

## DSA Transparency Report - April 2025

### Introduction

This report covers the content moderation activities of X's international entity X Internet Unlimited Company ("XIUC") (formerly known as Twitter International Unlimited Company ("TIUC")) under the Digital Services Act (DSA), during the date range **1 October, 2024 to 31 March, 2025**.

We may refer to "notices" as defined in the DSA as "user reports" and "reports".

### Description of our Content Moderation Practices

Our content moderation systems are designed and tailored to mitigate systematic risks without unnecessarily restricting the use of our service and fundamental rights, especially freedom of expression. Content moderation activities are implemented and anchored on principled policies and leverage a diverse set of interventions to ensure that our actions are reasonable, proportionate and effective. Our content moderation systems blend automated and human review paired with a robust appeals system that enables our users to quickly raise potential moderation anomalies or mistakes.

### Policies

X's purpose is to serve the public conversation. Violence, harassment, and other similar types of behaviour discourage people from expressing themselves, and ultimately diminish the value of global public conversation. Our Rules are designed to ensure all people can participate in the public conversation freely and safely.

X has policies protecting user safety as well as platform and account integrity. The [X Rules and policies](#) are publicly accessible on our Help Center, and we are making sure that they are written in an easily understandable way. We also keep our Help Center regularly updated anytime we modify our Rules.

For the purposes of the summary tables below, the X policy titles in use at the start of the reporting period have been retained, even if they changed throughout the period.

### Enforcement

When determining whether to take enforcement action, we may consider a number of factors, including (but not limited to) whether:

- The behaviour is directed at an individual, group, or protected category of people;
- The report has been filed by the target of the abuse or a bystander;
- The user has a history of violating our policies;
- The severity of the violation; and

- The content may be a topic of legitimate public interest.

When we take [enforcement actions](#), we may do so either on a specific piece of content (e.g., an individual post or Direct Message) or on an account. We may employ a combination of these options. In most cases, this is because the behaviour violates the X Rules.

To enforce our Rules, we use a combination of machine learning and human review. Our systems are able to surface content to human moderators who use important context to make decisions about potential violations. This work is led by an international, cross-functional team with 24-hour coverage and the ability to cover multiple languages. We also have a complaints process for any potential errors that may occur.

To ensure that our human reviewers are prepared to perform their duties we provide them with a robust support system. Each human reviewer goes through extensive training and refreshers, they are provided with a suite of tools that enable them to do their jobs effectively, and they have a suite of wellness initiatives available to them. For further information on our human review resources, see the section titled “Human resources dedicated to Content Moderation”.

### **Reporting violations**

X strives to provide an environment where people can feel free to express themselves. If abusive behaviour happens, we want to make it easy for people to report it to us. EU users can also report any violation of our Rules or their local laws, no matter where such violations appear.

### **Transparency**

We always aim to exercise moderation with transparency. Where our systems or teams take action against content or an account as a result of violating our Rules or in response to a valid and properly scoped request from an authorised entity in a given country, we strive to provide context to users. Our [Help Center article](#) explains notices that users may encounter following actions taken. We promptly notify affected users about legal requests to withhold content, including a copy of the original request, unless we are legally prohibited from doing so. We have also updated our global [transparency centre](#) covering a broader array of our transparency efforts.

## **Content Moderation Governance Structure**

### **Own Initiative Content Moderation Activities**

X employs a combination of heuristics and machine learning algorithms to automatically detect content that we believe violates the [X Rules and policies](#) enforced on our platform. We use combinations of natural language processing models, image processing models and other sophisticated machine learning methods to detect potentially violative content. These models vary in complexity and in the outputs they produce. For example, the model used to detect abuse on the platform is trained on abuse violations detected in the past. Content flagged by these machine learning models are either reviewed by human content reviewers before an

action is taken or, in some cases, automatically actioned, based on the historical accuracy of the model's output. Heuristics are typically utilised to enable X to react quickly to new forms of violations that emerge on the platform. Heuristics are common patterns of behaviours, text, or keywords that may be typical of a certain category of violations. Pieces of content detected by heuristics may also get reviewed by human content reviewers before an action is taken on the content. These heuristics are used to flag content for review by human agents proactively.

## **Testing, Evaluation, and Iteration**

Automated enforcements under the [X Rules and policies](#) undergo rigorous testing before being applied to the live product. Both machine learning and heuristic models are trained and/or validated on thousands of data points and labels (e.g., violative or non-violative) including those that are generated by trained human content moderators. For example, inputs to content-related models can include the text within the post itself, the images attached to the post, and other characteristics. Training data for the models comes from both the cases reviewed by our content moderators, random samples, and various other samples of pieces of content from the platform.

## **Use of Human Moderation**

Before any given algorithm is launched to the platform, we verify its detection of policy violating content or behaviour by drawing a statistically significant test sample and performing item-by-item human review. Reviewers have expertise in the applicable policies and are trained by our Policy teams to ensure the reliability of their decisions. Human review helps us to confirm that these automations achieve a level of precision, and sizing helps us understand what to expect once the automations are launched.

In addition, humans proactively conduct manual content reviews for potential policy violations. We conduct proactive sweeps for certain high-priority categories of potentially violative content both periodically and during major events, such as elections. Content moderators also proactively review content flagged by heuristic and machine learning models for potential violations of other policies, including our [adult content](#), [violent content](#), [child sexual exploitation](#) (CSE) and [violent and hateful entities](#) policies.

Once reviewers have confirmed that the detection meets an acceptable standard of accuracy, we consider the automation to be ready for launch. Once launched, automations are monitored dynamically for ongoing performance and health. If we detect anomalies in performance (for instance, significant spikes or dips against the volume we established during sizing, or significant changes in user complaint/overtake rates), our Engineering (including Data Science) teams - with support from other functions - revisit the automation to diagnose any potential problems and adjust the automations as appropriate.

## **Automated Moderation Activity Examples**

A vast majority of all accounts that are suspended for the promotion of terrorism and CSE are proactively flagged by a combination of technology and other purpose-built internal proprietary

tools. When we remove CSE content with these automated systems, we immediately report it to the National Center for Missing and Exploited Children (NCMEC). NCMEC makes reports available to the appropriate law enforcement agencies around the world to facilitate investigations and prosecutions.

Our current methods deploy a range of internal tools and third party solutions that utilise industry standard hash libraries (e.g., [PhotoDNA](#)) to ensure known CSAM is caught prior to any user reports being filed. We leverage the hashes provided by NCMEC and industry partners. We scan media uploaded to X for matches to hashes of known CSAM sourced from NGOs, law enforcement and other platforms. We also have the ability to block keywords and phrases from Trending and block search results for certain terms that are known to be associated with CSAM.

We [commit](#) to continuing to invest in technology that improves our capability to detect and remove, for instance, terrorist and violent extremist content online before it can cause user harms, including the extension or development of digital fingerprinting and AI-based technology solutions. Our participation in multi-stakeholder communities, such as the [Christchurch Call to Action](#), [Global Internet Forum to Counter Terrorism](#) and EU Internet Forum (EUIF), helps to identify emerging trends in how terrorists and violent extremists are using the internet to promote their content and exploit online platforms.

You can learn more about our commitment to eradicating CSE and terrorist content, and the actions we've taken [here](#). Our continued investment in proprietary technology is steadily reducing the burden on people to report this content to us.

### **Scaled Investigations**

These moderation activities are supplemented by scaled human investigations into the tactics, techniques and procedures that bad actors use to circumvent our Rules and policies. These investigations may leverage signals and behaviours identifiable on our platform, as well as off-platform information, to identify large-scale and/or technically sophisticated evasions of our detection and enforcement activities. For example, through these investigations, we are able to detect coordinated activity intended to manipulate our platform and artificially amplify the reach of certain accounts or their content.

### **Indications of Accuracy for Content Moderation**

The possible rate of error of the automated and human means used in enforcing [X Rules and policies](#) is represented by the number of Content Removal Complaints (appeals) received and the number of Content Removal Complaints that resulted in reversal of our enforcement decision (successful appeals) by remediation type and by country.

### **Closing Statement on Content Moderation Activities**

Our content moderation systems are designed and tailored to mitigate systematic risks without unnecessarily restricting the use of our service and fundamental rights, especially freedom of expression. Content moderation activities are implemented and anchored on principled policies and leverage a diverse set of interventions to ensure that our actions are reasonable, proportionate and effective. Our content moderation systems blend automated and human review paired with a robust appeals system that enables our users to quickly raise potential moderation anomalies or mistakes.

**Human resources dedicated to Content Moderation**

Today, we have 1486 people working in content moderation. Our teams work on both initial reports as well as on complaints of initial decisions across the world (and are not specifically designated to only work on EU matters).

**Linguistics Expertise of our Content Moderation Team**

X’s scaled operations team possesses a variety of skills, experiences, and tools that allow them to effectively review and take action on reports across all of our Rules and policies. X has analysed which languages are most commonly found in reports reviewed by our content moderators, and has hired content moderation specialists who have professional proficiency in these commonly spoken languages. The following table is a summary of the number of people in our content moderation team who possess professional proficiency in the languages that are most commonly contained in reported content in the EU on our platform:

Primary Language	People
Bulgarian	1
English	1,307
French	63
German	63
Italian	1
Polish	1
Portuguese	17
Spanish	33
Total	1,486

In addition to the primary language support, we have also have people supporting additional languages. The following is the list of secondary EU language support:

Secondary Language	People
--------------------	--------

Bulgarian	1
Romanian	1
French	73
German	64
Greek	1
Irish	1
Italian	3
Latvian	1
Polish	1
Portuguese	22
Spanish	53
Total	221

Please note that the numbers included in the secondary language support are not separate or distinct from the numbers included in the primary language support data. Additionally, the English language is not indicated as a secondary language category in the table above, since all agents with different primary language capability also speak English as well.

#### Qualifications of our Content Moderation Team

Content Moderation Team Qualifications	
Years in Current Role	Headcount
0 to 1	602
1 to 2	157
2 to 3	213
3 to 4	243
4 to 5	87
5 to 6	47
6 to 7	53
7 or more	84

The above table includes all moderators who support EU member state languages as of March 2025. The content moderation team collectively provides linguistic capacity in multiple languages. In situations where we need additional language support, we use translation services and/or machine translation tools, to investigate and address challenges in additional languages. Additionally, content moderators will leverage playbooks that contain colloquial terms and phrases that are consistently being updated to reflect various EU languages spoken within the region and trends.

Moderators are recruited using a standard job description that includes a language requirement which states that the candidate should be able to demonstrate written and spoken fluency in the language and have at least one year of work experience for entry-level positions. In the interview and application process, each agent candidate must meet certain linguistic standards to be considered “language qualified”. This determination is made through multiple tests (i.e. written, oral, etc.) of the candidate’s respective language, to determine their respective proficiency level. Candidates must also meet the educational and background requirements in order to be considered, as well as demonstrate an understanding of current events for the country or region of content moderation they will support.

### **Organisation, Team Resources, Expertise, Training and Support of our Team that Reviews and Responds to Reports of Illegal Content**

#### **Description of the team**

X has built a specialised team made up of individuals who have received specific training in order to assess and take action on illegal content that X becomes aware of via reports or other processes on our own initiative. This team consists of different tier groups, with higher tiers consisting of more senior, or more specialised, individuals.

When handling a report of illegal content or a complaint against a previous decision, content and senior content reviewers first assess the content under X’s Rules and policies. If no violation of X’s Rules and policies is determined warranting a global removal of the content, the content moderators will assess the content for potential illegality under Local Laws. If more detailed investigation is required, content moderators can escalate reports to experienced policy and/or legal request specialists who have also undergone in-depth training and/or have language expertise in the respective case’s language. These individuals take appropriate action after carefully reviewing the report and/or complaint in close detail. In cases where the specialist team cannot determine a final decision or action on a case, regarding the potential illegality of the reported content, the report will be discussed with in-house legal counsel. Everyone involved in this process works closely together with daily exchanges through meetings and other channels to ensure the timely and accurate handling of reports. Additionally, in the instance that a case warrants in-house legal counsel review, the lessons learned and actions made on this case will be disseminated to all relevant content moderator parties to ensure consistency in

review and an understanding of best practices made by the agent, if a similar case is encountered in the future.

Furthermore, all teams involved in solving these reports closely collaborate with a variety of other policy teams at X who focus on X Rules and policies. This cross-team effort is particularly important in the aftermath of tragic events, such as violent attacks, to ensure alignment, swift consistency in review, and the same potential remediation actions if the content is found violative.

Content moderators are supported by team leads, subject matter experts, quality auditors and trainers. We hire people with diverse backgrounds in fields such as law, political science, psychology, communications, sociology and cultural studies, and languages.

### **Training and support of persons processing legal requests**

All team members, i.e. all employees hired by X as well as vendor partners working on these reports, are trained and retrained regularly on our tools, processes, Rules and policies, including special sessions on cultural and historical context. Initially when joining the team at X, each individual follows an onboarding program and receives individual mentoring during this period, as well as thereafter through our Quality Assurance (QA) program, in house and external counsels (for internal employees).

All team members have direct access to robust training and workflow documentation for the entirety of their employment, and are able to seek guidance at any time from trainers, leads, and internal specialist legal and policy teams as outlined above, as well as managerial support.

Updates about significant current events or Rules and policy changes are shared with all content reviewers in real time, to give guidance and facilitate balanced and informed decision making. In the case of Rules and policy changes, all training materials and related documentation is updated. Calibration sessions are carried out frequently during the reporting period. These sessions aim to increase collective understanding and focus on the needs of the content reviewers in their day-to-day work, by allowing content moderators to ask questions and discuss aspects of recently reviewed cases, X's Rules and policies, and/or local laws.

The entire team also participates in obligatory X Rules and policies refresher training as the need arises or whenever Rules and policies are updated. These trainings are delivered by the relevant policy specialists who were directly involved in the development of the Rules and policy change. For these sessions we also employ the "train the trainer" method to ensure timely training delivery to the whole team across all of the shifts. All team members use the same training materials to ensure consistency.

QA is a critical measure to the business to help ensure that we are delivering a consistent service at the desired level of quality to our key stakeholders, both externally and internally as it pertains to our case work. We have a dedicated QA team within our vendor team to help us



identify areas of opportunity for training and potential defect detection in our workflow or Rules and policies. The QA specialists perform quality checks of reports to ensure that content is actioned appropriately.

The standards and procedures within the QA team ensure the team's QA is assessed equally, objectively, efficiently and transparently. In case of any mis-alignments, additional training is scheduled, to ensure the team understands the issues and can handle reports accurately.

In addition, given the nature and sensitivity of their work, the entire team has access to online resources and regular onsite group and individual sessions related to resilience and well-being. These are provided by mental health professionals. Content reviewers also participate in resilience, self-care, and vicarious trauma training as part of our mandatory wellness plan during the reporting period.

### **Training and Support provided to those Persons performing Content Moderation Activities for our XIUC Terms of Service and Rules**

Training is a critical component of how X maintains the health and safety of the public conversation through enabling content moderators to accurately and efficiently moderate content posted on our platform. Training at X aims to improve the content moderators' enforcement performance and quality scores by enhancing content moderators' understanding and application of X Rules through robust training and quality programs and a continuous monitoring of quality scores.

### **Training Process**

There is a robust training program and system in place for every workflow to provide content moderators with the adequate work skills and job knowledge required for processing user cases. All content moderators must be trained in their assigned workflows. These focus areas ensure that content moderators are set up for success before and during the content moderation lifecycle, which includes:

- Training analysis/design focused on agent and learning needs;
- Classroom training with expert trainers;
- Nesting period to apply new skills;
- Cross-skilling opportunities;
- Upskilling opportunities;
- Refresher programs;
- New launch/update roll-outs process; and
- Remediation plans.

### **Training Analysis and Design**

Before commencing design work on any content moderators program or resource, a rigorous learner analysis is conducted in close collaboration with training specialists and quality analysts to identify performance gaps and learning needs. Each program is designed with key stakeholder engagement and alignment. The design objective is to adhere to visual and learning design principles to maximise learning outcomes and ensure that agents can perform their tasks with accuracy and efficiency. This is achieved by making sure that the content is:

- Easy to experience;
- Easy to understand; and
- Easy to apply.

X's training programs and resources are designed based on needs, and a variety of modalities are employed to diversify the content moderators learning experience, including:

- Self-led learning: microlearning, scenario-based learning, e-learning modules, and gamification (where appropriate);
- Virtual live instructor-led trainings;
- Face-to-face classroom training; and
- Videos.

### **Classroom Training**

Classroom training is delivered either virtually or face-to-face by expert trainers. Classroom training activities can include:

- Instructor-led policy training;
- Interactive e-learnings;
- Scenario-based learning sets;
- Shadowing sessions with seasoned agents;
- Guided casework sessions with trainers; and
- Knowledge checks, quizzes and assessments.

### **Onboarding and Ramp Up**

When content moderators successfully complete their classroom training program, they undergo an onboarding period. The onboarding phase includes case study by observation, demonstration and hands-on training on live cases. Onboarding activities include content moderator shadowing, guided case work, Question and Answer sessions with their trainer, coaching, feedback sessions, etc. Quality audits are conducted for each onboarding content moderator and content moderators must be coached for any mis-action spotted in their quality scores the same day that the case was reviewed. Trainers conduct needs assessment for each onboarding content moderator and prepare refresher training accordingly. After the onboarding period, content is evaluated on an ongoing basis with the QA team to identify gaps and address potential problem areas. There is a continuous feedback loop with quality analysts across the different workflows to identify challenges and opportunities to improve materials and address performance gaps.

## Up-Skilling

When a content moderator needs to be upskilled they receive training of a specific workflow within the same pillar that the content moderator is currently working. The training includes a classroom training phase and onboarding phase which is specified above.

## Refresher Sessions

Refresher sessions take place when a content moderator has previously been trained, has access to all the necessary tools, but would need a review of some or all topics. This may happen for content moderators who have been on prolonged leave, transferred temporarily to another content moderation policy workflow, or ones who have recurring errors in the quality scores. After a needs assessment, trainers are able to pinpoint what the content moderator needs and prepare a session targeting their needs and gaps.

## New Launch/Update Roll-Outs

There are also processes that require new and/or specific product training and certification. These new launches and updates are identified by X and the knowledge is transferred to the content moderators.

## Remediation Plans

There are remediation plans in place to support content moderators who do not pass the training or onboarding phase, or are not meeting quality requirements.

## Relevant Data for the Reporting Period

### Member States Orders to Act Against Illegal Content

Removal Orders Received - 1 October to 31 March		
Illegal Content Category	France	Spain
Illegal or harmful speech	1	3
Unsafe and illegal products	11	

### Removal Orders Median Handle Time (Hours) - 1 October to 31 March



Unsafe and illegal products						1	5			1	1									1
Violence		6	4	1	5	164	305	5		14	14		33	13	1	2				31

**Information Request Median Handle Time (Hours) - October to 31 March**

Illegal Content Category	Austria	Belgium	Denmark	Estonia	Finland	France	Germany	Greece	Hungary	Ireland	Italy	Luxembourg	Netherlands	Poland	Portugal	Romania	Slovenia	Spain
Data protection & privacy violations	0.4					217.7	189.3	46.0							0.1		49.3	218.8
Illegal or harmful speech	380.6	115.9				137.8	139.7	361.1		159.3	217.1		442.6	23.2	2.7	74.3		96.6
Intellectual property infringements						0.6	1.6				451.5							
Issue Unknown						529.2	27.4			45.5			20.6	34.6				0.8
Negative effects on civic discourse or elections						51.9	74.4							0.4				
Non-consensual behavior						149.5	77.7				461.5							
Pornography or sexualized content						243.4	71.5				368.6		404.5					367.0
Protection of minors			1.4			4.8	2.6	4.2		70.3	35.9		201.6	6.7	22.0			2.1
Risk for public security	121.5	142.8			83.7	53.9	99.7	18.4	44.0	144.3	22.1		1.9	482.7	457.0			83.5
Scams and fraud	28.8	145.0				139.4	78.1	116.3	28.6	1.5	30.0	27.2	62.7	46.6	104.9			203.7
Scope of platform service																		104.6
Self-harm							83.2	0.1										233.6
Unsafe and illegal products						382.3	96.0			4.9	342.8							459.2
Violence		288.9	431.4	0.6	76.3	166.5	90.8	44.2		46.6	136.2		106.1	4.5	68.0	86.7		263.4

**Information Request Median Time To Acknowledge Receipt**

X provides an automated acknowledgement of receipt of information requests submitted by law enforcement through our [Legal Request submission portal](#). As a consequence of this immediate acknowledgement of receipt, the median time is zero.

**Important Notes about Information Requests:**

- To improve clarity, we have omitted countries and violation types with no legal requests from the tables above;
- The content category for each request is determined by the information law enforcement provides when submitting requests through the X online submission platform. If law enforcement does not provide sufficient information during form submission, the category is determined based on the allegations provided in the legal process. Where multiple illegal content categories were provided, only the gravamen offence was included;
- The median handling time is the time between receiving the order and either: 1) disclosing information to law enforcement if the order is valid; or 2) pushing back due to









	Terms of Service and/or X's Rules or Policies	Illegal or harmful speech	45	35	17	5		59	12	1340	9	16	708	739	14	3	29	52	6			1			65	112	35	23	1	5	295	34
	Terms of Service and/or X's Rules or Policies	Negative effects on civic discourse or elections	1			1		2		9		1	24	32	1			2							1			1		6	1	
	Terms of Service and/or X's Rules or Policies	Non-commercial behavior	1	7		2			8	75	2	4	55	59	6	4	6	9	1			1			10	3	6		1	37	2	
	Terms of Service and/or X's Rules or Policies	Pornography or sexualized content	31	33	8	4		11	23	615	14	22	395	470	8	12	26	82	6	7	2			91	66	16	14	3	1	176	32	
	Terms of Service and/or X's Rules or Policies	Protection of minors	35	62	5	7	1	15	26	1225	54	345	916	4486	18	24	69	114	1	25	2	7	730	576	29	12	17	4	3788	34		
	Terms of Service and/or X's Rules or Policies	Risk for public security	4	7		4	1	6	93	74		8	194	380	3	1	5	5	1	1	2	1	14	27	11	7	1		41	9		
	Terms of Service and/or X's Rules or Policies	Scams and fraud	3			1				27		3	38	9	4	1						1		2	5	2		1	29			
	Terms of Service and/or X's Rules or Policies	Scope of platform service			1			1	1	6		2	43	12	1							1		3	1		2	1		5	1	
	Terms of Service and/or X's Rules or Policies	Self-harm	1	5				2	87	1		28	31	1	1	2	4					1	5	10	1		1	2	22	5		
	Terms of Service and/or X's Rules or Policies	Unsafe and illegal products	3	5	1	1	1	11	34	108		2	123	114				8	6	1	2			6	1				24	5		
	Terms of Service and/or X's Rules or Policies	Violence	32	24	3	7		34	7	1828	2	120	361	782	11	38	29	84	1	2				64	80	41	24	1	2	383	15	
	Terms of Service and/or X's Rules or Policies	Animal welfare																														
	Terms of Service and/or X's Rules or Policies	Data protection & privacy violations	101	37	6	4	3	14	13	433		12	178	171	8	14	14	58	7	3	3	1	78	59	26	4	3	1	313	21		
	Terms of Service and/or X's Rules or Policies	Illegal or harmful speech	128	272	47	47	35	140	107	2213	18	44	1687	1671	170	26	162	397	29	23	15	5	600	679	199	79	6	18	1542	98		
	Terms of Service and/or X's Rules or Policies	Negative effects on civic discourse or elections	871	789	351	101	132	1279	238	24730	110	370	13889	13352	427	107	1008	2577	141	108	76	11	1376	2057	782	538	62	50	13188	461		
	Terms of Service and/or X's Rules or Policies	Non-commercial behavior	91	91	32	16	12	106	59	1274	20	53	685	3300	40	35	103	147	18	17	6		188	512	71	546	9	4	344	36		
	Terms of Service and/or X's Rules or Policies	Pornography or sexualized content	26	46	25	8	18	42	39	3031	12	38	666	358	26	26	60	111	5	12	5	3	171	173	40	10	6	1	458	151		
	Terms of Service and/or X's Rules or Policies	Protection of minors	148	277	110	11	33	209	150	2028	32	45	1168	915	83	152	196	822	14	23	14	4	313	321	225	112	10	9	1350	95		
Manual Closure			42	123	10	10	6	31	58	1078	47	140	688	1365	86	25	106	154	5	17	7	4	478	307	76	17	9	7	3402	55		

No Violation Found

		Terms of Service and/or X's Rules or Policies	Risk for public security	65	132	24	26	10	121	56	1119	25	55	1480	1574	44	14	94	162	45	17	6	1	170	302	60	231	19	5	517	70	
		Terms of Service and/or X's Rules or Policies	Scams and fraud	311	975	80	97	47	312	275	4058	114	180	1307	3455	237	254	603	1116	120	53	37	17	1059	802	655	411	43	49	3494	438	
		Terms of Service and/or X's Rules or Policies	Scope of platform service	8	7	2	7		8	7	544	14	5	70	100	6	7	9	17	2		1		21	21	5	8	2		93	2	
		Terms of Service and/or X's Rules or Policies	Self-harm	7	16	3		3	24	4	505	1	20	149	172	7	14	16	31	1	1		2	39	44	13	4	4		172	59	
		Terms of Service and/or X's Rules or Policies	Unsafe and illegal products	33	58	28	3	7	36	53	709	12	24	651	393	4	20	30	101	15	7	3	2	119	107	53	29	7	8	359	31	
		Terms of Service and/or X's Rules or Policies	Violence	145	197	80	74	48	180	86	6871	63	414	2294	3321	114	82	219	557	62	42	54	25	305	295	149	143	53	48	2314	229	
		Terms of Service and/or X's Rules or Policies	Illegal or harmful speech								2			2									1							1		
		Terms of Service and/or X's Rules or Policies	Pornography or sexualized content								1																					
		Terms of Service and/or X's Rules or Policies	Protection of minors											1																	2	
		Terms of Service and/or X's Rules or Policies	Risk for public security								1																				1	
		Terms of Service and/or X's Rules or Policies	Scams and fraud											1											1							
		Terms of Service and/or X's Rules or Policies	Self-harm	1	1			1	2		39			1	32	2	3		11		2			4	11	2				35	5	
		Terms of Service and/or X's Rules or Policies	Violence											1																		

Manual Closure

Reports of Illegal Content Median Handle Time (Hours) - 1 October to 31 March																															
Reason Code	Austria	Belgium	Bulgaria	Canada	Cyprus	Czechia	Denmark	Estonia	EU	Finland	France	Germany	Greece	Hungary	Ireland	Italy	Latvia	Lithuania	Luxembourg	Malta	Netherlands	Poland	Portugal	Romania	Slovakia	Slovenia	Spain	Sweden			
Animal welfare	4.4	7.4	6.1	4.7	2.5	3.8	3.2	10.1	1.4	2.3	16.7	1.8	8.4	3.1	1.2	3.3	2.3	2.9	0.3	13.5	8.6	9.2	5.1	2.6	5.9	2.7	10.7	2.3			
Data protection & privacy violations	3.2	3.5	4.6	1.9	4.3	3.5	5.7	4.4	4.5	4.1	6.9	3.3	3.7	1.4	2.7	4.8	10.5	12.5	2.1	0.9	3.0	9.3	6.5	3.1	8.8	0.7	3.5	4.9			
Illegal or harmful speech	1.8	2.5	3.5	1.9	2.0	1.9	3.7	1.5	2.6	2.5	5.8	1.5	2.7	2.6	2.9	1.8	1.2	1.5	3.0	6.1	3.0	2.2	3.9	1.8	1.8	1.1	1.9	2.9			







User Report	Manual Review	Child Sexual Exploitation	10	26	12	2	1	9	15	1	11	407	242	7	5	24	53	1	12	1	108	46	16	19	10	68	26
User Report	Manual Review	Discussed Individuals								1	1										1						3
User Report	Manual Review	Distribution of Hacked Materials										1															
User Report	Manual Review	Financial Scam	1										4	4	1										1		3
User Report	Manual Review	Hateful Conduct	9	16	8	2	3	8	8	4	8	217	107	9	3	19	24	6	7	2		34	38	22	13	3	5
User Report	Manual Review	Help with my compromised account											9														70
User Report	Manual Review	Illegal or certain regulated goods and services	288	323	546	224	140	924	221	112	378	16881	17173	296	249	279	3564	190	157	71	55	4012	1568	388	597	119	104
User Report	Manual Review	Intellectual property infringements	4	19	2	1	1	7	7	1	6	323	87	13	5	8	59	5	4	1	2	57	48	41	20	1	4
User Report	Manual Review	Misleading & Deceptive Identifies	1021	697	513	314	361	1109	459	237	347	7525	5814	1046	577	240	2712	307	398	21	28	3004	2109	640	2106	94	399
User Report	Manual Review	Non-Consensual Nudity	34	69	88	15	12	52	30	6	34	1937	692	26	38	53	233	28	18	12	2	512	345	61	130	19	9
User Report	Manual Review	Other	28	26	9	9	2	17	27	9	14	231	199	13	17	17	69	4	8	1	144	65	24	31	7	6	
User Report	Manual Review	Perpetrators of Violent Attacks	3	2	1	3		4	5		4	17	32	2		12	12	3	6	1		13	23	13	1	1	
User Report	Manual Review	Platform Manipulation & Spam	325	321	200	352	106	570	224	117	200	8206	4816	297	210	602	2161	139	110	44	34	2514	1151	410	640	89	
User Report	Manual Review	Private Information & Media	2	8	2		1	2			4	55	46	4	1	8	9				24	18	7	4	2		
User Report	Manual Review	Sensitive Media	1		1			2	3	3	59	22	3	1	2	6	1				9	4	1	2	2	1	
User Report	Manual Review	Suicide & Self Harm	8	7	1		2	5	2	1	5	18	39	2	1	10	17	1	3		18	25	2	6	1	29	
User Report	Manual Review	Unknown	1								1	2	2								1						
User Report	Manual Review	Username Squatting									1	1														1	
User Report	Manual Review	Violent & Hateful Endtias	27	21	12	2	13	8	12	5	17	185	193	12	7	10	46	4	6	4	1	170	32	7	39	1	
User Report	Manual Review	Violent Speech	131	204	69	71	21	146	97	24	124	1996	1303	103	77	258	537	18	56	28	13	695	722	201	202	35	

Overall Figures

COMPLAINTS OF ACTIONS TAKEN FOR ILLEGAL CONTENT RECEIVED - 1 October to 31 March		Austria	Belgium	Bulgaria	Croatia	Cyprus	Czechia	Denmark	Estonia	Finland	France	Germany	Greece	Hungary	Ireland	Italy	Latvia	Lithuania	Netherlands	Poland	Portugal	Romania	Slovakia	Slovenia	Spain	Sweden
Complaints Received		50	24	53	3	3	128	4	2	4	122	204	24	26	48	66	2	1	127	94	43	41	2	1	848	32
Overturned Appeals		7	3	6	0	0	23	1	0	0	12	66	6	4	8	14	1	0	14	11	12	9	0	0	165	1
Median Time to Respond (hours)		0.6	1.7	11.2	1.8	4.6	0.8	3.5	2.2	0.9	1.1	1.2	1.5	1.0	2.4	0.6	0.5	2.3	1.4	0.6	3.6	2.3	0.4	0.5	0.7	2.4

COMPLAINTS OF ACTIONS TAKEN FOR XUC TERMS OF SERVICE AND RULES VIOLATIONS RECEIVED - 1 October to 31 March		Austria	Belgium	Bulgaria	Croatia	Cyprus	Czechia	Denmark	Estonia	Finland	France	Germany	Greece	Hungary	Ireland	Italy	Latvia	Lithuania	Luxembourg	Malta	Netherlands	Poland	Portugal	Romania	Slovakia	Slovenia	Spain	Sweden
Appeal Category	Music																											
Content Action Category	Complaints Received	457	651	176	132	86	273	268	70	274	5439	6169	342	198	695	1714	84	113	82	61	2365	835	600	408	75	85	4134	603

Content Action Complaints	Overturned Appeals	43	79	20	13	3	29	30	4	18	667	449	31	11	106	67	11	12	3	5	171	78	58	35	4	8	569	44	
Content Action Complaints	Median Time to Respond (Hours)	2.3	2.8	10.2	4.0	381.5	1.1	11.2	3.1	162.0	2.1	13.7	230.3	364.0	0.5	268.1	246.7	0.5	217.3	183.8	366.5	3.8	1.4	2.8	347.7	3.0	0.8	227.7	
Live Feature Action Complaints	Complaints Received		6	2			10	4	3	24	174	209	9	13	8		6	4	2	1	140	47	14	18	2	3	105	36	
Live Feature Action Complaints	Overturned Appeals		0	0			0	0	0	16	13	1	3	1			0	2	0	0	7	3	3	1	0	0	16	1	
Live Feature Action Complaints	Median Time to Respond (Hours)						23.5	217.7	23.2	23.9	23.8	23.9	12.7	24.0	48.3		13.6	16.6	55.1	23.6	23.9	21.9	23.8	23.5	22.2	19.0	23.8	23.9	
Sensitive Media Action Complaints	Complaints Received	14	58	31	10	7	37	65	22	34	306	275	39	15	28	136	15	8	1		187	111	47	45	11	1	150	89	
Sensitive Media Action Complaints	Overturned Appeals	7	46	2	6	5	28	48	19	27	181	175	27	9	22	86	12	4	1		126	71	32	37	4	1	112	52	
Sensitive Media Action Complaints	Median Time to Respond (Hours)	0.9	1.4	674.5	0.8	1.2	1.6	0.7	0.3	2.2	1.4	1.3	2.3	3.1	2.1	1.0	0.8	0.3	2.6			1.2	1.0	2.4	0.9	0.1	1.0	1.8	1.4
Account Suspension Complaints	Complaints Received	2149	3184	1347	603	422	2034	1730	583	4186	28919	38332	1857	1806	2317	9546	738	883	381	133	23597	9317	3190	3711	713	389	15854	4523	
Account Suspension Complaints	Overturned Appeals	236	365	138	75	47	241	193	68	481	3404	4261	220	195	313	906	80	102	37	17	2821	1641	352	428	99	42	1877	614	
Account Suspension Complaints	Median Time to Respond (Hours)	4.2	2.5	14.1	6.3	2.8	3.9	3.2	6.1	37.1	3.5	4.7	2.4	1.1	3.1	2.5	14.7	8.2	8.2	2.5	9.4	2.3	6.4	1.5	6.1	1.2	6.0		
Restricted Reach Complaints	Complaints Received	235	316	134	120	30	251	256	87	277	1578	2307	220	119	848	840	49	72	19	23	1649	753	252	322	61	72	1478	561	
Restricted Reach Complaints	Overturned Appeals	71	93	42	33	10	72	84	33	82	499	757	66	39	296	223	16	26	6	8	558	227	73	94	16	22	493	218	
Restricted Reach Complaints	Median Time to Respond (Hours)	1.4	0.8	0.5	0.7	1.8	0.7	1.3	0.8	1.1	1.2	1.1	0.8	1.0	1.2	1.5	1.6	0.5	0.7	1.2	1.0	1.0	1.3	0.6	0.8	0.7	1.3	0.9	

INDICATORS OF ACCURACY FOR CONTENT MODERATION  
VISIBILITY FILTERING INDICATORS

Metric	Enforcement	Policy	Bulgarian	Czech	Danish	Dutch	English	Estonian	Finnish	French	German	Greek	Hungarian	Italian	Latvian	Lithuanian	Polish	Portuguese	Romanian	Slovenian	Spanish	Swedish	
Appeal Rate	Automated Means	Hateful Conduct					8.9%	2.0%	6.9%	5.2%	2.0%	7.3%	0.0%	0.4%	6.3%	5.5%	0.0%	5.4%	3.5%	1.2%	8.0%	2.5%	14.0%
	Manual Review	Abuse & Harassment	28.6%	23.1%	0.0%	7.0%	2.4%	0.0%	0.0%	9.0%	12.0%	7.7%	0.0%	12.4%	0.0%	0.0%	3.4%	3.0%	8.6%	0.0%	5.7%	2.3%	
	Manual Review	Hateful Conduct	2.7%	2.2%	0.9%	2.1%	1.1%	0.9%	2.4%	1.1%	3.0%	0.7%	0.0%	3.0%	0.0%	0.0%	0.8%	1.2%	1.8%	0.9%	1.9%	1.5%	
Appeal Rate	Manual Review	Violent Speech	0.0%	0.6%	0.0%	0.0%	0.0%	0.0%	0.0%	0.8%	1.3%	0.0%	0.0%	0.0%	0.0%	0.0%	1.5%	1.2%	1.2%	0.0%	1.5%	4.0%	

Note: Cells that are blank mean that there was no enforcement. For cells containing "0.0%" values, there were no cases of successful appeals or overturns.

Metric	Enforcement	Policy	Bulgarian	Czech	Danish	Dutch	English	Estonian	Finnish	French	German	Greek	Hungarian	Italian	Latvian	Lithuanian	Polish	Portuguese	Romanian	Slovenian	Spanish	Swedish	
Overturn Rate	Automated Means	Hateful Conduct		46.7%	19.3%	59.6%	40.4%	64.3%	27.3%	31.3%	50.4%			0.0%	68.0%	33.3%		52.4%	36.4%	33.3%	80.0%	39.7%	72.2%
	Manual Review	Abuse & Harassment	100.0%	100.0%		80.0%	63.4%		90.1%	113.3%	80.0%				76.9%			53.8%	33.3%	80.0%		80.0%	
	Manual Review	Hateful Conduct	0.0%			26.2%	36.9%		44.4%	32.7%	38.5%				31.0%			16.1%	33.3%	68.7%		25.3%	0.0%
Overturn Rate	Manual Review	Violent Speech		0.0%		50.0%	14.5%			17.4%	12.3%				18.2%			33.3%	33.3%	0.0%		19.2%	16.7%

Note: Cells that are blank mean that there was no enforcement. For cells containing '0.0%' value, there were no cases of successful appeals or overturns.

INDICATORS OF ACCURACY FOR CONTENT REMOVAL

Metric	Enforcement	Policy	Bulgarian	Croatian	Czech	Danish	Dutch	English	Estonian	Finnish	French	German	Greek	Hungarian	Irish	Italian	Latvian	Polish	Portuguese	Romanian	Slovak	Slovenian	Spanish	Swedish			
Appeal Rate	Automated Means	Sensitive Media			0.0%			0.4%			0.0%						0.0%		0.0%					0.0%			
		Non-Consensual Nudity	0.0%	0.0%	0.0%	0.0%	0.0%	0.4%	0.4%		0.0%	0.0%	0.0%	11.1%	0.0%			3.7%		0.0%	22.2%	0.0%			2.8%	0.0%	
		Abuse & Harassment							0.0%	0.0%				0.0%	0.0%					0.0%						0.0%	
		Private Information & media						0.0%	0.0%	2.2%			2.9%	4.6%				1.2%		0.0%						0.0%	0.0%
		Perpetrators of Violent Attacks							0.0%				0.0%													0.0%	0.0%
		Violent Speech	0.0%	0.5%	1.6%	1.1%	1.6%	4.4%	0.0%	0.0%	1.5%	5.0%	6.4%	1.7%	1.3%	0.0%	1.0%	0.0%	1.1%	4.3%	0.4%	1.4%			5.5%	2.2%	
		Child Sexual Exploitation							0.0%	0.0%			0.0%	0.0%					0.0%	0.0%	0.0%	0.0%	0.0%			0.0%	0.0%
		Other	0.0%			0.0%	0.0%	0.0%	0.0%		0.0%	0.0%	1.0%	0.0%	0.0%	0.0%			0.0%	0.0%	0.0%	0.0%	0.0%			0.0%	0.0%
		Deceased Individuals							0.0%	0.9%			0.0%	0.0%	0.0%					0.0%	0.0%	0.0%	0.0%			6.3%	0.0%
		Suicide & Self Harm	0.0%	0.0%	0.0%	0.0%	2.9%	5.0%		5.0%		0.0%	6.4%	11.3%	2.3%	0.0%		1.0%	0.0%	0.8%	5.5%	0.0%	0.0%	0.0%		6.0%	0.0%
	Sensitive Media	0.0%	0.0%	0.0%	0.0%	3.1%	2.9%		2.9%		0.0%	4.6%	8.0%	0.0%	0.0%		0.4%	0.0%	0.4%	6.9%	0.0%	0.0%			3.7%	0.0%	
	Hateful Conduct								0.0%			0.0%														0.0%	
	Abuse & Harassment	0.0%	2.3%	1.0%	0.0%	3.1%	0.8%		0.8%		1.2%	6.8%	8.6%	0.0%	0.0%	0.0%	0.6%	0.0%	1.0%	8.8%	0.0%	0.0%	0.0%		7.7%	0.0%	
	Non-Consensual Nudity	0.0%	0.0%	0.0%	0.0%	0.7%	0.8%		0.8%		6.3%	1.8%	2.1%	0.0%	0.0%		0.4%		0.0%	0.8%	0.0%	0.0%			1.0%	4.5%	
	Private Information & media	0.0%	33.3%	0.0%	16.7%	1.1%	3.7%		3.7%		6.0%	3.4%	10.3%	0.0%	0.0%		0.0%		2.6%	5.9%	0.0%				8.3%	5.6%	
	Perpetrators of Violent Attacks								0.0%			0.0%								0.0%						0.0%	0.0%
	Violent Speech	3.5%	0.0%	1.4%	1.2%	0.5%	3.5%		3.5%		1.2%	4.9%	6.1%	2.1%	0.7%	0.0%	0.6%	0.0%	0.7%	4.6%	0.0%	0.0%			5.1%	1.3%	
	Distribution of Hacked Materials								0.0%			50.0%														0.0%	
	Illegal or certain regulated goods and services				0.0%				0.0%			0.0%	2.4%	0.0%	0.0%				0.0%	0.0%	0.0%	0.0%				0.0%	0.0%
	Manual Review								0.0%			0.0%	0.0%							0.0%	0.0%					0.0%	
				0.0%				0.0%			0.0%	0.0%							0.0%	0.0%					0.0%		

Note: Cells that are blank mean that there was no enforcement. For cells containing '0.0%' value, there were no cases of successful appeals or overturns.



Metric	Enforcement	Policy	Bulgarian	Croatian	Czech	Danish	Dutch	English	Estonian	Finnish	French	German	Greek	Hungarian	Irish	Italian	Latvian	Polish	Portuguese	Romanian	Slovak	Slovenian	Spanish	Swedish								
Overturn Rate	Automated Means	Sensitive Media						0.0%																								
		Non-Consensual Nudity							100.0%					100.0%			0.0%			100.0%					0.0%							
		Abuse & Harassment																														
		Private Information & media							0.0%				100.0%	57.1%				0.0%														
		Perpetrators of Violent Attacks																														
		Violent Speech			0.0%		75.0%		33.3%		43.5%		32.0%		40.0%		31.8%		35.5%		66.7%		50.0%		30.0%	42.9%	29.6%	0.0%	0.0%		32.6%	33.3%
		Child Sexual Exploitation																														
		Other													100.0%																	
		Deceased Individuals									100.0%																					
		Suicide & Self Harm								25.0%		13.4%		27.3%		12.7%		0.0%		0.0%	0.0%	0.0%								0.0%	19.6%	
	Sensitive Media								0.0%		12.6%		6.7%		5.3%		0.0%		0.0%	0.0%									5.3%			
	Hateful Conduct																															
	Abuse & Harassment			0.0%		0.0%			50.0%		7.4%		0.0%		13.8%		8.5%			50.0%		20.0%		7.1%					19.7%			
	Non-Consensual Nudity								0.0%		23.6%		0.0%		7.7%		22.2%		0.0%			50.0%							50.0%	100.0%		
	Private Information & media			0.0%					0.0%		100.0%		12.7%							0.0%		13.6%							50.0%	14.7%	0.0%	
	Perpetrators of Violent Attacks																															
	Violent Speech			25.0%			14.3%		100.0%		55.6%		14.7%		66.7%		17.8%		16.2%		50.0%		0.0%		19.0%		24.1%		7.0%		18.0%	14.3%
Manual Review	Distribution of Hacked Materials																															
	Illegal or certain regulated goods and services											25.0%																				
	Intellectual property infringements																															
		Child Sexual Exploitation																														

Note: Cells that are blank mean that there was no enforcement. For cells containing 0.0% value, there were no cases of successful appeals or overturns.

INDICATORS OF ACCURACY FOR SUSPENSIONS

Metric	Enforcement	Policy	Bulgarian	Croatian	Czech	Danish	Dutch	English	Estonian	Finnish	French	German	Greek	Hungarian	Irish	Italian	Latvian	Lithuanian	Polish	Portuguese	Romanian	Slovak	Slovenian	Spanish	Swedish	
Appeal Rate	Automated Means	Abuse & Harassment						20.0%					0.0%													
	Automated Means	Ban Evasion			50.0%				15.0%			6.7%	17.4%				50.0%			0.0%	0.0%					0.0%
	Automated Means	Child Sexual Exploitation	39.7%	39.7%	45.2%		33.1%	47.6%	8.1%		27.3%	46.3%	62.5%	57.8%	54.7%		63.4%	22.9%		45.0%	40.6%	48.6%	58.2%		57.7%	36.0%



Manual Review	Sensitive Media								6.4%		0.0%	10.0%	42.9%					0.0%							0.0%	0.0%
Manual Review	Suicide & Self Harm			100.0%	100.0%	100.0%			47.9%			37.5%	79.6%					58.3%		40.0%		100.0%	0.0%		61.1%	50.0%
Manual Review	Username Squatting								0.0%									0.0%							0.0%	
Manual Review	Violent & Hateful Entities								133.3%	66.7%	26.1%		50.0%	16.7%	33.3%	50.0%	0.0%		72.2%		66.7%	100.0%	0.0%		50.0%	87.5%
Manual Review	Violent Speech	57.1%	29.4%	47.1%	63.6%	64.3%	43.2%				40.7%	48.6%	54.1%	54.5%	28.0%			54.8%	0.0%		42.0%	57.3%	60.0%	50.0%	55.6%	53.8%

Note: Cells that are blank mean that there was no enforcement. For cells containing '0.0%' value, there were no cases of successful appeals or overturns.

Metric	Enforcement	Policy	Bulgarian	Croatian	Czech	Danish	Dutch	English	Estonian	Finnish	French	German	Greek	Hungarian	Irish	Italian	Latvian	Lithuanian	Polish	Portuguese	Romanian	Slovak	Slovenian	Spanish	Swedish		
Automated Means	Abuse & Harassment								0.0%																		
Automated Means	Ban Evasion			100.0%					0.0%			0.0%	25.0%					0.0%									
Automated Means	Child Sexual Exploitation	3.2%	6.9%	2.9%	2.1%	17.2%	21.5%				14.3%	11.2%	7.0%	1.7%	3.7%			9.8%	0.0%	10.0%	37.1%	0.7%	154.4%		8.2%	0.0%	
Automated Means	CWC for various countries for illegal activity																										
Automated Means	Financial Scam							0.0%	0.0%			0.0%	0.0%												0.0%		
Automated Means	Help with my compromised account																										
Automated Means	Illegal or certain regulated goods and services														0.0%												
Automated Means	Misleading & Deceptive Identities								80.0%				100.0%												100.0%	50.0%	
Overturn Rate	Automated Means	Non-Consensual Nudity																									
Automated Means	Other							0.0%	3.4%				12.5%						0.0%	0.0%					0.0%		
Automated Means	Perpetrators of Violent Attacks								38.2%		0.0%	0.0%	50.0%					0.0%		100.0%					37.5%		
Automated Means	Platform Manipulation & Spam	12.8%	13.8%	19.4%	25.0%	15.0%	14.1%				31.9%	16.6%	13.5%	14.5%	20.7%		13.6%	0.0%		17.2%	13.5%	18.3%	3.1%		13.5%	26.7%	
Automated Means	Sensitive Media								50.0%																	0.0%	
Automated Means	Violent & Hateful Entities								50.0%				9.9%								0.0%					0.0%	42.9%
Automated Means	Violent Speech								0.0%																	0.0%	
Manual Review	Abuse & Harassment	14.3%	0.0%	0.0%	16.7%	6.5%	5.5%				0.0%	4.9%	2.6%	10.3%	11.1%	0.0%	5.0%			3.7%	6.8%	7.1%	0.0%		9.7%	11.8%	
Manual Review	Ban Evasion								2.3%				0.0%	0.0%							0.0%	0.0%				0.0%	

Manual Review	Child Sexual Exploitation						0.0%	4.5%		33.3%	0.0%	0.0%	0.0%	25.0%		0.0%		16.7%	0.0%	0.0%	0.0%	0.0%	100.0%			
Manual Review	CWC for various countries for illegal activity																						0.0%			
Manual Review	Deceased Individuals																						0.0%			
Manual Review	Distribution of Hacked Materials																						0.0%			
Manual Review	Financial Scam																						0.0%			
Manual Review	Harassment							66.7%		90.1%		100.0%	66.7%	110.0%	100.0%				28.6%		100.0%	75.0%		86.7%	100.0%	
Manual Review	Help with my compromised account																							0.0%		
Manual Review	Illegal or certain regulated goods and services							2.6%				0.0%	0.0%		0.0%				0.0%		0.0%			0.0%		
Manual Review	Intellectual property infringements						0.0%	5.9%				4.1%	0.0%	5.6%				2.8%		0.0%	8.0%			2.8%		
Manual Review	Misleading & Deceptive Identities							10.3%				0.0%									100.0%			0.0%		
Manual Review	Non-Consensual Nudity	0.0%	0.0%	0.0%	0.0%	0.0%	5.3%	7.1%		0.0%	2.5%	4.0%	0.0%	0.0%				0.0%		25.0%	0.0%	0.0%	0.0%	5.6%	0.0%	
Manual Review	Other						100.0%	7.2%				0.0%	0.0%	0.0%						0.0%	0.0%			0.0%		
Manual Review	Perpetrators of Violent Attacks						0.0%	0.0%		0.0%	0.0%	0.0%								0.0%	0.0%			0.0%		
Manual Review	Platform Manipulation & Spam	25.0%	0.0%	40.0%	0.0%	15.0%	17.1%			0.0%	21.4%	30.6%	0.0%	16.7%				30.8%		8.1%	0.0%	60.0%	0.0%	13.1%	50.0%	
Manual Review	Private Information & media						0.0%	54.5%				25.0%	40.0%					100.0%			0.0%				28.6%	
Manual Review	Sensitive Media							16.7%				100.0%	0.0%												0.0%	
Manual Review	Suicide & Self Harm				0.0%	0.0%	0.0%	14.9%				33.3%	9.1%						28.6%		25.0%		0.0%		27.3%	0.0%
Manual Review	Username Squatting																								0.0%	
Manual Review	Violent & Hateful Entities				25.0%	0.0%	6.3%				0.0%	0.0%	0.0%	0.0%						0.0%	0.0%			0.0%	28.6%	
Manual Review	Violent Speech	25.0%	20.0%	29.2%	21.4%	22.7%	30.7%				18.2%	29.7%	20.9%	50.0%	42.9%			16.7%		23.3%	29.4%	50.0%	20.0%	31.2%	34.3%	

Note: Cells that are blank mean that there was no enforcement. For cells containing '0%' value, there were no cases of successful appeals or overturns.

Ann. 34.2: Average Monthly Active Recipients - 1 October to 31 March			
Country Name	Logged In Users	Logged Out Users	Total
Austria	820,508	627,375	1,347,884
Belgium	1,395,023	871,258	2,266,281
Bulgaria	419,838	244,604	664,442
Cyprus	182,679	93,858	276,537
Czechia	163,143	927,332	1,090,475
Germany	10,490,789	5,107,639	15,598,427
Denmark	773,601	330,220	1,103,821
Estonia	186,397	97,151	283,549

Spain	9,623,028	5,432,943	15,055,971
Finland	1,504,663	698,343	2,203,007
France	11,643,813	5,777,497	17,421,310
Greece	868,983	767,844	1,636,827
Croatia	319,177	369,862	689,040
Hungary	701,480	471,939	1,173,419
Ireland	1,382,790	812,254	2,195,045
Italy	5,457,672	2,367,161	7,824,833
Lithuania	264,246	148,601	412,847
Luxembourg	110,635	61,402	172,038
Latvia	264,294	143,346	407,641
Malta	79,743	35,850	115,594
Netherlands	4,737,950	2,736,329	7,474,280
Poland	4,284,544	2,979,738	7,264,282
Portugal	1,550,643	704,858	2,255,502
Romania	1,086,831	538,532	1,625,363
Sweden	1,729,830	812,559	2,542,389
Slovenia	198,042	238,148	436,190
Slovakia	273,202	219,924	493,126

#### Further Information on Suspensions

During the applicable reporting period **1 October, 2024 to 31 March, 2025** there were zero actions taken for: provision of manifestly unfounded reports or complaints; or manifestly illegal content. While manifestly illegal content is not a category that we have taken action on during the reporting period, we suspended **132,155** accounts for violating our Child Sexual Exploitation policy and **4,626** for violating our Violent and Hateful Entity policy.

Disputes submitted to out-of-court dispute settlement bodies.

To date, zero disputes have been submitted to the out-of-court settlement bodies.

Reports received by trusted flaggers.

During the reporting period, we have received 271 reports from Article 22 DSA approved trusted flaggers. Once Article 22 DSA awarded trusted flaggers information is published, we immediately enrol them in our trusted flaggers program, which ensures prioritisation of human review, via their email, username, and account.

---