MICROSOFT TRANSPARENCY REPORT (REGULATION (EU) 2021/1232)

GENERAL INFORMATION (as of 31 Dec. 2022)

Microsoft takes seriously its responsibility to prevent child sexual exploitation and abuse imagery (CSEAI) from distribution through its services. Our service terms prohibit illegal activities, and as specified in our Code of Conduct, we prohibit activities that exploit, harm, or threaten to harm children.

Microsoft has a longstanding commitment to participating in multi-stakeholder approaches to prevent the spread of CSEAI. Its efforts include the development of PhotoDNA (https://www.microsoft.com/en-us/photodna), a technology it has shared with organizations around the world to fight CSEAI.

Microsoft also provides transparency to the public about the actions it takes on its services to address CSEAI in its Digital Safety Transparency Report (https://www.microsoft.com/en-us/corporate-responsibility/digital-safety-content-report).

In accordance with Article 3, Subsection (g)(vii) of Regulation (EU) 2021/1232, Microsoft provides the following report on its data processing activities specific to Microsoft Number-Independent Interpersonal Communications Services [NI-ICS] in connection with the use of technology to detect CSEAI for the period January–December 2022.

1. Type and volumes of data processed during the year-long reporting period

   a. Over 12.3 billion content items scanned, globally.

   b. Microsoft's transparency report is scoped to the services impacted by and the content at issue in EU Regulation 2021/1232, that is: Microsoft's number-independent interpersonal communications services (NI-ICS) that use technology for detection of CSEAI as described below.

      Content types scanned for CSEAI are images and videos. Microsoft relies on the hash matching technologies PhotoDNA and MD5 to detect matches of previously identified CSEAI. Note, the hashes themselves contain no data about the user or image that caused the hash to be created.

      Traffic data Microsoft collects is included in its Cybertip reports to the U.S. National Center for Missing and Exploited Children (NCMEC). This data includes the following items:

      I. User ID (i.e., Microsoft Account ID) and username;
      II. Event timestamp; and

III. IP address.

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

   Varies based on processing, including public interest under GDPR Article 6(1)(e).

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

   Microsoft relies on standard contractual clauses under GDPR Article 46(2)(c).

4. Number of cases of online child sexual abuse identified

   Over 50 000 content items identified as CSEAI globally in NI-ICS during the period, with over 12 800 of those content items from the European Union.

5. 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

   a. Microsoft tracks customer contacts to its privacy team and EU Data Protection Officer where its data processing activities in connection with the use of technology on NI-ICS to detect CSEAI are at issue. Civil litigation involving these activities, and their impact on the customer, would also be tracked.

   b. During the period January to December 2022, Microsoft tracked zero complaints for the European Union that address the use of this technology in NI-ICS.

6. Number and ratios of errors (false positives) of the different technologies used

   a. Microsoft uses hash-matching technology (including PhotoDNA and MD5 hashing) to detect known CSEAI imagery shared through NI-ICS. Hash-matching technology works by using a mathematical algorithm to create a unique signature (known as a "hash") for digital images and videos. The hashing technology then compares the hashes generated from user-generated content (UGC) with hashes of reported (known) CSEAI, in a process called "hash matching".

   b. Microsoft applies a layered approach to detection of CSEAI through NI-ICS combining both hash-matching technology and manual review. During the reporting period, Microsoft tracked reversals of its initial content moderation decisions connected with 17 content items with respect to

CSEAI detection in NI-ICS. This number reflects Microsoft's application of hash-matching technology.

7. Measures applied to limit the numbers and ratios of errors (false positives) of the different technologies used

   Microsoft implements our own hash verification process in which Microsoft trained analysts review and confirm images associated with hashes provided from non-profits and other industry partners. Microsoft also leverages an additional manual review process as an ongoing hash quality check.

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

   Data retention varies depending on the type of data, but in each case the retention period is limited to the time appropriate for the type of data and the purpose of processing. Data will be deleted at the end of the retention period. Data minimization and protection efforts include de-identification or pseudonymization techniques (e.g., masking, hashing, differential privacy). Privacy reviews are conducted to identify, assess, and mitigate potential privacy risks from the collection, processing, storing, and sharing of personal data when new system capabilities or processes are being designed. Finally, Microsoft minimizes the data involved in the scanning process by hashing it.

9. The names of organizations acting in the public interest against child sexual abuse with which data has been shared

   The U.S. National Center for Missing and Exploited Children (NCMEC). According to NCMEC's official statement, NCMEC staff review each report and work to find a potential location for the incident reported so that it may be made available to the appropriate law-enforcement agency for possible investigation.