

# SUMMARY OF RESPONSES FROM GAMING COMPANIES

February 2023

### **TABLE OF CONTENTS**

- 1. How are you assessing and mitigating the risks and harms of in-game harassment and extremism in your online games? What plans do you have to further address this issue?
- 2. What mechanisms do you have in-game and externally (such as official websites, forums, and support pages) for players to report in-game harassment?
- 3. How are player reports of in-game harassment handled, how large is the team that reviews those reports, and what automated tools are used in-game and by your review team? What investments are you making in ensuring the development and improvement of these systems?
- 4. How do you integrate feedback from groups that represent communities most impacted by online hate and harassment into the process of improving reporting in your games?
- 5. How do you identify extremist content in your games? Can you specify which of your currently existing policies address extremist content?
- 6. What data do you collect on in-game player reporting mechanisms and automatic bans for inappropriate behavior? Will you consider releasing those data in regular transparency reporting?

# 1. How are you assessing and mitigating the risks and harms of in-game harassment and extremism in your online games? What plans do you have to further address this issue?

Activision Blizzard, Inc.	<ul> <li>In-game acknowledgement of code of conduct (Call of Duty, World of Warcraft)</li> <li>Discourages disruptive play and encourages positive play in-game (Overwatch 2)</li> <li>Partnerships with online safety organizations to improve and deploy relevant tools</li> <li>No information on specifically addressing extremism</li> </ul>
Electronic Arts Inc.	<ul> <li>Positive Play Charter explains in accessible language behavioral expectations in-game and consequences for violating those expectations</li> <li>Automated tools combined with human oversight to sweep in-game content</li> <li>No information on specifically addressing extremism</li> </ul>
Epic Games, Inc.	<ul> <li>In-game mechanisms for players to limit communications</li> <li>Extra protections for younger players via Cabined Accounts</li> <li>No information on specifically addressing extremism</li> </ul>
Innersloth LLC	<ul> <li>Extra protections for children under 13 such as randomized player names, banned third-party links in-game, and Quick-Chat Mode</li> <li>Players sanctioned for harassment, sexual misconduct, or illegal activity are permanently banned</li> <li>No information on specifically addressing extremism</li> </ul>
Krafton Inc.	<ul> <li>Player tools to control communication features; proximity chat is off by default</li> <li>Reputation scores assigned to players, encouraging them to act reasonably and responsibly</li> <li>Proactively monitors situations where they anticipate potential disturbances</li> </ul>
Microsoft Gaming	<ul> <li>Players can customize settings like message filters and parental controls</li> <li>Special protections to address bullying and harassment for children on Xbox Network</li> </ul>
Riot Games, Inc.	<ul> <li>Automated systems that reduce problematic behavior in real-time like auto-mute</li> <li>Central Player Dynamics team dedicated to building gaming structures that promote rewarding social interaction and reduce harmful interaction</li> <li>Collaborates with online safety organizations and other game companies on in-game safety</li> </ul>

Roblox Corporation	<ul> <li>Has a Safety Advisory Board composed of leading experts in children's online safety</li> <li>Has a dedicated team reviewing Terrorism and Violent Extremism content</li> <li>Built a "Trust by Design" process to ensure new products meet their safety and security standards</li> <li>Collaborates with online safety organizations on in-game safety</li> </ul>
Sony Interactive Entertainment	<ul> <li>Educates players and their parents (for minors) of their safety features</li> <li>Shared Commitment to Safer Gaming initiative with Xbox and Nintendo based on prevention, responsibility, and partnership</li> </ul>
Square Enix Holdings Co., Ltd.	<ul> <li>In-game chat filter enabled by default; no voice chat is offered in any of their games</li> <li>Community guidelines outline prohibited activities like discrimination and harassment</li> <li>No information on specifically addressing extremism</li> </ul>
Take-Two Interactive Software, Inc.	<ul> <li>Regularly updates filtering systems to prevent inappropriate content and communications</li> <li>Investing in new technologies and in-house teams to address trust and safety</li> <li>Partners with external organizations to develop and support positive gaming communities</li> <li>No information on specifically addressing extremism</li> </ul>
Tencent Holdings Ltd.	<ul> <li>Restrictions on features for minors to reduce the risks of cyberbullying and harassment</li> <li>Notices to players outlining acceptable and unacceptable behavior</li> <li>Automated tools to help detect and prevent inappropriate language from displaying</li> <li>No information on specifically addressing extremism</li> </ul>
Ubisoft Entertainment SA	<ul> <li>Uses proactive (i.e. profanity filters) and reactive (i.e. blocking and muting) systems</li> <li>Enhances solutions based on artificial intelligence (AI) to prevent harmful player interactions</li> <li>Fair Play Program that educates players on in-game disruptive behaviors and their effects</li> <li>Reputation system in Rainbow Six Siege shows players how their ingame actions are perceived</li> <li>No information on specifically addressing extremism</li> </ul>
Valve Corporation	<ul> <li>Automatic filtering for text communications, muting for voice communications</li> <li>Temporary auto-mute for players that receive enough reports for communications abuse</li> <li>No information on specifically addressing extremism</li> </ul>

### 2. What mechanisms do you have in-game and externally (such as official websites, forums, and support pages) for players to report in-game harassment?

Activision Blizzard, Inc.	<ul><li>In-game menus</li><li>Contacting player support</li></ul>
Electronic Arts Inc.	<ul> <li>In-game reporting</li> <li>Report posts in Answers HQ and game forums</li> <li>In-game via EA app or Origin</li> <li>Contact form accessible via EA Help</li> </ul>
Epic Games, Inc.	<ul> <li>In-game options on other players' profiles allow users to report them</li> <li>Player support portals</li> </ul>
Innersloth LLC	<ul> <li>In-game reporting feature conspicuously below the chat button</li> <li>Contacting Innersloth via email, mail, or social media</li> </ul>
Krafton Inc.	<ul> <li>In-game reporting</li> <li>Customer support center at support.pubg.com</li> </ul>
Microsoft Gaming	<ul> <li>Features built into Xbox One, and Xbox Series X S consoles</li> <li>Xbox Social on Windows devices</li> <li>Xbox Windows and mobile apps</li> <li>Xbox 360 console allows reporting recent players</li> </ul>
Riot Games, Inc.	<ul><li>Inside the game client</li><li>Player support website</li></ul>
Roblox Corporation	<ul> <li>In-game reporting</li> <li>Via Roblox Help Center (non-account holders can use this)</li> </ul>
Sony Interactive Entertainment	<ul> <li>Reporting mechanisms for content on PlayStation Network (PSN) are proximate to where the activity takes place</li> <li>Reporting system is available on PS4, PS5, and the PlayStation mobile app</li> </ul>
Square Enix Holdings Co., Ltd.	<ul><li>In-game</li><li>Customer Support website</li></ul>
Take-Two Interactive Software, Inc.	<ul> <li>Reporting tools built directly into in-game menus</li> <li>Dedicated support websites for each game</li> </ul>
Tencent Holdings Ltd.	<ul><li>In-game report button</li><li>Email service@pubgmobile.com</li></ul>
Ubisoft Entertainment SA	<ul><li>Ubisoft Connect (their PC online gaming platform)</li><li>Support portal</li></ul>

	• In-game
Valve Corporation	<ul><li>In-game on a player's profile</li><li>Via their Steam platform</li></ul>

# 3. How are player reports of in-game harassment handled, how large is the team that reviews those reports, and what automated tools are used in-game and by your review team? What investments are you making in ensuring the development and improvement of these systems?

Activision Blizzard, Inc.	<ul> <li>Internal tools and licensed technology from partners to review reports</li> <li>Appeals are reviewed by staff</li> <li>Research investments, such as their two-year project with the California Institute of Technology</li> </ul>
Electronic Arts Inc.	<ul> <li>Dedicated team to review and respond to abuse reports, supported by automation tools and Legal and Security teams</li> <li>Primary focus is to remove the objectionable content and take disciplinary action</li> <li>Updating and improving moderation tools to keep up with the evolving tech, player needs, and nature of disruptive behavior</li> </ul>
Epic Games, Inc.	Trust and Safety Team that has over 1500 specialists to investigate and address reports
Innersloth LLC	<ul> <li>Gathers relevant data surrounding reported action and evaluates through human moderators to account for language and cultural nuances</li> <li>Engages ESRB and gathers community feedback to improve their systems</li> </ul>
Krafton Inc.	<ul> <li>Dedicated teams that vary in size depending on the region and platform</li> <li>Combination of customer service representatives in collaboration with other teams</li> <li>Improving workflows and systems to expedite response and resolution times</li> </ul>
Microsoft Gaming	<ul> <li>Report reviewers provide global coverage 24/7/365</li> <li>Use a variety of tools and techniques tailored to the particular game environment</li> </ul>
Riot Games, Inc.	<ul> <li>Combination of human reviewers and automated systems</li> <li>Manual review process for voice chat abuse</li> <li>Developing an automated voice chat evaluation tool</li> </ul>
Roblox Corporation	<ul> <li>Uses AI to scan reports for actionable content and forward reports for human review</li> <li>Thousands of human reviewers working 24/7/365</li> <li>Investing in new technology to assist automated scanners</li> </ul>
Sony Interactive Entertainment	<ul> <li>All reports undergo review by human moderators</li> <li>Players can access their most recent five minutes of voice chat, and a 20-second clip of this can be attached to the report</li> </ul>

Square Enix Holdings Co., Ltd.	<ul> <li>Investigates chat logs to verify and assess the severity of the infraction</li> <li>Approximately 100 customer support staff dedicated to North America (team size depends on the number of people playing the game)</li> </ul>
Take-Two Interactive Software, Inc.	<ul> <li>Dedicated moderation staff reviews reports</li> <li>Moderation team can remove the violative user-generated content</li> </ul>
Tencent Holdings Ltd.	<ul> <li>Human moderation through community managers and customer services teams</li> <li>Prioritize reports relating to accounts used by minors</li> </ul>
Ubisoft Entertainment SA	<ul> <li>Reports are handled by the Consumer Relationship Center and/or Game Production teams</li> <li>Use internal systems and tools</li> <li>Review player feedback to track trends relating to disruptive behaviors, and to understand the impacts of new tools and systems to mitigate such behaviors</li> </ul>
Valve Corporation	<ul> <li>On Counter-Strike: Global Offensive (CS:GO), a temporary auto-mute is triggered for players that receive enough reports for communications abuse</li> <li>On Steam, a dedicated moderator examines relevant information</li> </ul>

# 4. How do you integrate feedback from groups that represent communities most impacted by online hate and harassment into the process of improving reporting in your games?

Activision Blizzard, Inc.	<ul> <li>Engaged with the Anti-Defamation League (ADL) on their report on hate, harassment, and extremism in online games</li> <li>Supports the Fair Play Alliance and sits on their Executive Steering Committee</li> </ul>
Electronic Arts Inc.	<ul> <li>Convened a summit that brought together an international group of players with online safety leaders and EA's community teams to discuss disruptive behavior in games</li> <li>Engages player feedback and conducts research on game experiences for positive play</li> </ul>
Epic Games, Inc.	Welcomes diversity in race, ethnicity, color, religion, gender identity, sexual orientation, ability, national origin, and other groups
Innersloth LLC	<ul> <li>Gets feedback from the Among Us community, including those most impacted by online hate and harassment, when developing new features</li> <li>Development team is made up of people with diverse backgrounds and perspectives</li> </ul>
Krafton Inc.	Gathers feedback from multiple sources to improve systems
Microsoft Gaming	<ul> <li>Xbox Ambassador community provides updates on player concerns</li> <li>Co-creates with the Gaming for Everyone community to ensure that content is relatable and resonant, which includes getting feedback and data from underrepresented communities</li> </ul>
Riot Games, Inc.	Collaborated with the ADL and Fair Play Alliance on the Disruption and Harms in Online Gaming Framework
Roblox Corporation	Has streamlined reporting channels from subject matter experts
Sony Interactive Entertainment	Integrated feedback from these communities to create their zero tolerance Hate Speech Policy
Square Enix Holdings Co., Ltd.	Gathers feedback from all sources and shares them with internal teams, who determine whether and how to implement features
Take-Two Interactive Software, Inc.	Partners with external organizations to foster safe online communities through education
Tencent Holdings Ltd.	Works with NGOs to make online experiences safer

Ubisoft Entertainment SA	Regularly reviews player feedback from player reports and communities
Valve Corporation	Relies on their employees and users for their input on features

## 5. How do you identify extremist content in your games? Can you specify which of your currently existing policies address extremist content?

Activision Blizzard, Inc.	<ul> <li>No information on specifically identifying extremism</li> <li>Codes of conduct that prohibit hate speech</li> </ul>
Electronic Arts Inc.	<ul> <li>No information on specifically identifying extremism</li> <li>User Agreement that has a section on the Rules of Conduct prohibiting all abusive, offensive, harassing, threatening, or unlawful activity, including hate speech and racially, ethnically, religiously or otherwise offensive content</li> </ul>
Epic Games, Inc.	No information on specifically identifying extremism or policies addressing extremist content
Innersloth LLC	<ul> <li>No information on specifically identifying extremism</li> <li>Code of Conduct and Terms of Use prohibits exploitative, harmful, and objectionable content</li> </ul>
Krafton Inc.	<ul> <li>Reviews player reports for violating content</li> <li>Terms of Service and Rules of Conduct prohibit the use of their games that is illegal or negatively affects other players</li> </ul>
Microsoft Gaming	<ul> <li>Content filtering technology that specifically addresses extremist content</li> <li>Participates in GIFCT's Hash-Sharing Database to identify extremist imagery</li> <li>Code of Conduct in the Microsoft Services Agreement and Xbox Community Standards</li> </ul>
Riot Games, Inc.	<ul> <li>Communications filtering technology and name checking system prevent extremists from communicating and using screen names linked to extremist ideology</li> <li>Terms of Service prohibits harmful content</li> </ul>
Roblox Corporation	<ul> <li>Technology scans and flags material against their Community Standards</li> <li>Community Standards explicitly prohibits terrorism and violent extremism</li> </ul>
Sony Interactive Entertainment	<ul> <li>Communication filtering technology detects and blocks profanity, hate speech, and terms associated with extremist activity in many languages</li> <li>Community Code of Conduct and Terms of Service and User Agreement prohibit offensive user-generated content, including extremist content</li> </ul>
Square Enix	No information on specifically identifying extremism

Holdings Co., Ltd.	User Agreement, Prohibited Activities, Account Penalty Policy, and Forum Guidelines policies prohibit political and religious activities along with other inappropriate behavior
Take-Two Interactive Software, Inc.	<ul> <li>No information on specifically identifying extremism</li> <li>Codes of Conduct in their terms of service prohibits harmful material like hate speech, harassment, and espousing hateful ideologies</li> </ul>
Tencent Holdings Ltd.	<ul> <li>No information on specifically identifying extremism</li> <li>Terms of Service addresses illegal and harmful content</li> </ul>
Ubisoft Entertainment SA	<ul> <li>No information on specifically identifying extremism</li> <li>Terms of Use and Code of Conduct set forth requirements of players playing and/or communicating in games</li> </ul>
Valve Corporation	<ul> <li>No information on specifically identifying extremism</li> <li>Steam Online Conduct rules prohibit harmful behavior</li> </ul>

# 6. What data do you collect on in-game player reporting mechanisms and automatic bans for inappropriate behavior? Will you consider releasing those data in regular transparency reporting?

Activision Blizzard, Inc.	Blog posts contain Call of Duty transparency reports that they are committed to growing
Electronic Arts Inc.	No answer
Epic Games, Inc.	No answer
Innersloth LLC	<ul> <li>Reviews chat snapshot and player names in the room at report time</li> <li>No answer on transparency reporting</li> </ul>
Krafton Inc.	<ul> <li>Information in player reports, in-game events like gameplay patterns, chat, and access logs</li> <li>No answer on transparency reporting</li> </ul>
Microsoft Gaming	<ul> <li>Xbox Transparency Report, goal is to share this work bi-annually</li> <li>Digital Safety Content Report, published bi-annually, covers actions         Microsoft has taken against harmful content and their conduct across         consumer products and services, including Xbox</li> </ul>
Riot Games, Inc.	<ul> <li>Collects comprehensive player behavior data sets</li> <li>Shares learnings and strategies with other companies, and shares progress with the public through published reports like their Update on Player Dynamics</li> </ul>
Roblox Corporation	<ul> <li>Preparing for transparency reporting under regulations both nationally and abroad</li> </ul>
Sony Interactive Entertainment	<ul> <li>Collects data in connection with their safety and moderation practices</li> <li>Open to a dialogue on how data transparency can further player safety and more generally enhance positive experiences for players</li> </ul>
Square Enix Holdings Co., Ltd.	<ul> <li>Looks at logs, posts, and relevant communications</li> <li>Publishes weekly announcements on actions taken against accounts (but does not differentiate between types of violations)</li> </ul>
Take-Two Interactive Software, Inc.	No answer
Tencent Holdings Ltd.	No answer
Ubisoft Entertainment SA	<ul> <li>In-game data for training AI-based moderation tools</li> <li>Making plans relating to online safety, including new projects and transparent reporting</li> </ul>

#### **Valve Corporation**

- Examines Steam profile and recent activity on Steam Community
- No answer on transparency reporting