

Code of Practice on  
Disinformation – Report of  
Twitch for the period  
Jan–Dec 2023

## Table of Content

|  |           |
|--|-----------|
| <b>Executive summary</b>                     | <b>4</b>  |
| <b>Guidelines for filling out the report</b> | <b>5</b>  |
| <b>II. Scrutiny of Ad Placements</b>         | <b>1</b>  |
| Commitment 1                                 | 1         |
| Commitment 2                                 | 4         |
| Commitment 3                                 | 5         |
| <b>III. Political Advertising</b>            | <b>5</b>  |
| Commitment 4                                 | 5         |
| Commitment 5                                 | 6         |
| Commitment 6                                 | 6         |
| Commitment 7                                 | 7         |
| Commitment 8                                 | 8         |
| Commitment 9                                 | 8         |
| Commitment 10                                | 8         |
| Commitment 11                                | 9         |
| Commitment 12                                | 9         |
| Commitment 13                                | 10        |
| <b>IV. Integrity of Services</b>             | <b>10</b> |
| Commitment 14                                | 10        |
| Commitment 15                                | 12        |
| Commitment 16                                | 12        |
| <b>V. Empowering Users</b>                   | <b>13</b> |
| Commitment 17                                | 13        |
| Commitment 18                                | 13        |
| Commitment 19                                | 14        |
| Commitment 20                                | 15        |
| Commitment 21                                | 15        |
| Commitment 22                                | 16        |
| Commitment 23                                | 17        |
| Commitment 24                                | 17        |
| Commitment 25                                | 18        |
| <b>VI. Empowering the research community</b> | <b>19</b> |
| Commitment 26                                | 19        |
| Commitment 27                                | 19        |
| Commitment 28                                | 20        |

|  |           |
|--|-----------|
| Commitment 29  | 20        |
| <b>VII. Empowering the fact-checking community</b>                   | <b>21</b> |
| Commitment 30  | 21        |
| Commitment 31  | 22        |
| Commitment 32  | 23        |
| Commitment 33  | 23        |
| <b>VIII. Transparency Centre</b>                                     | <b>24</b> |
| Commitment 34  | 24        |
| Commitment 35  | 24        |
| Commitment 36  | 25        |
| <b>IX. Permanent Task-Force</b>                                      | <b>25</b> |
| Commitment 37  | 25        |
| <b>X. Monitoring of Code</b>   | <b>26</b> |
| Commitment 38  | 26        |
| Commitment 39  | 26        |
| Commitment 40  | 26        |
| Commitment 41  | 27        |
| Commitment 42  | 27        |
| Commitment 43  | 28        |
| Commitment 44  | 28        |
| <b>Reporting on the service's response during a period of crisis</b> | <b>29</b> |
| [Name of crisis]   | 30        |
| <b>Reporting on the service's response during an election</b>        | <b>34</b> |
| European Elections   | 35        |

## Executive summary

Twitch is a live streaming service, where creators engage in a wide variety of different activities, including video games, music, cooking, and creating creative content.

At Twitch, we strive to create a space that supports and sustains streamers' ability to express themselves, and provides a welcoming and entertaining environment for viewers, free of illegal and harmful interactions. This starts with Twitch's Community Guidelines, which seek to balance user expression with community safety, and set the rules for the behaviour of everyone on Twitch. Our Community Guidelines are developed by a dedicated team of policy professionals in consultation with external safety, human rights, and policy experts, and we review and update them regularly to respond to the community's evolving needs.

We identify potential violations of our Community Guidelines using a combination of machine detection, proactive human review, and user reporting. Our global Customer Trust operations team works to quickly review content and accounts flagged by users and by our machine detection models. The speed at which we can respond to user reports is critical given the live nature of Twitch, and in H1 2023, we responded to 70.0% of reports in under 10 minutes and 89.1% of reports in under an hour. We prioritise having a human in the review process to ensure that decisions are accurate and fair for our community members.

We take pride in how Twitch fosters community and brings people together, and we believe that individuals who use online services to spread false, harmful information do not have a place in our community. This is why we have a Harmful Misinformation Actor policy. Harmful misinformation actors account for a disproportionate amount of damaging, widely debunked misinformation online. These actors share three characteristics: their online presence—whether on or off Twitch—is dedicated to (1) persistently sharing (2) widely disproven and broadly shared (3) harmful misinformation topics, such as conspiracies that promote violence. We prohibit harmful misinformation actors who meet all three of these criteria since taken together they create the highest risk of harm, including inciting real world harm.

Our enforcement numbers under this policy are relatively low due to several factors. (i) The mechanics of Twitch are not conducive to spreading misinformation or investing in large-scale disinformation campaigns. It is extremely difficult for a new streamer to garner large numbers of concurrent viewers; it takes time to grow an audience on Twitch. Most Twitch content is also long-form and ephemeral. Since this means that most content is gone the moment it is created, it is not shared and does not go viral in the same way that it does on other UGC video and social media services, where videos are uploaded and can be viewed and shared by users on demand. (ii) Our targeted policy only applies to those who persistently share harmful misinformation. Due to the long-form nature of Twitch's content, we are focused on a streamer's aggregated content rather than a specific, isolated statement within a longer piece of content. (iii) When we launched our Harmful Misinformation Actor policy, we took swift action against accounts that posed a threat to our community. We believe enforcement of our policy—particularly upon its adoption in H1 2022—has been an effective deterrent to harmful misinformation actors; we have not seen large numbers of them attempt to join our service.

Even if someone is not a Harmful Misinformation Actor, Twitch enforces on misinformation that targets specific communities under our Hateful Conduct & Harassment policies, and we take action

on content that encourages others to engage in physically harmful behaviour under our Self-Destructive Behaviour policy.

In addition to misinformation, Twitch invests significant resources to ban bots, spammers, impersonators, and other types of bad actors to combat inauthenticity on our service. We have proactive detection to work alongside our reporting system to programmatically remove bots, known bad actors, and those who are trying to evade a suspension or ban.

While misinformation is not currently prevalent on Twitch, we recognize the harm that this content can cause, particularly when it is related to an election. We have assembled a cross-functional working group to ensure Twitch is prepared and can stay ahead of potential harms related to elections occurring in the EU and the US in 2024.

We are always evolving our approach to safety in accordance with expert guidance and trends in our community. We understand that the prevalence of harmful misinformation can change, and we will continue to engage with industry, academia, and civil society to adapt our approach as necessary to ensure its continuing effectiveness. We became a signatory to the Australian Code of Practice on Disinformation and Misinformation at the end of 2023 to demonstrate our ongoing commitment to fighting misinformation, and to facilitate better information sharing across experts and industry. We also participate in a variety of other industry knowledge-sharing initiatives, including the New Zealand Code of Practice for Online Safety and Harms (which also addresses disinformation), the EU Hate Speech Code, the EU Internet Forum, and the Global Internet Forum to Counter Terrorism (GIFCT).

The EU Code of Practice on Disinformation provides a useful framework for information sharing and collaboration that will help strengthen industry's abilities to react quickly to the spread of misinformation. As a signatory to the Code, Twitch's goal is to make meaningful contributions and learn from expert organisations and industry peers. We are committed to combating misinformation on Twitch in an effective yet targeted manner that balances freedom of expression with keeping our communities safe.

## II. Scrutiny of Ad Placements

### Commitment 1

Relevant signatories participating in ad placements commit to defund the dissemination of disinformation, and improve the policies and systems which determine the eligibility of content to be monetised, the controls for monetisation and ad placement, and the data to report on the accuracy and effectiveness of controls and services around ad placements.

Measure 1.1

**[adapted]** Relevant Signatories involved in the selling of advertising, inclusive of media platforms, publishers and ad tech companies, will deploy, disclose, and enforce policies with the aims of:

- first avoiding the publishing and carriage of harmful Disinformation to protect the integrity of advertising supported businesses
- second taking meaningful enforcement and remediation steps to avoid the placement of advertising next to Disinformation content or on sources that repeatedly violate these policies

QRE 1.1.1

From Twitch's [Community Guidelines](#):

"We remove users whose online presence is dedicated to (1) persistently sharing (2) widely disproven and broadly shared (3) harmful misinformation topics.

This policy is focused on Twitch users who persistently share harmful misinformation. It will not be applied to users based upon individual statements or discussions that occur on the channel. We will evaluate whether a user violates the policy by assessing both their on-platform behaviour as well as their off-platform behaviour. You can report these actors by sending an email to our internal investigations team with the account name and any available supporting evidence.

Under this policy we cover the following topic areas, and will continue to update this list as new trends emerge:

- Misinformation that targets protected groups, which is already prohibited under our Hateful Conduct & Harassment Policy
- Harmful health misinformation and wide-spread conspiracy theories related to dangerous treatments, COVID-19, and COVID-19 vaccine misinformation
  - Discussions of treatments that are known to be harmful without noting the dangers of such treatments
  - For COVID-19—and any other WHO-declared Public Health Emergency of International Concern (PHEIC)—misinformation that causes imminent physical harm or is part of a broad conspiracy
- Misinformation promoted by conspiracy networks tied to violence and/or promoting violence
- Civic misinformation that undermines the integrity of a civic or political process
  - Promotion of verifiably false claims related to the outcome of a fully vetted political process, including election rigging, ballot tampering, vote tallying, or election fraud
- In instances of public emergencies (e.g., wildfires, earthquakes, active shootings), we may also act on misinformation that may impact public safety"

|   |  |
|---|--|
| SLI 1.1.1 – [adapted] actions taken to enforce each of the policies mentioned in the qualitative part of this service level indicator. This could include, for instance, actions to remove, to block, or to otherwise restrict advertising on pages and/or domains that disseminate harmful Disinformation. | This data represents the total number of indefinite suspensions issued between January and December 2023. Twitch’s enforcement is focused on indefinitely suspending all dedicated misinformation actors and any of their related accounts, and removing any associated content.<br>Indefinite suspensions   |
| Level   | Page/Domain  |
| Data  | 10   |
| Measure 1.2   |  |
| QRE 1.2.1   | Please refer to QRE 1.1.1 which outlines Twitch’s Harmful Misinformation Actor Policy in the Community Guidelines. Actors that systematically provide harmful misinformation are prohibited and are not eligible for monetization.   |
| Measure 1.5   |  |
| QRE 1.5.1   | We are currently accredited by TAG for Brand Safety ( <a href="https://www.tagtoday.net/registry">https://www.tagtoday.net/registry</a> ). As part of this accreditation, Twitch is audited each year by external auditor BPA WW to ensure that all of our digital advertising agreements and practices adhere to brand safety and anti-piracy principles, and we have robust policies and procedures in place to minimise, monitor, and detect ad misplacement. |
| QRE 1.5.2   | Current industry body accreditations do not evaluate specific components under 1.5.<br><br>Twitch does not conduct independent accreditation via audits.   |

## II. Scrutiny of Ad Placements

### Commitment 3

Relevant Signatories involved in buying, selling and placing digital advertising commit to exchange best practices and strengthen cooperation with relevant players, expanding to organisations active in the online monetisation value chain, such as online e-payment services, e-commerce platforms and relevant crowd-funding/donation systems, with the aim to increase the effectiveness of scrutiny of ad placements on their own services.

Measure 3.1

QRE 3.1.1

In 2023, Twitch became a signatory/member to the following initiatives to facilitate information sharing across industry and civil society:

- Australian Code of Practice on Disinformation and Misinformation
- The EU Hate Speech Code
- The Global Internet Forum to Counter Terrorism (GIFCT)

Twitch also continues to participate in the following cooperative initiatives:

- New Zealand Code of Practice for Online Safety and Harms (which also addresses mis/disinformation)

|             |  |
|-------------|--|
|             | <ul style="list-style-type: none"> <li>• The EU Internet Forum (EUIF)</li> <li>• Global Alliance for Responsible Media (GARM)</li> <li>• MediaWise partnership (to develop an array of educational materials that teach Twitch streamers and viewers how to better identify and avoid spreading mis/disinformation online)</li> </ul>  |
| Measure 3.2 |  |
| QRE 3.2.1   | <p>Twitch participated in the Tactics, Techniques, and Procedures working group in 2023 to learn new tactics that the industry is seeing arise from misinformation actors, and used this information to inform our continuous review of our policy and operational guidance.</p> <p>Twitch has only recently become a signatory of the AU Code of Practice, but through our membership in GIFCT and the EUIF, we participated in a number of cross-industry information-sharing sessions, including after the start of the Israel-Hamas conflict in October 2023.</p> <p>As a participant in GARM, Twitch regularly engages with social media platforms and advertising industry members through community and working group meetings.</p> |
| Measure 3.3 |  |
| QRE 3.3.1   | <p>Through the information-sharing initiatives discussed above, Twitch has attended several cross-industry information-sharing sessions. The gathered insights—particularly those related to the political unrest in Brazil and the Israel-Hamas conflict—were converted into training materials for internal operations teams to ensure we were effectively capturing potential harmful misinformation actors that might share such content.</p>  |

## IV. Integrity of Services

### Commitment 14

**[adapted]** In order to limit impermissible manipulative behaviours and practices across their services, Relevant Signatories commit to put in place or further bolster policies to address both misinformation and disinformation across their services, and to agree on a cross-service understanding of manipulative behaviours, actors and practices not permitted on their services. Such behaviours and practices, which should periodically be reviewed in light with the latest evidence on the conducts and TTPs employed by malicious actors, such as the AMITT Disinformation Tactics, Techniques and Procedures Framework, include:

The following TTPs pertain to the creation of assets for the purpose of a disinformation campaign, and to ways to make these assets seem credible:

- 1. Creation of inauthentic accounts or botnets (which may include automated, partially automated, or non-automated accounts)
- 2. Use of fake / inauthentic reactions (e.g. likes, up votes, comments)
- 3. Use of fake followers or subscribers
- 4. Creation of inauthentic pages, groups, chat groups, fora, or domains
- 5. Account hijacking or impersonation

The following TTPs pertain to the dissemination of content created in the context of a disinformation campaign, which may or may not include some forms of targeting or attempting to silence opposing views. Relevant TTPs include:



|              |   |
|--------------|---|
|              | <ul style="list-style-type: none"> <li>● 6. Deliberately targeting vulnerable recipients (e.g. via personalized advertising, location spoofing or obfuscation)</li> <li>● 7. Deploy deceptive manipulated media (e.g. “deep fakes”, “cheap fakes”...)</li> <li>● 8. Use “hack and leak” operation (which may or may not include doctored content)</li> <li>● 9. Inauthentic coordination of content creation or amplification, including attempts to deceive/manipulate platforms algorithms (e.g. keyword stuffing or inauthentic posting/reposting designed to mislead people about popularity of content, including by influencers)</li> <li>● 10. Use of deceptive practices to deceive/manipulate platform algorithms, such as to create, amplify or hijack hashtags, data voids, filter bubbles, or echo chambers</li> <li>● 11. Non-transparent compensated messages or promotions by influencers</li> <li>● 12. Coordinated mass reporting of non-violative opposing content or accounts</li> </ul>   |
| Measure 14.1 |   |
| QRE 14.1.1   | <p>In addition to the Harmful Misinformation Actor policy, as discussed above, Twitch’s policies state that “Any content or activity that disrupts, interrupts, harms, or otherwise violates the integrity of Twitch services or another user’s experience or devices is prohibited.” This includes the creation of inauthentic and malicious bots, engaging in viewership tampering (such as artificially inflating follow or live viewer stats), and selling or sharing user accounts, services, or features. Additionally, Twitch prohibits any “content or activity meant to impersonate an individual or organisation.” Taken together, these policies are aimed at preventing inauthentic or manipulated behaviours/actors on Twitch.</p>   |
| QRE 14.1.2   | <p>Through the mechanisms discussed above, Twitch regularly engages in information sharing with experts and industry to stay abreast of upcoming trends that directly/indirectly involve Twitch. Internal teams also monitor social media for potential bad actors proclaiming a move to Twitch, particularly after being removed from some of the larger services.</p> <p>We use historical enforcement data to proactively identify patterns associated with bots and spammers. Depending on the level of confidence, we can take several actions against a suspected bot account, including requesting that the account verify a mobile phone, auto-reporting the account to be reviewed by our operations team, and adding client-side friction that increases the cost of automation.</p> <p>Most cases of impersonation on Twitch are phishing attempts, where a fraudulent channel is trying to get a user to click on a malicious link. We scan the text on our channel pages for these malicious URLs and then report the channel for review by our operations team and takedown. We also actively monitor channels for viewership tampering, using a combination of handcrafted filters based on ASN and IP reputation, as well as a machine learning model based on past examples.</p> <p>We also conduct audits of suspicious users who may be evading suspensions for botting or spam. If we find that these suspicious users are linked via account metadata, we will take mass action to remove them from our service.</p> |
| Measure 14.2 | <p>[adapted] Relevant Signatories will keep a detailed, up-to-date list of their publicly available policies that clarifies behaviours and practices that are prohibited on their services, and will outline in their reports how their respective policies and their implementation address the above set of TTPs, threats and harms as well as other relevant</p>   |

|   |   |
|---|---|
|   | threats. The list of TTPs will serve as the base for the TTPs to be reported upon and relevant signatories will work within the Permanent Task-force to further develop and refine related indicators on the impact/effectiveness of their related actions.   |
| <b>QRE 14.2.1 [adapted] Report on actions taken to implement the policies they list in their reports and covering the range of TTPs identified/employed</b> | <p>The most relevant TTPs for Twitch are the following:</p> <ul style="list-style-type: none"> <li>• Creation of inauthentic accounts or botnets</li> <li>• Use of fake followers or subscribers</li> <li>• Account hijacking or impersonation</li> </ul> <p>The other TTPs are not tracked by Twitch and do not present a threat on our service. However, Twitch will continue to participate in the working group and monitor for these potential threats.</p> <p>When our content moderation team verifies any of the TTPs above, or a violation of our Harmful Misinformation Actor policy, they issue an enforcement action against the violator's account. The type of enforcement depends on the nature of the violation, and can be a warning, a temporary account suspension (up to 30 days), or for the most serious offences, an indefinite suspension from Twitch. If a record of any content that contains the violation has been saved on our service, we remove it. Violation of our Harmful Misinformation Actor Policy always results in an indefinite suspension.</p> |

**SLI 14.2.1 – SLI 14.2.4**

|  |  |                            |
|--|--|----------------------------|
| <b>TTP OR ACTION 1</b><br>[replicate for number of TTPs or actions reported] | <p>This table will cover:</p> <ul style="list-style-type: none"> <li>• Creation of inauthentic accounts or botnets</li> <li>• Use of fake followers or subscribers</li> <li>• Account hijacking or impersonation</li> </ul> <p>The data provided is the monthly average from January to December 2023. The 'Nr of instances of identified TTPs' are the same as the 'Nr of account enforcements' since these types of violations always result in an account enforcement action.</p> |                            |
|  | <b>SLI 14.2.1 [adapted]</b> Number of instances of identified TTPs and actions taken under policies addressing each of the TTPs as well as information on the type of content.   |                            |
| <b>Data</b>  | Nr of instances of Creation of Inauthentic Accounts or Botnets   | Nr of account enforcements |
|  | 1,133,719  | 1,133,719                  |
|  | Nr of instances of Use of Fake Followers or Subscribers  | Nr of account enforcements |
|  | 8  | 8                          |
|  | Nr of instances of Account Hijacking or Impersonation  | Nr of account enforcements |
|  | 2  | 2                          |

**IV. Integrity of Services**

**Commitment 16**

|   |  |  |                          |
|---|--|--|--------------------------|
| Relevant Signatories commit to operate channels of exchange between their relevant teams in order to proactively share information about cross-platform influence operations, foreign interference in information space and relevant incidents that emerge on their respective services, with the aim of preventing dissemination and resurgence on other services, in full compliance with privacy legislation and with due consideration for security and human rights risks. |  |  |                          |
| Measure 16.1  |  |  |                          |
| QRE 16.1.1  |  | As discussed above, Twitch participates in a number of information-sharing initiatives—in addition to the working groups under the EU Disinformation Code—including the Australian Code of Practice on Disinformation and Misinformation, the New Zealand Code of Practice for Online Safety and Harms, the EU Hate Speech Code, the EU Internet Forum (EUIF), the Global Internet Forum to Counter Terrorism (GIFCT), and the Global Alliance for Responsible Media (GARM).<br><br>These fora have provided useful insights into trending misinformation narratives, language, and keywords; key misinformation actors to monitor if they attempt to make the move to Twitch; and the overlap between mis/disinformation, hate speech, and violent extremist content, particularly far-right extremist content. |                          |
| SLI 16.1.1 – Numbers of actions as a result of information sharing  |  | Twitch has not made any specific changes to our policy, nor have we been able to identify any Twitch accounts that violate our Harmful Misinformation Actor policy based on cross-platform information sharing. However, these conversations have informed the guidance that we provide to our operational teams and the trends that we investigate.   |                          |
|   |  | Nr of actions taken (total)  | Type of detected content |
| Data  |  | 0  | N/A                      |
| Measure 16.2  |  |  |                          |
| QRE 16.2.1  |  | We monitor misinformation actors closely across the online ecosystem, and this year, identified at least one misinformation actor attempting to move to our service after having their accounts suspended elsewhere. We have also seen a number of harmful misinformation actors sharing misinformation narratives common across other services. Over half of the misinformation actors suspended in the past year promoted the core tenets of the violent conspiracy network, QAnon.  |                          |

## V. Empowering Users

### Commitment 17

In light of the European Commission’s initiatives in the area of media literacy, including the new Digital Education Action Plan, Relevant Signatories commit to continue and strengthen their efforts in the area of media literacy and critical thinking, also with the aim to include vulnerable groups.

|              |  |
|--------------|--|
| Measure 17.1 |  |
|--------------|--|

|   |  |
|---|--|
| <p><b>QRE 17.1.1 [adapted] Relevant Signatories will outline the tools they develop or maintain that are relevant to this commitment and report on their deployment.</b></p>  | <p>Twitch collaborated with media literacy expert MediaWise to develop an array of educational materials that teach Twitch streamers and viewers how to better identify, and avoid spreading, misinformation and disinformation online. These materials are hosted on a page on the Twitch Safety Center.</p> <p>Additionally, in 2023, Twitch partnered with UC Irvine's Connected Learning Lab (CLL) to do a research study on prosocial behaviour. The research aimed to support and enable prosocial behaviour and the growth of positive communities through outreach, education, and recognition. Internal teams are currently processing the report recommendations and determining the appropriate next steps.</p> |
| <p><b>SLI 17.1.1 - [adapted] Relevant Signatories will report on metrics pertinent to assessing the effects of the tools described in the qualitative reporting element for Measure 17.1, which will include: the total count of impressions of the tool; and information on the interactions/engagement with the tool.</b></p> | <p>We measure interactions or engagement as the number of users who visited the page with the media literacy educational materials on the Twitch Safety Center between January and December 2023.</p>  |
| <p><b>Data</b></p>  | <p>Interactions/ engagement with the tool</p>  |
| <p>Measure 17.2</p>   | <p>1,459 interactions</p>  |
| <p><b>QRE 17.2.1 [adapted] Relevant Signatories will describe the activities they launch or support. Relevant signatories will further report on actions taken to promote the campaigns to their user base.</b></p>   | <p>Twitch collaborated with MediaWise to develop an array of media literacy materials, and is continuing to explore educational and research opportunities related to misinformation.</p>  |

|  |  |
|--|--|
| <p><b>V. Empowering Users</b></p>  |  |
| <p><b>Commitment 18</b></p>  |  |
| <p>Relevant Signatories commit to minimise the risks of viral propagation of Disinformation by adopting safe design practices as they develop their systems, policies, and features.</p> |  |
| <p>Measure 18.2</p>  |  |
| <p><b>QRE 18.2.1</b></p>   | <p>Persistent sharing of harmful misinformation is a violation of our Harmful Misinformation Actor policy (<a href="#">link</a>), and will result in a permanent ban from the service. If alternate accounts are created to avoid this ban, those accounts will be immediately removed.</p>  |
| <p><b>SLI 18.2.1 - actions taken in response to policy violations</b></p>  | <p>Any violation of our Harmful Misinformation Actor policy results in a permanent ban from the service. The accounts actioned under this policy in 2023 were smaller accounts that did not have a meaningful number of followers, subscribers, or engagement on Twitch. This does not mean these actors were not larger on other platforms but their Twitch presence was small. As such, we are not able to track meaningful metrics to measure the impact of removing these actors from our service.</p> <p>The numbers below reflect the total from January to December 2023.</p> |

|      | Total no of violations    | Metric 1: indicating the impact of the action taken | Metric 2: indicating the impact of the action taken | Metric 3: indicating the impact of the action taken |
|------|---------------------------|---|---|---|
| Data | 10 Indefinite Suspensions | N/A   | N/A   | N/A   |

| <b>V. Empowering Users</b>   |  |
|--|--|
| <b>Commitment 19</b>   |  |
| <p><b>[adapted]</b> Relevant Signatories using recommender systems commit to make them transparent to the recipients regarding how the algorithms are expected to behave and the outcomes prioritised and deprioritized on behalf of users, and provide options to users about recommender systems, and make available information on those options.</p> |  |
| Measure 19.1   | <b>[adapted]</b> Relevant Signatories will make available to their users, in a clear, accessible and easily comprehensible manner, information outlining the main behaviours recommenders use and optimise for.  |
| <b>QRE 19.1.1 [adapted] Relevant Signatories will provide details of the policies and measures put in place to implement the above-mentioned measures accessible to EU users, especially by publishing information outlining the main behaviours recommenders use and optimise for.</b>  | Twitch published a detailed summary of our various recommendation systems, outlining the main parameters used by each system ( <a href="#">link</a> ). The page also provides information for how users can influence our recommendations and control what they see on Twitch.   |
| Measure 19.2   |  |
| <b>SLI 19.2.1 – user settings</b>  | <p>Users can customise their recommendations on Twitch by letting Twitch know if they are “not interested” in a streamer or content category that is recommended to them. At any time, users can navigate to their settings page and review what they have marked as “not interested” and then edit those selections. Below is the number of times users navigated to the settings page per month, on average, between January and December 2023.</p> <p>When browsing, viewers can sort by ‘Recommended for You’ or by other channel attributes, such as the number of viewers. We are not able to report on the number of users who chose to browse under the various options.</p> |
|  | No of times users actively engaged with the settings page  |
| Data   | Approximately 733,00 views per month   |

| <b>V. Empowering Users</b>   |  |
|--|--|
| <b>Commitment 23</b>   |  |
| <p>Relevant Signatories commit to provide users with the functionality to flag harmful false and/or misleading information that violates Signatories policies or terms of service.</p> |  |

|                   |  |
|-------------------|--|
| Measure 23.1      |  |
| <b>QRE 23.1.1</b> | Users are able to report potential harmful misinformation actors with an easy, direct reporting mechanism. Users can submit a report by clicking on the three vertical dots icon, which is shown in the bottom right below the video player on the channel page; on the bottom right of a clip, highlight or past broadcast; or on the bottom right when you click on a username to report. An option to “Report Live Stream” or “Report [username]” will appear for the user to enter our reporting interface. The interface prompts the user to select the most relevant category for the violation, which in this case is “Misinformation.” Alternatively, a user can search for the appropriate reporting reason. Before submitting the report, an open text field allows users to submit additional text evidence if desired. Users can also submit a report to the specialised off-service investigations team through the team’s email alias. |
| Measure 23.2      |  |
| <b>QRE 23.2.1</b> | Twitich employs extensive human review to help ensure any enforcement action we take remains accurate and fair for our community members. Our content moderation team reviews content and accounts reported by users and flagged by our machine detection tools. This team, which is distributed around the world, provides 24/7/365 capacity to review and quickly respond to reports. Reports are prioritised so that the most harmful behaviour can be handled most quickly.  |

## V. Empowering Users

### Commitment 24

Relevant Signatories commit to inform users whose content or accounts has been subject to enforcement actions (content/accounts labelled, demoted or otherwise enforced on) taken on the basis of violation of policies relevant to this section (as outlined in Measure 18.2), and provide them with the possibility to appeal against the enforcement action at issue and to handle complaints in a timely, diligent, transparent, and objective manner and to reverse the action without undue delay where the complaint is deemed to be founded.

|   |   |
|---|---|
| Measure 24.1  |   |
| <b>QRE 24.1.1 [adapted] Relevant Signatories will report on the availability of their notification and appeals systems and provide details on the steps of the appeals procedure.</b> | When a user violates our Harmful Misinformation Actor Policy—or any of our policies—they receive a detailed email notification. The notification includes the action taken, whether a suspension is permanent or temporary, the reason for the suspension, examples of violating content, a link to the Community Guidelines to learn more about the policy, where the violation occurred, and a link to the Appeals Portal if they disagree with the decision.<br><br>Appeals can be filed via our appeals portal. The appeals portal provides visibility into the enforcements that are eligible for appeal, and displays the status and outcome of ongoing and prior requests. |
| <b>SLI 24.1.1 - enforcement actions</b>   | To qualify as a harmful misinformation actor, you must meet the criteria outlined in our Harmful Misinformation Actor policy. Your online presence must be dedicated to: (1) persistently sharing (2) widely disproven and broadly shared (3) harmful misinformation topics.  |

|             |   |                        |                               |   |
|-------------|---|------------------------|-------------------------------|---|
|             | The data below is the total from January to December 2023. Note that the appeals number includes those that were filed for enforcements issued in 2022, which is why the number of appeals is higher than the number of enforcements. |                        |                               |   |
|             | Nr of enforcement actions   | Nr of actions appealed | Metrics on results of appeals | Metrics on the duration and effectiveness of the appeal process |
| <b>Data</b> | 10 Indefinite Suspensions   | 22 appeals submitted   | 1 appeal accepted             | 4.55% acceptance rate   |

| <b>VI. Empowering the research community</b>  |  |
|---|--|
| <b>Commitment 27</b>  |  |
| [adapted] Relevant Signatories commit to provide vetted researchers with access to data necessary to undertake research on Disinformation |  |
| Measure 27.4  |  |
| <b>QRE 27.4.1</b>   | We support misinformation research through our membership in the Australian Digital Industry Group (DIGI), who last year organised a misinformation-focused research symposium. We also remain open to supporting independent research if approached. At this time, Twitch does not directly support any third-party research. |

| <b>VIII. Transparency Centre</b>  |  |
|---|--|
| <b>Commitment 34</b>  |  |
| To ensure transparency and accountability around the implementation of this Code, Relevant Signatories commit to set up and maintain a publicly available common Transparency Centre website. |  |
| Measure 34.1  |  |
| Measure 34.2  |  |
| Measure 34.3  |  |
| Measure 34.4  |  |
| Measure 34.5  |  |

| <b>VIII. Transparency Centre</b>  |   |
|---|---|
| <b>Commitment 35</b>  |   |
| Signatories commit to ensure that the Transparency Centre contains all the relevant information related to the implementation of the Code's Commitments and Measures and that this information is presented in an easy-to-understand manner, per service, and is easily searchable. |   |
| Measure 35.1  |   |
| Measure 35.2  | <b>[adapted]</b> Signatories provide information on the implementation and enforcement of their policies per service.   |
| Measure 35.3  |   |
| Measure 35.4  |   |
| Measure 35.5  |   |
| Measure 35.6  | <b>[adapted]</b> The Transparency Centre will enable users to easily access and understand the Service Level Indicators and Qualitative Reporting Elements tied to each Commitment and Measure of the Code for each service, in a standardised and searchable way. The Transparency Centre should also enable users to easily access and understand Structural Indicators for each Signatory. |

| <b>VIII. Transparency Centre</b>  |   |
|---|---|
| <b>Commitment 36</b>  |   |
| Signatories commit to updating the relevant information contained in the Transparency Centre in a timely and complete manner.   |   |
| Measure 36.1  |   |
| Measure 36.2  |   |
| Measure 36.3  |   |
| <b>QRE 36.1.1 (for the Commitments 34-36)</b>   | The administration of the Transparency Centre website has been transferred fully to the community of the Code's signatories, with VOST Europe taking the role of developer.   |
| <b>QRE 36.1.2 (for the Commitments 34-36)</b>   |   |
| <b>SLI 36.1.1 - (for Measures 34 and 36) meaningful quantitative information on the usage of the Transparency Centre, such as the average monthly visits of the webpage</b> | The common Transparency Center has received around 35.000 views between July 1st 2023 and December 31st 2023. The average engagement time on the website is 1m11s and reports have been downloaded more than 9.000 times. |

| <b>IX. Permanent Task-Force</b> |  |
|---------------------------------|--|
| <b>Commitment 37</b>            |  |



|  |   |
|--|---|
| Signatories commit to participate in the permanent Task-force. The Task-force includes the Signatories of the Code and representatives from EDMO and ERGA. It is chaired by the European Commission, and includes representatives of the European External Action Service (EEAS). The Task-force can also invite relevant experts as observers to support its work. Decisions of the Task-force are made by consensus. |   |
| Measure 37.1   |   |
| Measure 37.2   | <p><b>[adapted]</b> Signatories agree to work in the Task-force in particular – but not limited to – on the following tasks:</p> <ul style="list-style-type: none"> <li>- Establishing a risk assessment methodology and a rapid response system to be used in special situations like elections or crises.</li> <li>- Cooperate and coordinate their work in special situations like elections or crisis</li> <li>- Agree on the harmonised reporting templates for the implementation of the Code’s Commitments and Measures, the refined methodology of the reporting, and the relevant data disclosure for monitoring purposes.</li> <li>- Review the quality and effectiveness of the harmonised reporting templates, as well as the formats and methods of data disclosure for monitoring purposes, throughout future monitoring cycles and adapt them, as needed.</li> <li>- Contribute to the assessment of the quality and effectiveness of Service Level and Structural Indicators and the data points provided to measure these indicators, as well as their relevant adaptation.</li> <li>- Refine, test and adjust Structural Indicators and design mechanisms to measure them.</li> <li>- Agree, publish and update a list of TTPs employed by malicious actors, and set down baseline elements, objectives and benchmarks for Measures to counter them, in line with the Chapter IV of this Code.</li> <li>- Seek out and discuss research, expert input and up-to-date evidence relevant to the Code’s commitments, such as, <i>inter alia</i>, emerging best practices in safe design, retroactive flagging, repository of fact-checks, provenance tools.</li> <li>- Discuss and provide guidance on the adequate quantitative information to be provided by signatories to fulfil their reporting obligations regarding agreements with fact-checking organisations across different services.</li> <li>- Regularly discuss whether the Code’s Commitments and Measures need updating in view of technological, societal, market and legislative developments, as well as in view of accommodating new signatories and, where the Task-force agrees to be necessary, carry out such updates.</li> <li>- Review the appropriateness and consistency of adapted Measures for smaller or emerging services.</li> <li>- Promote the Code among relevant peers and integrate new Signatories to the Code.</li> </ul> |
| Measure 37.3   |   |
| Measure 37.4   |   |
| Measure 37.5   |   |

|              |  |
|--------------|--|
| Measure 37.6 |  |
| QRE 37.6.1   | Twitch focuses on actor-level misinformation, and our data and monitoring suggest that misinformation does not show up on our service the way it does on other services. As a Non-VLOP, Twitch participates in the working groups in a learning capacity to be aware of what the rest of the signatories are facing so we can take proactive action if needed. |

|   |  |
|---|--|
| <b>X. Monitoring of Code</b>  |  |
| <b>Commitment 38</b>  |  |
| The Signatories commit to dedicate adequate financial and human resources and put in place appropriate internal processes to ensure the implementation of their commitments under the Code. |  |
| Measure 38.1  |  |
| QRE 38.1.1 [adapted] Relevant Signatories will outline the teams and internal processes they have in place, per service, to comply with the Code.   | <ul style="list-style-type: none"> <li>• Internal Teams <ul style="list-style-type: none"> <li>○ Customer Trust: Policy, Trust &amp; Safety Operations, Risk &amp; Strategic Response</li> <li>○ Product: Community Health</li> <li>○ Legal: Public Policy</li> </ul> </li> <li>• Twitch has committed to and implemented regular internal audits of yearly transparency reports to inform scoping of resources needed to complete various measures outlined in the code.</li> </ul> |

|   |  |
|---|--|
| <b>X. Monitoring of Code</b>  |  |
| <b>Commitment 39</b>  |  |
| Signatories commit to provide to the European Commission, within 1 month after the end of the implementation period (6 months after this Code's signature) the baseline reports as set out in the Preamble. |  |

|                              |  |
|------------------------------|--|
| <b>X. Monitoring of Code</b> |  |
| <b>Commitment 40</b>         |  |

Signatories commit to provide regular reporting on Service Level Indicators (SLIs) and Qualitative Reporting Elements (QREs). The reports and data provided should allow for a thorough assessment of the extent of the implementation of the Code's Commitments and Measures by each Signatory, service and at Member State level.

|              |   |
|--------------|---|
| Measure 40.2 | <b>[adapted]</b> Other Signatories will report yearly on the implementation of the Commitments and Measures taken under the present Code, including on the relevant QREs and SLIs.  |
| Measure 40.3 |   |
| Measure 40.4 |   |
| Measure 40.5 |   |
| Measure 40.6 | <b>[adapted]</b> Signatories will cooperate with the European Commission, respond to its reasonable requests and provide the European Commission with reasonable information, data and further input necessary to assess the implementation of the Code, allowing for the Code's efficient and thorough monitoring. |

## X. Monitoring of Code

### Commitment 42

Relevant Signatories commit to provide, in special situations like elections or crisis, upon request of the European Commission, proportionate and appropriate information and data, including ad-hoc specific reports and specific chapters within the regular monitoring, in accordance with the rapid response system established by the Taskforce.

## X. Monitoring of Code

### Commitment 43

Relevant Signatories commit to provide, in special situations like elections or crisis, upon request of the European Commission, proportionate and appropriate information and data, including ad-hoc specific reports and specific chapters within the regular monitoring, in accordance with the rapid response system established by the Taskforce.