

**WebGroup Czech Republic, a.s. (xvideos.com)**

# **TRANSPARENCY REPORT**

Date of the report publication	<b>29.08.2025</b>
Date of the preceding report publication	<b>28.02.2025</b>
Starting date of the reporting period	<b>01.01.2025</b>
Ending date of the reporting period	<b>30.06.2025</b>

# Contents

<b>1. Introduction.....</b>	<b>3</b>
1.1. Purpose of the Report .....	3
1.2. Compliance statement with Regulation (EU) 2022/2065 (Digital Services Act).....	3
1.3. Reporting Period .....	3
<b>2. Article 15(1)(a) DSA: Orders from Public Authorities.....</b>	<b>4</b>
2.1. Orders to act against illegal content .....	4
2.2. Orders to provide information about recipients.....	4
<b>3. Article 15(1)(b) DSA: Information on Notices Submitted .....</b>	<b>6</b>
3.1. Notices submitted through notice and action mechanisms .....	6
3.2. Notification submitted by trusted flaggers .....	8
<b>4. Article 15(1)(c) DSA: Own-Initiative Content Moderation.....</b>	<b>9</b>
4.1. General overview of content moderation process .....	9
4.2. Types of measures.....	10
4.3. Measures against platform misuse.....	13
4.4. Transparency and accountability.....	13
<b>5. Article 15(1)(d) DSA: Internal Complaint-Handling Systems .....</b>	<b>14</b>
5.1. Description of the complaint-handling system.....	14
5.2. Ensuring fair and transparent moderation .....	15
<b>6. Article 15(1)(e) DSA: Automated Tools.....</b>	<b>16</b>
6.1. Qualitative information on the automated tools used .....	16
6.2. Indicators of accuracy, error rates, safeguards .....	16
<b>7. Article 42(2)(a)(b): Human resources.....</b>	<b>19</b>
7.1. Human resources involved in content moderation .....	19
7.2. Training and support .....	20
<b>8. Article 24(1)(a) DSA: Out-of-court dispute settlement submissions.....</b>	<b>22</b>
<b>9. Article 42(3) DSA: Information on monthly active users.....</b>	<b>23</b>

# 1. Introduction

## 1.1. Purpose of the Report

This transparency report was prepared by WebGroup Czech Republic, a.s. ("WGCZ") the provider of XVideos.com ("XVideos" or "platform"), the Very Large Online Platform designated by the European Commission under Article 33 of the European Union's Digital Services Act (Regulation (EU) 2022/2065) ("DSA"). The primary purpose of this transparency report is to provide a comprehensive overview of WGCZ's content moderation activities and compliance and fulfil transparency reporting requirements in accordance with Articles 15, 24 and 42 of the DSA. This report aims to enhance transparency and accountability by detailing the steps WGCZ takes to detect, address, and report illegal content and violations of its Terms of Service ("ToS") on the platform. By disclosing these activities, WGCZ aims to build trust with its users, regulatory authorities, and the broader public, ensuring that XVideos operates in a safe, legal, and ethical manner.

By providing detailed information on these aspects, WGCZ demonstrates platform's commitment to adhering to the DSA's requirements and platform's dedication to fostering a transparent and secure online environment.

## 1.2. Compliance statement with Regulation (EU) 2022/2065 (Digital Services Act)

WGCZ is fully committed to complying with the requirements set forth by the DSA. This regulation mandates comprehensive transparency and accountability measures for Very Large Online Platforms (VLOPs) to ensure a safer and more secure digital environment for users within the European Union. As an adult content platform, WGCZ recognizes the importance of adhering to these stringent regulations to prevent the dissemination of illegal content and to protect the rights and privacy of XVideos' users. This transparency report is a testament to WGCZ's ongoing efforts to meet and exceed the obligations outlined in the DSA.

In alignment with the DSA, WGCZ has implemented robust content moderation practices that include both automated tools and human oversight. It has established procedures to promptly respond to government orders and to provide detailed reports on platform's moderation activities. This includes the removal or restriction of illegal content, the termination of user accounts in violation of platform's ToS, and the measures taken to ensure the accuracy and fairness of platform's automated detection systems. By adhering to these regulations, WGCZ ensures compliance and reinforces platform's commitment to creating a safe, transparent, and accountable platform for all users.

## 1.3. Reporting Period

This report covers the period from 1 January 2025 to 30 June 2025 ("Reporting Period").

## 2. Article 15(1)(a) DSA: Orders from Public Authorities

In accordance with Article 15(1)(a) of the DSA, this section provides detailed information about the orders received from EU Member States' authorities related to illegal content and recipient information requests, including orders issued in accordance with Articles 9 and 10 of the DSA. These orders require the service provider to act against illegal content or provide specific information about platforms' users.

### 2.1. Orders to act against illegal content

This subsection focuses on the orders to act against illegal content issued by the relevant Member State authorities.

Table 1: Statistics on the Member States' orders to act against illegal content

Category of illegal content	Scope - Member State	Number of orders to act against illegal content	Number of specific items of information	Time to inform the authority of the receipt of the order [hrs]	Time to give effect to the order [hrs]
Fake profile deletion request	Portugal	1	1	192	192

### 2.2. Orders to provide information about recipients

This subsection includes orders issued by the relevant Member States' authorities requiring the provision of specific information about one or more platform recipients. The comprehensive breakdown of all Member States' orders to provide information, categorised by type of illegal content (see Table 2), indicates that 30 orders were issued during the Reporting Period.

Table 2: Statistics on the Member States' orders to provide information categorised by the type of illegal content concerned

Category of illegal content	Description of the sub-category "Other"	Scope - Member State	Number of orders to provide information	Time to inform the authority of the receipt of the order [hrs] <sup>1</sup>	Time to give effect to the order [hrs]
Defamation	n/a	Finland	1	9	9
Drug trafficking	n/a	Poland	1	3	3
Drug trafficking	n/a	Germany	1	6	6
Extortion	n/a	Germany	1	5	5
Extortion	n/a	Germany	1	26	26
Extortion	n/a	Germany	1	22	22
Fraud	n/a	Germany	1	23	23
Hate speech	n/a	Germany	1	80	80
Child pornography	n/a	Germany	1	5	5
Child pornography	n/a	Germany	1	2	2
Non consensual	n/a	Germany	1	3	3
Non consensual	n/a	Spain	1	16	16
Non consensual	n/a	Germany	1	3	3
Other	investigation related to several alleged fraud crimes	Spain	1	3	3
Other	not specified by the issuing authority	Czech Republic	1	0	24
Other	not specified by the issuing authority	Austria	1	0	24

<sup>1</sup> The time to inform the authority of receipt of the order is 0 when orders are issued via the Data Box system (for example, in the Czech Republic "Datová schránka"). The Data Box is a state-operated electronic communication tool used primarily for secure interactions between public authorities, businesses, and individuals. It ensures immediate confirmation of receipt: once the recipient accesses the message, the system generates a delivery report confirming the exact date and time the message was received.

Other	Dissemination without the persons consent of a recording image of a sexual nature. Moral harassment, violation of privacy by transmitting the image of a person representing a sexual nature	France	1	0	192
Other	not specified by the issuing authority	Czech Republic	1	0	624
Other	not specified by the issuing authority	Belgium	1	0	144
Other	Drug related offences and thefts with display of weapons	Belgium	1	0	24
Other	not specified by the issuing authority	Czech Republic	1	0	192
Other	not specified by the issuing authority	Czech Republic	1	0	312
Other	not specified by the issuing authority	Czech Republic	1	0	24
Other	not specified by the issuing authority	Czech Republic	1	0	24
Other	not specified by the issuing authority	Czech Republic	1	0	24
Other	not specified by the issuing authority	Czech Republic	1	0	192
Other	not specified by the issuing authority	Czech Republic	1	0	216
Other	not specified by the issuing authority	Czech Republic	1	0	192
Other	not specified by the issuing authority	Czech Republic	1	0	192
Violence	n/a	Germany	1	2	2

Additionally, data from the Reporting Period indicate that the largest number of orders to provide information were obtained from Czech Republic authorities (11 orders in total) and German authorities (11 orders in total), as shown in Table 3.

Table 3: Member States' orders to provide information, broken down by relevant Member State

Category of illegal content	Scope - Member State
Austria	1
Belgium	2
Czech Republic	11
Finland	1
France	1
Germany	11
Poland	1
Spain	2

In accordance with Article 15(1)(a) of the DSA, XVideos calculated the median time required to notify the issuing authority of the order's receipt and to give effect to the order (see Table 4). The starting point for calculating this median time is the moment XVideos receives the order.

Table 4: Median processing time of orders to provide information

Median time to inform the authority of the receipt of the order to provide information [hrs]	Median time to give effect to the order to provide information [hrs]
1	24

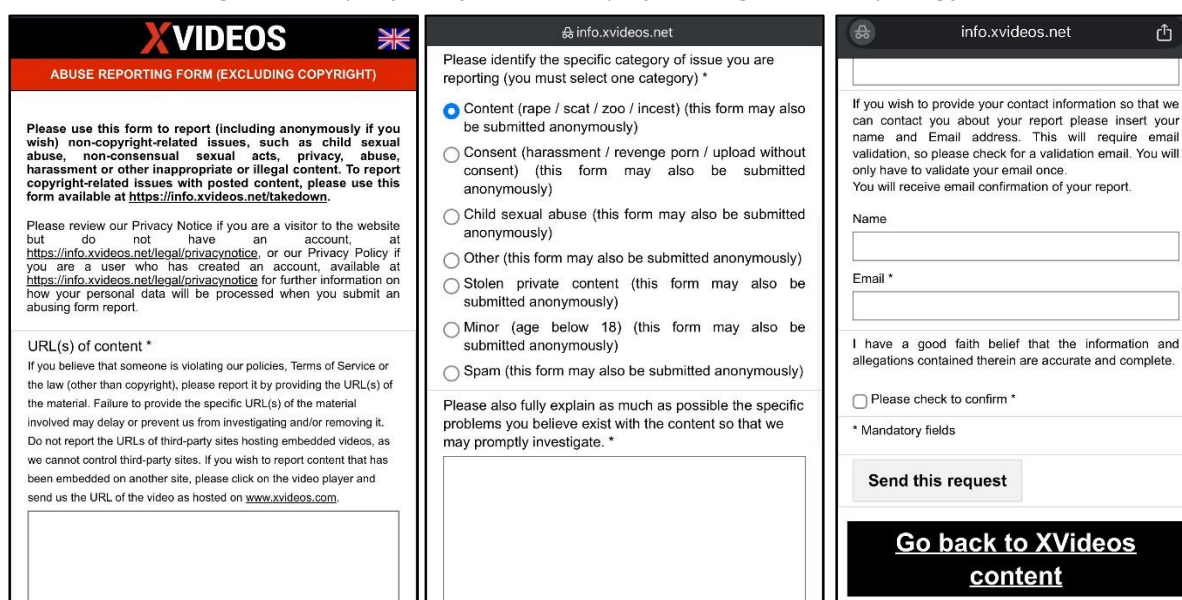
## 3. Article 15(1)(b) DSA: Information on Notices Submitted


### 3.1. Notices submitted through notice and action mechanisms

WGCZ has implemented notice and action mechanisms in accordance with Article 16 of the DSA, allowing users and third parties (individuals or legal entities) to report suspected or alleged illegal content. These mechanisms are designed to be user-friendly, with action buttons directly accessible from the respective content, enabling quick and efficient reporting. This approach enhances the safety and trustworthiness of the platform by providing a straightforward process for addressing potentially illegal content or content incompatible with the ToS. The ease of access to these reporting tools underscores WGCZ's commitment to maintaining a safe online environment for all users.

Specifically, XVideos makes available to its users an "abuse reporting form" (see Figure 1) under the "content removal" section, which allows users to flag potentially illegal/doubtful material they identify when viewing content hosted on XVideos. The "abuse reporting form" is a tool that, in accordance with Article 16(1) of the DSA, allows anyone (be it an individual or a legal entity) to notify XVideos of the presence on their service of specific items of information that is considered to be illegal content. The abuse reporting form of XVideos is easily accessible, user-friendly and accessible exclusively by electronic means.

Figure 1: Example of a notification on the platform using the abuse reporting form



**XVIDEOS** 

**ABUSE REPORTING FORM (EXCLUDING COPYRIGHT)**

Please use this form to report (including anonymously if you wish) non-copyright-related issues, such as child sexual abuse, non-consensual sexual acts, privacy, abuse, harassment or other inappropriate or illegal content. To report copyright-related issues with posted content, please use this form available at <https://info.xvideos.net/takedown>.

Please review our Privacy Notice if you are a visitor to the website but do not have an account, at <https://info.xvideos.net/legal/privacynotice> or our Privacy Policy if you are a user who has created an account, available at <https://info.xvideos.net/legal/privacynotice> for further information on how your personal data will be processed when you submit an abusing form report.

**URL(s) of content \***

If you believe that someone is violating our policies, Terms of Service or the law (other than copyright), please report it by providing the URL(s) of the material. Failure to provide the specific URL(s) of the material involved may delay or prevent us from investigating and/or removing it. Do not report the URLs of third-party sites hosting embedded videos, as we cannot control third-party sites. If you wish to report content that has been embedded on another site, please click on the video player and send us the URL of the video as hosted on [www.xvideos.com](http://www.xvideos.com).

Please identify the specific category of issue you are reporting (you must select one category) \*

- ☒ Content (rape / scat / zoo / incest) (this form may also be submitted anonymously)
- ☐ Consent (harassment / revenge porn / upload without consent) (this form may also be submitted anonymously)
- ☐ Child sexual abuse (this form may also be submitted anonymously)
- ☐ Other (this form may also be submitted anonymously)
- ☐ Stolen private content (this form may also be submitted anonymously)
- ☐ Minor (age below 18) (this form may also be submitted anonymously)
- ☐ Spam (this form may also be submitted anonymously)

Please also fully explain as much as possible the specific problems you believe exist with the content so that we may promptly investigate. \*

If you wish to provide your contact information so that we can contact you about your report please insert your name and Email address. This will require email validation, so please check for a validation email. You will only have to validate your email once. You will receive email confirmation of your report.

Name

Email \*

I have a good faith belief that the information and allegations contained therein are accurate and complete.

☐ Please check to confirm \*

\* Mandatory fields

**Send this request**

**Go back to XVideos content**

The abuse reporting form facilitates the submission of sufficiently precise and adequately substantiated notices containing all the elements required by Article 16(2) of the DSA. Specifically, the form includes a section allowing the user to provide a substantiated explanation of the reasons why the information in question constitutes illegal content (Article 16(2)(a) of the DSA). The abuse reporting form includes a clear indication of the exact electronic location of the allegedly illegal/doubtful content as it requests that the reporting user provide the URL of the relevant video (Article 16(2)(b) of the DSA).

The abuse reporting form also includes a section for the reporting individuals/entities to provide their names and email addresses. If the individual/entity submitting the abuse reporting form fills in their email address, XVideos sends them, without undue delay, a confirmation of receipt of the notice as specified in Article 16(4) of the DSA. Thereafter, XVideos processes the notice in a timely, diligent, non-arbitrary and objective manner and notifies the individual/entity who submitted the notice of its decision (e.g., whether it has deleted the content or assessed it and kept it online). In addition, XVideos also provides them information on the possibilities for redress in respect of the specific decision made (including information about the possibility of using the internal complaint-handling system).

In addition, each video hosted on XVideos features a reporting button underneath it to allow the immediate reporting of suspected illegal/doubtful content, as shown in Figure 2 below. By simply clicking on the reporting button or using the sites' other reporting mechanisms, users and site visitors can report any video featuring actual or perceived Non-Consensual

Intimate Imagery (“NCII”), Child Sexual Abuse Material (“CSAM”), or any other reason (such as, e.g., copyright infringement). Complaints relating to CSAM and NCII are prioritized and immediately investigated.

All notices, pursuant to Article 16 of the DSA, are reviewed and processed exclusively by human moderators. WGCZ does not utilize any fully or partially automated systems to resolve issues highlighted in these notices. Instead, all aspects of the process, including assessing the notice, gathering relevant information, collecting data, and communicating outcomes, are carried out manually.

Figure 2: Example of a notification on the platform using the report button

Pursuant to Article 15(1)(b) of the DSA, the Table 5 reports notices submitted by third parties through the "Notice and Action" mechanism under Article 16 of the DSA.

Table 5: Statistical data on notices submitted in accordance with Article 16 of the DSA

Category of illegal content	Number of notices received	Number of specific items of information	Number of actions taken – content taken down	Non-correct notices	Median time to take action [hrs] (time between request and review)
Cartoon Zoophilia	6	6	6	0	48,04
Consent	1160	1382	32	1316	0,3
Content	691	746	15	731	43,14
Copyright	232	683	353	281	0,28
Extortion or Blackmail	317	317	317	0	0,2
Fake	39	39	16	23	20,07
Harassment	33	34	34	0	0,17
Incest	4	4	4	0	0,25
Consent paperwork issues (lost, changed, only oral)	3	3	3	0	48,05
Non-consensual	403	423	423	0	0,22
Non-consensual (Simulated)	7	7	7	0	2,39
Non-consensual Candid Filming	31	31	31	0	0,18
Other	4103	4135	982	3151	9,52
Revenge Porn	3	5	5	0	0,08
Scatophilia	19	19	19	0	0,27
Spam	12816	12816	342	12466	0,29
Spammer	18	18	4	14	22,71
Stolen private content	157	176	79	92	0,33
Underage	3183	3183	207	2976	0,28

Unknown	986	986	337	649	18,05
Violence (Real)	1	1	1	0	0,15

Table 5, among others, presents the total number of actions taken in response to notices based simultaneously on legal requirements and the ToS. It is important to address the differentiation between actions taken according to notices based on legal requirements and those based on the platform's ToS. The ToS are closely aligned with applicable laws and regulations. The ToS are designed and regularly updated to prohibit illegal content explicitly. For instance, Section 7.B of the ToS requires the users not to submit contents related to (or that is):

- Minors or material involving child sexual exploitation i.e., “underage”, “child sexual abuse”,
- Acts of non-consensual sexual behaviour, rape, bestiality, or controlled substances,
- Content deemed unlawful, threatening, harassing, or hateful,
- Material protected by copyright or trade secret laws without proper authorisation.

Content falling into these categories is generally both illegal under the laws of EU Member States and in breach of the ToS. As a result, actions to remove such content are often rooted in both legal obligations and ToS violations, making a distinction between the two bases of action largely irrelevant in most cases.

Additionally, WGCZ's existing database systems currently lack the functionality to automatically record whether a specific action was taken on legal grounds or solely for breaching the ToS. Adding this level of tagging would require significant resources and infrastructure changes. However, WGCZ is actively working to address this limitation by improving its technical capabilities to enable better data collection and categorization for future transparency reports.

## 3.2. Notification submitted by trusted flaggers

In compliance with Article 22 of the DSA, WGCZ has established a procedure that ensures notices submitted by trusted flaggers are treated with priority over other notices. This system is designed to expedite the review and removal of flagged content, leveraging the expertise and reliability of trusted flaggers to quickly address issues of concern. The breakdown of notices received from trusted flaggers during the Reporting Period is presented in Table 6 below.

Table 6: Statistical data on notices submitted by trusted flaggers in accordance with Article 16 of the DSA

Category of illegal content	Number of notices received	Number of specific items of information	Number of actions taken – content taken down	Non-correct notices	Median time to take action [hrs] (time between request and review)
Other	1	1	1	0	74,06
Consent	1	2	0	2	0

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

### 4.1. General overview of content moderation process

For a long time, WGCZ employs a proactive approach to moderating various types of content violations to maintain a safe and respectful environment for all users. The following categories (types) of content are moderated:

- text content: this includes comments, wall posts, and "about me" sections, where inappropriate or harmful language is identified and addressed;
- images: gallery and profile are reviewed based on samples to ensure they comply with platform's community standards;
- videos: uploaded video content is subject to rigorous checks to prevent the dissemination of incompatible or illegal material.

#### (a) Approach to content moderation

WGCZ employs a comprehensive and proactive approach to content moderation, focusing on identifying and addressing (via below described automated tools) illegal or incompatible (i.e. not compatible/compliant with platform's ToS or rules it declare and enforce) content. The moderation team assess content to mitigate the chance it includes illegal or incompatible material, such as CSAM, NCII, gender-based violence, and more. It should be noted that identified CSAM content not only is deleted from the platform but investigated internally and underlying respective user data are provided to local enforcement agencies.

Additionally, it should also be noted that when a piece of content, especially high-priority content such as CSAM, non-consensual, or gender-based violence, is flagged for suspected illegality or incompatibility, the URL is typically immediately unindexed by the platform's systems, rendering it unavailable on XVideos. It is then reviewed by the platform's dedicated team specializing in handling notices and complaints. In such cases, the content remains inaccessible until the notice is resolved.

Regarding the spread of misinformation, political views, and agendas XVideos notes that these topics are rarely present. Nonetheless, platform's automated tools are designed to identify such content, particularly textual. These tools prevent publication if users employ blacklisted words.

To ensure accuracy and effectiveness, XVideos utilizes a combination of advanced automated tools (which include AI functionalities), and human moderation. The platform's automated tools flag content based on specific settings tailored to each content type. Flagged items are then reviewed by human moderators to ensure an accurate and context-sensitive assessment. This process allows us to focus human efforts on potentially illegal or incompatible content and take appropriate actions based on the type of violation and the expertise of the platform's moderators.

The moderation team is responsible for reviewing all potentially illegal content and confirming that the content is NCII or CSAM. The general criteria moderators use to review content include (i) assessing the age of the persons featured in the video, including signs of physical development; (ii) assessing the activity/conduct depicted in the video; (iii) assessing whether there is a difference in the size of the persons depicted in the video; (iv) assessing whether a crime has been committed; (v) obtaining, if needed, a second opinion on potential NCII and CSAM content; (vi) erring on the side of the potential victim; (vii) erring on the side of removing content as NCII and CSAM when there is any doubt; and (viii) identifying typical objects suggesting illegality via the use of software tool recognizing various objects in the pictures which may indicate the illegality of the content (e.g., weapons, blood).

The moderation team also conducts voluntary, own-initiative investigations in good faith and diligently to identify any potentially doubtful material, in line with Article 7 of the DSA.

WGCZ has also created an internal interface tool to facilitate and expedite the review of all videos reported or identified as containing NCII or CSAM. This tool captures a screenshot of the video in question every set number of seconds, thereby allowing a speedy and effective review without the need for the reviewer to watch the video in its entirety. Where closer scrutiny is required, the full video is obviously available for review.

## (b) Moderation process

When a piece of content is flagged for review, it enters a review pool. This pool is periodically reviewed by the moderation team daily. Most actions are executed automatically through the review interface once a human reviewer assesses the case. Some types of violations require additional actions such as banning comments, initiating direct communication with the user, deleting the account, or conducting further investigation of the user profile.

### Platform back-end and moderation interface

The moderation interface includes a backlog where flagged content is distributed among moderators. This distribution can be automatic, based on moderators' skill sets, or manual. Content from the pool is allocated to reviewers until the pool is empty or the designated review time for the day is completed. The whole team usually processes reviews collectively, sharing their findings in a group chat to improve accuracy, such as by identifying spam profiles, new jargon or code language that is suspected to be used by users that are suspected to publish incompatible or illegal content.

Flagged content is assigned to human moderators through automated systems, ensuring that potentially illegal material is not copied, shared, or further disseminated. The moderation interface generates a content pipeline containing links to newly uploaded material alongside selected historical content. This system prioritizes new uploads, especially videos, over older archived content.

Content is distributed randomly to moderators, but it is structured to provide a balanced mix of new and historical material for review. Additionally, a manual process allows moderators to re-assign specific content to more experienced team members when needed.

### Handling errors and false positives

To ensure legitimate content is not wrongly removed, XVideos has several measures in place. Content can be reinstalled during a pending delete period unless it is set for immediate deletion (e.g. CSAM content). Users can contact support through platform's ticketing system, and if a mistake in manual moderation is identified, it is flagged for additional review and final decision is made. Such an interaction between the moderation team and the user is recognized as a so-called extended communication, i.e. extended communication might entail complaints, appeals etc.

Main activities carried out to account for false positives:

- false positives are reviewed with the same rigor as any other flagged content; platform's experienced moderation team ensures an efficient review process;
- training focuses on differentiating between incompatible and illicit content, users are informed about the redress process, which is available should a mistake occur.

The performance of content moderation efforts is monitored through the observation of instances of human error, i.e. reviewer accuracy. If a reviewer repeatedly makes mistakes, targeted training is provided to ensure they understand the process and avoid similar errors in the future. Mistakes are noticed and corrected through direct feedback and mentoring.

At present, the accuracy of content moderators is not measured using a standardized, quantitative metric, as the often-nuanced nature of content moderation does not easily lend itself to such measurement. Instead, the evaluation process primarily depends on identifying moderators' errors retrospectively. When a user or a third party challenges a content moderation decision, the contested decision is re-examined. In such cases, a secondary review is initiated to identify any mistakes made by the original moderator.

Content wrongly flagged and deleted is restored during the pending delete period. If the error is due to a process failure, the workflow is updated and improved. If it is a personal error, the reviewers receive additional guidance and training to enhance their accuracy.

By combining programmed workflows, regular team communication, and continuous training, WGCZ maintains an effective and reliable content moderation process.

## 4.2. Types of measures

This section details the measures taken by WGCZ to affect the availability, visibility, and accessibility of information provided by recipients of the service. These measures are categorised by the type of illegal content or violation of the ToS (more information is provided in Section 4.4.), the detection method used, and the type of restriction applied to content or accounts (see Table 7).

## Detection Method:

- **Automated detection:** content identified as illegal or incompatible using automated tools (see set of automated tools in place below) was automatically identified and flagged for review by the moderation team; examples include copyrighted material detected through fingerprinting and illegal content such as CSAM identified through AI classifiers.
- **Manual review:** content flagged by users or automated systems was reviewed by human moderators; if found to be in violation, it was removed from the platform; most flagged content undergoes human review, though some comments review may be fully automated.

## Content Restrictions

- **Content ghosted (i.e. unindexed content):** the content remains on the platform's backend but becomes invisible to the public. It does not appear in searches or listings, although the uploader can still see it under their account. This approach effectively functions as a temporary suspension. If, after review, the content is found acceptable, it can be restored and made visible on the platform.
- **Content taken down (i.e. pending deletion):** the content hidden from the public but not yet permanently removed. This phase acts as an intermediate checkpoint, ensuring that content moderation decisions remain potentially reversible. During the pending deletion period the affected user can lodge complaints or provide evidence demonstrating that the removal was incorrect. If, after this period and any additional reviews, the initial decision is upheld, the content will move on to the permanent deletion stage.
- **Content deleted (i.e. permanent deletion):** the content is irreversibly removed from the platform's systems. The data cannot be retrieved or restored in any way. Permanent deletion represents a final and conclusive sanction applied to content.

## Account Restrictions

- **Account terminated:** an account is inaccessible to the user, effectively functioning like a temporary suspension. While the user cannot access their account, the underlying account data is preserved, and the user is given an opportunity to submit a complaint or additional evidence to challenge the decision. If the complaint is successful, the account can be fully restored. If not, it may move toward permanent deletion after a designated review period.
- **Account deleted:** an account is permanently removed from the platform's systems, including all associated data and content. Once deleted, the account cannot be recovered or restored in any form. This measure represents a final, irreversible step, resulting in the complete loss of the user's account presence on the platform.

Table 7: Statistics on the types of measures taken that impact content or account

Category of illegal content	Content restriction - Content ghosted	Content restriction - Content taken down	Content restriction - Content deleted	Account restriction - Account terminated	Account restriction - Account deleted
Active Zoophilia (Real)	0	0	11	0	11
Cartoon Zoophilia	55	0	32	0	6
Consent	29	0	0	0	0
Copyright	64	2	139	0	99
Extortion or Blackmail	0	0	9	0	9
Harassment	23	0	454	0	446
Incest	10	1	7	0	6
Minor consent paperwork issues (lost, changed, only oral)	2	0	0	0	0
Non-consensual*	304	2	244	0	137
Non-consensual (Simulated)	108	1	34	0	12
Non-consensual Candid Filming	546	3	158	0	115
Other	906	21	240	0	147
Passive Zoophilia (Real)	18	0	61	0	34
Rape (Simulated)	16	0	4	0	1
Revenge Porn	9	0	1	0	1
Scatophilia	77	0	27	0	9

Spam	123	0	9939	0	465
Stolen private content	1603	98	1315	0	903
Underage**	142	8	150	0	102
Uploader request	0	0	53039	0	0
Violence (Real)	6	0	3	0	2
Violence (Simulated)	175	29	8	0	34

\* Contained miscategorised copyright infringements

\*\* Contained non-sexual material, cartoons and 3D animations

### (c) Combination of human and automated moderation

XVideos employs a suite of automated tools designed to flag potentially illegal or incompatible content, in case of text input such as comments. Keyword search and match automated tool is used when the comment is submitted for publishing. Once content is flagged by platform's automated tools, it undergoes a human review. This process allows to apply nuanced judgment and context-sensitive decisions that automated tools alone cannot achieve. Platform's human moderators are tasked with the following:

- **Ghosting (i.e., unindexed content):** Ghosting refers to temporarily rendering content invisible or inaccessible to platform users without removing it from the system entirely. Such content remains available to the uploader under their account but is not indexed or searchable by others. Ghosting serves as a form of temporary suspension, particularly for content awaiting further review. This method allows for a cautious approach, ensuring that potentially inappropriate material is not immediately visible while moderators make a final determination.
- **Advanced Review:** For ambiguous or borderline cases, content undergoes advanced review by experienced moderators. This process involves a deeper evaluation of the content's context, legality, and adherence to the platform's ToS. Moderators assess cultural, regional, and linguistic nuances that automated tools might overlook. For example, expressions or terms may have varying meanings depending on the cultural context, and experienced moderators ensure such subtleties are accounted for. Advanced reviews act as a bridge between basic moderation and final decision-making.
- **Validation:** Validation serves as the final step in the human review process. Moderators assess flagged content to determine its compatibility with platform guidelines and applicable laws. Content deemed appropriate is approved and made publicly accessible. If further scrutiny or user input is needed, the content may be temporarily ghosted to allow for necessary corrections. In cases where content breaches the law or platform's ToS, it is deleted. This step ensures only compliant material is indexed and visible to users.
- **Deletion:** Content or accounts determined to be illegal or incompatible with the platform's ToS are removed. In cases of manifestly illegal uploads, such as NCII, CSAM, or terrorist content, both the content and the uploader's account are permanently deleted. This ensures such material is irreversibly removed and inaccessible to all users, including the uploader.

Moderation teams' mission is not only to review uploaded content, but they also focus their efforts on content, which was uploaded historically and are searching for, and acting upon potentially illegal or incompatible content. These pre-emptive reviews focus especially in CSAM, gender-based violence and NCII imagery. Automated actions are applied to textual content using platform's keyword search and match tool to prevent the publication of blacklisted words and expressions. However, we do not maintain statistical data on automated deletions, such as comments containing blacklisted words or expressions.

#### Effectiveness of automated tools

The implementation of these automated tools has significantly increased the effectiveness of the platform's content moderation efforts. Automated systems excel at identifying large volumes of potentially problematic material, allowing human moderators to concentrate on high-risk content requiring nuanced judgment. The table below illustrates the impact of automated tools.

Table 8: Effectiveness of automated tools in content moderation process

Indicator	Number of items identified by automated means	Number of items identified by automated means + action taken	Number of items identified by automated means +no action taken
Visual Content (video, images)	5127462	2378	5125084
Text Content	385144	11554	373590

The data demonstrate that while automated tools identify very large volumes of potentially illegal or incompatible content, the majority of flagged items ultimately require no enforcement action after human review. For visual content, automated systems flagged a total of 5,127,462 items, of which only 0.05% resulted in enforcement action, while the vast majority (99.95%) required no actions. For text content, the systems identified 385,144 items, with 3% leading to enforcement action and 97% requiring no further intervention. These figures indicate that automated tools tend to generate a high number of “false positives,” particularly for visual material, which underscores the importance of robust human oversight to ensure contextual accuracy and prevent erroneous removal of lawful content.

### 4.3. Measures against platform misuse

When a user uploads clearly illegal content, such as suspected CSAM, NCII, or terrorist material, both the content in question and the associated account are promptly removed from XVideos (see Table 9).

Currently, the platform cannot provide statistics that differentiate account terminations resulting from manifestly unfounded notices or complaints. This limitation exists because, during the Reporting Period, the database was not designed to track such distinctions. The lack of these specific system features means that terminations linked to unfounded notices or complaints were not separately recorded or categorised. XVideos is actively working on upgrading its systems to enable the collection of this data.

Table 9: Statistics on the measures taken against platform misuse according to Article 23 of the DSA

Category of illegal content	Number of account termination enacted for the provision of manifestly illegal content
Active Zoophilia (Real)	2
Rape (Real)	6
Underage	268
Scatophilia	2

### 4.4. Transparency and accountability

To ensure transparency and maintain trust with the platform’s users, WGCZ actively communicates XVideos’s content moderation policies and practices through the ToS. This document outlines all the rules and guidelines that govern user behaviour and content on the platform. It is readily accessible to all users and provides detailed information on what constitutes a violation and the consequences thereof.

## 5. Article 15(1)(d) DSA: Internal Complaint-Handling Systems

### 5.1. Description of the complaint-handling system

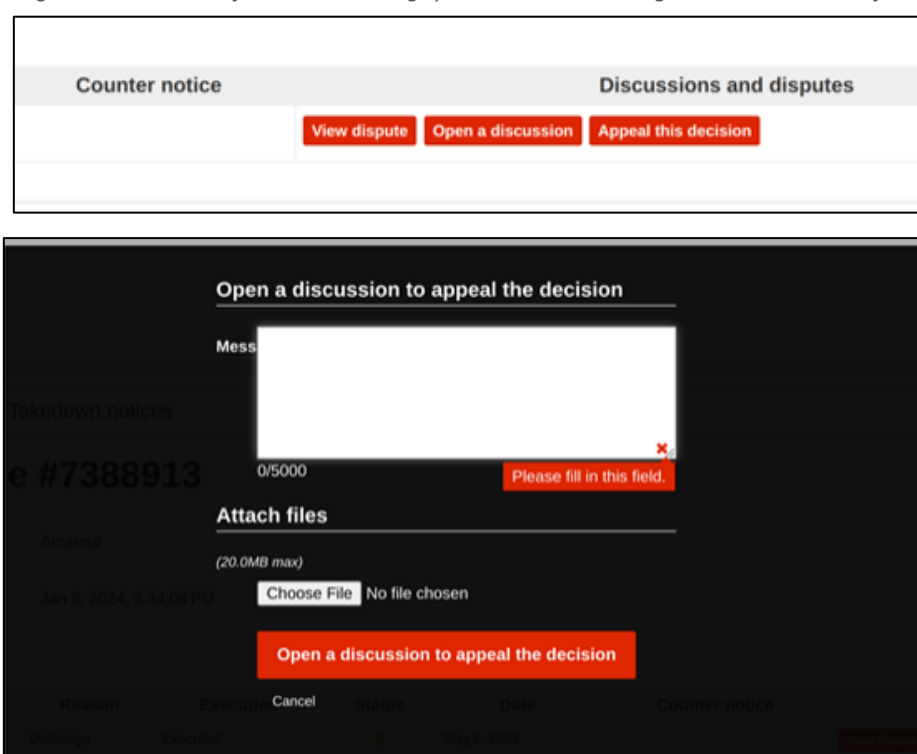
We believe in providing platform's users with a fair opportunity to appeal content moderation decisions (see the term "extended communication" in sub-section "Handling errors and false positives"). XVideos has a structured mechanism in place for users to contest actions taken against their content or accounts.

#### Ticketing system

The ticketing system is an operational tool used by the support team to manage and track all communications with users regarding complaints (appeals) and disputes related to content or account restrictions, as well as other policy enforcement actions. It serves as a centralized platform where structured discussion threads, known as "tickets," are created whenever a user challenges a decision or when additional information or documentation is required. Consequently, the "ticketing system" functions as a means to receive, organize, and monitor user inquiries and messages, thereby enabling more efficient management of their requests.

Users may lodge a complaint directly through the "ticketing system" available in their account interface. By selecting the "Appeal this decision" button, users can initiate the complaint process. This interface also displays the notices that prompted the moderation team's initial content review (see Figure 3 below). Even if a user's profile is suspended, affected users can still access their profile to some extent and use the ticketing system to communicate with platform's support team, which ensures that users can express their concerns and request a review of the decision; while suspended, users are restricted from interacting with content or other users, but they retain the ability to seek redress. As an alternative, the complaint process can be initiated via email by clicking the link included in the notification message sent to the user, which outlines the decision connected to the respective notice (see Figure 4 below).

Figure 3: Screenshot of XVideos' ticketing system accessible through user's account interface



The screenshot displays the XVideos user interface for initiating a dispute. At the top, a navigation bar includes a "Counter notice" tab and a "Discussions and disputes" tab. Below the tabs, three red buttons are visible: "View dispute", "Open a discussion", and "Appeal this decision".

The main content area is titled "Open a discussion to appeal the decision". It features a "Message" input field with a character count of "0/5000" and a red error message "Please fill in this field." Below the message field is an "Attach files" section with a "(20.0MB max)" limit and a "Choose File" button. A "No file chosen" status is displayed. At the bottom of the form is a large red button labeled "Open a discussion to appeal the decision".

Below the form, a table lists various notices. The first row shows a "Takedown notice" with the ID "e #7388913", a date of "Jan 9, 2024, 3:34:08 PM", and a status of "Cancelled". The table has columns for "Reason", "Execution", "Cancel", "Status", "Date", and "Counter notice".

Figure 4: Screenshot of Initiating a complaint through the notification email

Ground for Decision	Incompatible with terms and conditions
Possibilities for redress	You can submit a complaint against this decision through our internal complaint-handling mechanism here: <a href="https://www.xvideos.com/account/takedown-notice/">https://www.xvideos.com/account/takedown-notice/</a>

## Handling of complaints

Once a complaint has been filed, WGCZ promptly conducts an investigation. The review process verifies the validity of any imposed restrictions, such as content taken down or account termination. If needed, extra documentation is requested to determine whether the complaint lacks merit or warrants further action. If the complaint is deemed valid, corrective measures are taken — this may involve restoring the content or reversing the account termination. The “ticket” is then updated to reflect the final decision. For complaints found to be clearly unfounded, the ticket is closed, and a corresponding notification is sent to inform the user of the outcome.

Table 10: Statistics on internal-complaints mechanism decisions

Indicator	Total number	Decisions upheld	Decisions reversed	Decision omitted	Median time [hrs]
Number of complaints submitted to the internal-complaints mechanism	1030	800	230	0	169,17

## 5.2. Ensuring fair and transparent moderation

To maintain fairness and accountability in platform’s content moderation processes, we take two key steps:

- all complaints/appeals within the process of extended communication are thoroughly reviewed by dedicated team handling notices and complaints; if it is determined that a mistake was made, appropriate corrective actions are taken, such as restoring wrongly flagged content or lifting account termination;
- feedback from user appeals is invaluable for improving platform’s moderation systems; we use this information to refine systems and human review processes, ensuring that similar mistakes are less likely to occur in the future.

By clearly communicating platform’s content moderation policies and practices, providing mechanisms for users to complain about/appeal decisions, and continuously improving based on user feedback, XVideos is committed to maintaining transparency and accountability in content moderation practices.

## 6. Article 15(1)(e) DSA: Automated Tools

### 6.1. Qualitative information on the automated tools used

WGCZ employs several automated tools and AI technologies to identify illegal or incompatible content and the moderation process (see Table 11). These tools are tailored to handle different types of content, ensuring comprehensive coverage and enforcement of platform's content policies.

Table 11: Qualitative description and purposes of use of the automated means

Automated means	Qualitative description of the automated means	Purpose of the use of automated means
<b>MD5 hash</b>	MD5 hashing creates a fixed-size fingerprint of a file to verify integrity, detect duplicates, or track changes.	Used to detect exact duplicates of known files. If a match is found, the content is flagged for review.
<b>Vercury (fingerprinting of video and image content)</b>	Vercury is a fingerprint system that relies on an internal signature database. It compares video signatures against the Vercury database to detect matches.	Vercury generates a match score by comparing video or image content signatures to its internal database. Based on the score, content may be blocked or sent for review, primarily targeting illegal materials.
<b>Hive</b>	Hive uses AI to analyse images and videos, attempting to identify behaviour, context, or potentially harmful activity in the scene.	Hive is employed mainly to identify content that violates terms of service (e.g., alcohol, violence, firearms). Actions include blocking or flagging for review.
<b>Google SafetyNet API</b>	Google SafetyNet API uses AI to analyse images and assigns an underage risk score based on its assessment of the content.	Used to flag potentially harmful content for review, focusing on violations like malware or abusive material.
<b>Safer</b>	Safer maintains a fingerprint database of known content to detect and flag matches against its database efficiently.	Functions similarly to Google SafetyNet API to flag potentially harmful or inappropriate content for review.
<b>Keyword search and match</b>	Text content is managed using blocklists and grey lists to identify and flag problematic keywords or phrases.	The block list prevents users from using specific words or phrases entirely, with such content not being saved. The grey list is used to review flagged content for further analysis.

### 6.2. Indicators of accuracy, error rates, safeguards

Table 12 provides information on the accuracy rates and possible errors of the automated tools involved in content moderation activities. In addition, the table includes information about the scanned parts of the online interface to clarify the scope of each automated tool, as the accuracy rates and error rates might vary significantly depending on the type of content being moderated.

Table 12: Indicators of accuracy and error rate for automated content moderation means

Automated means	Indicator of accuracy	Possible rate of error	Parts of online interface scanned
<b>MD5 hash</b>	Data not available for this means.	Data not available.	Any file type (binary data)

<b>Vercury (fingerprinting of video and image content)</b>	Vercury achieves: 83% accuracy for identifying blocked content; and initially 2.5% accuracy for flagged-only content, which since June 2024 improved up to 10%.	Around 15–20% error rate for blocked content, pending review	Videos only
<b>Hive</b>	Hive's accuracy is currently not quantifiable, as it is used primarily as an aid in manual review. However, 10% of highly problematic Hive reports are confirmed to be correct upon review.	Approximately 90% error rate for problematic content; review required to include context.	Videos and images (pictures)
<b>Google SafetyNet API</b>	For videos: Google High is 100% accurate; Google Medium is 1.6% accurate. For pictures: Google Very High is 100% accurate; High is 75% accurate; Google Medium is 15% accurate.	98.5% error rate for videos; and 70% error rate for pictures	Videos and images
<b>Safer</b>	Safer has an accuracy rate of 10%.	90% error rate	Videos only
<b>Keyword search and match</b>	Data not available.	Data not available.	Text only (comments)

The automated tools are third-party solutions. Thus, WGCZ's influence on improvements by providing feedback to improve them is limited. However, Table 13 shows a series of safeguards applied alongside these automated means. For example, content flagged by these tools is frequently subjected to manual review or additional verification steps before any enforcement action is taken.

Additionally, WGCZ constantly improves its proprietary systems based on moderation team feedback. The development team continually fine-tunes systems based on human reviewers' feedback and experiences, enhancing the overall efficiency and accuracy of content moderation systems.

The databases employed to search and match textual content are updated regularly with terms, jargon and invented words or speech used by users who were sources of illegal or incompatible content.

Given the importance of issues, e.g., the global fight against CSAM, paedophilia, and violence (including gender-based violence), the automated tools WGCZ uses are, to knowledge and monitoring, improved continuously.

WGCZ's approach to moderation (as described above) leverages the advantages of both automated tools and human moderation, providing an effective, accurate and efficient form of moderation fighting, e.g. the above-mentioned issues.

*Table 13: Safeguards applied alongside the use of automated means in the content moderation process*

<b>Automated means</b>	<b>Safeguards applied</b>
<b>MD5 hash</b>	Used for basic file integrity checks; flagged matches are manually reviewed before any action.
<b>Vercury (fingerprinting of video and image content)</b>	All content flagged by the Vercury system is manually reviewed to determine whether it should be removed.
<b>Hive</b>	Content flagged as "problematic" by Hive is manually reviewed for final decision-making
<b>Google SafetyNet API</b>	All flagged content is manually reviewed before any enforcement action is taken.
<b>Safer</b>	All content flagged by the system is subjected to manual review to confirm issues

<b>Keyword search and match</b>	Content flagged by the block list displays an error message to the user without retention of data. Content flagged by the grey list is manually reviewed.
---------------------------------	---

Article 42(2)(c) the indicators of accuracy and related information referred to in Article 15(1), point (e), broken down by each official language of the Member States.

WebGroup is unable to provide the requested information broken down by the official languages of the Member States due to technical limitations that make it impossible to collect such data. The existing data collection systems are not designed to differentiate or categorise data by the official languages of the Member States.

## 7. Article 42(2)(a)(b): Human resources

### 7.1. Human resources involved in content moderation

The human moderation team at WGCZ is composed of a dedicated and skilled workforce committed to ensuring that XVideos remains compliant with its content policies.

Table 14 provides the number of employees performing specific work functions at each moderation stage, where 1 FTE (full-time equivalents) equals one employee primarily responsible for implementing activities related to the moderation process. However, it is essential to note that, if necessary, some moderation activities may overlap. For example, the Head of the Moderation Team participates in advanced review or provides information for processing notices and complaints. Similarly, moderators responsible for advanced review may perform a basic review when operationally required. Therefore, the number of human resources allocated to a moderation process is difficult to precisely determine because the functions overlap.

Table 14: Human resources dedicated by WGCZ to content moderation on the XVideos

Moderation stage	Dedicated FTEs	Moderation scope
Head of the Moderation Team	1	oversees the entire moderation team, establishes policies and workflows, and handles complex content moderation cases that require higher expertise and judgment;
Basic review	20	moderators responsible for the initial review of content, including monitoring and moderating channels;
Advanced review	4	moderators responsible for more complex content moderation cases that require a higher level of expertise and judgment, they are also responsible for their own initiative moderation, identifying problematic areas, and improving current procedures;
Channel oversight	25	this specialized team focuses on overseeing specific channels, ensuring compliance and addressing any issues that arise;
Complaints per Art. 20 DSA Trusted flaggers notices per Art. 22 DSA Third party notices and content take-down	3	this specialized team focuses on processing notices and executing content take- down requests as well as handling complaints

The moderation process, as outlined above, involves the participation of the following teams:

- The moderation team, which consists of 25 FTEs,
- The team handling notices and complaints, which consists of 3 FTEs,
- The channel oversight team, which consists of 25 FTEs.

The moderation team is further divided into internal and external moderators. The internal team consists of 16 members based in the Czech Republic. This group forms the core of the moderation workforce, handling most of the platform's content reviews and policy enforcement activities from the local office. Complementing the internal team is a group of 9 external moderators who work remotely. These external moderators are located across three countries, with five members in Peru, two in the Philippines, and two in Italy.

#### Language skills and qualifications

According to Article 42(2)(b) of the DSA, table 15 specifies the linguistic expertise of internal and external FTEs carrying out content moderation activities (i.e. moderation team and team handling notices and complaints) broken down by applicable official language of the Member States.

Table 15: Linguistic expertise of the FTEs carrying out activities dedicated to content moderation

EU official languages	Code	Internal FTEs	External FTEs	FTEs with sufficient linguistic expertise
Czech	cs	37	0	37
English	en	44	9	53
French	fr	1	0	1
Italian	it	0	2	2
Polish	pl	1	0	1
Slovak	sk	17	0	17
German	de	1	0	1
Spanish	es	0	6	6

Additionally, two external moderation team members are fluent in Tagalog, the official language of the Philippines, and one member of the team handling notices and complaints speaks Japanese. The linguistic expertise of the channel oversight team also includes languages that are not official languages of the European Union, specifically languages such as Russian, Ukrainian, Kazakh, Serbian, Hindi, and Gujarati. For languages not directly covered, translation tools are implemented to ensure effective moderation. Furthermore, the keyword databases include variations in all EU languages, enabling comprehensive moderation across diverse linguistic contexts.

Advanced English proficiency and strong PC skills are mandatory for all moderators. While there are no specific requirements for formal educational qualifications, moderators are expected to demonstrate ability to exercise judgment and particular sensitivity when working with adult content. Qualification is primarily based on experience, measured by the time spent monitoring content. The moderation team members have long years of experience and maintain a minimal employee turnover rate.

## 7.2. Training and support

### Training given to the moderation team

The training provided to the WGCZ moderation team are designed to maintain high content moderation standards and ensure the team is equipped to handle the nuances of moderating adult content. Training is ongoing and tailored to address general and specific content moderation challenges.

Training is conducted by the Head of the Moderation Team and an experienced moderator. This ensures that both new and existing team members receive guidance from professionals who are well-versed in the complexities of moderating adult content.

From the outset, newly hired moderators are introduced to WGCZ's moderation process through a structured mentorship approach. On their first day, they observe the tasks and techniques experienced moderators employ, gaining insight into the operational framework, rules, and classification methods for various categories of illegal content. In the following days, they gradually take on responsibilities under close supervision, receiving ongoing feedback to help them build confidence and proficiency. Support remains available once they begin working independently, with experienced colleagues ready to provide consultation and advice.

The training emphasizes the responsible handling of sensitive and explicit content while offering clear guidance on detecting and reporting illegal material. It is designed to prepare moderators for various scenarios they may encounter during their tasks, including identifying illegal content, understanding moderation processes, and applying them effectively.

To support both internal moderators and external full-time equivalents, a Q&A reference table has been developed to address frequently asked questions. Additionally, a dedicated group chat enables moderators to seek real-time advice and guidance from one another. A shared communication platform further enhances collaboration, allowing team members to exchange best practices, clarify doubts promptly, and maintain continuous dialogue.

Through this comprehensive training and support system, WGCZ ensures that moderators are well-prepared to manage the complexities and challenges of effectively monitoring adult content.

### Quality assurance and review

Moderation decisions are periodically reviewed by the Head of Moderation Team to ensure alignment with the platform's policies and standards.

Moderation decisions at WGCZ are regularly reviewed to ensure accuracy and consistency across the platform:

- decisions made by the moderation team are periodically reviewed by the Head of Moderation Team to ensure they align with the platform's policies and standards.

- performance of checks for consistency in applying moderation guidelines, identifying any discrepancies, and providing feedback to respective moderation team members for continuous improvement.
- the Head of Moderation Team conduct these reviews, leveraging their extensive experience to maintain high standards of content moderation and to address any areas needing improvement.

### **Mental and physical support given to the moderation team**

WebGroup strongly emphasises its content moderation team's physical and mental well-being, recognising the demanding nature of its responsibilities. To safeguard physical health, each moderator is provided with an ergonomically designed workstation featuring supportive chairs, adjustable desks, and monitors that minimise eyestrain. Beyond the immediate work environment, the company subsidises access to diverse fitness and wellness facilities, such as gyms, swimming pools, and sports fields, as well as personalised guidance on maintaining a healthy lifestyle.

Equally integral is the organisation's commitment to mental well-being. WebGroup cultivates a culture where moderators feel confident discussing personal and work-related challenges, from stress management to emotional resilience. Management encourages and facilitates open communication, which actively addresses concerns and helps identify early signs of burnout. Support mechanisms include the possibility of confidential consultations at any time and the option to adjust working hours on a case-by-case basis when needed.

Training activities are designed to empower moderators with the skills and knowledge to cope effectively with the psychological demands of their role. In addition to flexible shift arrangements and the opportunity to work remotely, which helps individuals balance personal commitments and avoid excessive overtime, the company provides regular breaks throughout the workday to aid in rest and recovery. The ongoing dialogues with senior team members ensure that everyone remains informed and supported.

A combination of these measures has proven successful. Many moderators have chosen to remain with WebGroup for years, and the company's turnover rate remains exceptionally low. This enduring stability is a testament to the positive, supportive environment that WebGroup continually strives to foster.

## 8. Article 24(1)(a) DSA: Out-of-court dispute settlement submissions

Users, individuals, and entities have the right to challenge platform's decisions in a relevant court, and they may also be able to refer the decision to a certified dispute settlement body. During the Reporting Period, WGCZ received no disputes from certified out-of-court settlement bodies according to Article 21 of the DSA.

## 9. Article 42(3) DSA: Information on monthly active users

Table 16: Numbers of average monthly active users on the platform, broken down by each Member State

Member States	Number of AMR
Austria	440043
Belgium	760882
Bulgaria	270457
Croatia	310672
Cyprus	33228
Czech Republic	753822
Denmark	542395
Estonia	6438
Finland	490238
France	6287839
Germany	2570759
Greece	774917
Hungary	779998
Ireland	520326
Italy	4272228
Latvia	135496
Lithuania	97085
Luxembourg	38729
Malta	45104
Netherlands	1196569
Poland	2048746
Portugal	1168205
Romania	784082
Slovakia	287774
Slovenia	40569
Spain	8767891
Sweden	958808

To be clear, when calculating the average number of monthly service recipients, WGCZ does not consider users who use Virtual Private Network (“VPN”) connections to view content on XVideos. It is common for Internet users to visit adult content websites also through VPN connections to encrypt Internet traffic and disguise data about their actual geographical location. By using a VPN connection, a user effectively causes a website to believe that they are in another country. This is because instead of sending that user’s Internet traffic directly to the user’s Internet Service Provider (“ISP”), the VPN first routes traffic through a VPN server. That way, when the data is finally transmitted to the platform, it appears to originate from the VPN server and not the user’s personal device. VPN is used predominantly to bypass geolocation restrictions imposed in certain jurisdictions that ban access to adult-rated content websites/platforms (such as India or the United Arab Emirates). It follows that the overwhelming majority of users who rely on VPNs are located outside the EEA and, therefore, should be excluded from WGCZ calculations of its average monthly active recipients of service.