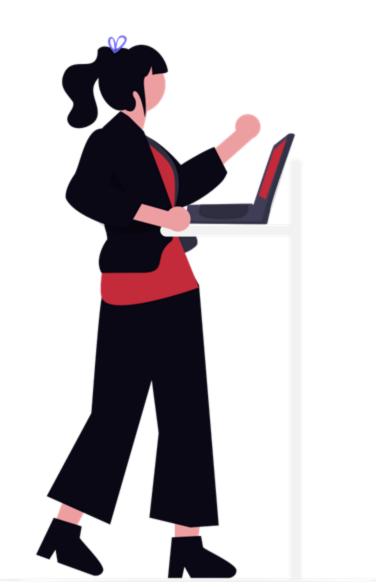


HACK THE DSA

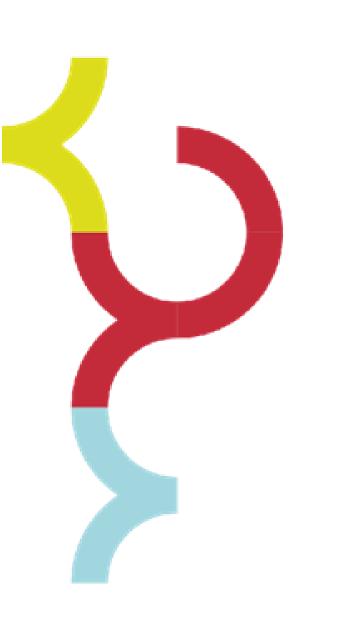
Les rapports sur les risques

systémiques









Problèmes de base

- > Comparaison des rapports
- > Manquements de ces rapports
- > Voies d'amélioration





RAPPORTS RISQUES SYSTEMIQUES



PROBLEMES DE FOND

PROBLEMES DE FORME

Analyse plateformes 🛍 ☆ 🕸 🛆 File Edit View Insert Format Data Tools Extensions Help ▼ fx I J K N 0 Réseaux/ Médias sociaux Ø Instagram 3 STRUCTURE Résumé Détail 5 Longueur 93 pages 93 pages 6 Calendrier date de publication : 01/08/2024 août 2024 Rapport qui semble extrêmement dense avec un sommaire de deux pages. Il y a un "executive summary" de quatre pages. La typo est petite et les interlignes ressérées, le texte ne respire pas, il y a quelques graphiques, des encadrés et beaucoup de tableaux. Il y a des notes en bas de page ainsi que de nombreux liens externes et internes. Les informations sont noyées dans une mer de verbiage ce qui rend assez difficile la recherche rapide d'informations Apparence / Design Rapport dense avec tableaux et schémas. précises. Il est possible de cliquer sur les entrées du sommaire pour aller à la page correspondante ce qui facilite la lecture, l'absence de respiration dans la mise en page ainsi que la petite taille de la police de caractère rend la lecture ardue, le texte est en anglais, sans traduction dans les autres langues utilisées par la Commission européenne. Il n'est pas possible de Sommaire cliquable, résumé, langue anglaise revenir facilement à la page sur laquelle on était uniquement, liens internes et externes, navigation lorsque l'on suit un lien interne pour aller, par exemple, Accessibilité peu aisée dans le document dans les pages d'annexe. 9 Méthode de détection 10 Signaleurs de Confiance Oui Meta responds to requests from law enforcement for preservation or disclosure of user data in Oui accordance with applicable laws and its Terms of Use. Inionctions 12 Utilisateurs Oui 13 Détection automatique 14 Confidentialité du rapport Pas de mention visible Pas de mention visible 15 17 FOND Le rapport reprend les risques systémiques listés par Reprise des catégorie du DSA et ajout de facteurs l'article 34 du DSA et y ajoute des risques de risques spécifiquement identifiés par Meta. 3 spécifiquement identifiés par Meta qu'ils appellent modalités : risk areas, influencing factors et "influencing factors" afin de définir les risques Définition "detailled risk observation" systémiques propres à Instagram Deceptive and misleading / Civic discourse and election / Public Health / Public security / Gender-based violence / Protection opf minors / Fundamental rights / Illegal content. 8 "risk areas" "Risk areas" (reprise catégories DSA) : reprises de l'article 34 et 19 risques identifiés spécifiquement et "influencing factors" identifiés. Les Deceptive and misleading / Civic discourse and election / Public Health / Public security / méthodes de détection ne sont pas clairement Gender-based violence / Protection of minors / expliquées, le rapport fait état d'un "risk assesment Fundamental rights / Illegal content. "Influencing process" en six étapes dont une étape "Identify and factors": Recommender systems / Content qualify" qui consiste en "Meta leveraged a diverse set moderation systems / Terms of use and their of signals and inputs to scope the risk assessment. enforcement / Ads systems / Data related These inputs were used to define the in-scope practices / Intentional manipulation / Generative Problem Areas and the associated risks that artificial intelligence. META indique dans son collectively create a systemic risk to users in the EU." rapport que certains de ces risques peuvent se Ce qui est une manière élégante de ne rien dire sur les croiser ou se superposer. Peu d'informations sur methodes de détection tout en utilisant beaucoup de Identification le protocole d'identification des risques. Le rapport reconnait qu'il y a des pistes d'amélioration en ce qui concerne l'atténuation des risques externes et l'atténuation des "facteurs d'influence" et donc aussi des risques internes. Le rapport reconnait qu'il y a des pistes d'amélioration Un très grand nombre de mesures d'atténuation en ce qui concerne l'atténuation des risques externes et l'atténuation des "facteurs d'influence" et donc aussi sont listées en regard de chaque risque Atténuatior systémique détaillé des risques internes. Meta evaluated the inherent risk of 122 risks that cut across the 19 Problem Areas and determined the effectiveness of controls for Instagram to determine the residual risk for each in-scope risk. These controls were evaluated individually to determine how effectively they mitigated each risk, as applicable. 72 All residual Protocole détaillé de méthode d'évaluation des risk scores were ranked and rated using an ordinal tiering scale from Tier 1 to 5 that allows us to measure Evaluation risques risques avec des degrés de sévérité identifiés Problem Areas consistently. 22 Risques externes Certains risques internes sont listés, cependant ils 23 ne sont pas qualifiés de risques mais de "facteurs" Les risques internes ne sont pas qualifiés de risques mais de facteurs d'influence Risques internes 24 Données Les données ne sont pas détaillées en chiffres, mais en images : par exemple un schéma montre des points gris et des points roses pour montrer la quantité de contenus liés à des risques systémiques sur la + = STRUCTURE - FOND - Facebook - Instagram - TikTok - SnapChat - X - LinkedIn - YouTube - Pinterest - Zalando - Amazon - AliExpress - Pornhub - Xvideos - Stripchat - Bing - Google - Wikipédia - Apple -

Constats et observations - Forme

NEGATIF

- 1. Dense et peu digeste
- 2. Absence de standardisation dans la méthode d'analyse et la présentation
- 3. Caviardage

POSITIF

- 1. Amélioration de la mise en page
- 2. Présence de matrices
- 3. Mention des articles DSA









IX. Annex: Matrices

1. Probability matrix

	1	2	3	4	5
Probability	Very unlikely	Unlikely	Possible	Likely	Almost certain
Frequency of incident or event occurring	May occur within a year Rare but could occur	May occur within 6 months Has occurred for comparable platforms	May occur within a month Has occurred for X and / or commonly occurs for comparable platforms	Likely to occur within the next 2 weeks Has occurred for X regularly	Immediately or within days Occurs for X every day

Fig.13: Probability scale for the purpose of the DSA risk assessment

2. Severity matrix

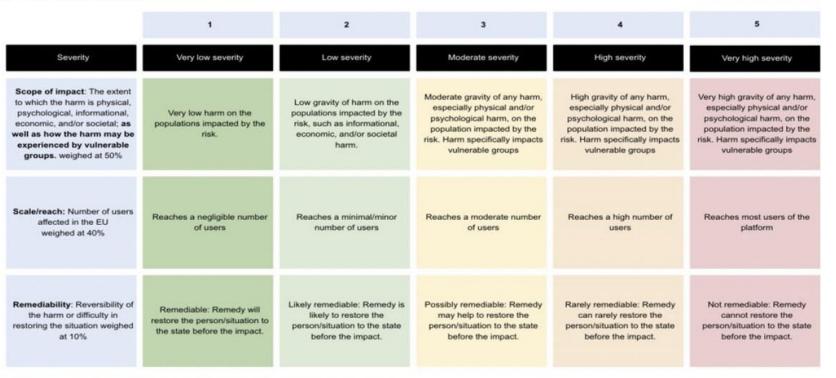


Fig.14: Severity scale for the purpose of the DSA risk assessment

3. Inherent risk matrix

	Se	verity		
Very low severity	Low severity	Moderate severity	High severity	Very high severity
Low	Medium	High	Critical	Critical
Negligible	Low	Medium	High	Critical
Negligible	Low	Low	Medium	High
Negligible	Negligible	Low	Low	High
Negligible	Negligible	Negligible	Negligible	High
	Low Negligible Negligible Negligible	Very low severity Low Medium Negligible Low Negligible Low Negligible Negligible	Low Medium High Negligible Low Medium Negligible Low Low Negligible Negligible Low	Very low severity Low Medium High Critical Negligible Low Medium High Negligible Low Low Medium High Negligible Low Low Medium Medium Low Low Medium Medium

Fig.15: Residual risk matrix for the purpose of the DSA risk assessment

Constats et observations - Fond

1. Autopromotion / Plaidoyer

- Absence d'autocritique sur les risques internes au service
- Valorisation des mécanismes d'atténuation sans tierce vérification possible (données non disponibles)

2. Données

- Absence
- Superficialité
- Subjectivité





Constats et observations - Fond

3. Présentation trompeuse

Présentation manipulatoire (saturation d'information, cherry picking





Risk Category	Risk Scenario	Controls & Measures in place	Additional (mitigation) Measur to implement
Protection of Minors	Inadequate access verification systems allowing minors access to adult content	- warnings about adult content - age confirmation - guidance for guardians for protection of minors - security measures and parental controls setup - Terms of Service - Point 2 - Access (https://info.xvideos.net/legal/tos/)	 consider liaising with UNICEF and other globally operating NGOs with similar impacts to fund programs focused on raising awareness of onsafety and security of minors consider expanding and improving wording of the disclaimer about the inappropriateness of content for minor on entry and page closely monitor the latest developments in age verification tool with a view to finding an instrument the strikes a sufficient balance between effective protection of minors and the fundamental right to do business support the development and raising awareness of effective age verification tools, such as those integrated into the operating system of individual devices consider liaising with experts and NGOs dealing with the protection of minors and age verification tools

There are more than 100,000 app submissions in an average week. In 2022, App Review reviewed 6,101,913 submissions (including app updates), of which over 25 % were rejected by the App Review team for various compliance issues.²⁰ In 2023, App Review reviewed 6,892,500 submissions (including app updates); again, over 25 % were rejected.²¹ App

While TikTok does not believe that use of platforms (in general or of TikTok in particular), is automatically harmful in and of itself, it recognises that some users (in particular Younger Users) could at times and in specific circumstances face serious negative mental and physical health impacts as a result of the way in which they use such platforms.

Given the volume of content on the Pinterest platform, detecting and taking appropriate action on harmful content cannot completely eliminate these four risk areas. There are inherent risks on any platform that deals with vast quantities of content and data. Pinterest takes a risk-based approach when it comes to content moderation, prioritising harms that pose the greatest potential risk and balancing fundamental rights with keeping users safe.

Recommandations

- 1. Standardisation des méthodes
- 2. Ajout de données sourcées et vérifiables
- 3. Contrôle de la Commission Européenne (lignes rouges)
- 4. Partage de bonnes pratiques inter-acteurs (société civile services numériques régulateurs)



