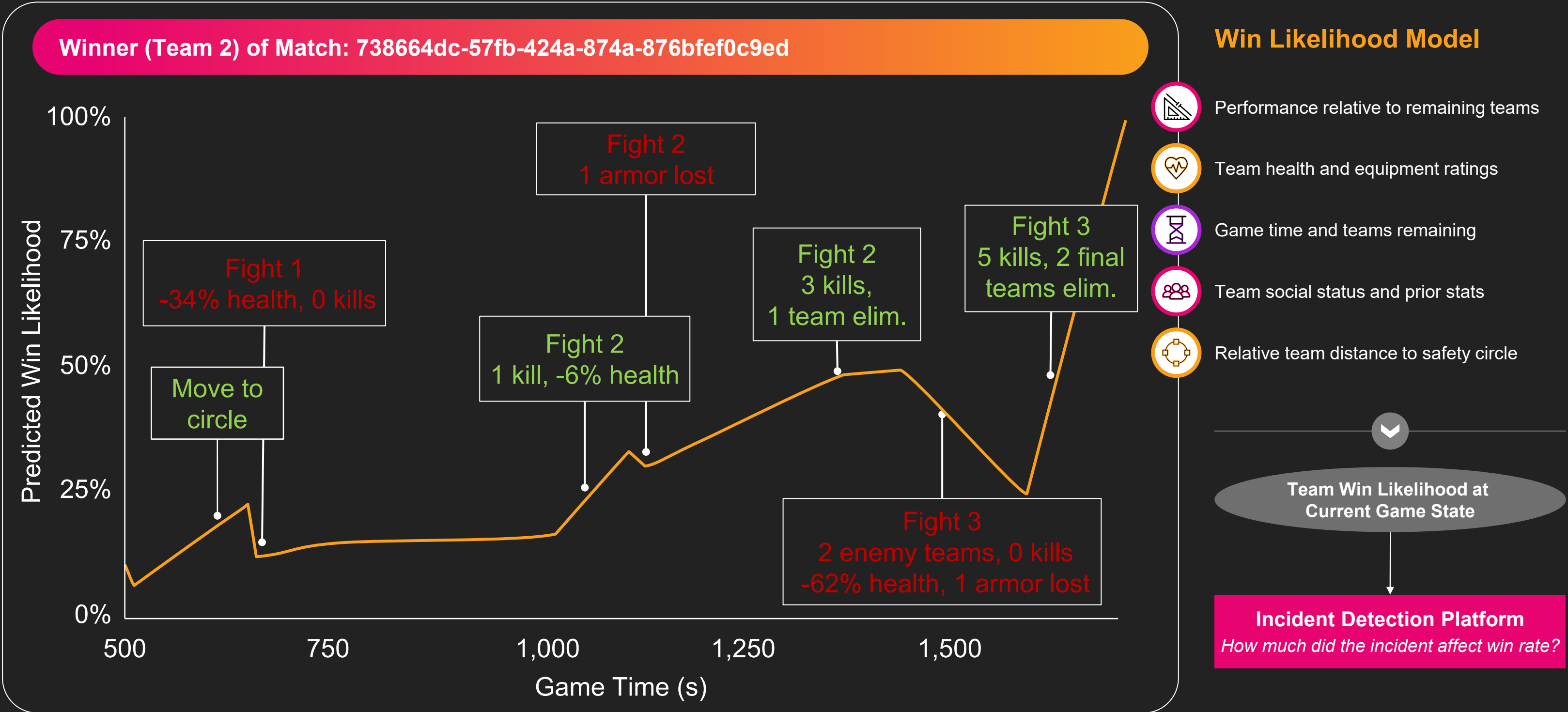Friendly Fires Detected - <u>When enemy team has been eliminated</u>

Stat	Value
Match_Id	1702bcf7-11fd-4bf5-9ffc-2da1d69691a8
Ranked	False
Round	12
Team	Red - Defender
Kill_Time	35.971
Killer_Rank	Unranked
Killer_Agent	Raze
Ability	PaintShells
Victim_Rank	Unranked
Victim_Agent	Viper
Score	2-9
Record	LWLLLLLLWLL
Round_Winner	Red
Killer Round Stats	Enemy_Kills: 0, teammate_kills: 2, deaths: 0, assists: 0
Killer_Match_Stats	Enemy_Kills: 6, teammate_kills: 2, deaths: 17, assists: 2
Victim_Round_Stats	Enemy_Kills: 3, teammate_kills: 0, deaths: 1, assists: 0
Victim_Match_Stats	Enemy_Kills: 13, teammate_kills: 0, deaths: 16, assists: 2

Timing in Game – Ex. Valorant Friendly Fire <u>by Round</u>

Side Switch

1	2	3	4	5	6	7	8	9	10	11	12	13
0.067	0.049	0.032	0.037	0.032	0.033	0.032	0.033	0.029	0.035	0.035	0.073	0.047

14	15	16	17	18	19	20	21	22	23	24	25	OT
0.032	0.029	0.033	0.038	0.042	0.033	0.041	0.036	0.04	0.03	0.032	0.048	0.03

Win Likelihood Model – Ex. Following a Team in a PubG

Winner (Team 2) of Match: 738664dc-57fb-424a-874a-876bfef0c9ed

Win Likelihood Model

Predicted Win Likelihood

- 100%
- 75%
- 50%
- 25%
- 0%

Fight 1
-34% health, 0 kills

Move to circle

Fight 2
1 armor lost

Fight 2
1 kill, -6% health

Fight 2
3 kills,
1 team elim.

Fight 3
5 kills, 2 final
teams elim.

Fight 3
2 enemy teams, 0 kills
-62% health, 1 armor lost

Game Time (s)

500   750   1,000   1,250   1,500

- Performance relative to remaining teams
- Team health and equipment ratings
- Game time and teams remaining
- Team social status and prior stats
- Relative team distance to safety circle

Team Win Likelihood at Current Game State

Incident Detection Platform
How much did the incident affect win rate?

Operationalizing Models – Impact of Errors

Impact of False Positives

Movie Recommendation

Loan Underwriting

Possible Threat Identification

Does Woody like dinosaur movies?

Will Carl default on his loan?

Confirm the target's location?

More Conservative Detection

Penalty Applications

Impact of False Positives →

Warning vs Penalty

Duration

Cooldown

Suspension

Ban

Channel

XP Gain

Match-making

Chat, Audio

PvP, Ranked, Event, Leaderboard
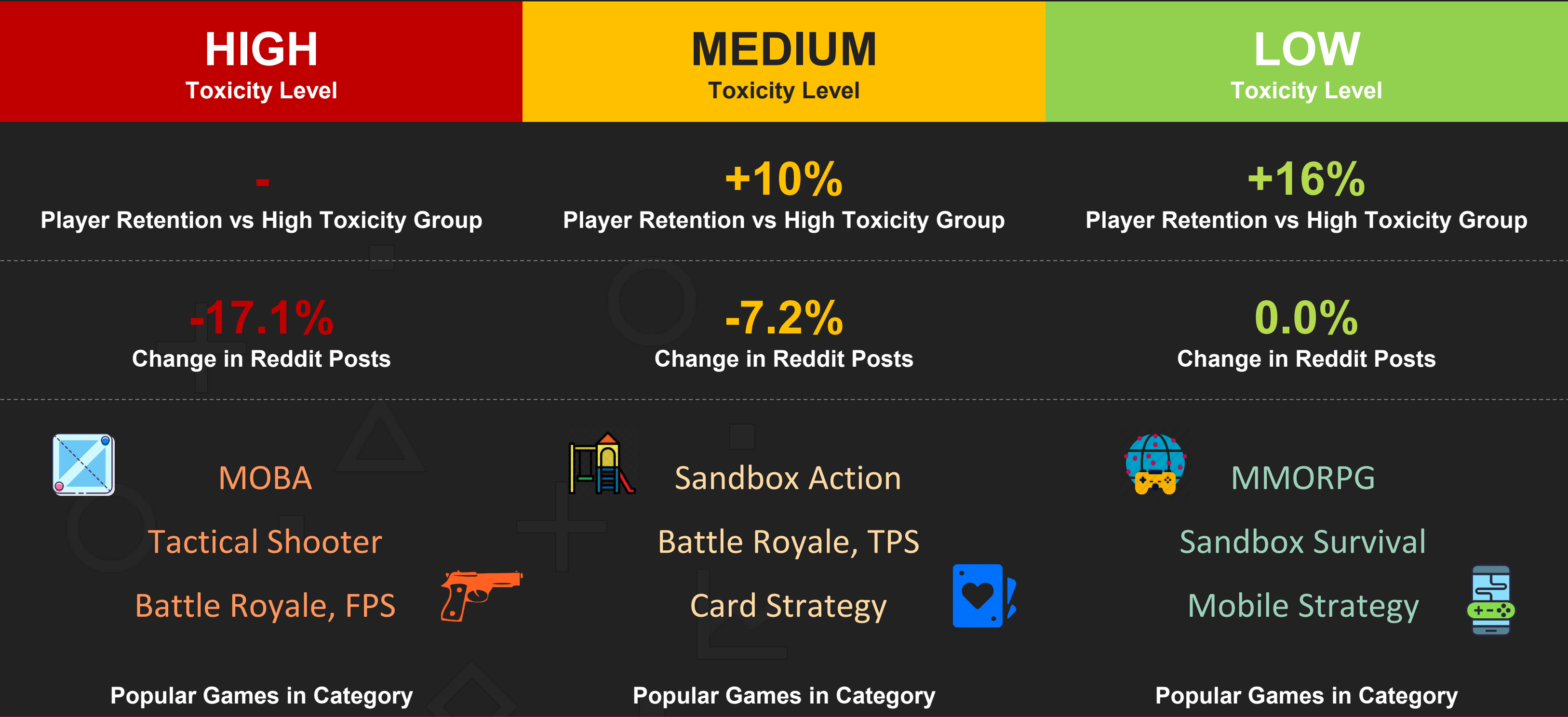
All

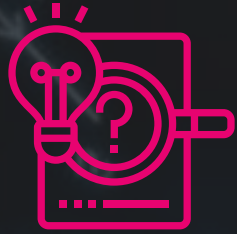**More Conservative Detection** →
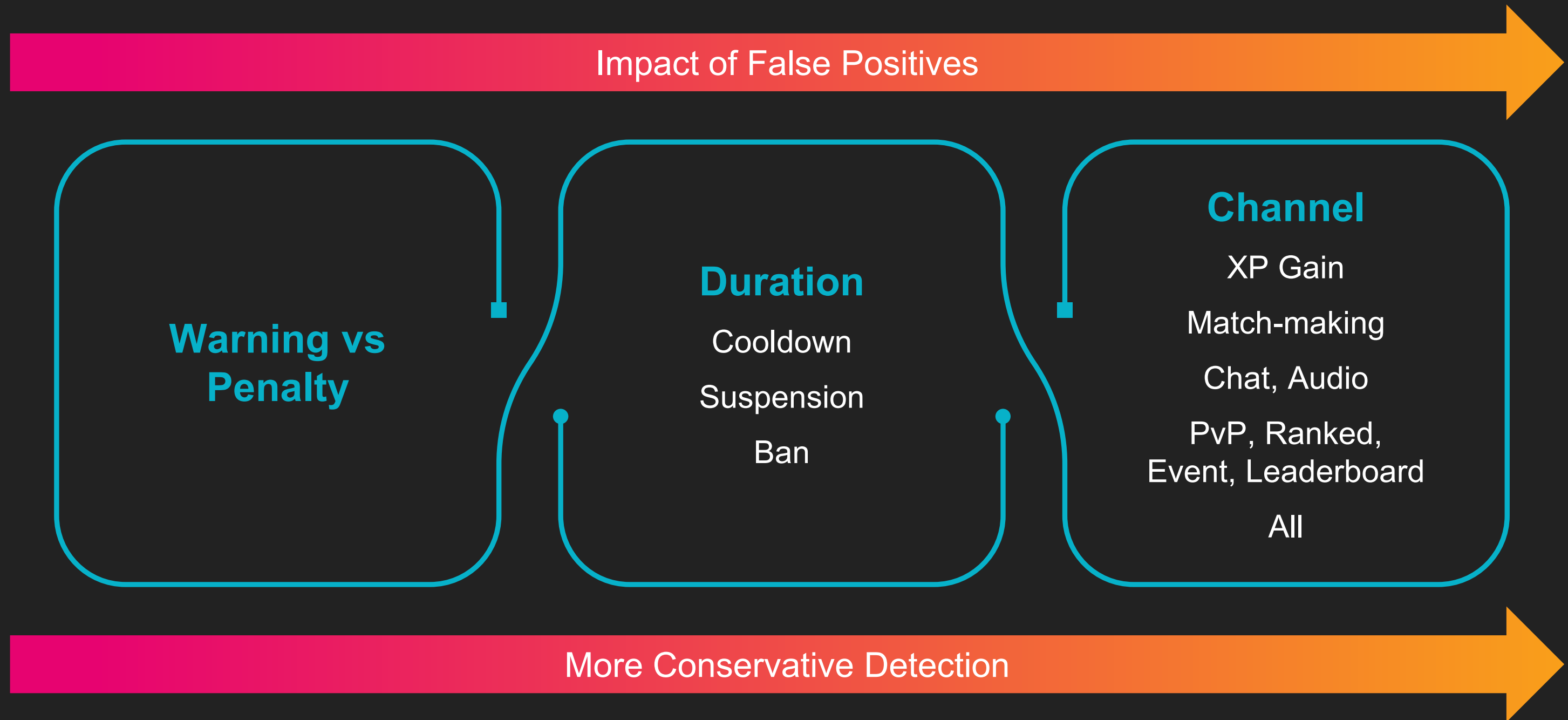
Comprehensive
Multiple Events,
Multiple Games

Consistent
Less Sensitive to
Individual Errors

Interpretable
Single Holistic
Number

Quantifying the Impact of Toxicity

HIGH	MEDIUM	LOW
Toxicity Level	**Toxicity Level**	**Toxicity Level**
-	**+10%**	**+16%**
Player Retention vs High Toxicity Group	**Player Retention vs High Toxicity Group**	**Player Retention vs High Toxicity Group**
-17.1%	**-7.2%**	**0.0%**
Change in Reddit Posts	**Change in Reddit Posts**	**Change in Reddit Posts**
MOBA	Sandbox Action	MMORPG
Tactical Shooter	Battle Royale, TPS	Sandbox Survival
Battle Royale, FPS	Card Strategy	Mobile Strategy
Popular Games in Category	**Popular Games in Category**	**Popular Games in Category**

Summary

Automate Anomalous Event Detection

Add Contextual Factors

Tune Based on Application

ggwp

George Ng
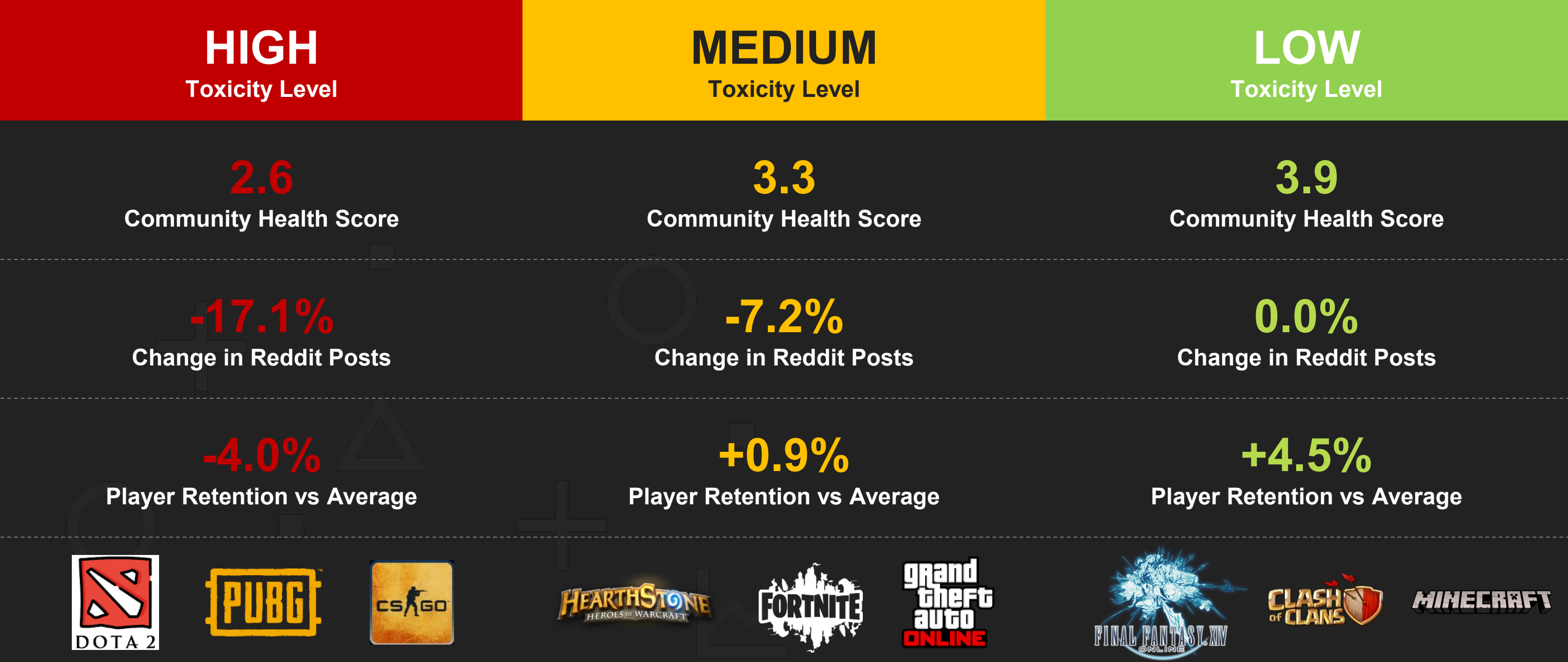
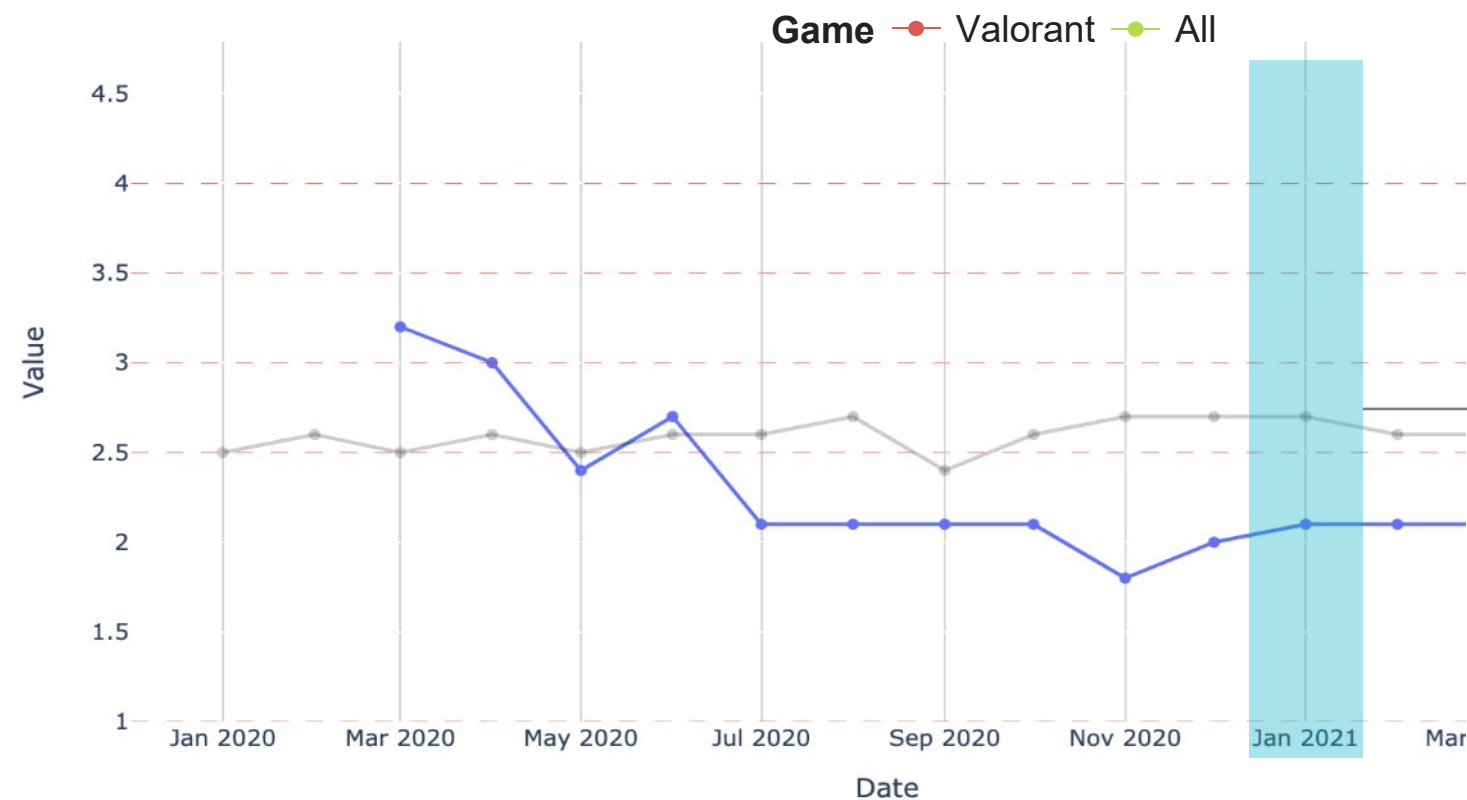Chief Technology Officer

✉ george@ggwp.com

in https://www.linkedin.com/in/georgeng/

Appendix

Quantifying the Impact of Toxicity

HIGH Toxicity Level	MEDIUM Toxicity Level	LOW Toxicity Level
2.6 Community Health Score	**3.3** Community Health Score	**3.9** Community Health Score
-17.1% Change in Reddit Posts	**-7.2%** Change in Reddit Posts	**0.0%** Change in Reddit Posts
-4.0% Player Retention vs Average	**+0.9%** Player Retention vs Average	**+4.5%** Player Retention vs Average

GGWP Game Health Scores

Community Health score

Game ● Valorant ● All

GGWP Community Health Score –

Single value to summarize the "healthiness" level of a game's community based on social media data

	Bronze	1.0 – 2.5		Platinum	3.5 – 4.0
	Silver	2.5 – 3.0		Diamond	4.0 – 5.0
	Gold	3.0 – 3.5			

Valorant – Risk factors 2021-03

Game
Valorant
All

Risk Factors – 10 toxicity components that explain the Community Health Score. They represent

- **Gameplay Toxicity:** In-game toxic behaviors like AFK, cheating, griefing, being mean to other players, etc.
- **Community Toxicity:** Level of hatred within the Reddit community
- **Community Sentiment:** General sentiment of posts and comments

GGWP Game Health Snapshot – 2021 Q1

Game	Avg Daily Posts	AFK	Cheating	Smurfing	Rage-Quitting	Griefing	Teaming	Player Toxicity	Reddit Toxicity	Reddit Identity Hate	Reddit Negativity	Community Health Score
DOTA 2	195	1.4%	3.3%	3.5%	0.9%	4.4%	0.0%	3.3%	6.8%	0.1%	34.6%	1.9
Valorant	475	1.8%	2.7%	1.9%	0.5%	1.5%	0.0%	2.5%	4.7%	0.0%	39.0%	2.1
League of Legends	628	1.4%	1.0%	1.2%	0.5%	2.1%	0.0%	2.6%	6.3%	0.0%	36.0%	2.3
Apex Legends	778	0.2%	2.2%	0.3%	0.7%	0.3%	0.2%	0.7%	7.4%	0.1%	33.8%	2.5
Escape from Tarkov	407	0.1%	3.9%	0.0%	1.1%	0.6%	0.1%	0.3%	9.3%	0.1%	36.9%	2.5
Rainbow Six	346	0.2%	2.1%	0.7%	0.4%	0.6%	0.0%	1.5%	6.6%	0.2%	32.2%	2.6
Rust	137	0.2%	3.1%	0.0%	0.5%	0.8%	0.1%	1.7%	8.5%	0.1%	33.2%	2.6
CS: GO	195	0.3%	4.9%	0.4%	0.5%	0.8%	0.0%	1.1%	6.0%	0.1%	32.7%	2.6
Overwatch	203	0.1%	1.5%	0.8%	0.5%	0.8%	0.0%	1.6%	5.8%	0.1%	31.7%	2.7
Wild Rift	161	2.0%	1.0%	0.4%	0.4%	2.1%	0.0%	2.1%	4.1%	0.0%	31.6%	2.7
World of Warcraft	281	0.3%	1.8%	0.0%	0.5%	1.2%	0.0%	1.3%	5.9%	0.1%	32.8%	2.8
COD: Warzone	380	0.1%	5.2%	0.0%	0.3%	0.2%	0.1%	0.2%	9.3%	0.1%	35.2%	2.9
COD: Black Ops - Cold War	446	0.1%	1.7%	0.0%	0.6%	0.3%	0.0%	0.3%	9.0%	0.1%	36.5%	3.0
Arena of Valor	35	1.7%	1.0%	0.5%	0.2%	2.2%	0.0%	0.9%	4.8%	0.0%	28.1%	3.0
PUBG	35	0.2%	4.5%	0.0%	0.2%	0.5%	0.2%	0.6%	6.4%	0.1%	35.4%	3.0
PUBG Mobile	69	0.1%	4.5%	0.0%	0.4%	0.2%	0.3%	0.6%	4.9%	0.1%	29.5%	3.2
GTA Online	488	0.3%	1.4%	0.0%	0.3%	1.8%	0.0%	0.3%	6.4%	0.1%	29.1%	3.3
Super Smash Bros	31	0.0%	0.8%	0.0%	0.4%	0.6%	0.1%	1.0%	5.5%	0.1%	34.8%	3.3
COD Mobile	615	0.1%	2.4%	0.0%	0.1%	0.2%	0.0%	0.4%	7.2%	0.1%	30.5%	3.4
Warframe	114	0.2%	1.8%	0.0%	0.1%	0.9%	0.0%	0.6%	4.0%	0.0%	31.6%	3.4
Fortnite: Battle Royale	612	0.1%	0.6%	0.0%	0.1%	0.4%	0.1%	0.4%	4.3%	0.1%	28.7%	3.6
Final Fantasy XIV	60	0.2%	1.2%	0.0%	0.2%	0.6%	0.0%	0.7%	3.3%	0.0%	32.0%	3.7
Minecraft	1561	0.4%	0.9%	0.0%	0.2%	0.4%	0.0%	0.1%	3.3%	0.0%	23.8%	4.1
Clash of Clans	187	0.0%	0.8%	0.0%	0.1%	0.2%	0.0%	0.1%	2.6%	0.0%	30.0%	4.5

25

Player Profile - Ex. Valorant Teammate Damages and Kills

Teammate Damage Distribution
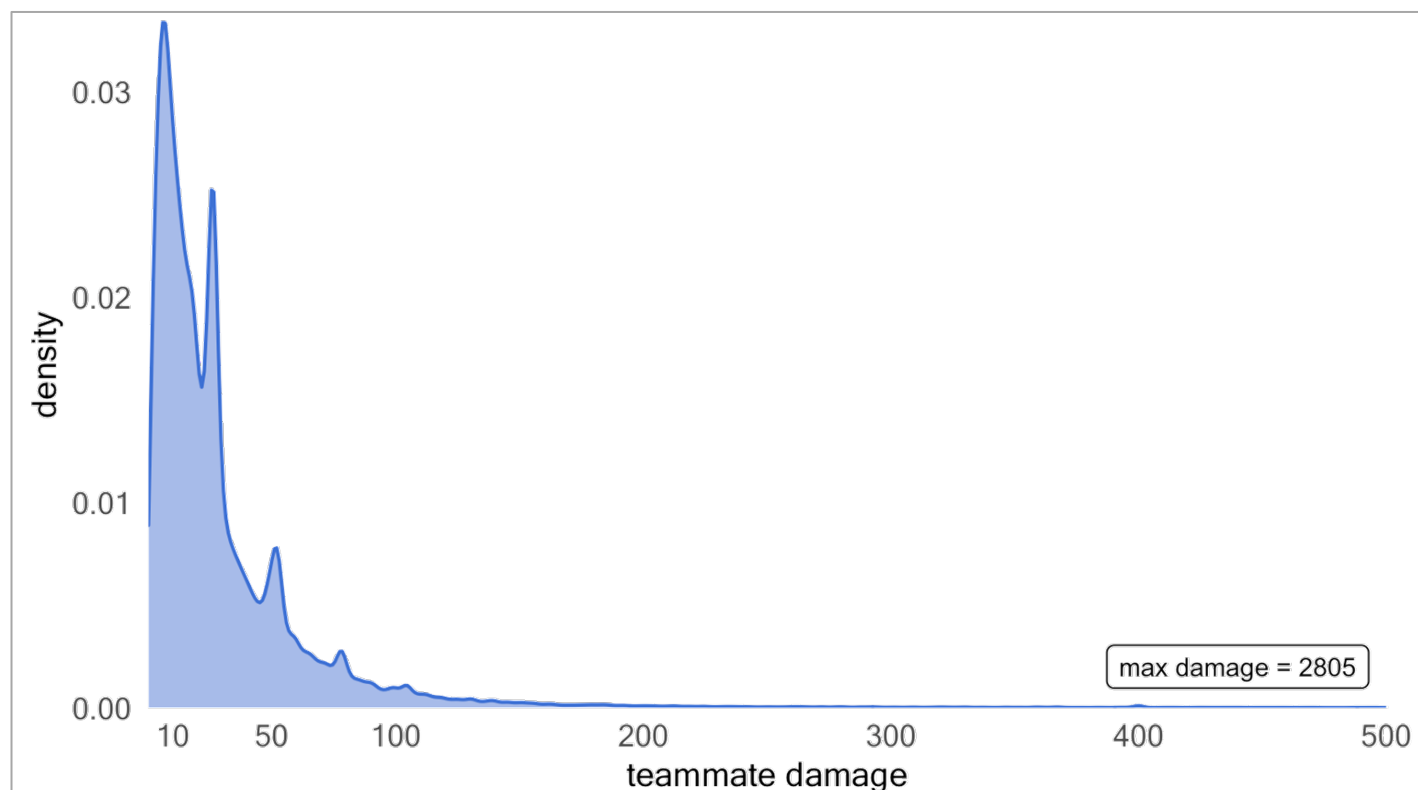
Number Friendly Fires per Player per Match
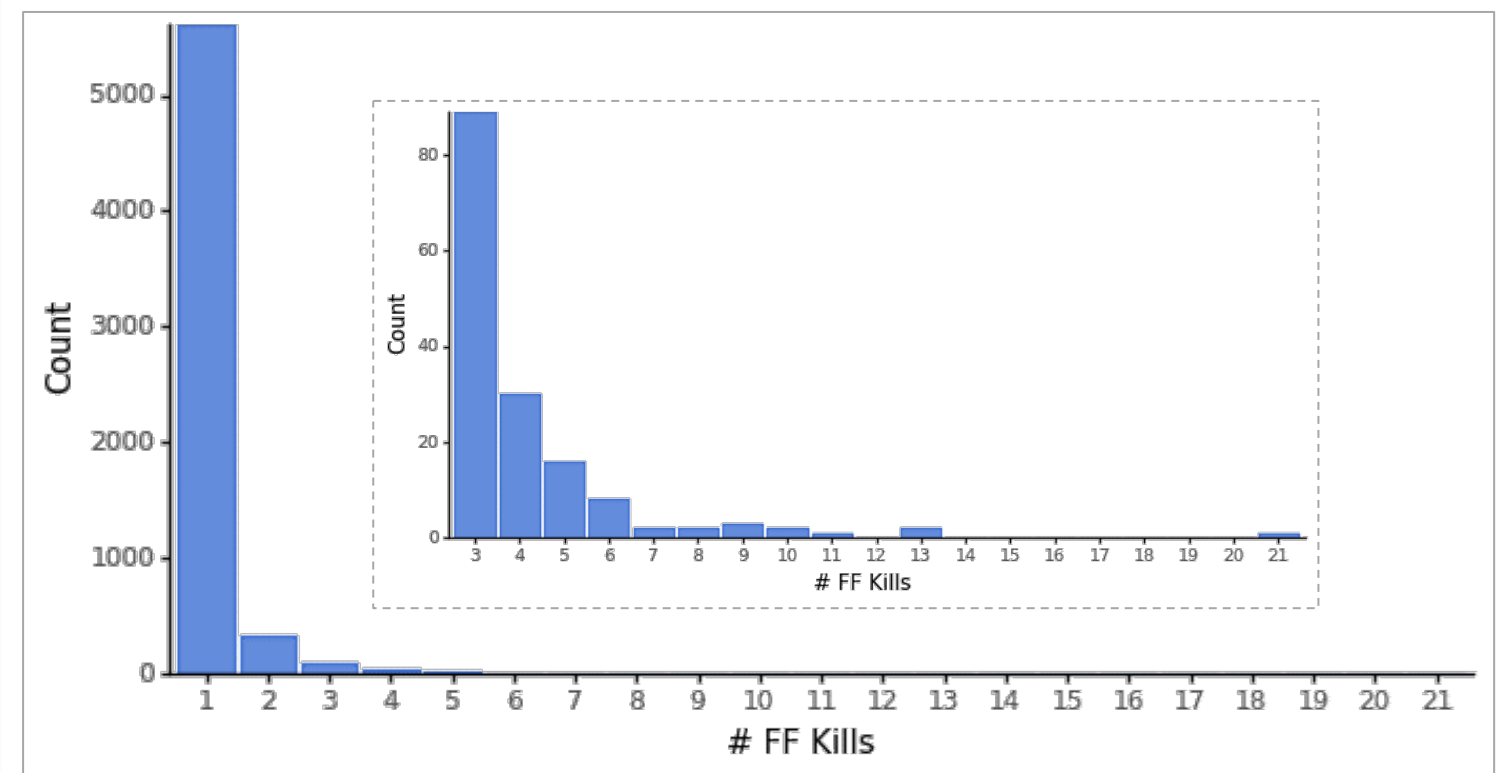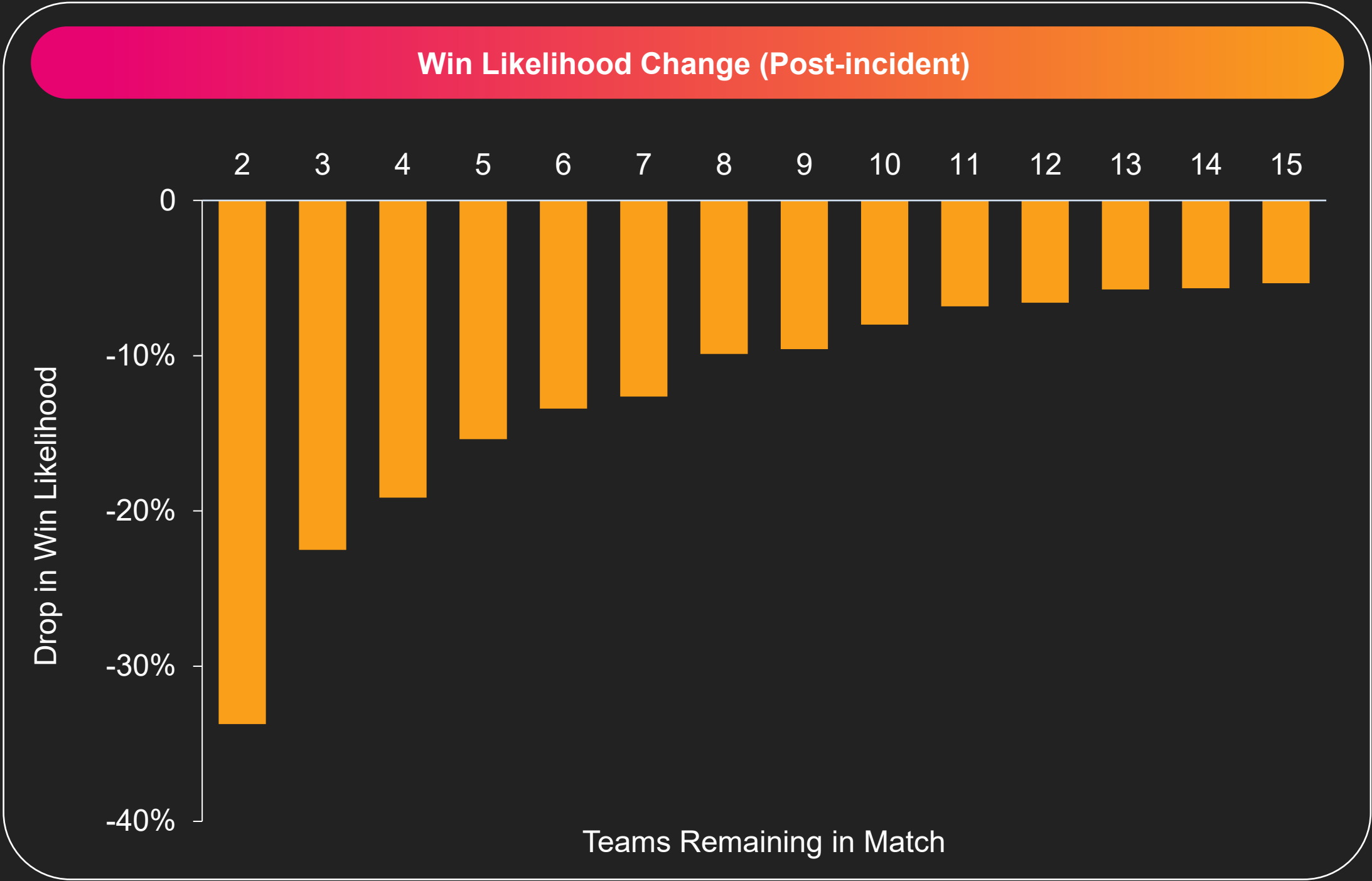
Reputation Subscores

Legend:
- Positive contribution
- Negative contribution
- Cancel-worty
- Descriptive

Net Community Contribution Score: v1

Negative Engagement (PvP)
- Trolling
- Griefing
- AFK
- …
- Toxic chat (via audio transcription)

Cr. Mode: Negative Engagement
- Destruction of Assets
- Trolling

Cr. Mode: Creator Toxicity
- Mature content
- Negative/trolling/toxic content
- PEGI ratings***

Positive Engagement
- Incident-free play
- Encouragement
- Friendliness / inclusiveness

Social Influence & Assertiveness
- Network cluster score
- % of times started party
- # of friends imported
- # of friends made in-game

Red Flag Behavior
- Cheating
- Extreme identity hate
- Doxxing, stalking, grooming

Cr. Mode: Affinity for Toxic Content
- % time toxic content chosen

Cr. Mode: Creator Contribution +
- Others' engagement in content [assuming non-toxic]
- Amount contributed (tier)
- Content Epic wants to encourage (e.g. extra marks for gathering places for concerts)

Engagement Style Segmentation
- Instigator (wolf)
- Passive (sheep)
- Tit-for-Tat (bird)

Social Affinity
- % times solo queue vs in party

Spend
- [for analytics purposes]

Engagement
- How many games, how long, games/session, frequency

Behavior Base Weights

Severity Tier	Time Decay	Example Behaviors	Cross-Game Weight
Extreme (-)	None	Grooming, Credible Threat	High
Gameplay High (-)	Slow	Cheating, Exploits	Med
Social High (-)	Slow	Identity Hate, Scam, Doxxing	High
Medium (-)	Med	Griefing, AFK/DC, Trolling	Med
Low (-)	Fast	Vulgarity, Not following teammates	Low
Very Low (+) [Capped]	Fast	Non-negative engagement	Med
Low Positive (+)	Med	Sharing items, unexpected revives	Med
Med Positive (+)	Slow	Unique positive Commendations	High