



Truecaller Annual Transparency Report Under the DSA

This transparency report has been prepared by True Software Scandinavia AB (“Truecaller”) in accordance with Article 15(1) of Regulation (EU) 2022/2065 (the “DSA”), outlining Truecaller’s content moderation activities during the period of January 1 to December 31, 2024.

1. Reporting period

- **Date of the Publication of the Report:** 2025-02-14
- **Date of the Publication of the Latest Previous Report:** This is the first transparency report.
- **Starting Date of Reporting Period:** 2024-01-01
- **Ending Date of Reporting Period:** 2024-12-31

2. Article 15(1)(a). Member State Orders

- **Number of Orders to Act Against Illegal Content Received:** During the relevant period, Truecaller did not receive any orders from authorities in EU Member States to take action against illegal content.
- **Number of Orders Processed in Total:** N/A
- **Median Time to Inform the Authority of the Receipt of the Order:** N/A
- **Median Time to Give Effect to the Order:** N/A
- **Categorization by Type of Illegal Content:** N/A

3. Article 15(1)(b). Notices Submitted

- **Number of Notices Received:** A total of 8,722 profiles were reported by Truecaller users.
- **Number of Notices Received from Trusted Flaggers:** During the relevant period, Truecaller did not receive any notices from trusted flaggers.
- **Median Time to Take Action:** 24 hours
- **Categorization by Type of Alleged Illegal Content:**
 - **Nudity or Sexual Content:** One profile reported under nudity or sexual content.
 - **Violence:** No reports involved violent content.

truecaller

- **Terrorism:** No reports involved terrorism.
- **Famous Person:** No reports involved impersonation of a famous person.
- **Impersonating Someone:** No reports involved impersonation.
- **False Information:** No reports involved false information.
- **Misleading or Scam:** 8,721 profiles were reported as misleading or scam.
- **Actions Taken Pursuant to Notices:** All actions were carried out in accordance with our Terms and Content Guidelines. We conducted a thorough assessment of the reported profiles to ensure accuracy in our moderation decisions.
 - 3,524 of the reported profiles were associated with business numbers. No violating images or comments were identified.
 - 743 of the reported profiles were related to individual profile images, and after careful consideration, no instances of profanity or legal concerns were detected.
 - 4,454 of the reported profiles were flagged as potential spam, but after careful consideration, none fell under any legal category.
 - One image was identified as containing nudity and has been removed in line with our Terms and Content Guidelines.
- **Number of Notices Processed Using Automated Means:** During the relevant period, Truecaller did not process any notices using automated means. All notices are processed using manual human review.

4. Article 15(1)(c). Own-Initiative Content Moderation

- **Number of Measures Taken at the Truecaller's Own Initiative:** A total of 5 profile pictures were identified and removed due to the violation of our Terms and Community Guidelines.
- **Number of Measures Taken at the Truecaller's Own Initiative solely based on the Use of Automated Tools:** A total of 346 images were automatically blocked due to the violation of our Terms and Community Guidelines.



- **Visibility Restriction (Removal, Disable, Demoted, etc.):** A total of 227,403 comments were automatically removed due to the violation of our Terms and Community Guidelines.

5. Article 15(1)(d). Complaints Handling

- **Number of Complaints Received Through Internal Complaint-Handling Systems:** During the relevant period, Truecaller did not receive any complaints through its internal complaint handling systems.
- **Basis for Complaints:** N/A
- **Decisions Taken in Respect of Complaints:** N/A
- **Median Time to Take Decisions on Complaints:** N/A
- **Number of Decisions Reversed:** N/A

6. Article 15(1)(e). Meaningful and Comprehensible Information on Content Moderation and Automated Means

We strive to enforce Truecaller Content Guidelines consistently across our platform. In addition to responding to user reports, we leverage a combination of technology and human review to proactively detect and prevent prohibited content. Our approach prioritizes user privacy, ensuring that moderation efforts are both proportionate and targeted. We avoid unnecessary monitoring, focusing only on specific types of content that may violate our policies, as outlined below.

6.1. Public Comments

Detection Methods and Purpose: We use automated tools to analyze public comments, identifying terms or patterns that may indicate prohibited content, such as hate speech or harassment. This ensures a safer and more respectful environment for our users.

Accuracy and Error Handling: Our system is designed to minimize false positives by continuously refining detection models. However, since we rely on third-party APIs, we cannot always anticipate changes or improvements they implement. To maintain accuracy,



we regularly monitor system performance and adjust our models as needed to reduce the risk of legitimate comments being flagged incorrectly.

6.2. Image Detection

Detection Methods and Purpose: Truecaller is committed to fostering a safe, positive, and respectful community, and profile images play an essential role in that effort. As an extension of our broader Content Guidelines, our [Profile Image Policy](#) ensures that the images users choose to represent themselves align with values of inclusivity, kindness, and mutual respect.

To uphold these principles, we employ image moderation to review all profile images uploaded to the platform. Our system is designed to detect and block images that violate our policies, including those featuring brand logos, government insignias, law enforcement symbols, or references to financial and insurance institutions. We prevent such images from being visible to other users and notify the uploader, prompting them to submit a different image. This detection occurs in real time using proprietary image recognition models.

Accuracy and Error Rates: To minimize false positives and maintain accuracy, our detection model applies strict thresholds when identifying prohibited images. We continuously monitor the system's performance and make necessary adjustments to improve reliability and reduce errors.

6.3. Spam and Fraud Detection

Detection Methods and Purpose: To protect users from spam and fraudulent activities, we employ automated systems that analyze calling patterns. These methods incorporate user feedback, including spam reports, to enhance detection and prevent disruptive behaviors.

Accuracy and Error Handling: Our models adapt based on evolving spam tactics and user reports. By leveraging real-time data and machine learning, we continuously refine our detection capabilities to reduce false positives and improve efficiency.

6.4. Collecting User Feedback

Detection Methods and Purpose: We provide expanded layouts in the after-call screen interface and the details view of the app, allowing users to share feedback on data accuracy and report inappropriate content. This helps us continuously improve our services based on user input.

Accuracy and Error Handling: User feedback is analyzed to refine our data models and improve content moderation. We ensure that the system prioritizes relevant reports while reducing the likelihood of false alerts.

6.5. Verification Questions

Detection Methods and Purpose: To streamline user participation in moderation, we present simple verification questions, such as “Was this caller a spammer?” This low-threshold approach encourages engagement while maintaining ease of use.

Accuracy and Error Handling: Responses are aggregated to improve spam and fraud detection, with safeguards in place to prevent misuse. Repeated or inconsistent reports are filtered to maintain data integrity.

Anyone can report content on Truecaller, even without a Truecaller account. If users do not have an account or cannot access the content they wish to report, they can submit a report [here](#). If users come across content that they believe breaches these guidelines, they can report it directly through the “Report” link available in the detailed view of the



content within the Truecaller app. Truecaller users can learn more about identifying, blocking, and reporting spam calls and messages [here](#). Please note that reporting content does not guarantee its removal.

All users involved will be notified of any action taken, including the specific reasons for the decision, with references to the relevant policy or legal provisions. If content is removed or restricted, users will receive an explanation, and affected users will have the right to appeal the decision. Further information can be found [here](#).

For reports related to illegal content, we assess whether the content violates applicable laws and take necessary action, which may include restricting access within certain jurisdictions. In cases where content is removed due to legal requirements, users will be informed, along with details about how to seek redress or challenge the decision.

7. Human Resources Dedicated to Content Moderation

At Truecaller, we are dedicated to maintaining a safe and respectful platform by combining automated detection with human review. We enforce our Content Guidelines responsibly, ensuring a balance between effective moderation, user privacy, and appropriate safeguards for both users and reviewers.

Our human reviewers receive extensive training to uphold our guidelines with accuracy and fairness. Their training also incorporates legal considerations and evolving content risks, ensuring policies are applied correctly and consistently.

To maintain high standards, we provide ongoing coaching and conduct periodic reviews, allowing us to adapt to emerging challenges and continuously improve our moderation processes.