

X's report in compliance with Regulation (EU) 2021/1232

Sunday, 19 May 2024 X f in e

Submitted and uploaded on 19 May 2024

X REPORT (PERIOD OF JANUARY 1 2023 - DECEMBER 31 2023)

Under Article 3(1)(g)(vii) of Regulation (EU) 2021/1232 concerning a temporary derogation from certain provisions of Directive 2002/58/EC regarding the use of technologies by providers of number-independent interpersonal communication services for processing personal and other data to combat online child sexual abuse ("EU CSAM Derogation"), which entered into force on 2 August 2021, Twitter International Unlimited Company ("TIUC") is required to compile, publish, and submit this report to both the Irish Data Protection Commission ("IDPC") and the European Commission ("EC"). This report specifically covers the period from 1 January 2023 to 31 December 2023, inclusive.

Please note that the content of this report is specifically tailored to address measures within the purview of the EU CSAM Derogation. As such, it focuses on the actions and initiatives directly related to the regulation and its requirements for combating online Child Sexual Exploitation ("CSE"). This report does not encompass the entirety of efforts and measures employed by our platform to protect children.

Overview

X is committed to facilitating the public conversation in a responsible manner. Integral to this commitment is X's [zero tolerance policy towards CSE on our platform](#). We strictly prohibit any content that features, promotes or glorifies CSE, including but not limited to, media, text, illustrated content, or computer-generated images.

Our policy underscores that the act of viewing, sharing, or linking to CSE material not only contributes to the re-victimisation of the depicted minors but also violates our platform's guidelines. This stance extends to any content that could further contribute to the victimisation of children by promoting or glorifying CSE.

Our Approach

We are deeply committed to protecting children globally from CSE. Our approach encompasses the development of advanced technological solutions, comprehensive training of our content moderators, continued support for law enforcement, and ongoing partnerships to address and prevent CSE effectively.

Our approach integrates machine learning algorithms with human oversight to efficiently identify and assess content that potentially violates our policies against CSE. Our systems flag content for review, enabling our human moderators to consider crucial contextual information in their decision-making process. This work is led by an international, cross-functional team that provides 24-hour monitoring in multiple languages, ensuring rapid and effective responses to emerging threats.

Upon identification of CSE media, including images, videos, or any content that promotes child exploitation, we remove such material from our platform without further notice and report it to The National Center for Missing & Exploited Children ("NCMEC"). NCMEC plays a pivotal role in coordinating with law enforcement agencies worldwide to support investigations and legal actions. Benefiting from robust partnerships with law enforcement bodies, NGOs, and the INHOPE network, NCMEC is instrumental in our collective efforts to eradicate CSE.

In December 2022, we embarked on a significant partnership with Thorn, utilising its Safer product to substantially increase our capacity to identify, remove, and report violative content. Further solidifying our stance against CSE, we are further partnering with the Tech Coalition and WeProtect. These partnerships facilitate critical information sharing on emerging threats and behavioural patterns associated with CSE, enabling us to stay ahead of potential risks and adapt our strategies accordingly.

Our dedication to safeguarding children extends to continuous investment in both technology and talent. We are committed to enhancing our detection and response mechanisms, actively seeking out advanced technologies from third-party developers that can enhance our protective measures.

We also have instituted an appeal process. This mechanism ensures that decisions to remove content or suspend accounts can be reviewed, safeguarding against inaccuracies and maintaining our commitment to fairness and transparency.

For further information on our approach please visit [this page](#).

(1) the type and volumes of data processed;

During the reporting period of 1 January 2023, to 31 December 2023, TIUC took significant action in its fight against CSE. We suspended 12.4 million accounts for violations related to our CSE policies globally, a substantial increase from 2.3 million accounts in 2022. In the EU, we suspended more than 700K accounts during the reporting period. Furthermore, we submitted 870,000 reports to NCMEC globally, making a significant uptick in our reporting efforts, including our first fully-automated report. This volume is over eight times the number reported in 2022. It is important to note that TIUC does not currently track accounts reviewed but not actioned for policy violations. However, we may process personal data, including account details, text, and media, to investigate potential CSE policy violations.

(2) the specific ground relied on for the processing pursuant to Regulation (EU) 2016/679;

The specific grounds for processing personal data by TIUC are detailed in X's [Privacy](#)

[Policy](#) and [Additional information about data processing](#), aligning with Regulation (EU) 2016/679 ("GDPR").

(3) the ground relied on for transfers of personal data outside the Union pursuant to Chapter V of Regulation (EU) 2016/679, where applicable;

Consistent with Chapter V of the GDPR, X relies on the European Commission's adequacy decision or Standard Contractual Clauses ("SCCs") for the transfers of personal data outside the European Union, as detailed in [X's Privacy Policy](#).

(4) the number of cases of online child sexual abuse identified, differentiating between online child sexual abuse material and solicitation of children;

While all reported violations contravene [X's CSE policy](#), we currently lack the capability to categorically distinguish between online sexual abuse material and the specific context of each piece of material, such as the solicitation of children. Nonetheless, the 12.4 million account suspensions during this period reflect our commitment to combating all forms of CSE.

(5) The number of cases in which a user has lodged a complaint with the internal redress mechanism or with a judicial authority and the outcome of such complaints

In 2023 we received 224k appeals in the EU for CSE-related actions.

(6) The numbers and ratios of errors (false positives) of the different technologies used

In 2023 we reversed 1,721 CSE suspensions in the EEA. Of these reversals, 210 of the original suspensions were made by automation and 1,511 were applied manually. In 2023 we made 500k automated and 200k manual CSE suspensions in the EEA, which translates to a precision of over 99% for each method.

(7) The measures applied to limit the error rate and the error rate achieved

We continually provide training to our agents to ensure consistent enforcement of our policies, thereby guaranteeing a high level of accuracy. For automated defences we sample potential actions and label to determine error rate before launch. We deploy a variety of stopgaps to ensure we don't over-enforce or enforce on high trust accounts. Based on our suspensions and appeals volume, we achieved an error rate of less than 0.1%.

(8) the retention policy and the data protection safeguards applied pursuant to Regulation (EU) 2016/679;

TIUC's data retention policies are outlined in our [Privacy Policy](#). We adhere to ISO standards for security and privacy, undergoing regular third-party audits as needed to ensure the robust protection of the processed data.

(9) the names of the organisations acting in the public interest against child sexual abuse with which data has been shared pursuant to this Regulation

The National Center for Missing and Exploited Children ("NCMEC") in the United States of America (<https://www.missingkids.org/>).