

Les fournisseurs de services intermédiaires mettent à la disposition du public, dans un format lisible par une machine et d'une manière facilement accessible, au moins une fois par an, des rapports clairs et facilement compréhensibles sur les éventuelles activités de modération des contenus auxquelles ils se sont livrés au cours de la période concernée.

L'objectif est d'assurer la transparence de nos pratiques de modération de contenus, protéger la liberté d'expression et lutter contre les contenus illicites.

Période couverte : [01/01/2025 – 31/12/2025]

Date de publication du rapport : [05/03/2026]

1. Données générales d'utilisation

- Nombre moyen de destinataires actifs mensuels dans l'UE : 12.5 millions
(Conformément à l'art. 24(2) du DSA)

2. Signalements reçus

- **Nombre total de signalements reçus :** [77,000]
- **Origine des signalements :**
 - Utilisateurs (via formulaire / bouton signaler) : [42,000]
 - Intelligence Artificielle : [35,000]
- **Motifs invoqués (typologie des contenus modérés):**

NB : la répartition des motifs n'étant disponible qu'à partir du 18/12/2025, la répartition suivante est présentée en pourcentage, sur une base de 2,600 commentaires.

- Discours malveillants ou illégaux : [29%]
- Violence : [23%]
- Cyberviolence : [16%]
- Effet négatif sur le débat public ou les élections : [14%]
- Risque pour la sécurité publique : [4%]
- Discours de haine : [3%]
- Autres (préciser) : [10%]

4. Mesures prises suite aux signalements :

- **Nombre de contenus retirés ou rendus inaccessibles :** [75,000]
- **Nombre de contenus maintenus après examen :** [2,000]
- **Délai moyen de traitement des signalements :** [30 jours]
- **Modalités de modération appliquées :**
 - Retrait du commentaire : [33,000]
 - Restriction (masquage, limitation de visibilité) : [42,000]
 - Avertissement de l'auteur : [0]
 - Suspension de compte : [0]

5. Mesures proactives

- Utilisation d'outils automatisés : [Oui]
- Description : L'ensemble des commentaires sont modérés par des outils automatisés en première analyse. L'outil d'automatisation, OpenAI, mesure un degré de toxicité du commentaire, ainsi qu'une probabilité à être un spam. La publication du commentaire dépendra du résultat de cette évaluation. Sont également masqués par défaut les commentaires incluant des images, des liens et/ou un numéro de téléphone.

Enfin, une liste de mots prohibés est également mise en place dans chaque langue, plaçant automatiquement tout commentaire en attente d'évaluation par un modérateur (action humaine).
- Proportion de retraits effectués automatiquement : [44 %]

6. Recours et contestations

- **Nombre total de recours déposés par les utilisateurs :** [1]
- **Issue des recours :**
 - Décision confirmée (contenu maintenu retiré) : [1]
 - Décision annulée (contenu rétabli) : [0]
 - En cours de traitement : [0]
- **Délai moyen de traitement des recours :** [34 jours]

7. Coopération avec les autorités et signaleurs de confiance

- Demandes officielles reçues d'autorités judiciaires ou administratives : [0]
- Demandes émanant de signaleurs de confiance : [0]

8. Ressources et politiques internes

- Effectifs dédiés à la modération : [1 ETP]
- Formation spécifique des modérateurs :
 - Les modérateurs sont formés à utiliser notre plateforme de modération. Une vidéo de démonstration et des liens vers des tutoriels sont mis à disposition. Nous leur expliquons les règles de modération automatiques.
 - 13 langues étant couvertes, différents modérateurs ont été désignés dans chaque équipe éditoriale.
- En cas de recours, nous faisons appel aux modérateurs pour savoir s'il faut ou non modérer le commentaire. La décision finale appartient aux équipes éditoriales.
- Engagements pour améliorer la transparence, la rapidité et l'équité de la modération : formation et sensibilisation continues.

9. Conclusion et engagements

- Analyse des tendances observées durant la période : augmentation du nombre de commentaires hostiles, menant à une plus forte proportion de commentaires rejetés.
- Au vu du nombre de signalements constatés, nous avons modifié la règle du nombre de signalements nécessaires avant de masquer les commentaires (de 2 à 3 signalements), cela afin de garantir la liberté d'expression.