TheFork

DSA Transparency Report

Reporting period: 17/02/2024 – 31/01/2025 *(included)*

Published: 16 February 2025

Table of Contents

Report 1 - Member State orders

TheFork has not received any Member State order under the DSA.

Report 2 - Notices submitted in accordance with article 16 of Regulation (EU) 2022/2065

Category label	Category description	Category of illegal content / incompatible with the terms and conditions	Number of notices received	Number of specific items of information included in the total number of notices	Median time to take action (minutes)	Number of actions taken on the basis of the terms and conditions of the service
		TOTAL	8471	8471	14960	1535
Category 3	Cyber violence	STATEMENT_CATEGORY_CYBER_VIOLENCE	717	717	16969	104
Category 3c	Cyber incitement to hatred or violence	KEYWORD_CYBER_INCITEMENT	314	314	13791	36
Category 3g	Not captured by any other sub-category	KEYWORD_OTHER	403	403	20323	68

Category						
Category 5	Data protection and privacy violations	STATEMENT_CATEGORY _DATA_PROTECTION_ AND_PRIVACY_VIOLATI ONS	**127**	**127**	**19490**	**56**
Category 5e	Not captured by any other sub-category	KEYWORD_OTHER	127	127	19490	56
Category 6	Illegal or harmful speech	STATEMENT_CATEGORY _ILLEGAL_OR_HARMFU L_SPEECH	**293**	**293**	**14655**	**71**
Category 6b	Discrimination	KEYWORD_DISCRIMINA TION	184	184	15589	53
Category 6c	Illegal incitement to violence and hatred based on protected characteristics (hate speech)	KEYWORD_HATE_SPEE CH	109	109	12908	18
Category 7	Intellectual property infringements	STATEMENT_CATEGORY _INTELLECTUAL_ PROPERTY_INFRINGEM ENTS	**5**	**5**	**18619**	**3**
Category 7a	Copyright infringements	KEYWORD_COPYRIGHT_ INFRINGEMENT	5	5	18619	3
Category 11	Scams and/or fraud	STATEMENT_CATEGORY _SCAMS_AND_FRAUD	**3379**	**3379**	**16603**	**306**
Category 11d	Inauthentic user reviews	KEYWORD_INAUTHENTI C_USER_REVIEWS	3100	3100	16076	238
Category 11g	Not captured by any other sub-category	KEYWORD_OTHER	279	279	19637	68

Category 15	Other violation of provider's terms and conditions	STATEMENT_CATEGORY _OTHER_VIOLATION_TC	3950	3950	12759	995
Category 15a	Adult sexual material	KEYWORD_ADULT_SEXUAL_MATERIAL	16	16	15836	4
Category 15g	Not captured by any other sub-category	KEYWORD_OTHER	3934	3934	12726	991

Report 3 - Own-initiative content moderation on the basis of illegality

No measures have been taken on our own initiative on the basis of illegality.

Report 4 - Own-initiative content moderation on the basis of incompatibility with the Terms and Conditions

Category label	Category description	Category of illegal content / incompatible with the terms and conditions	Number of measures taken at the provider's own initiative	Number of measures taken after detection with solely automated means	Visibility restriction Removal	Visibility restriction Disable	Account restriction Termination
		TOTAL	521799	163785	510221	11578	7084
Category 3	Cyber violence	STATEMENT_CATEGORY_CYBER_VIOLENCE	545	0	0	545	0
Category 3a	Cyber bullying and intimidation	KEYWORD_CYBER_BULLYING_INTIMIDATION	290	0	0	290	0
Category 3c	Cyber incitement to hatred or violence	KEYWORD_CYBER_INCITEMENT	255	0	0	255	0
Category 6	Illegal or harmful speech	STATEMENT_CATEGORY_ILLEGAL_OR_HARMFUL_SPEECH	153	0	0	153	0
Category 6b	Discrimination	KEYWORD_DISCRIMINATION	153	0	0	153	0
Category 11	Scams and/or fraud	STATEMENT_CATEGORY_SCAMS_AND_FRAUD	16169	15612	16134	35	7084
Category 11d	Inauthentic user reviews	KEYWORD_INAUTHENTIC_USER_REVIEWS	16169	15612	16134	35	7084
Category 15	Other violation of provider's terms and conditions	STATEMENT_CATEGORY_OTHER_VIOLATION_TC	504932	148173	494087	10845	0
Category 15g	Not captured by any other sub-category	KEYWORD_OTHER	504932	148173	494087	10845	0

Report 5 - Number of complaints received through the internal complaints mechanism and out-of court dispute settlement bodies

Indicator	Total number	Contextual information
Number of complaints submitted to the internal-complaints mechanism	31	To determine this number, we analysed cases escalated from Customer Care to the relevant Moderation team, applying a filtering approach to identify the cases that would be complaints concerning content moderation restrictions.
Number of disputes submitted to out-of-court dispute settlement bodies	0	

Report 6 - Use of automated means for content moderation

Indicator	Scope	Value
Number of measures solely taken by automated means	Own-initiative	163785
Number of notices solely processed by automated means	NAM Total	0
Number of notices not processed by automated means	NAM Trusted Flagger	0
Total number		**163785**
Number of measures not taken by automated means	Own-initiative	358014
Number of notices not processed by automated means	NAM Total	8471
Number of notices not processed by automated means	NAM Trusted Flagger	0
Total number		**366485**

Report 7 - Qualitative Template

1. Summary of the content moderation engaged in at the providers' own initiative

Content moderation at our own initiative is based on the moderation of photos uploaded by our users, reviews and restaurants' replies to the reviews. The majority of content is disabled when it fails to comply with our Terms and Conditions or Content Moderation Guidelines. In such cases, we request users to reformulate their content to ensure compliance with our internal rules. Additionally, certain content, such as photos, may be removed if they do not meet our platform's quality standards or do not comply with our Content Moderation Guidelines. Content posted by users who have been suspended for fraudulent activities is also subject to removal. Finally, accounts that repeatedly publish fraudulent content in violation of our Terms and Conditions may face account termination.

2. Meaningful and comprehensible information regarding content moderation engaged in at the providers' own initiative

Diners' reviews and restaurant replies are moderated using a two-step process combining automated filtering and manual review. First, we automatically scan text-based content for prohibited words, phrases, or patterns that may indicate harmful, offensive, or inappropriate language. If flagged, the content is sent for further assessment. The moderation team then manually reviews flagged content to ensure accuracy, reducing false positives and allowing for contextual evaluation.

For photos, an automated system analyzes images to detect specific content not allowed, such as photos focused on alcohol. However, a large amount of photos uploaded by Diners are manually moderated based on quality criteria, and photos uploaded by restaurants require a full analysis of the restaurant's profile before making the decision to take down content.

3. Qualitative description of the automated means

Content detected by automated means is applied to reviews and replies, utilizing an automatic mechanism to flag reviews for further revision.

Content removed by automated means is only limited to photos that do not comply with the content requirements established in Article 9 of our Content Moderation Guidelines, with a confidence rate higher than 68%.

4. Qualitative description of indicators of accuracy and possible rate of error of automated means

The content removed is based on a confidence rate which needs to be higher than 68%.

5. Specification of the precise purposes to apply automated means

Automatic moderation of photos is implemented with the objective of efficiently managing large volumes of images in real time, ensuring rapid detection and filtering of inappropriate content, allowing for quicker response times and consistent enforcement of guidelines.

6. Safeguards applied to the use of automated means

Safeguards applied to the use of automated means include implementing a confidence rate of 68% for content removed by automated means, as well as giving users the possibility to appeal our content moderation decisions.