# Strava Terms of Service Report 3rd Quarter 2023

Strava values transparency, including in its content moderation policies and practices, and welcomes the opportunity to provide this information to the California Attorney General's office in line with Assembly Bill No. 587. As Strava builds the platform for active people, we optimize for connecting people to what motivates them and helps them find their personal best. We created terms and standards to support an inclusive, safe, and welcoming space for active people to strive, and aim to keep Strava accessible to the entire Strava community provided they respect those rules.

Below is information relevant to Strava regarding policies, reporting, and enforcement under those rules, corresponding to AB 587, <u>section 22677</u>:

# (1) Current terms of service

Strava's Terms of Service are available at <a href="https://www.strava.com/legal/terms">https://www.strava.com/legal/terms</a>. They state, "For questions regarding the Services and these Terms, please contact us via <a href="https://support.strava.com">https://support.strava.com</a>."

# (2) First report

This is Strava's first Terms of Service report.

# (3) Specified categories of content in terms of service

The current version of the terms of service defines the following categories of content as follows:

#### (A) Hate speech or racism.

- is illegal or promotes illegal activities, including in relation to export control or sanctions laws
- is threatening, harassing, abusive, hateful, discriminatory, or that advocates violence
- may be harmful to minors
- significantly or repeatedly disrupts the experience of other users

Strava's Community Standards further state

- "Be inclusive and anti-racist. Strava is committed to actively dismantling and eradicating racism and discrimination in all forms. If you use Strava, you're joining us in this commitment, without exception."
- "Don't use hate speech. Hate speech is a direct attack based on race, ethnicity, age, national origin, religious affiliation, sexual orientation, caste, sex, gender, gender identity, serious disease, disability, body type, or immigration status. We will remove hate speech and may suspend or terminate the accounts of anyone who uses it."

#### (B) Extremism or radicalization.

Although our terms and policies do not specifically use the words *extremism* or *radicalization*, if this content were to appear on the platform, it would violate one or more of the following prohibitions in our terms:

- is illegal or promotes illegal activities, including in relation to export control or sanctions laws
- is threatening, harassing, abusive, hateful, discriminatory, or that advocates violence
- may be harmful to minors
- significantly or repeatedly disrupts the experience of other users

#### (C) Disinformation or misinformation.

- is fraudulent or contains false, deceptive or misleading statements, claims or representations—including misinformation, or impersonation such as of a person, company, or brand
- is inauthentic, including manipulating data used for leaderboards
- is intended to or likely to deceive, mislead or defraud others through phishing or social engineering techniques

#### (D) Harassment.

- is threatening, harassing, abusive, hateful, discriminatory, or that advocates violence
- may be harmful to minors
- significantly or repeatedly disrupts the experience of other users

#### (E) Foreign political interference.

Although our terms and policies do not specifically use the words *foreign political interference*, if this content were to appear on the platform, it would violate one or more of the following prohibitions in our terms:

- is inauthentic, including manipulating data used for leaderboards
- is intended to or likely to deceive, mislead or defraud others through phishing or social engineering techniques.
- collecting or harvesting any information relating to an identified or identifiable individual, including account names and information about users of the Services, from the Services

#### (4) Content moderation practices

- (A) Any existing policies intended to address the categories of content described in paragraph (3).
  - Strava's Acceptable Use Policies describe the content and conduct restrictions under Strava's Terms of Service. Strava's Community Standards further elaborate on some of those restrictions, including Strava's commitment to Anti-Racism, which our users must respect.
  - Please see previous section ((3) Specified categories of content in terms of service) for relevant excerpts from those terms and policies.
- (B) How automated content moderation systems enforce terms of service of the social media platform and when these systems involve human review.
  - All content moderation to enforce our Terms of Service, including our Acceptable
    Use Policies, currently happens by humans. Where we use automation, this is solely
    to flag content for internal, human review.
- (C) How the social media company responds to user reports of violations of the terms of service.
  - As we explain in our <u>Community Standards</u>, as well as in our <u>Help Center</u>, if users believe content on Strava violates our terms of service, they can follow these <u>instructions</u> on how to report content directly from a particular piece of content (comment, post, etc.) for review, or they can <u>contact us directly</u> from the support site, mobile app, or by email. As we note in our Community Standards and Help Center, we take user reports seriously and will do our best to understand context and take appropriate actions based on our policies and the results of our investigation.
- (D) How the social media company would remove individual pieces of content, users, or groups that violate the terms of service, or take broader action against individual users or against groups of users that violate the terms of service.

- As noted in Strava's Terms of Service, we may remove any Content that violates the Terms or is otherwise objectionable, and we may suspend or terminate a user's account and/or access to the Strava services for violations of the Terms or other incorporated agreements, policies, or guidelines. This includes Strava's Acceptable Use Policies, which also state that Strava may remove content and/or to suspend or block account access due to a violation of this Acceptable Use Policy. In addition, we note in our <u>instructions</u> for reporting content that we may delete or hide content found to be in violation of our Terms and Community Standards, as well as take action on violating athletes, if necessary.
- Relatedly, in our Acceptable Use Policy, as well as in the notification of violation sent to users, we inform them of their right to appeal: If you believe your account or content has been disabled in error and you wish to appeal that decision by providing information explaining why you believe there was a mistake, you can contact us via https://support.strava.com.
- (E) The languages in which the social media platform does not make terms of service available, but does offer product features, including, but not limited to, menus and prompts.
  - Strava's Terms of Service are available in the same languages in which we offer product features, including, but not limited to, menus and prompts.

# (5) Information on content in the specified categories of content terms of service

Below is data for reports, actions, and appeals in the third quarter (July to September) of 2023 for the relevant content violations:

(i) Flagged (reported) content

#### **CATEGORY OF CONTENT**

| (A) Hate speech or racism            | 1480 |
|--------------------------------------|------|
| (B) Extremism or radicalization      | n/a  |
| (C) Disinformation or misinformation | 0    |
| (D) Harassment                       | 1761 |

| (E) Foreign political interference | n/a  |
|------------------------------------|------|
| TOTAL                              | 3241 |

# **TYPE OF CONTENT**

| Activity | 5    |
|----------|------|
| Club     | 2    |
| Comment  | 1497 |
| Other    | 24   |
| Post     | 290  |
| Profile  | 9    |
| Segment  | 795  |
| Image    | 566  |
| Untagged | 7    |
| Video    | 46   |
| TOTAL    | 3241 |

# **TYPE OF MEDIA**

| Image   | 566  |
|---------|------|
| Text    | 2622 |
| Unknown | 7    |
| Video   | 46   |
| TOTAL   | 3241 |

# **FLAGGED BY**

| Users             | 3229 |
|-------------------|------|
| Company employees | 12   |
| TOTAL             | 3241 |

# (ii) Actioned content

# **CATEGORY OF CONTENT**

| (A) Hate speech or racism            | 249 |
|--------------------------------------|-----|
| (B) Extremism or radicalization      | n/a |
| (C) Disinformation or misinformation | 0   |
| (D) Harassment                       | 218 |
| (E) Foreign political interference   | n/a |
| TOTAL                                | 467 |

# **TYPE OF CONTENT**

| Activity | 3   |
|----------|-----|
| Club     | 1   |
| Comment  | 325 |
| Other    | 1   |
| Post     | 64  |
| Profile  | 2   |
| Segment  | 36  |
| Image    | 6   |
| Untagged | 29  |
| Video    | 0   |
| TOTAL    | 467 |

# **TYPE OF MEDIA**

| Image   | 6   |
|---------|-----|
| Text    | 432 |
| Unknown | 29  |

| Video | 0   |
|-------|-----|
| TOTAL | 467 |

In all cases, the content was actioned by company employees.

Please note, where we receive multiple reports of the same content/user, it will show as one item of content actioned.

(iii) Actioned items of content that resulted in action taken against the user or group of users responsible for the content.

In all cases where content was actioned, we took action against the user responsible for the content. Thus, the numbers in (ii) Actioned content would apply here.

- (iv) Actioned items of content removed, demonetized, or deprioritized In all cases where we took action, we removed content (for example, removing a comment or suspending access to an account). Thus, the numbers in (ii) Actioned content would apply here.
- (v) The number of times actioned items of content were viewed by users. In the past, we looked into using views (impressions) as a metric to assess abusive content on our platform. Our analyses showed that impressions and abuse reports were highly correlated, so we shifted our focus to abuse reports. Thus, the numbers in (i) Flagged (reported) content would provide an indication of this metric.
  - (vi) The number of times actioned items of content were shared, and the number of users that viewed the content before it was actioned.

As noted above, our past analyses indicated a high correlation between impressions and abuse reports, so we shifted our focus to abuse reports. The numbers in (i) Flagged (reported) content would provide an indication of this metric.

## (vii) Appeals and reversals

#### CATEGORY OF CONTENT

| (A) Hate speech or racism       | 0   |
|---------------------------------|-----|
| (B) Extremism or radicalization | n/a |
| (C) Disinformation or           | 0   |

| misinformation                     |     |
|------------------------------------|-----|
| (D) Harassment                     | 4   |
| (E) Foreign political interference | n/a |
| TOTAL                              | 4   |

## **TYPE OF CONTENT**

| Activity | 1 |
|----------|---|
| Club     | 0 |
| Comment  | 1 |
| Other    | 0 |
| Post     | 0 |
| Profile  | 1 |
| Segment  | 1 |
| Image    | 0 |
| Untagged | 0 |
| Video    | 0 |
| TOTAL    | 4 |

# **TYPE OF MEDIA**

| Image | 0 |
|-------|---|
| Text  | 4 |
| Video | 0 |
| TOTAL | 4 |

Thank you again for the opportunity to provide this information to you regarding our content moderation policies, reporting, and enforcement.