

Déclaration annuelle à l'Autorité de régulation de la communication audiovisuelle et numérique en application du titre III de la loi du 22 décembre 2018 relative à la lutte contre la manipulation de l'information

Le 1 juillet 2022,

Propos introductif

Snapchat est une plateforme de **communication visuelle et de messagerie interpersonnelle** créée en 2011 qui offre à ses utilisateurs, appelés Snapchatters, la possibilité de communiquer plus facilement, de vivre l'instant présent et de découvrir le monde tout en s'amusant. L'application est intentionnellement différente des médias sociaux traditionnels, son modèle étant fondé sur les **principes de confidentialité et de sécurité par défaut** ("*privacy and safety by design*").

Nous avons **construit Snapchat autour de l'appareil photo** parce que nous voulions donner à notre communauté un moyen de partager la totalité de leurs expériences, avec leurs vrais amis. Notre objectif est de démontrer que les réseaux sociaux ne reposent pas sur une compétition entre utilisateurs. Ainsi, **aucun critère de vanité n'est proposé, ce qui limite tout risque de viralité** : pas d'option j'aime, de partage, ou de possibilité live. Nous concevons des produits et des technologies qui permettent de naviguer dans un environnement sain, sûr et amusant.

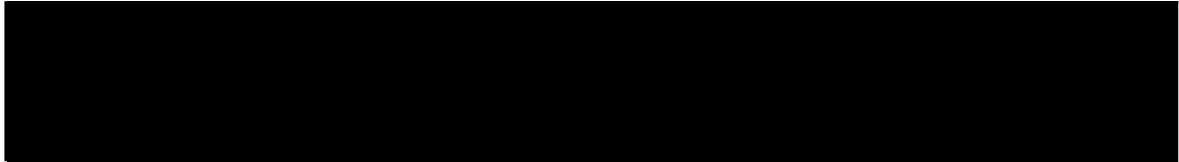
Notre **engagement à mettre l'humain au cœur de nos technologies** a influencé toutes les décisions que nous avons prises concernant nos produits. Nous adoptons une approche de l'innovation centrée sur l'humain, ce qui signifie que nous partons d'un **ensemble de valeurs universelles et que nous prenons délibérément en considération la sécurité, la vie privée et le bien-être des Snapchatters avant les données**.

Ces décisions de conception nous ont aidées à protéger notre communauté contre la désinformation et la toxicité. Nous ne proposons pas de fil d'actualité ouvert, mais une **plateforme de contenu fermée qui ne propose que des informations et des divertissements provenant d'éditeurs et de créateurs de confiance**. Nous ne donnons à personne la possibilité de partager du contenu non vérifié avec un large public sur Snapchat. Nous limitons également la taille des discussions de groupe et nous ne permettons pas qu'elles soient consultables ou recommandées. Cette vision, que nous portons au quotidien, résulte dans un nombre **très faibles de cas confirmés de violation de nos politiques contre les fausses informations en 2021 : 971, à rapporter aux 18 millions d'utilisateurs actifs quotidiens en France**.

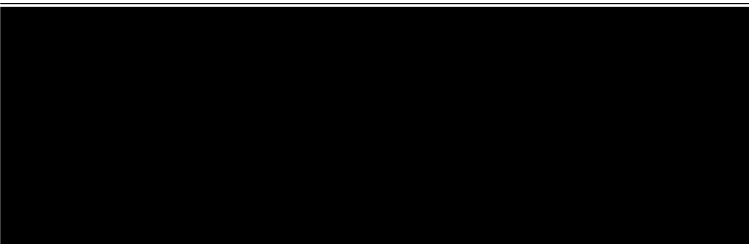
Pour renforcer ces efforts, Snapchat a mené un certain nombre d'améliorations et d'approfondissements de son outil de modération. En effet, depuis 2018, de nombreuses améliorations ont été apportées sur notre plateforme grâce aux recommandations de l'Arcom, comme **l'insertion d'une catégorie dédiée aux fausses informations dans notre menu de signalement en 2020**. Plus récemment, nous avons mis en ligne une **nouvelle version de notre système de signalement**. Il intègre désormais, pour l'ensemble des composantes de la plateforme, une **catégorie « fausses informations » plus détaillée (4 sous-catégories)**. Les conditions d'utilisation ont été mises à jour afin de souligner l'importance du signalement. Nous avons également mis en place, en janvier 2022, notre mécanisme de recours (*feedback loop*).

Enfin, Snapchat collabore avec un certain nombre d'institutions françaises et internationales pour approfondir son travail de lutte contre les fausses informations. Nous nous sommes par exemple associés à l'Organisation mondiale de la santé (OMS) pour faciliter les questions-réponses entre les utilisateurs et les experts sur le coronavirus, ou encore nous avons collaboré avec le Gouvernement dans le cadre des élections présidentielles pour offrir une information claire sur les conditions du vote.

Dans un contexte géopolitique et sanitaire incertain, nous continuons à être particulièrement vigilants quant à notre responsabilité de fournir à notre communauté des informations fondées sur des faits avérés. Notre priorité est de nous assurer que Snapchat fournit des informations de la plus haute qualité.



Questions liminaires : présentation du service

Nom du service:	Snapchat
Nom, raison sociale et domiciliation de la société opératrice :	Snap Inc., domicilié au 3000 31St, Santa Monica, CA 90405, États-Unis.
Identité et fonction du représentant légal exerçant les fonctions d'interlocuteur référent sur le territoire français au sens de l'article 13 de la loi du 22 décembre 2018 :	
Nombre de visiteurs uniques par mois du service en France et dans le monde, calculé sur la base de l'année précédant l'année d'exercice (moyenne mensuelle en <u>2020</u>) :	Pour l'année 2020, la moyenne mensuelle de visiteurs uniques par mois en France était de 21 881 000.
Modalités de mise à disposition du service auprès du public français (web, applications, assistants vocaux, etc.) :	Snapchat est une application disponible principalement sur téléphone portable. Quelques fonctionnalités sont disponibles <i>via</i> le web, notamment les services <u>Map</u> , <u>Spotlight</u> et les <u>contenus</u> de nos partenaires médias.
Langue du service proposée par défaut sur le territoire français :	Français
Le service disponible sur le territoire français constitue-t-il une version spécifique du service?	Non
Description du service (objet, population(s) ciblée(s), types de contenus publiés et consommés, thématiques, etc.) :	<p>Snapchat est une plateforme de communication visuelle et de messagerie interpersonnelle qui permet aux utilisateurs d'échanger avec leur entourage de manière simple et créative.</p> <p>Chez Snapchat, rien n'est plus important que la sécurité de nos utilisateurs, appelés <i>Snapchatters</i>. Nous concevons des produits et développons des technologies qui favorisent et soutiennent des amitiés authentiques dans un environnement sain, sûr et divertissant. Cet objectif guide nos décisions pour structurer et exploiter notre plateforme, développer de nouveaux produits, anticiper nos développements futurs et, en définitive, penser notre rôle dans la société. Dès le départ, nous avons conçu Snapchat de manière à limiter la viralité sur l'ensemble de notre plateforme et éviter toute incitation de diffusion massive de contenus préjudiciables. Nous utilisons des processus de développement de produits qui tiennent</p>

compte des exigences en matière de confidentialité, de sécurité et d'éthique, et ce sur l'ensemble de notre plateforme et de nos fonctionnalités. Snapchat se présente ainsi en cinq fonctionnalités :

- **Appareil photo** : L'application s'ouvre directement sur une caméra, et non sur un flux de contenu où n'importe quelle personne peut diffuser du contenu non approuvé au monde entier. Cet outil encourage la créativité des *Snapchatters* avec la mise à disposition de *lens* (filtre de réalité augmentée) et de filtres.
- **Messagerie privée** : Nous donnons la priorité aux communications privées entre amis et, par défaut, demandons aux *Snapchatters* de s'accepter mutuellement comme contacts avant de pouvoir commencer à communiquer. Les groupes sont limités en taille et ne sont pas accessibles et visibles pour les personnes n'en étant pas membre.
- **Discover** : Plateforme fermée sur laquelle nos partenaires médias de confiance publient leurs contenus. Ces partenaires (une quarantaine en France) doivent se conformer à nos règles communautaires et nos règles de contenus médias qui interdisent explicitement la manipulation de l'information, ou bien encore l'usurpation d'identité. Discover propose également du contenu provenant de créateurs certifiés appelés "*Snap stars*", et des utilisateurs populaires.
- **Spotlight** : Lancé fin 2020, Spotlight est un espace axé sur le contenu populaire généré par nos utilisateurs, essentiellement des scènes courtes de vie quotidienne. Il ne comporte pas de commentaires publics et les profils sont privés par défaut. Tous les contenus sont modérés par une analyse humaine avant d'être largement diffusés.
- **Map** : Carte présentant des *stories* publics en temps réel d'utilisateurs du monde entier, qui permet également de partager de manière optionnelle sa localisation avec des amis.

Modèle économique du service (type d'activités rémunératrices pour l'opérateur tirées de l'exploitation du service en France) :

Revenu publicitaire.

Remplir le tableau suivant :

Données chiffrées (France, pour l'année 2021)	
Chiffre d'affaires du service en France (€)	
Produits d'exploitation (€)	
Charges d'exploitation (€)	
Résultat net (€)	
Effectifs employés aux activités du service en France	

Questions générales

Retour sur la mise en œuvre de la loi du 22 décembre 2018 depuis son entrée en vigueur

L'Arcom invite l'opérateur à présenter les enseignements qu'il tire de la mise en œuvre de son devoir de coopération au titre du titre III de la loi du 22 décembre 2018 depuis qu'il y est soumis, notamment en période électorale. Il pourra préciser quelles ont été les éventuelles difficultés rencontrées dans son application.

Chez Snap, nous considérons qu'un contenu relève d'une fausse information dès qu'il s'agit d'une allégation ou d'une imputation inexacte ou trompeuse qui pourrait altérer la vérité.

A cet effet, nous interdisons la diffusion de fausses informations qui seraient malveillantes ou qui causeraient des dommages, comme la négation de l'existence d'événements tragiques (par exemple, l'attaque terroriste du Bataclan ou l'Holocauste), les allégations médicales non fondées (par exemple, "la covid-19 est un canular"), ou l'atteinte à l'intégrité des processus civiques (par exemple, la diffusion d'une fausse date d'élection).

Pour déterminer si un contenu français peut être qualifié de fausse information, nous nous alignons sur les critères définis par la loi *relative à la lutte contre la manipulation de l'information* de 2018 et les principes de la loi de 1881 *sur la liberté de la presse* pour éviter tout retrait excessif et protéger les droits de nos utilisateurs. Dans ce cadre, nous prenons en compte plusieurs critères précisés par la loi – notamment le risque que son maintien représente pour la paix publique et pour les échanges entre utilisateurs, l'aspect évident de son caractère erroné, etc. - et évaluons non seulement la véracité du contenu, mais aussi sa nocivité potentielle, ce qui contribue à garantir une réponse proportionnelle en matière d'application de la loi.

Nos modérateurs reçoivent une formation spécifique et évaluent, au cas par cas, les contenus signalés, en tenant compte du contexte national et géopolitique. Nous fournissons à notre équipe des directives et une liste d'exemples, tels que les cas cités ci-dessus, qui leur permettent d'identifier les fausses informations. Plus de détails sont fournis dans nos [règles communautaires](https://www.snap.com/fr-FR/community-guidelines), disponibles à l'adresse suivante : <https://www.snap.com/fr-FR/community-guidelines>.

Comme indiqué lors de nos précédentes contributions, la conception de notre plateforme, fondée sur les principes de sécurité et confidentialité par défaut, permet d'assurer un rempart contre les risques de diffusion massive de contenus haineux ou de manipulation de l'information.

Nous savons que les périodes électorales peuvent présenter d'importants risques. A ce titre, en période d'élections, nous mettons en place des réunions régulières en interne afin de suivre et évaluer la situation en France, veiller sur les temps forts (débat télévisés, meetings, etc.) ou les phénomènes perçus sur le terrain pouvant avoir un impact potentiel sur notre plateforme. Nos équipes sur le terrain assurent également un dialogue continu avec les autorités locales afin de prévenir tout phénomène pouvant mettre en péril les élections et altérer le vote.

Depuis 2018, de nombreuses améliorations ont été apportées sur notre plateforme grâce aux recommandations de l'Arcom, comme l'insertion d'une catégorie dédiée aux fausses informations dans notre menu de signalement en 2020, ou la mise à jour de nos règles communautaires pour cibler les pratiques de *deep fakes* par exemple.

Nous avons particulièrement apprécié l'accompagnement de l'Arcom lors de la dernière élection présidentielle française avec la mise en place d'entretiens réguliers. Malgré le fait que nous ayons une faible quantité de contenus politiques sur notre plateforme, et peu de personnalités

politiques (lors de la dernière élection présidentielle - seuls deux candidats étaient sur Snapchat), Snapchat est soucieux de protéger le processus démocratique et souhaite prendre part aux efforts collectifs.

Lutte contre la manipulation de l'information sur le service

Définition des fausses informations (ou de toute notion voisine adoptée) retenue par l'opérateur :

Sur l'ensemble de la plateforme Snapchat, la définition retenue concernant les fausses informations est la suivante : **“Usurpation d'identité, pratiques trompeuses et fausses informations :**

- *Il est interdit de se faire passer pour quelqu'un (ou quelque chose) que vous n'êtes pas, ou de tenter de tromper les utilisateurs sur votre identité. Il s'agit notamment de l'usurpation de l'identité de vos amis, des célébrités, des marques ou d'autres organisations.*
- *Nous interdisons la diffusion de fausses informations qui causent préjudice à autrui ou qui sont malveillantes, telles que la négation de l'existence d'événements tragiques, les déclarations non fondées en matière de santé, l'atteinte à l'intégrité des processus civils et la manipulation de contenu à des fins fausses ou trompeuses.*
- *Nous n'autorisons pas les spams et les pratiques trompeuses, notamment l'imitation de Snapchat ou de Snap Inc.”*

Cette définition diffère-t-elle selon les zones géographiques de mise à disposition du service ? :

NON

Interdiction des fausses informations (ou notion voisine) sur le service ? :

OUI

Le cas échéant, où l'utilisateur peut-il trouver l'information relative à cette interdiction ?

L'utilisateur peut trouver directement cette information sur notre page dédiée aux règles communautaires, qui est également disponible depuis l'application.

Remplir le tableau suivant :

Données chiffrées et indicateurs (en France, en 2021)

Nombre de contenus identifiés comme comportant une fausse information (ou notion voisine adoptée)	971 contenus et comptes sur nos parties publiques
Parmi ces contenus identifiés, nombre détectés par des outils automatisés	0%
Parmi ces contenus identifiés, nombre détectés à la suite de signalements utilisateurs	100%

Nombre de contenus comportant une fausse information (ou notion voisine adoptée) qui ont fait l'objet de mesures (recontextualisation, apposition d'un libellé, déréférencement, suppression, etc.), en déclinant par types de mesures

971. Nous tenons toujours compte du contexte lorsque nous appliquons la loi sur les fausses informations.

Nombre d'émetteurs de ces contenus identifiés et part qui a fait l'objet d'une mesure (avertissement, restrictions etc.)

971 contenus et comptes sur nos parties publiques

Audience moyenne (impressions ou nombre de vues) des contenus identifiés comme comportant une fausse information (ou notion voisine adoptée)

Notre taux total de visionnage de contenus non-conformes pour 2021 était de ~0,09%. Nous ne distinguons pas ce taux d'affichage entre les catégories afin de fournir une image globale des violations sur la plateforme.

Parmi les contenus détectés comme contenant une fausse information (ou notion voisine adoptée), fournir la liste des 20 ayant suscité le plus d'interactions¹ de la part des utilisateurs en France en 2021 sur le service :

- 20 premiers contenus relevant d'une communication commerciale (publicitaires, sponsorisés et en partenariat) :
- 20 premiers contenus « organiques » (hors communication commerciale) :

La conception de Snapchat repose sur les principes de "vie privée et confidentialité par défaut" auxquels s'accompagne une modération accrue sur les parties publiques.

A cet effet, nous n'avons pas de fil d'actualité ouvert où n'importe quel utilisateur peut diffuser un contenu au monde entier et le rendre viral. Notre plateforme de contenu, Discover, ne propose que du contenu provenant d'éditeurs de médias et de créateurs de contenu approuvés et Spotlight, notre plateforme de divertissement pour le contenu généré par les utilisateurs est modérée *a priori* avant que le contenu puisse atteindre un large public.

Par conséquent, notre plateforme ne générant aucune viralité, nous ne pouvons pas répondre à cette question.

Respect de la liberté d'expression

En se fondant sur des cas de fausses informations regardées comme telles par ses équipes en 2021, l'opérateur est invité à exposer **les éventuelles difficultés d'articulation entre sa politique de modération et le respect de la liberté d'expression et de communication sur son ou ses services.**

Les principes de liberté d'expression et d'application responsable des politiques de contenu peuvent dans certains cas être en conflit les uns avec les autres. Par exemple, alors que la pandémie de coronavirus continuait à faire rage tout au long de l'année 2021, il était nécessaire que notre politique de modération trouve un équilibre entre l'importance de permettre aux utilisateurs de débattre et de remettre en question les réponses en matière de santé publique, et les risques de contenus susceptibles de porter atteinte à la santé et sécurité de nos utilisateurs et de notre société. Notre équipe gère ces situations de manière responsable grâce à une approche

¹ Il est entendu par « interaction » toute action effectuée par les utilisateurs vis-à-vis de ces contenus : consultation, commentaire ou réponse, mise en favori, « like », partage etc.

qui évalue non seulement la véracité du contenu, mais aussi le risque de préjudice qu'il présente, ce qui permet de garantir que toute suppression est nécessaire et proportionnelle au préjudice.

Existence d'un traitement différencié pour certains comptes en matière de modération (ex. : personnalités publiques) :

Il est important pour nous que nos politiques soient appliquées de manière cohérente et équitable, indépendamment de la stature de l'utilisateur ou des implications en termes de revenus. Veuillez noter que les différents produits de notre application peuvent avoir des politiques et des processus de modération différents.

Le cas échéant :

- o critères d'identification de ces comptes :

N/A, comme décrit ci-dessus, nous appliquons nos politiques de manière cohérente pour tous les comptes d'un même produit.

- o expliquer en quoi le traitement diffère :

N/A, voir réponse au-dessus.

Préciser si les efforts de lutte contre la manipulation de l'information dans le contexte de la crise sanitaire ont eu un impact sur l'exercice de leur liberté d'expression par les utilisateurs, en explicitant les raisons le cas échéant (modération automatisée, multiplication de fausses informations sur son service, montée des attaques coordonnées, etc.)

En matière de lutte contre la manipulation de l'information, nous nous appuyons essentiellement sur la modération humaine chez Snapchat. Nous pensons que ce type de contenu doit être examiné au cas par cas en prenant en considération le contexte afin de préserver de manière appropriée la liberté d'expression, et de rendre compte de manière responsable des préjudices potentiels pour les utilisateurs. Nous évaluons constamment nos efforts de modération pour nous assurer qu'ils représentent une approche efficace et responsable pour atténuer les risques de manipulation de l'information sur notre plateforme.

Il convient néanmoins d'indiquer que la crise sanitaire de la covid-19 a eu un impact sur notre plateforme en matière de modération et ce, non pas en raison des fausses informations à ce sujet, mais en raison de la lutte contre la fraude aux pass sanitaires.



Transparence des mesures

- Moyens pour fournir aux **utilisateurs des explications claires et accessibles** sur les mesures mises en œuvre face aux risques de manipulation de l'information :
- Si ces informations sont **personnalisées et/ou contextuelles, fournir des illustrations.**
- Actions pour **davantage de transparence** envers l'ensemble des publics (utilisateurs, médias, régulateur, citoyens, etc.) notamment hors service (rapports de transparence, études internes, communiqués, billets de blog, etc.) :

Bien que nos règles communautaires² continuent d'évoluer, notre principe clé est de nous assurer que celles-ci soient claires et lisibles pour l'ensemble de notre communauté d'utilisateurs. Nous avons apporté plusieurs changements majeurs sur la manipulation de contenu à des fins trompeuses qui inclut désormais explicitement l'usage de *deep fakes*, et inséré plusieurs illustrations de types de fausses informations, comme la négation de l'existence d'événements tragiques, les allégations médicales non fondées ou bien encore l'atteinte à l'intégrité des processus civiques.

En parallèle, Snap prend un certain nombre de mesures afin de mieux accompagner ses utilisateurs, et les aider à mieux comprendre les mécanismes que nous utilisons pour lutter contre les fausses informations :

- Publication chaque semestre d'un rapport de transparence³, affichant le nombre de contenus et comptes supprimés pour cause de désinformation ;

Lutte contre la propagation de fausses informations

Nous avons toujours pensé que lorsqu'il s'agit de contenu préjudiciable, il ne suffit pas de penser aux politiques et à leur mise en œuvre, mais que les plateformes doivent tenir compte de leur architecture fondamentale et de la conception de leurs produits. Dès le début, Snapchat a été construit différemment des plateformes de médias sociaux traditionnelles, sans un fil d'actualité ouvert où n'importe qui peut diffuser à une large audience sans modération.

Nos règles interdisent clairement la propagation de fausses informations susceptibles de causer des dommages, notamment les fausses informations visant à porter atteinte aux processus civiques, les allégations médicales non fondées et la négation d'événements tragiques. Nos Règles et leur mise en œuvre s'appliquent systématiquement à tous les Snapchatters. Nous n'aménageons aucune exception spéciale en faveur des politiciens ou autres personnalités publiques.

Au cours de cette période, Snapchat a pris des mesures à l'encontre d'un total combiné de 14 613 comptes et contenus pour violation de nos règles relatives aux fausses informations.

Nombre total de contenus et de comptes sanctionnés	14 613
--	--------

Illustration de notre rapport de transparence

² <https://snap.com/fr-FR/community-guidelines>

³ <https://snap.com/fr-FR/privacy/transparency>

- Mise à disposition d'une bibliothèque en ligne dédiées aux publicités politiques, permettant aux utilisateurs d'identifier l'origine de ces messages ;

Bibliothèque des publicités politiques de Snap



La confiance. C'est ce qui rend l'expérience "authentique".

Chez Snap, nous nous efforçons de créer une plateforme de confiance, sûre et transparente pour nos utilisateurs. Notre bibliothèque des publicités politiques et promotionnelles est l'un des nombreux efforts que nous avons fournis pour y arriver. Cela donne au public la possibilité de trouver des informations sur toutes les publicités politiques et idéologiques diffusées sur notre plateforme.

Illustration de notre Bibliothèque des publicités politiques

- Mise en place d'un site d'actualité fournissant des informations sur les nouveautés, fonctionnalités, produits ou prises de positions de Snapchat :

We Support Ukraine

Par **Team Snap**
le mardi 01 mars 2022 à 21:30



Illustration de notre site d'actualité

1. Dispositif de signalement des fausses informations

Note Snap : L'ensemble des éléments énoncés ci-dessous porte sur notre version du dispositif de signalement de fausses informations mis en place en 2021.

1.1 Existence **d'un dispositif de signalement** permettant de signaler de fausses informations susceptibles de troubler l'ordre public ou d'altérer la sincérité du scrutin :

Oui.

1.2 **Libellé du ou des motifs** adéquats : **Raisons** qui sous-tendent le choix de ce ou ces libellés

“information fausse, trompeuse ou malveillante”

Tout d'abord un objectif de clarté, en reprenant explicitement en premier le terme d'information fausse, pour coller au mieux avec la définition la plus communément donnée. Par ailleurs, la pratique nous a montré qu'il était parfois difficile pour un utilisateur de caractériser ce qui relève réellement et concrètement d'une fausse information. À cet effet, afin d'encourager au signalement, nous avons élargi cette catégorie aux informations trompeuses et/ou malveillantes de manière à englober explicitement toutes les hypothèses et exemples de manipulation de l'information.

Description du dispositif de signalement


1.3 Moyens permettant d'assurer la **visibilité, l'accessibilité et l'ergonomie** de ce dispositif :

- présentation visuelle :
- **affichage** (visibilité par défaut ou non) :
- **proximité** du contenu :
- nombre d'actions nécessaires pour accéder au dispositif de signalement :
- nombre d'actions et d'hyperliens nécessaires à l'envoi du signalement d'une fausse information depuis la page du contenu :
- accessibilité aux personnes **en situation de handicap** :
- accessibilité à un **internaute non authentifié** sur la plateforme (et les éventuels différences avec le dispositif de signalement de ceux qui le sont) :
Fournir des éléments visuels pour illustrer la réponse.

Chez Snapchat, nous accordons une grande importance au signalement. A cet effet, nous avons mis en place des dispositifs de signalement de contenus disponibles directement sur l'application, notre site d'assistance (<https://support.snapchat.com/fr-FR/i-need-help>) et notre centre de sécurité (<https://www.snap.com/fr-FR/safety/safety-center>) afin de permettre aux personnes inscrites et non-inscrites/connectées sur Snapchat d'effectuer un signalement. Ces outils de signalement sont facilement accessibles (visibles) et conviviaux (pas de jargon juridique).

Processus de signalement intégré à l'application Snapchat

Les utilisateurs de Snapchat peuvent signaler un contenu en suivant ces étapes simples :

1. Faites un appui long sur le contenu, ou appuyez sur l'icône à trois points située dans le coin supérieur droit du contenu.
2. Appuyez sur l'icône de signalement « drapeau  » ou l'icône « i » ou à trois points pour les *Lens*, les publicités et le contenu sponsorisé.
3. Envoyez un signalement en suivant le processus de signalement.

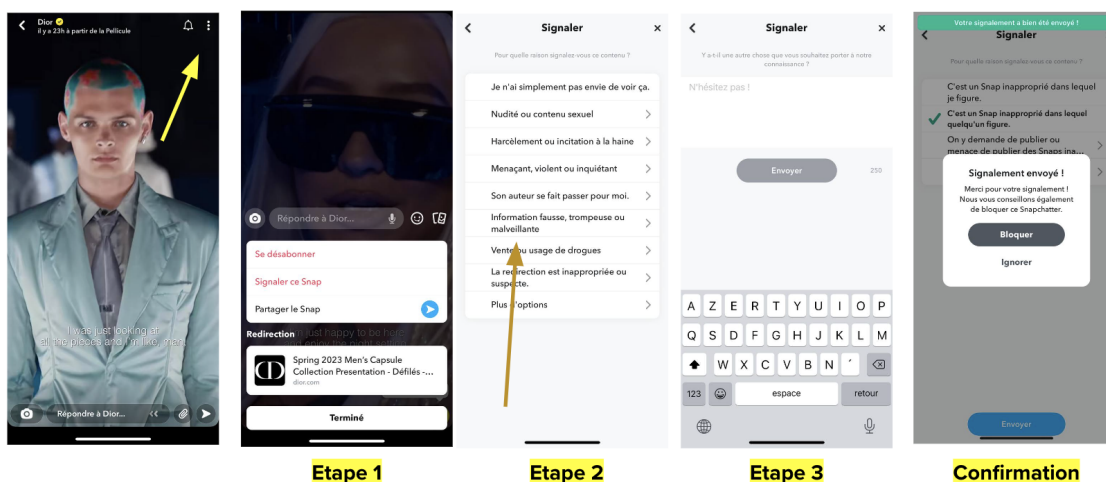
L'icône « drapeau » ou « i » mène à un menu de signalement, qui liste différentes catégories de motifs de signalement proposés à l'utilisateur, pour presque toutes les typologies de contenu de Snap (*Snap*s partagés en privé et en groupes, *Stories* privées, *Stories* officielles et populaires, *Stories* d'éditeur, y compris les publications d'actualités, *Shows* (émissions), *Stories* sélectionnées, Publicités). La plupart de nos menus de signalement comprennent aussi une catégorie « Autres » ou « plus d'option ». La présentation du menu de signalement permet également aux utilisateurs d'étoffer et d'argumenter leur signalement dans une case « commentaires » à compléter. Cette dernière étape est essentielle en matière de lutte contre le cyberharcèlement ou la désinformation, étant donné qu'elle permet d'apporter plus de contexte à notre équipe de modération.

Pour la partie Découvrir (Discover) les catégories de signalement suivantes sont possibles :

- Je n'ai simplement pas envie de voir ça
- Nudité ou contenu trop sexuel
- Harcèlement ou incitation à la haine
- Violence excessive
- On y promeut l'utilisation de substances illégales
- Son contenu est offensant ou choquant
- Autre
- **Plus d'options**

Lorsque les utilisateurs sélectionnent « Plus d'options », les options suivantes leur sont proposées, dont une option permettant de signaler les fausses informations :

- C'est du spam ou une arnaque
- C'est une violation de mon droit de propriété intellectuelle
- Terrorisme ou extrémisme
- Vente ou utilisation d'armes
- **information fausse, trompeuse ou malveillante (amélioration réalisée en 2020)**

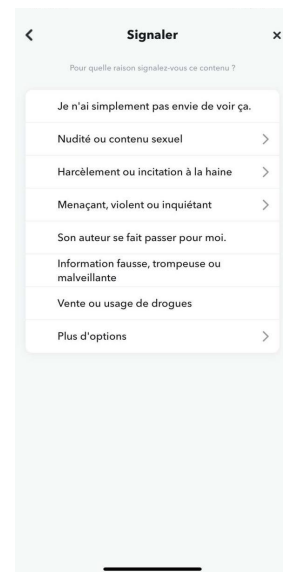


[Illustration des étapes de signalement](#)

Pour les parties Map et Spotlight, si le mécanisme de signalement reste le même que pour Discover, la partie “*information fausse, trompeuse ou malveillante*” apparaît directement en raison du fait que Spotlight est davantage axé sur le contenu généré par les utilisateurs.

Aussi, lorsqu’un utilisateur poursuit sa procédure, il lui est proposé de détailler la raison de son signalement afin d’apporter plus de contexte. Comme indiqué précédemment, nous estimons que cette étape est primordiale afin de recueillir le plus d’informations possibles, surtout en matière de désinformation.

Les membres de notre équipe dédiée dite « Confidentialité et Sécurité » ou « Trust and Safety » examinent tous les signalements des utilisateurs et prennent la décision de supprimer ou non le contenu en question et, dans certains cas, le compte de l’auteur du contenu. Ces dernières agissent dans les heures suivant le signalement.



1.4 Le dispositif décrit est-il **commun** :

- à toutes les **déclinaisons** du service (site internet, application) ? :

Ce système est mis en œuvre sur l'application Snapchat. Sa mise en œuvre sur le web, pour la catégorie “fausse information” est prévue dans un avenir proche.

- à tous les **types de contenus** (texte, vidéo, image, commentaires, contenus sponsorisés, compte utilisateur, contenus en direct, contenus éphémères) accessibles sur le service ? :

Nous sommes actuellement en train d’affiner notre mécanisme afin de le rendre plus lisible et simple d’utilisation [REDACTED]

1.5 Snapchat indiquera pourquoi les motifs de signalement ne sont pas uniformisés sur l'ensemble des types de contenus du service (Discover, Spotlight, publicités, *Snap Stars*, contenus réalisés en partenariat).

La procédure de signalement est uniforme sur l’ensemble de l’application. Seul point de différence parmi les différentes fonctionnalités publiques, l’emplacement de la catégorie “fausse information” apparaît soit en première page du menu de signalement, soit en seconde page via l’option “plus d’options”.

Nous sommes actuellement en train de d’affiner notre mécanisme, afin de le rendre plus lisible et simple d’utilisation [REDACTED]

1.6 Les utilisateurs sont-ils consultés sur l’accessibilité, la visibilité et l’ergonomie du dispositif de signalement ? :

Le cas échéant, comment et quels sont leurs retours ? :

Non applicable.

La procédure de traitement des signalements

1.7 Décrire la **procédure de traitement du signalement d'une fausse information** en fournissant toute illustration permettant une meilleure compréhension de la procédure.

- **Conditions** de mise en instruction d'un signalement :

Notre équipe de modération examine l'ensemble des signalements d'utilisateurs (et non-utilisateurs) indiqués comme étant une fausse information.

- **Étapes** de l'instruction :

Dès réception du signalement par le biais de l'application ou notre site d'assistance, celui-ci est attribué à un agent de l'équipe de modération. En fonction de la complexité du signalement, il peut être transmis à des niveaux supérieurs d'examen avant qu'une décision finale ne soit prise.

- **Issues possibles** de l'instruction (type d'actions prises) :

Plusieurs issues sont possibles : notre équipe peut approuver le contenu et dans cette hypothèse, le laisser en place, supprimer le contenu et en avertir l'auteur, ou dans des cas très graves, justifier la suppression du compte de l'utilisateur.

- **Information apportée aux utilisateurs** (lesquels sont informés ? quels stades de la procédure donnent lieu à une information ? sous quelle forme ?) :

Si le contenu est approuvé, l'auteur n'en sera pas informé et le contenu sera maintenu. Si le contenu est supprimé, l'auteur recevra un avertissement par mail indiquant qu'il a enfreint nos conditions d'utilisation et/ou nos règles communautaires. Si son compte est supprimé, il en sera également informé par mail avec les raisons de cette suppression.

- **Procédure spécifique** en période électorale ou en cas de crise

Nos équipes locales (dans ce cas, les équipes françaises) envoient une notification spécifique sur les élections ou autres événements prioritaires qui peuvent avoir une incidence élevée sur notre plateforme - que cela soit en matière de désinformation, ou autres contenus préjudiciables. Notre équipe de modération se charge par la suite de distribuer ces informations en interne aux agents afin qu'ils puissent être en alerte et crée également une boucle de rétroaction pour informer les équipes locales des phénomènes (ou non) perçus (le cas échéant).

- Recours à des **signaleurs de confiance** (*trusted flaggers* (préciser lesquels et à quelles fins)

Bien que nous disposions d'un programme de signaleurs de confiance, en raison du modèle de notre plateforme (absence de viralité, espaces publics modérés), nous ne disposons pas pour l'instant de signaleurs de confiance français en matière de lutte contre la désinformation.

1.8 **Voies de recours** internes à l'issue d'une décision :

- **décisions** de la plateforme susceptibles d'être contestées :
- **critères** de recevabilité des recours :
- **présentation à l'utilisateur** de ses voies de recours possibles contre une décision :
- **équipes** en charge de l'instruction des recours :

- **modalités d'instruction** du recours (notamment : possibilité pour l'utilisateur de faire valoir ses explications lors de l'instruction) :
- **exceptions** à cette procédure : NON / OUI (préciser)

Nous avons mis en place notre mécanisme de recours⁴ (*feedback loop*) en janvier 2022. De ce fait, nous ne pouvons pas répondre à cette question pour l'année 2021. Nous nous tenons néanmoins à disposition du régulateur pour apporter plus d'informations sur ce mécanisme.

1.9 Snapchat détaillera sa politique en matière de modération pour chaque type de comptes (médiàs, SnapStars, utilisateurs en profil public, utilisateurs en profil privé, créateurs, etc.).

Chez Snapchat, nous interdisons la diffusion de fausses informations sur l'ensemble de notre plateforme. À cette fin, nos règles communautaires et nos conditions d'utilisation du service s'appliquent à tous nos utilisateurs, qu'ils soient certifiés (*Snap stars*) ou non (utilisateur privé, ou profil public).

Il est important de noter que nos parties publiques font l'objet d'une modération accrue de la part de nos équipes. Certains acteurs sont donc soumis à des règles plus strictes et plus explicites :

- *Snap stars* (<https://support.snapchat.com/fr-FR/article/lost-snap-star-status>) :

Pourquoi ne suis-je plus une Snap Star ?

Nous nous réservons le droit de révoquer la vérification ou le statut de Snap Star à notre seule et entière discrétion et sans préavis. Les raisons pour lesquelles nous pouvons révoquer le statut de Snap Star sont les suivantes :

- Le compte ne répond plus aux critères d'éligibilité décrits ci-dessus ;
- Violations de nos Conditions d'utilisation du service ou de nos Règles communautaires ;
- Comportement dangereux qui met des tiers en danger ;
- Promotion de contenus, de comptes, d'applications ou de sites web (par des liens, des mentions ou autres) qui violeraient nos Conditions d'utilisation du service ou nos Règles communautaires ;
- Tentatives d'achat ou de vente d'abonnés ou d'augmentation artificielle du nombre d'abonnés ;
- Tout harcèlement envers Snap ou ses employés.

Nous effectuons régulièrement des audits de notre programme Snap Star. En définitive, nous nous fions à notre discrétion éditoriale pour évaluer les comptes qui doivent continuer à faire partie du programme.

Les comptes qui ne respectent pas nos Conditions d'utilisation du service et nos Règles communautaires peuvent être supprimés, en plus d'être exclus du programme Snap Star. Dans certains cas, un comportement en dehors de Snapchat non conforme à nos Règles communautaires ou à nos Conditions d'utilisation du service risque de vous rendre inéligible au statut Snap Star.

Illustration des règles pour les Snap stars

⁴ <https://support.snapchat.com/fr-FR/co/content-takedown-appeal>

- Partenaires médias (<https://publish-snapchat.zendesk.com/hc/en-us/articles/5191519553683-Content-Guidelines>)

Content Guidelines

Updated Guidelines are effective October 4, 2021

Snapchat values Content that is entertaining, informative, credible, compelling and tasteful. We select best-in-class partners to ensure that this Content is of the highest quality. Snapchatters represent a wide range of ages, cultures and beliefs, we want to provide a safe and healthy experience for each one of them, especially those as young as 13.

Snapchat rewards content that drives time well spent. Our goal is to delight Snapchatters each and every time they watch partnered Content on Snapchat. We do this through a mix of partners and channels that are welcoming, creative, credible, informative, inclusive and personable. We strongly discourage Content whose intent is to be click-bait, offensive, insensitive, salacious, upsetting, shocking, exaggerated, or in poor taste.

Ultimately, we will use our editorial discretion to determine whether your Content complies with these Guidelines.

Your Content must always comply with the Guidelines set out below and Snapchat's Community Guidelines, so please review them carefully. The Content you publish must be properly substantiated and fact-checked. Without limiting any rights or remedies under our agreement, we reserve the right to remove any Content that does not comply with these guidelines. Capitalized terms used but not defined in these guidelines will have the same meaning as in your agreement with us.

SECTION 1 - Rules For All Content

In addition to complying with our **Community Guidelines**, you may not include Content that:

[Illustration des règles applicables aux partenaires médias.](#)

- Spotlight (<https://support.snapchat.com/fr-FR/article/spotlight-guidelines>)

Prudence

- Outre le respect de nos [Règles communautaires](#), veillez à ce que vos publications dans Spotlight soient adaptées à une audience de plus de 13 ans.
- Les jeux de hasard, le tabac, les armes, les substances réglementées et la consommation d'alcool excessive ou par des mineurs sont interdits.
- N'essayez pas de tromper les gens, surtout quand il s'agit d'argent, de science, de santé, de politique ou de votre identité (faire bonne impression, c'est bien, à condition que ce soit sincère 😊).

Aucune sollicitation

- N'essayez pas de vendre ou de promouvoir des produits ou des services.
- Ne publiez pas de Snaps sponsorisés ou rémunérés.
- Les marques doivent consulter [Snapchat pour les entreprises](#) ou demander la [version bêta du Profil de marque](#).
- N'utilisez pas de redirections ou d'URL dans les Snaps publiés sur Spotlight.

[Illustration des règles sur Spotlight](#)

- Annonceurs publicitaires (<https://www.snap.com/fr-FR/ad-policies>)

Politiques de Snap relatives à la publicité

1. [Introduction](#)
2. [Conditions générales](#)
 - 2.1 [Ciblage et conformité](#)
 - 2.2 [Divulgations](#)
 - 2.3 [Confidentialité : Collecte et utilisation des données](#)
 - 2.4 [Propriété intellectuelle](#)
 - 2.5 [Qualité créative et page de renvoi](#)
 - 2.6 [Promotions](#)
3. [Exigences spécifiques à la catégorie](#)
 - 3.1 [Contenu adulte](#)
 - 3.2 [Marchandises réglementées](#)

[Illustration des règles relatives à la publicité](#)

Les moyens alloués au traitement des signalements

1.10 Moyens humains déployés, et notamment :

- statut des modérateurs (collaborateurs ou prestataires de la société opératrice) :
- formation (initiale et continue) :
- localisation géographique :
- langue(s) de travail (notamment, langues régionales françaises prises en charge) :
- modération à tout moment (24/24 et 7/7) : OUI / NON (préciser)
- sources dont disposent les modérateurs (juridique, journalistique, académique) :
- mesures pour réduire les risques professionnels pour les modérateurs :

Lorsque nous recevons un signalement d'un utilisateur concernant un contenu particulier, nous l'évaluons en fonction de nos règles communautaires, à la lumière de la loi locale. Si le contenu n'enfreint pas ces règles, nous ne prenons pas de mesure et le maintenons en ligne. Toutefois, en cas de violation, nous pouvons supprimer le contenu et avertir l'utilisateur qui l'a posté/envoyé. Dans le cas où un utilisateur participerait à des violations répétées de nos règles communautaires, ou en cas de compte dédié à un but tel que la pornographie, nous supprimons le compte. Avant les grands événements mondiaux et locaux, comme les élections ou d'importants événements sportifs, nous informons les membres de l'équipe afin qu'ils se préparent à une potentielle augmentation du volume de contenus illicites. S'agissant des fausses informations, nous faisons appel à des sources d'information de confiance, comme les autorités locales, pour nous alerter sur les tendances en matière de fausses informations, que nous communiquons à l'ensemble de notre équipe afin de nous tenir prêts à réagir rapidement à de nouvelles violations de nos règles communautaires.



L'équipe de modération a connaissance des interdictions et des règles applicables en France. L'équipe est également composée de plusieurs membres qui parlent couramment français.

Nous travaillons continuellement au développement et à la mise en œuvre d'outils automatisés pour gagner en efficacité. Comme indiqué ci-dessus, le contenu peut faire l'objet d'un examen manuel, d'un examen automatisé ou d'une combinaison des deux. Le cas échéant, nous publierons les données pertinentes dans notre rapport de transparence bi-annuel. Toutefois, en matière de fausses informations, aucun outil automatisé n'a été mis en place à cette heure.

Certains contenus populaires peuvent être mis en avant sur notre espace Discover pour une plus grande visibilité. Ce contenu est modéré avant d'être promu. Ainsi, si un compte populaire publiait du contenu trompeur ou susceptible de causer un préjudice dans sa Story, il serait exclu de manière proactive de toute promotion plus large de contenu sur Discover dans la partie « Pour vous ». Cette mesure s'ajoute à un suivi réactif de l'ensemble des signalements que nous recevons concernant ces comptes.

Les recours concernant les décisions relatives aux comptes sont examinés après leur envoi *via* notre outil d'assistance « contactez-nous ». Les processus sont structurés de manière à garantir le respect des politiques et des pratiques, ce qui permet un examen cohérent et équitable. Des contrôles d'assurance qualité sont ensuite utilisés pour garantir que ces pratiques sont minutieusement suivies à grande échelle.

1.11 Existence de moyens humains et/ou automatisés de **détection proactive et dédiée** aux contenus diffusant de fausses informations :
Processus d'instruction des contenus détectés identique à celui des contenus signalés par les utilisateurs :

Non applicable. Snapchat ne fait pas appel à la détection proactive en matière de lutte contre la désinformation.

1.12 **Commenter la préconisation du Conseil supérieur de l'audiovisuel (CSA) aux opérateurs**, dans son précédent bilan sur la lutte contre la manipulation de l'information, **de maintenir une intervention humaine dans le processus de décision d'une action** à l'égard d'un contenu ou d'un compte.

Snap s'appuie sur la modération humaine pour lutter contre la désinformation. En effet, les fausses informations nécessitent une réponse au cas par cas. Snap reconnaît depuis longtemps l'importance du contexte pour évaluer les risques de manipulation de l'information et les modérateurs humains continuent de jouer un rôle essentiel dans le fonctionnement de nos politiques, conformément à la recommandation du Conseil supérieur de l'audiovisuel français (CSA).

Questions complémentaires

1.13 Remplir le tableau suivant.

Données chiffrées et indicateurs (France, 2021)

Nombre de signalements pour fausses informations (ou notion voisine adoptée) par les utilisateurs	35 617
Nombre de procédures de traitement mises en œuvre	35 617
Délai moyen de traitement d'un signalement	4,3 heures
Part des signalements traités : - en moins de 24h - en moins d'une semaine	96,3 % des rapports de fausses informations ont été traités dans les 24 heures ; 99,8 % ont été traités dans les 7 jours.
Nombre de contenus considérés comme étant de fausses informations (ou notion voisine adoptée) à l'issue de la procédure de traitement	971
Nombre de recours reçus par l'opérateur à l'issue de la procédure de traitement	N/A
Délai moyen du traitement d'un recours	N/A
Part des recours traités : - en moins de 24h - en moins d'une semaine	N/A
Budget alloué au traitement des signalements et des recours des utilisateurs (€)	N/A
Nombre de personnes en charge du traitement des signalements susceptibles d'agir sur des contenus du service en France	

1.14 **Impact de la crise sanitaire sur le dispositif de signalement** de fausses informations (ou notion voisine adoptée) et mesures spécifiques prises dans ce contexte en 2021 :

- Augmentation des chiffres relatifs aux contenus supprimés pour fausses informations : nous avons commencé à permettre à nos utilisateurs de signaler directement les fausses informations sur l'application plutôt que sur un formulaire situé dans notre site support. Les utilisateurs ont ainsi pu signaler plus facilement cette catégorie et ces signalements sont devenus plus exploitables car ils étaient accompagnés du contenu signalé.
- Vente de faux pass sanitaires : L'introduction du pass sanitaire en France a entraîné un véritable phénomène de vente frauduleuse sur notre plateforme. Nous avons collaboré activement avec les autorités françaises pour lutter rapidement et efficacement contre ce problème. Cela a également eu un impact sur le nombre de signalement qui nous a été remonté - cependant, nos utilisateurs ont compris qu'en l'espèce il s'agissait de vente de contrefaçon (catégorie : arnaque), et non de fausses informations.

1.15 **Actions futures et nouvelles fonctionnalités envisagées à l'avenir :**

[Redacted text]

[Redacted text]

2 Transparence des algorithmes

Les questions suivantes s'adressent aux opérateurs qui auraient mis en œuvre des mesures relatives à la transparence des algorithmes au titre des mesures complémentaires prévues par l'article 11 de la loi du 22 décembre 2018.

2.1 Commenter les définitions de la **loyauté, de l'explicabilité et de l'équité** indiquées en annexe du questionnaire :

L'opérateur indiquera s'il s'y réfère (ou à des notions équivalentes) notamment dans le cadre des mesures qu'il met en œuvre afin de lutter contre la manipulation de l'information, et les éventuelles limites identifiées à leur application.

Il est invité à formuler d'éventuelles propositions afin d'affiner ces définitions.

Conformément à nos règles communautaires, Snap interdit la diffusion de fausses informations, comme l'existence d'événements tragiques, les allégations médicales non fondées ou l'atteinte à l'intégrité de processus civiques. Lien vers les règles communautaires : <https://www.snap.com/fr-FR/community-guidelines>



Nous partageons également les éléments suivants dans notre Centre de confidentialité (<https://snap.com/fr-FR/privacy/privacy-by-product>) pour aider les utilisateurs à comprendre comment nous personnalisons le contenu :

Personnalisation

Il n'y a pas deux personnes identiques, c'est pourquoi l'application peut personnaliser l'expérience Snapchat en fonction des interactions de chacun avec ses amis. Par exemple, les amis avec lesquels un *Snapchatter* interagit le plus souvent sont généralement les plus faciles à trouver.

Nous essayons généralement de montrer des publicités qui sont pertinentes pour chacun. Pour ce faire, nous faisons des suppositions sur ce que les utilisateurs aiment et n'aiment pas. Par exemple, si un utilisateur regarde constamment des photos de chiots sur Discover ou Spotlight, il peut être un "amateur d'animaux". Ou, s'il publie régulièrement des photos de ses souvenirs de sport sur Ma Story, notre système pourrait deviner qu'il est un "fan de sport". Nous appelons ces suppositions "catégories de style de vie" et nous les utilisons pour essayer de montrer des publicités et du contenu qui correspondent aux intérêts des *Snapchatters*. S'ils le souhaitent, ils peuvent toujours choisir leurs propres catégories de style de vie dans leurs paramètres.

Nous personnalisons le Snap Map. Par exemple, si un utilisateur visite un lieu ou applique un filtre ou un autocollant de lieu sur une publication dans My Story, Our Story, Spotlight ou Snap Map, nous pouvons utiliser ces informations de lieu pour créer une expérience de carte plus personnalisée pour lui et ses amis.

Nous essayons également de personnaliser le contenu qu'il voit en utilisant ce que nous appelons des "tags d'intérêt de contenu". Ces *tags* sont des suppositions sur le contenu qui peut intéresser les utilisateurs en fonction de leur activité, comme les choses qu'ils recherchent, les stories qu'ils regardent, les utilisateurs populaires auxquels ils sont abonnés et les types de lieux qu'ils visitent. Ainsi, s'ils regardent beaucoup de stories sur le basket-ball, nous pouvons créer des marqueurs d'intérêt pour le contenu "sports" et "basket-ball", afin de lui montrer plus de snaps

sur le basket-ball. Il peut réinitialiser les marqueurs d'intérêt de contenu à tout moment dans les paramètres de Snapchat.

Systèmes de modération de contenus

2.2 **Décrire chaque système de modération** de contenus utilisés sur le service pour lutter contre la manipulation de l'information en détaillant :

- ses **fonctionnalités** ;
- son **objectif principal** (tâche ou classe de problèmes à résoudre) ;
- le fait qu'il repose ou non sur un ou plusieurs algorithmes d'**apprentissage automatique** et, le cas échéant, s'ils sont utilisés avec d'autres types d'algorithmes (ex. algorithmes déterministes) ;
- lorsqu'il repose sur un ou plusieurs algorithmes d'apprentissage automatique, la nature de(s) **jeu(x) de données d'apprentissage** utilisé(s) ;
- pour la reconnaissance de texte, les **langues utilisées** (notamment régionales) ;
- la liste la plus exhaustive possible des **données prises en entrée** de ce système ;
- la liste la plus exhaustive possible des **données de sortie** de ce système ;
- les conditions mises en œuvre par l'opérateur qui permettent que ses services et leurs algorithmes respectent les **principes de loyauté et d'équité**, et qu'ils soient explicables ;
- le cas échéant, la manière dont ce système est **personnalisé** en fonction de l'utilisateur ;
- le cas échéant, les modalités de recours à une **intervention humaine** sur le système à partir de sa mise en œuvre et si recours est le fait de personnes en interne, de personnes employées par des prestataires, ou bien des deux.

Les opérateurs pourront recourir à des **représentations graphiques** afin de répondre aux questions⁵, faisant figurer clairement l'ensemble des éléments demandés.

Il est important de rappeler que ce sont des êtres humains qui écrivent les algorithmes et qui peuvent les optimiser pour tenir compte des comportements humains. Cela signifie qu'un algorithme peut être conçu pour fournir plusieurs sources de contenu et différents points de vue.

La conception initiale de l'application Snapchat, notamment l'accent mis sur la communication non virale, avec des amis proches et le contenu sélectionné, n'implique qu'un besoin limité d'employer d'autres techniques pour réduire la menace de manipulation de l'information. Le cas échéant et conformément à notre engagement à l'égard de la vie privée des utilisateurs, nous avons mis en place des outils automatisés pour supprimer rapidement certains contenus illégaux et interdits, notamment les discours haineux, la pornographie et les images d'abus sexuels sur des enfants.

Notre moteur de classement Discover offre aux utilisateurs une expérience personnalisée fondée sur nos valeurs :

- Contenu crédible et faisant autorité : nous valorisons le contenu des partenaires professionnels qui sont crédibles et font autorité dans leur domaine. Les partenaires ont une liberté éditoriale sur le contenu qu'ils publient, bien sûr, en restant dans le cadre de nos règles communautaires et de nos règles dédiées aux éditeurs. Les contenus médias sont présentés dans des éditions à part entière, et non dans des articles individuels, et le nom du média est toujours identifié dans l'interface utilisateur.
- Des perspectives multiples et des opinions diverses : cela nous aide à éviter de devenir une caisse de résonance. La consommation, les intérêts et le comportement dans

⁵ La [norme ISO 5807](#) pourra par exemple être utilisée pour établir des logigrammes.

l'application sont les principaux facteurs qui déterminent comment le flux d'un utilisateur peut différer de celui d'un autre.

- Des données et des humains au cœur de l'application : notre équipe de programmation média est issue du monde de la rédaction et veille à créer un contenu représentatif sur tous nos marchés.

D'une manière générale, l'application peut adapter l'expérience Snapchat de chaque *Snapchatter*. Nous y parvenons en devinant ce qu'un utilisateur aime ou n'aime pas (*cf.* plus haut).

De plus, Snap met en pratique la minimisation des données et la “*purpose limitation*”, ainsi que le respect de la vie privée dès la conception (“*privacy by design*”), en exigeant une finalité précise de la conception du produit pour les fonctionnalités qui collectent ou traitent des données personnelles. Elle doit non seulement décrire clairement les données personnelles qui seront collectées et utilisées, mais aussi les objectifs spécifiques et la période de conservation requise (avec des limitations de conservation strictes par défaut). Ces documents font l'objet d'une collaboration entre les équipes chargées des produits, les juristes et ingénieurs chargés de la protection de la vie privée au cours du processus de développement, et l'examen doit être achevé avant le lancement ou la mise en œuvre de la fonctionnalité. À titre d'exemple, en 2020, les avocats chargés de la protection de la vie privée/des produits de Snap ont examiné plus de 1 500 spécifications de conception.

Snap met l'accent sur la transparence dans tous les aspects de sa communication avec ses utilisateurs, qu'il s'agisse de sa politique de confidentialité, de son centre de confidentialité ou, surtout, de l'utilisation de notifications sur l'application. Snap propose aux utilisateurs différents moyens de personnaliser leur expérience et de supprimer / bloquer / ajuster le contenu qu'ils ne veulent pas voir dans l'application. Les utilisateurs peuvent également afficher et personnaliser les catégories de style de vie qui contribuent à déterminer le contenu qu'ils voient dans le menu des paramètres de Snapchat.

Il n'existe pas d'algorithme spécifique pour lutter contre la manipulation de l'information. Cela est dû au fait que la manipulation de l'information est extrêmement limitée sur Snapchat (comme indiqué dans le présent questionnaire). La diffusion de fausses informations qui causent des dommages ou qui sont malveillantes, telles que la négation de l'existence d'événements tragiques, les allégations médicales non fondées ou l'atteinte à l'intégrité des processus civiques, est clairement décrite comme contraire à nos lignes directrices communautaires et est supprimée lorsqu'elle est détectée.

2.3 Décrire la **politique de lutte contre les biais** de l'opérateur pouvant exister dans ses algorithmes d'apprentissage automatique :

Indiquer **si des biais ont pu être observés dans la modération** des contenus sur le service. Le cas échéant, détailler les procédures qui ont permis l'identification de ces biais et les modifications apportées pour les supprimer.

2.4 Décrire les **modalités d'évaluation** des algorithmes de modération utilisés sur le service.

Réponses aux questions 2.3 et 2.4

Pour nos algorithmes de modération Discover, nous mettons également en œuvre une part de régulation humaine pour nous assurer que l'automatisation ne mène pas à oublier nos politiques

de modération, ainsi que la réalité des interactions humaines. Ce mode de régulation est également utilisé pour contrôler les algorithmes afin d'améliorer leur précision et d'éviter les biais algorithmiques. Nous procédons également à un échantillonnage aléatoire.

2.5 Si des **modifications des systèmes de modération** sont survenues en 2021, en préciser les raisons, détailler les modifications apportées et indiquer si elles ont eu un caractère temporaire :

Nous affinons constamment les algorithmes utilisés pour la modération des contenus. Nous nous appuyons sur les données relatives au signalement, aux enquêtes et aux mesures prises à l'égard du contenu. Par exemple, en 2021, nous avons investi dans un système élargi de notification et d'appel pour les contenus signalés.

2.6 Indiquer si **des outils automatisés de modération pour lutter contre des pratiques de manipulation de l'information en particulier** (ex : outils de lutte contre les *deepfakes*, contre la détection de *bots*, etc.) ont été développés :

Il n'existe pas d'algorithmes spécifiques liés à la lutte et la détection en matière de manipulation de l'information. Ceci est dû au fait que la manipulation de l'information est extrêmement limitée sur Snapchat (comme détaillé dans ce questionnaire).

Systèmes de recommandation de contenus

2.7 **Décrire les systèmes de recommandation** de contenus utilisés sur le service, en détaillant les mêmes points que ceux énoncés à la question 2.2. :

Nous utilisons des systèmes de recommandation de contenu pour nos espaces Discover et Spotlight. Nous n'utilisons pas de systèmes de recommandation de contenu pour afficher du contenu sur la Map.

Comme nous l'avons indiqué précédemment, Discover est une plateforme fermée destinée aux sociétés médiatiques de confiance que Snap a préalablement examinées, au contenu sélectionné par Snap et aux comptes d'influenceurs et de célébrités que nous appelons *Snap Stars* ou Utilisateurs populaires. La page Discover présente des Stories d'amis, des Stories abonnées et la section "Pour vous". Afin de s'assurer que le contenu non sélectionné est également approprié pour être visionné sur notre fil Pour vous, notre équipe modère tout le contenu issu des Snapchatters qui est promu.

Notre page Spotlight présente des Snaps d'utilisateurs, d'influenceurs et de célébrités qui choisissent de publier des Snaps sur Spotlight.

La façon dont le contenu apparaît sur Discover et Spotlight pour chaque utilisateur dépend de certains facteurs, comme les stories auxquelles il s'est abonné, les stories et snaps qu'il a consultés, les stories et snaps qui sont actuellement populaires, ainsi que les préférences et les données démographiques de l'utilisateur.

2.8 Si des **modifications de ces systèmes de recommandation** sont survenues en 2021, en préciser les raisons, détailler les modifications apportées et indiquer si elles ont eu un caractère temporaire :

Nous affinons en permanence les algorithmes d'optimisation et de recommandation de contenu en fonction de divers facteurs, comme l'engagement envers le contenu.

2.9 Actions mises en place pour lutter, via les systèmes de recommandation de contenus sur le service, contre la manipulation de l'information (décrire chaque type d'actions et préciser dans quelles conditions elles sont appliquées (critères de déclenchement)) :

Il n'existe pas d'algorithmes spécifiques liés à la lutte contre la manipulation de l'information. Ceci est dû au fait que la manipulation de l'information est extrêmement limitée sur Snapchat et nécessite une analyse humaine au cas par cas.

Informations des utilisateurs

2.10 Manière dont les **systèmes de recommandation et de modération** utilisés dans le cadre de la lutte contre la manipulation de l'information sont **décrits à l'utilisateur** :

Nous avons une page sur notre [site support](#) qui détaille comment nous luttons contre la manipulation de l'information à travers la structure de notre plateforme, nos règles communautaires et leur application, ou bien encore notre collaboration avec nos partenaires médias de confiance.

2.11 Pour chacun de ces systèmes, préciser **de quelle manière les utilisateurs** :

- sont **informés des mises à jour** de ces systèmes :
- peuvent demander des **informations supplémentaires** sur ces systèmes :
- bénéficient d'une **explication des résultats des systèmes** qui les affectent :

Une description générale du classement algorithmique est fournie aux utilisateurs dans notre Centre de confidentialité, avec les descriptions déjà existantes de la façon dont les *Snapchatters* peuvent contrôler leur expérience sur la plateforme. Le Centre de confidentialité donne un aperçu de [l'approche générale](#) de Snapchat en matière de confidentialité, ainsi qu'une description par fonctionnalité des données collectées et de la manière dont les utilisateurs peuvent personnaliser leur expérience.

2.12 Modifications qu'un utilisateur peut apporter aux paramètres relatifs aux algorithmes de recommandation sur le service.

- **liste** des systèmes de recommandation paramétrables par l'utilisateur :
- comment ces réglages (2.11) sont rendus facilement **accessibles et compréhensibles** :

Nos utilisateurs peuvent ajuster leurs paramètres pour voir des contenus qui leur sont plus pertinents depuis la page paramètres de l'application, sous la rubrique "Gérer". En outre, les utilisateurs sont en mesure de gérer les chaînes Discover auxquelles ils sont abonnés, y compris la possibilité de masquer les *Stories* directement dans la page Discover.

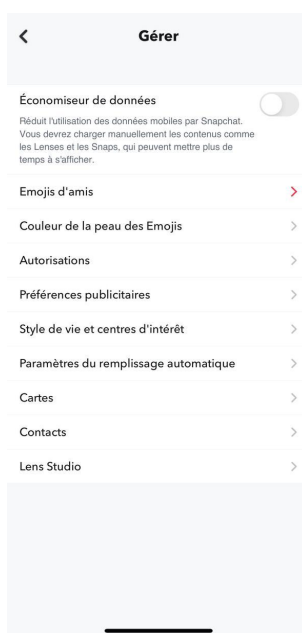
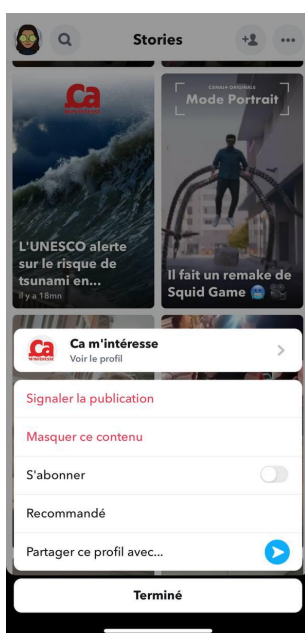


Illustration question 2.12

2.13 Commenter la préconisation du CSA aux opérateurs, dans son précédent bilan sur la lutte contre la manipulation de l'information, **de proposer aux utilisateurs des fonctionnalités leur permettant de comprendre, si possible de manière personnalisée et contextuelle, les effets des systèmes algorithmiques de recommandation et de modération.**

Ce sujet est particulièrement important, c'est pourquoi Snap a proactivement prévu des outils, facilement accessible et ce, directement depuis les paramètres de l'application ou sur le contenu lui-même, permettant à ses utilisateurs de comprendre et personnaliser leur expérience.

Questions complémentaires

2.14 Remplir le tableau suivant.

Données chiffrées et indicateurs (2021)	
Parts des accès à des fausses informations (ou notion voisine adoptée) imputables : <ul style="list-style-type: none"> - à l'algorithme du moteur de recherche interne du service - aux autres algorithmes de recommandation, classement ou référencement du service 	Snap ne collecte pas ces données.
Nombre de contenus de fausse information (ou notion voisine adoptée) ayant fait l'objet d'une recommandation	Snap ne collecte pas ces données.
Part des utilisateurs du service ayant recours à la personnalisation des paramètres relatifs aux algorithmes de recommandation des contenus (organiques et commerciaux) sur le service	Snap ne collecte pas ces données.

2.15 Impact de la crise sanitaire sur le fonctionnement des systèmes de recommandation ou de modération pour lutter contre la manipulation de l'information et mesures spécifiques prises dans ce contexte en 2021 :

Comme expliqué dans ce questionnaire, il n'existe pas d'algorithmes spécifiques liés à la lutte contre la manipulation de l'information. Ceci est dû au degré extrêmement limité de manipulation de l'information sur Snapchat (comme détaillé dans ce questionnaire). La diffusion de fausses informations qui causent des dommages ou qui sont malveillantes, telles que la négation de l'existence d'événements tragiques, les allégations médicales non fondées ou l'atteinte à l'intégrité des processus civiques, est clairement décrite comme contraire à nos directives communautaires et est supprimée lorsqu'elle est détectée. La crise sanitaire n'a pas eu d'impact significatif sur la présence de fausses informations sur notre plateforme, nous n'avons donc pas à adapter notre système de modération ni à prendre des mesures spécifiques en 2021.

2.16 Actions futures et nouvelles fonctionnalités envisagées à l'avenir

N/A.

3 Promotion des contenus issus d'entreprises et d'agences de presse et de services de communication audiovisuelle

Les questions suivantes s'adressent aux opérateurs qui auraient mis en œuvre des mesures relatives à la promotion des contenus issus d'entreprises et d'agences de presse et de services de communication audiovisuelle au titre des mesures complémentaires prévues par l'article 11 de la loi du 22 décembre 2018.

Identification des sources des contenus et labellisation

3.1 Identification des sources des contenus issus d'entreprises et d'agences de presse et de services de communication audiovisuelle. Merci d'illustrer ou de donner des exemples.

- **Identification réalisée par des moyens humains et / ou automatisés ?** : Oui
- **Sources et critères utilisés pour cette identification** : Humain
- **Visibilité de cette identification sur le service** : Partenariats directs avec des médias
- **Spécificité du dispositif à ces entreprises ou agences** : Les partenaires sont obligés de mettre leur logo afin d'éviter une confusion et une désintermédiation
- **Éléments complémentaires d'information et de contexte apportés** : Titre des émissions.
- **Espaces où apparaissent ces éléments d'identification sur le service** : Discover



Les logos des partenaires médias sont mis en évidence dans le fil Discover sur des illustrations conçues pour attirer l'attention de l'utilisateur. Nous pensons qu'une visibilité plus forte contribue à promouvoir la crédibilité de l'information. Le contenu des éditeurs partenaires porte le logo de la marque de ces partenaires, par exemple "Le Monde". Le logo et/ou le nom de la source, ainsi qu'un point d'entrée vers le profil de la source, sont disponibles partout où le contenu apparaît, que ce soit dans les parties "Recherche", "Pour vous" ou ailleurs.

3.2 Snapchat apportera des précisions quant aux **certifications nécessaires pour apparaître sur « Discover »**.

Le contenu de nos partenaires médias et des éditeurs professionnels est présenté sur Snapchat dans l'espace Discover. Avec d'autres contenus, cette section propose des contenus d'éditeurs professionnels de confiance sélectionnés et approuvés par l'équipe éditoriale de Snapchat.

L'approche de Snapchat en matière de contenu est différente de celle des autres plateformes. Grâce à la manière dont nous avons construit notre plateforme de contenu Discover, nous avons pu nous protéger contre la diffusion massive de fausses informations. Nos contenus d'information ne sont disponibles que grâce à un partenariat direct avec Snap. Notre équipe éditoriale est responsable de notre stratégie d'intégration des partenaires et nous examinons et sélectionnons soigneusement et individuellement les entreprises avec lesquelles nous établissons des partenariats.

Il s'agit d'une sélection mutuelle. Nos partenaires nous contactent d'abord pour devenir un partenaire Discover, et nous évaluons par la suite leur demande. Nous travaillons avec des médias qui sont connus pour leur forte empreinte éditoriale et qui sont reconnus comme des

professionnels dans le milieu journalistique. Pour être sélectionnés, nos partenaires doivent produire du contenu dans le cadre de nos [règles communautaires](#) et nos règles dédiés aux éditeurs (notamment pas de fausses informations, pas d'incitation à la haine ou à la discrimination, etc.). Ces directives sont appliquées par des agents humains. Ces derniers identifient le contenu qui ne respecte pas ces directives, le masquent ou le privent de priorité dans le flux de contenu «Pour vous» des utilisateurs. Des sanctions peuvent également être prises contre nos partenaires, dont la certification peut être rompue. En outre, les utilisateurs peuvent masquer ou signaler du contenu dans «Pour vous». Ces signalements sont également examinés par des agents humains et des mesures appropriées sont prises à l'encontre du contenu jugé contraire à nos politiques, y compris nos directives communautaires (<https://www.snap.com/fr-FR/community-guidelines>).

Les entreprises avec lesquelles nous passons un contrat ont accès à nos outils de contrôle avancé, de production et de gestion de contenu. Ces outils permettent aux partenaires de produire, d'éditer et de télécharger du contenu, ainsi que de mieux comprendre, grâce aux analyses, les performances du contenu produit.

Les logos de nos partenaires apparaissent explicitement sur les couvertures des éditions, et restent visibles tout au long des Snaps en haut de l'écran de l'utilisateur. Le contenu est parfois recommandé en fonction des intérêts personnels des utilisateurs, y compris leurs abonnements. Le contenu de type "*click-bait*" (appât à clics) est expressément interdit en vertu de nos directives relatives aux éditeurs. Les publications des éditeurs sont présentées comme des éditions entières (les articles ou les émissions sont à visionner dans l'ordre chronologique), et non comme des articles individuels. Aussi, notre nouveau format, "*Dynamic Stories*" permet à un éditeur d'utiliser son flux RSS pour créer automatiquement des stories à partir d'un contenu déjà publié sur le web - celui-ci est présenté comme un article individuel.

L'une des valeurs qui sous-tend notre algorithme de classement est la promotion d'un contenu faisant autorité et crédible, provenant de nos partenaires de confiance. Nous pensons que donner la possibilité de faire entendre sa voix va de pair avec un haut niveau de responsabilité. Nos partenaires sont tenus de vérifier les faits et de s'assurer que leur contenu n'est pas faux ou trompeur.

3.3 Commenter la préconisation du CSA aux opérateurs, dans son précédent bilan sur la lutte contre la manipulation de l'information, de maintenir une intervention humaine dans la vérification des informations fournies au sujet des organes de presse lorsqu'ils font l'objet d'une identification particulière.

Notre modération sur Discover comprend une modération humaine. Nos règles communautaires sont strictes et visibles par tous.

3.4 Décrire les espaces et processus de mise en avant de contenus provenant de ces sources sur le service et les critères de sélection de ces sources et contenus :

Discover est une plateforme fermée. Nos partenaires signent un contrat dans lequel ils acceptent de s'aligner sur nos règles communautaires et nos règles éditoriales (voir ci-dessus). Le contenu est mis en avant en fonction des affinités des utilisateurs.

3.5 Moyens mis en œuvre pour tenir compte des démarches de labellisation et de bonnes pratiques réalisées par des tiers, notamment par les entreprises et agences de presse et les services de communication audiovisuelle (ex : Trust Project, Journalism Trust Initiative, etc.) :

Snap gère le programme dit “*Trusted Flagger*”. Ce programme très efficace a été développé pour fournir aux organisations d'aide internationale, aux ONG, aux agences gouvernementales et autres partenaires de sécurité approuvés, un lieu confidentiel pour signaler les abus et les contenus qui violent nos règles communautaires. Ces agences gouvernementales et/ou organisations non gouvernementales (ONG) font remonter les signalements à Snap. Ce service de qualité est contrôlé 24 heures sur 24 et 7 jours sur 7 par l'équipe Confidentialité et Sécurité de Snap. Nous n'avons pas encore pu développer ce programme en matière de fausses informations, pour des raisons liées à des priorisations internes expliquées plus haut.

3.6 Description des processus de labellisation réalisés par l'opérateur pour identifier les comptes ou contenus de **médias affiliés à un État** (ou équivalent), en précisant :

- **critères** de labellisation :
- **informations fournies** :
- **modalités de mise en avant** de ces informations sur le service :
- **partenariats** éventuels pour réaliser cette labellisation :

Nous ne labellisons pas les comptes ou le contenu des médias affiliés à l'État.

3.7 Décrire les **éventuels freins** à la mise en œuvre des mesures d'identification et de labellisation :

Non applicable.

Dispositifs de vérification des faits (« *fact-checking* »)

3.8 Recours à un dispositif de ***fact-checking*** (confié à des tiers ou interne) :

Nous respectons la liberté éditoriale de nos partenaires ; les partenaires médias sur notre plateforme Discover doivent attester que leur contenu est soumis à des pratiques de vérification des faits, et d'autres contenus de créateurs à forte audience sont soumis aux pratiques de modération décrites dans ce questionnaire (et détaillées dans la section ci-dessous). Bien que nous ayons une relation établie avec une organisation tierce de vérification des faits mondialement reconnue, l'architecture de notre plateforme permet de limiter la portée des fausses informations à un point tel que très peu d'interventions ont été nécessaires, et le partenariat n'a pas été utilisé pour du contenu en France.

- **contenus susceptibles d'être soumis au *fact-checking*** : UNE SÉLECTION - Publicités politiques
- **moyens technologiques mis à la disposition des *fact-checkers* (outil de suggestion des contenus à vérifier et selon quels critères, mise à disposition d'une API, etc.)** : Non applicable.
- **existence d'un dispositif d'urgence ou en cas d'actualité particulière** : En cas d'urgence ou d'événement important (par exemple, la guerre en Ukraine), notre équipe chargée des partenariats avec les médias se met directement en lien avec nos partenaires médiatiques pour recueillir leurs impressions et assurer une meilleure coopération, compréhension et visibilité de l'information.

3.9 En cas de recours à des ***fact-checkers* tiers**, indiquer :

- **les critères de sélection des *fact-checkers*** : Nous avons examiné les partenaires potentiels dans ce domaine pour leur leadership et leur crédibilité dans le domaine de l'intégrité des médias et de l'information

- **la liste des structures de *fact-checking* partenaires de la plateforme** : Nous sommes en collaboration avec [Poynter Institute](#) - où aucun contenu comportant une fausse information n'a été constatée pour la France.
- **la nature des liens juridiques unissant l'opérateur aux *fact-checkers*** : Partenariat contractuel rémunéré.
- **la nature des contreparties qui leur sont offertes** : Payé pour les services facturés.

3.10 Procédures de traitement des contenus dont le *fact-checking* (interne ou externe) a montré qu'ils contenaient ou étaient susceptibles de contenir une fausse information :

- **décrire la chaîne de décision** amenant l'opérateur à agir sur un contenu ou un compte :
- **types de mesures** prises et leur **gradation** :
- **critères d'application** de chacune d'entre elles :
- **variations des mesures** selon le type ou la nature du contenu vérifié et pourquoi :
- **visibilité** de l'intervention des *fact-checkers* sur le service :
- **voies de recours** internes pour les auteurs des contenus modérés : OUI / NON
- **recours** traités par l'opérateur ou par le tiers partenaire ? :

Nos équipes de modération évaluent toutes les allégations de fausses informations en fonction de nos conditions d'utilisation ou de nos directives communautaires. S'il s'avère qu'un contenu contient de fausses informations qui violent nos conditions d'utilisation ou nos directives communautaires, nous appliquerons la politique à l'encontre de ce contenu ou de cet utilisateur, selon le cas.

3.11 Modalités de communication de l'opérateur ou moyens permettant aux *fact-checkers* de communiquer sur l'action prise auprès :

- des utilisateurs ayant publié ou partagé le contenu ou ayant interagi⁶ avec lui :
- de l'ensemble des utilisateurs (visibilité sur le service) :

Réponse aux questions 3.8 à 3.11

La section Discover contient exclusivement du contenu provenant d'entreprises média de premier plan, du contenu modéré issu des comptes publics de personnalités connues et de créateurs approuvés ("*Snap stars*"). Grâce à ce contrôle exacerbé, les risques de diffusion de fausses informations sur notre plateforme sont faibles.

Au vu de la nature même de Discover (voir ci-dessus), les « *fact-checkers* » apparaissent moins nécessaires que sur d'autres plateformes. En effet, cet espace constitue la meilleure illustration de la promotion par Snap d'un contenu fiable :

- Tous les contenus publics qui apparaissent sur Discover sont sélectionnés. La meilleure façon d'assurer la promotion d'un contenu « sûr » est de veiller à ce que les utilisateurs ne voient que des informations vraies et de qualité.
- L'essentiel du contenu provient d'entreprises média professionnelles connues (par exemple pour l'Europe : Spiegel, Bild, Telegraph, The Economist, Le Monde, Vogue, l'Express etc.) qui ont toutes leurs propres chartes et processus éditoriaux professionnels et rigoureux. Nous n'avons volontairement qu'une liste relativement restreinte de partenaires médias à travers l'Europe.
- Ces partenaires ont déjà mis en place leurs propres processus de production dans leurs salles de rédaction ou leurs studios, mais doivent également se conformer à nos lignes

⁶ Cf. note 3.

directrices en matière de contenu, qui précisent clairement que les informations inexactes et l'usurpation d'identité ne sont pas autorisées sur Snapchat.

- Nos partenaires médias acceptent aussi des lignes directrices supplémentaires (règles communautaires), comprenant l'obligation de publier un contenu correctement vérifié et exact.

Nous avons volontairement créé Découvrir comme une plateforme fermée, afin de garantir cet environnement et faire en sorte que nos utilisateurs sachent qu'ils peuvent se fier au contenu qu'ils vont y trouver.

Lorsqu'il s'agit de contenu provenant d'influenceurs très suivis, nous disposons d'équipes de modérateurs qui surveillent activement le respect de nos règles communautaires et nos conditions d'utilisation. Ces dernières précisent clairement que l'usurpation d'identité, les discours haineux et les contenus susceptibles d'encourager la violence et le terrorisme constituent des violations et entraîneront des sanctions, allant jusqu'à la suppression d'un compte utilisateur ou, le cas échéant, un signalement aux autorités compétentes.

Les équipes Éditoriales de Découvrir et de modération d'influenceur

L'équipe éditoriale de Découvrir est composée de personnes disposant d'une grande expérience de la curation de contenu dans les médias, notamment numériques, et ayant reçu une formation en matière de création de contenu, de curation, de production vidéo, de montage vidéo, de rédaction et de collecte d'informations. Notre équipe de modération de contenu est composée de personnes compétentes qui évaluent rigoureusement les contenus publiés sur la base des dispositions légales, de nos politiques et lignes directrices. Nos budgets alloués à la modération de contenu sont pensés pour nous permettre de répondre aux exigences applicables et à l'échelle concernée.

Partenaires media

Nous travaillons uniquement avec des partenaires média de confiance pour fournir des informations sur notre plateforme, et nous disposons de lignes directrices dédiées pour nos éditeurs partenaires sur Discover, qui prévoient l'interdiction des fausses informations. Notre plateforme de contenu n'est pas ouverte - nous pensons que notre communauté doit savoir d'où proviennent ses informations et nous sélectionnons les contenus provenant d'éditeurs de médias de confiance. Si un partenaire média publiait des fausses nouvelles, nous supprimerions rapidement ce contenu. Si un problème de publication de fausses nouvelles par un partenaire média perdurait, nous interrompons notre partenariat. Comme indiqué précédemment, nous ne sommes pas une plateforme ouverte, et nous n'avons en conséquence pas rencontré de problème lié à des fausses nouvelles avec un partenaire média.

Snap a toujours adopté une approche proactive pour protéger sa plateforme des fausses informations, et nous continuons à être vigilants sachant que le périmètre, le nombre et la viralité des fausses informations augmentent.

De manière générale, nos lignes directrices interdisent aux *Snapchatters* et à nos partenaires de partager du contenu trompeur ou de diffuser délibérément de fausses informations susceptibles de causer un préjudice. En interne, nous disposons de politiques et de procédures pour traiter les fausses informations et nous veillons à les mettre en œuvre de manière appropriée. En outre, Snapchat est conçu différemment des autres plateformes, et tant notre architecture que notre stratégie en termes de contenu, qui consiste à privilégier la collecte d'informations auprès d'entreprises média fiables, contribuent à protéger notre communauté contre les fausses informations.

3.12 Si aucun dispositif interne ou externe œuvrant au *fact-checking* n'est mis en œuvre, l'opérateur en précisera les raisons.

Non applicable.

Questions complémentaires

3.13 Remplir le tableau suivant.

Données chiffrées et indicateurs (France, 2021)	
Nombre de contenus soumis à la vérification de <i>fact-checkers</i>	Non applicable.
Nombre de contenus ayant fait l'objet d'une intervention de <i>fact-checkers</i> sur le service (décliné par types d'interventions possibles)	Non applicable.
Délai moyen de traitement d'un contenu par un <i>fact-checker</i>	Non applicable.
Nombre de contenus identifiés comme comportant une fausse information (ou notion voisine adoptée) par des <i>fact-checkers</i>	Non applicable.
Nombre de recours effectués sur des contenus ayant fait l'objet de mesures à la suite de l'intervention de <i>fact-checkers</i>	Non applicable.
Budget consacré aux initiatives ou aux actions concourant à la vérification de l'information sur le service (€)	Non applicable.

3.14 Impact de la crise sanitaire sur la promotion des contenus issus d'entreprises et d'agences de presse et de services de communication audiovisuelle et mesures spécifiques prises dans ce contexte en 2021 :

En réponse à la pandémie, nous continuons à être particulièrement vigilants quant à notre responsabilité de fournir à notre communauté des nouvelles et des informations fondées sur des faits. Comme indiqué ci-dessus, notre plateforme de contenu Discover est modérée et nous ne travaillons en étroite collaboration qu'avec un ensemble restreint de partenaires et de gouvernements, dont certains des organismes de presse les plus fiables au monde.

Notre priorité est de nous assurer que Snapchat fournit des informations de la plus haute qualité sur la covid-19 et que nous n'offrons pas un fil d'actualité ouvert où des éditeurs ou des individus non approuvés ont la possibilité de diffuser des informations erronées. Nos directives interdisent aux utilisateurs de Snapchat et à nos partenaires éditeurs de partager des contenus trompeurs ou diffusant délibérément de fausses informations.

Dans cette optique, nous avons créé un espace et un canal dédiés dans Discover, nous nous sommes associés à l'Organisation mondiale de la santé (OMS) pour faciliter les questions-réponses entre les utilisateurs et les experts sur le coronavirus et nous avons mis en place des canaux Discover provenant des communautés les plus établies dans le monde.

En outre, nous nous efforçons de fournir à notre communauté des informations factuelles et actualisées, ainsi que des ressources pour leur sécurité et leur santé mentale. Ainsi en 2021, Snap a poursuivi la campagne lancée en 2020, en produisant plusieurs Filtres et *Lens* présentant des conseils de sécurité et les meilleures pratiques, qui ont rapidement atteint des centaines de millions de personnes dans le monde entier.

3.15 **Actions futures et nouvelles fonctionnalités envisagées à l'avenir :**

N/A.

4 Lutte contre les comptes propageant massivement de fausses informations

Les questions suivantes s'adressent aux opérateurs qui auraient mis en œuvre des mesures relatives à la lutte contre les comptes propageant massivement de fausses informations au titre des mesures complémentaires prévues par l'article 11 de la loi du 22 décembre 2018.

Définitions et pratiques de diffusion massive de fausses informations

4.1 Appréhension par l'opérateur de la notion de comptes propageant massivement de fausses informations, telle que mentionnée à l'article 11 de la loi du 22 décembre 2018 :

Comme nous l'avons expliqué dans notre précédente contribution, la structure et le modèle de notre plateforme assure un rempart contre la diffusion massive de fausses informations. La viralité des contenus est limitée sur Snapchat.

A ce titre, l'application s'ouvre directement sur l'appareil photo, et non sur un flux de contenu. Sur l'ensemble de la plateforme, nous limitons la viralité - pas de partage possible, pas de critères de vanité -, ce qui restreint toute tentation ou incitations à publier des contenus qui font appel aux pires instincts humains, et limite donc les préoccupations liées à la diffusion de contenus dangereux, y compris les fausses informations.

Snap a tenu des réunions avec les fonctionnaires concernés de la Commission européenne (DG JUST, DG HOME) et du gouvernement britannique (Foreign & Commonwealth Office) pour expliquer en détail notre modèle de fonctionnement et les risques posés par la désinformation et les risques d'ingérences étrangères.

La Commission européenne et le FCO britannique ont tous deux été convaincus que Snapchat n'était pas une plateforme adéquate pour la diffusion de fausses informations, et qu'il n'y a pas eu de preuve à ce jour de tentatives de la part de gouvernements étrangers de manipuler la plateforme pour diffuser de la désinformation. Pour ce dernier point, au début de 2018, nous avons mené un audit interne approfondi à la recherche de contenu associé au gouvernement russe ; il n'y en avait pas.

Comme expliqué plus en détail ci-dessus, notre application a été conçue de manière réfléchie pour éviter la diffusion massive de fausses informations. Nos politiques, y compris nos directives communautaires, interdisent la diffusion délibérée de fausses informations. Outre les solides protections de conception mentionnées dans plusieurs réponses précédentes, les mesures que nous prenons en cas de fausses informations comprennent la suppression rapide du contenu incriminé lorsqu'il est détecté et, dans certains cas, la suppression du compte responsable.

Comptes populaires

Nous appelons « comptes populaires » les comptes qui publient un grand nombre de *Stories* publiques (*Stories* rendues publiques lorsque le paramètre « Ma Story » est réglé pour être visible par tous) qui sont très suivies. Nous appliquons nos pratiques de modérations habituelles aux comptes populaires et avons également mis en place des mesures anti-spam pour ces comptes afin de détecter et de traiter une éventuelle activité anormale.

Au-delà de nos comptes populaires, nous investissons des ressources significatives pour éradiquer les comportements automatisés de masse sur Snapchat. En combinant algorithmes propriétaires et méthode heuristique, nous mettons en œuvre des moyens ciblés pour supprimer activement les comptes présentant une activité organisée et non authentique. Les équipes et les processus concernés sont dotés des ressources nécessaires nous permettant de nous adapter rapidement à toute nouvelle tendance, en prenant en compte le retour d'expérience des situations gérées tant par la modération humaine que par celle résultant de l'utilisation d'outils automatisés.

4.2 Décliner les **différentes catégories de pratiques** de diffusion massive de fausses informations identifiées sur le service, en en donnant une **définition** et en les caractérisant, au regard notamment de leur **origine** (types d'acteur), **de leurs moyens de diffusion** (exclusivement algorithmiques, exclusivement humains ou hybrides) et de la **nature de l'objectif poursuivi par leurs émetteurs** (exclusivement économique et/ou politique, ou autre) :

Non applicable, en raison de la conception de la plateforme Snapchat et de l'absence de diffusion de masse, comme expliqué ci-dessus.

4.3 Pratiques détectées sur le service consistant à **générer un trafic ou un engagement artificiel massif** pour propager fortement un contenu en augmentant sa viralité (achat/vente de faux abonnés, fraude aux clics...) :

Mesures prises à leur encontre :

Non applicable, en raison de la conception de la plateforme Snapchat et de l'absence de diffusion de masse, comme expliqué ci-dessus.

4.4 Fournir la liste ou des exemples :

- **d'opérations de propagation de fausses informations soutenues par des entités étatiques ou institutionnelles** identifiées et mesures prises à leur égard :
- de **coopération** avec les autorités et/ou d'autres opérateurs dans le but de mieux détecter et de prévenir ces phénomènes :
- **de pratiques individuelles ou à l'échelle d'un groupe d'utilisateurs** de comportements trompeurs et abusifs massifs (ex. : multidiffusion, publication à répétition) identifiées et les moyens mis en œuvre pour lutter contre :

Non applicable, en raison du mode de conception de la plateforme Snapchat et de l'absence de diffusion de masse, comme indiqué ci-dessus.

Moyens de lutte contre la diffusion massive de fausses informations et informations aux utilisateurs

4.5 Pour chacune des catégories identifiées en réponse à la question 4.2, indiquer :

- les **moyens** (humains et technologiques) pour les **détecter** et y **faire obstacle** :
- les **mesures** prises et critères déclenchant l'application de chacune d'entre elles :

Non applicable, en raison de la conception de la plateforme Snapchat et de l'absence de diffusion de masse, comme expliqué ci-dessus.

4.6 Information des utilisateurs sur les pratiques liées à l'utilisation de comptes propageant massivement de fausses informations susceptibles d'entraîner une intervention et les mesures susceptibles d'être prises : NON / OUI (préciser)

Non applicable, en raison de la conception de la plateforme Snapchat et de l'absence de diffusion de masse, comme expliqué ci-dessus.

4.7 Commenter la préconisation du CSA aux opérateurs, dans son précédent bilan sur la lutte contre la manipulation de l'information, **d'informer davantage les utilisateurs sur les pratiques coordonnées d'influence et les risques qui en découlent.**

Non applicable, en raison de la conception de la plateforme Snapchat et de l'absence de diffusion de masse, comme expliqué ci-dessus.

4.8 Décrire les mesures et moyens mis en œuvre pour **la détection de bots et les outils développés** pour lutter contre ceux utilisés à des fins de manipulation de l'information. L'opérateur précisera notamment :

- s'il possède une **définition** du terme « *bot* » et quels en sont les usages qu'il autorise et proscrit sur son service :
- les **estimations du trafic généré** par ces derniers sur son service :
- s'il permet à des **applications tierces** d'utiliser certaines fonctionnalités des comptes des utilisateurs (ex. : publication et partage de contenus, demande de connexion entre utilisateurs) et comment il s'assure que ces pratiques ne soient pas utilisées à des fins de manipulation de l'information.

Non applicable, en raison de la conception de la plateforme Snapchat et de l'absence de diffusion de masse, comme expliqué ci-dessus.

4.9 S'agissant des **hypertrucages (« deepfakes »)** :

- politique du service (si autorisation ou interdiction sous conditions, lesquelles ?) :

Nous interdisons aux utilisateurs de se faire passer pour quelqu'un (ou quelque chose) qu'ils ne sont pas, ou de tenter de tromper les gens sur leur identité. Cela inclut l'usurpation de l'identité des amis, de célébrités, de marques ou d'autres organisations.

Nous nous appuyons sur nos sources de signalement pour identifier les fausses informations, y compris les allégations et les tendances concernant les *deep fakes* (manipulation d'images et de contenu) pour tenir nos équipes informées et agir en conséquence en réponse au contenu qui est porté à notre attention. En cas de violation avérée, nous supprimons directement le contenu, et envoyons un message d'avertissement à l'utilisateur.

Questions complémentaires

4.10 Remplir le tableau suivant.

Données chiffrées et indicateurs (France, 2021)	
Comptes propageant massivement de fausses informations (ou notion voisine adoptée) :	
- nombre de comptes détectés	N/A.
- part sur le nombre total de comptes sur le service	
Nombre de ces comptes détectés :	
- par des outils algorithmiques	N/A.
- par des signalements d'utilisateurs	
Nombre de comptes détectés qui ont fait l'objet de mesures (en déclinant par types de mesures)	N/A.
Montant des rémunérations tirées par ces comptes	N/A.
Nombre de contenus identifiés comme comportant une fausse information (ou notion voisine adoptée) ayant bénéficié d'une pratique consistant à générer un trafic ou un engagement artificiel massif	N/A.
Faux compte automatisés (<i>bots</i>) :	
- nombre de signalements	N/A.
- nombre de mesures prises	
Nombre de <i>deepfakes</i> sur lesquels des mesures ont été prises	N/A.
Moyens financiers mis en œuvre pour lutter contre les comptes propageant massivement de fausses informations (ou notion voisine adoptée) (en €)	N/A.

4.11 Impact de la crise sanitaire la lutte contre les comptes propageant massivement de fausses informations (ou notion voisine adoptée) (notamment, types de pratiques accrues dans ce contexte) et mesures spécifiques prises dans ce contexte en 2021 :

N/A.

4.12 Actions futures et nouvelles fonctionnalités envisagées à l'avenir :

N/A.

5. Mesures de lutte contre les fausses informations en matière de communications commerciales et de promotion des contenus d'information se rattachant à un débat d'intérêt général

5. a. Lutte contre la manipulation de l'information en lien avec des communications commerciales

L'Arcom estime que parmi les mesures complémentaires que les opérateurs de plateforme peuvent prendre au titre de l'article 11 de la loi du 22 décembre 2018 pour lutter contre la diffusion de fausses informations susceptibles de troubler l'ordre public ou d'altérer la sincérité des scrutins figure la suppression des flux financiers générés, directement ou non, par des contenus propageant de fausses informations. Ces flux financiers peuvent en effet être un facteur, d'une part, d'augmentation de la viralité des fausses informations et, d'autre part, d'incitation à les diffuser. Si ces phénomènes dépassent le seul cadre des services de plateformes en ligne, ces derniers peuvent y participer.

Catégorisation des communications commerciales

5.1. Qualifier les types de communications commerciales présentes sur le service à partir des catégories **d'annonces publicitaires, de contenus sponsorisés et de contenus en partenariat** proposées en annexe du questionnaire.

Si des pratiques commerciales sur le service n'entrent pas dans ce cadre, expliquer :

Chez Snapchat, nous travaillons avec des partenaires et des annonceurs de confiance pour permettre une variété de contenus publicitaires sur les parties publiques de notre plateforme. Ceux-ci sont disponibles sous plusieurs formes ;

- **Snap Ads** : Une seule image ou vidéo affichée en plein écran.
- **Annonces publicitaires** : Une publicité vidéo en plein écran qui n'est pas "skippable" pendant les 6 premières secondes.
- **Filtres** : Superpositions d'image que les utilisateurs peuvent choisir d'appliquer sur un Snap pour obtenir un effet artistique.
- **Lens** : Filtre de réalité augmentée qui permet aux utilisateurs de placer des objets et des personnages animés dans leurs Snaps.
- **Story Ads let** : Des affichages de marque dans la section Discover de Snapchat qui s'ouvrent sur une collection de 1 à 20 images ou vidéos.
- **Annonces de produits** : Formats publicitaires qui permettent aux utilisateurs de choisir parmi plusieurs produits, services ou expériences proposés par l'annonceur.

Annonces publicitaires et contenus sponsorisés : modalités de vente et informations apportées aux utilisateurs

5.2. Décrire les **modalités de vente des espaces publicitaires** sur le service, notamment :

- placement régi par des mécanismes programmatiques (enchères ouvertes ou privées, programmation garanti ou « *preferred deal* ») ou des accords de gré à gré :

Si mixte, indiquer la part de chaque procédé :

- modèle propriétaire ou traitements externalisés (indiquer les intermédiaires utilisés) :

Toutes les publicités publiées par nos annonceurs sont affichées comme des contenus sponsorisés. Ces derniers peuvent utiliser notre plateforme publicitaire libre-service, *Ads Manager*, ou travailler avec un partenaire en utilisant notre API pour diffuser des annonces sur notre plateforme. Les annonceurs enchérissent sur les impressions dans le cadre d'une vente

aux enchères privée. Aucune donnée n'est partagée avec les annonceurs pour leur permettre d'enchérir sur des utilisateurs spécifiques. Nos conditions complètes pour les annonceurs peuvent être consultées ici : <https://www.snap.com/en-US/terms/business-services>.

5.3. Snapchat décrira les modalités qui permettent aux entreprises et gestionnaires des lieux de figurer sur le produit « Map » : est-ce à leur initiative ? Cet affichage fait-il l'objet d'une contrepartie ?

Snap décide des lieux qui apparaissent sur la carte. Nous nous procurons ces emplacements auprès de partenaires tiers, comme Yext. Les annonceurs peuvent également demander à ce que leur lieu d'affaires soit ajouté à la Map, mais aucune compensation n'est requise pour que Snap inclut le lieu. Les annonceurs peuvent choisir de promouvoir leur emplacement sur la carte au moyen d'annonces standard, mais cela n'est pas nécessaire pour que Snap inclut un emplacement.

5.4. Possibilité pour les annonceurs de cibler un public en particulier :

Les annonceurs peuvent indiquer un souhait de ciblage particulier, que Snap tentera de faire correspondre à ses utilisateurs, ou cibler des publics prédéfinis identifiés par Snapchat en fonction de l'activité des utilisateurs sur la plateforme. Par exemple, si un annonceur choisit de cibler notre public de Fans de Sport, sa publicité apparaîtra auprès des utilisateurs ayant regardé beaucoup de contenus sportifs sur Snapchat.

Il convient de mentionner que pour certains annonceurs, nous imposons un ciblage obligatoire afin de protéger nos utilisateurs - comme par exemple pour les paris sportifs, ou la promotion de produits bancaires, nous imposons à l'annonceur de cibler uniquement les +18 ans.

Plus d'informations sur les audiences personnalisées et prédéfinies sur les liens attachés.

5.5. Possibilité de promouvoir des contenus diffusés en direct :

- possibilité d'**apposer une annonce publicitaire** sur de tels contenus :

Non, Snapchat ne propose pas de fonctionnalités en direct sur l'ensemble de sa plateforme, et pour l'ensemble de ses utilisateurs.

5.6. Existence d'une base de données relative aux communications commerciales diffusées sur le service

Snap ne dispose pas d'une base de données générale des contenus sponsorisés. Toutefois, nous disposons d'une bibliothèque publique listant les publicités à caractère politique et militant qui fournit une base de données sur les contenus d'informations sponsorisés relatifs à un sujet d'intérêt général. Cette base de données peut être filtrée spécifiquement pour les contenus destinés à la France. Comme Snapchat est essentiellement un service mobile, et que les données sont beaucoup plus volumineuses et complexes que ce que la plupart des appareils mobiles peuvent traiter, les utilisateurs n'ont pas accès à ces données à partir de l'application Snapchat. Elles peuvent cependant être téléchargées à tout moment sur notre site à l'adresse <https://www.snap.com/fr-FR/political-ads> .

5.7. Informations fournies à l'utilisateur sur les annonces publicitaires et contenus sponsorisés présents sur le service :

- **moyens de les distinguer des autres contenus** : Toutes les publicités payantes sont affichées explicitement comme des publicités.
- **identification de l'origine de l'annonce ou du contenu sponsorisé (et de l'annonceur, s'il diffère)** : Les annonceurs sont tenus d'inclure dans leurs publicités un marquage permettant d'identifier l'origine de la publicité.
- **indications sur les modalités de diffusion de chaque annonce ou contenu sponsorisé (existence et montant de contreparties financières, nombre de vues, population ciblée, manière dont le contenu a été généré (automatisée ou non), etc.)** : Nous ne fournissons pas ce service.
- **politiques publicitaires** : <https://snap.com/fr-FR/ad-policies>

5.8. Gestion des préférences publicitaires :

- **décrire les moyens d'informer l'utilisateur des critères pris en compte dans l'affichage des annonces publicitaires et contenus sponsorisés** : Les utilisateurs peuvent cliquer sur l'icône à trois points figurant sur chaque publicité et sélectionner l'icône « i » pour obtenir plus d'informations sur la raison pour laquelle ils voient la publicité et sur la façon de personnaliser les réglages utilisés pour afficher les publicités.
- **possibilité pour l'utilisateur de paramétrer ses préférences publicitaires** : Oui
- **possibilité de masquer des annonces publicitaires et contenus sponsorisés** : Oui

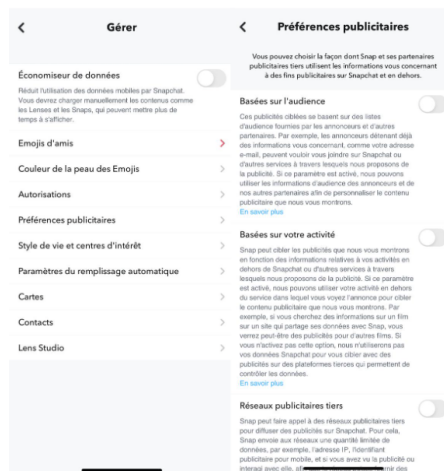


Illustration pour la question 5.8

5.9. Commenter les préconisations du CSA aux opérateurs, dans son précédent bilan sur la lutte contre la manipulation de l'information :

- de proposer des **politiques publicitaires en français** et comportant une **partie relative à la manipulation de l'information** :

Snapchat cible dans ses politiques publicitaires les pratiques frauduleuses et mensongères. Celles-ci sont disponibles directement en français à l'adresse suivante : <https://snap.com/fr-FR/ad-policies>

“Contenus mensongers

Nous attachons du prix à la lutte contre les publicités frauduleuses. La fraude comprend toute une série d'arnaques et de pratiques commerciales frauduleuses qui abusent de la confiance de la communauté ou incitent les utilisateurs à effectuer des achats ou à s'engager dans des

activités sous des prétextes fallacieux.

Sont prohibés :

- Les publicités fausses ou fallacieuses, notamment les allégations, offres, fonctionnalités ou pratiques commerciales mensongères ;
- Les contenus sponsorisés non autorisés ou non divulgués ;
- La promotion de biens ou de services frauduleux, notamment de documents ou de certificats contrefaits, ou de produits contrefaits ;
- La création ou le partage des contenus qui reproduisent l'apparence ou les fonctions des formats ou des fonctionnalités Snapchat ;
- Les annonces qui contiennent des invitations à l'action trompeuses, ou qui mènent à des pages de destination n'ayant aucun rapport avec la marque ou le contenu faisant l'objet de la publicité ;
- Le cloaking, la restriction de l'accès aux pages de destination ou la modification du contenu de l'URL après l'envoi dans le but d'échapper à toute vérification ;
- Les publicités qui encouragent les comportements malhonnêtes (par exemple, les publicités sur les fausses cartes d'identité, le plagiat, les services de rédaction) ;
- La non-livraison de marchandises, les retards d'expédition ou les contraintes de stock présentés de manière erronée.
- Vous pouvez aussi consulter les spécifications du secteur en matière de produits et de services financiers.”

- de mettre en place des **outils faciles d'accès et d'utilisation** permettant à l'utilisateur de **comprendre pourquoi il a été ciblé** dans une situation donnée et de **paramétrer ses préférences publicitaires** :

Snapchat met à disposition des outils disponibles directement depuis l'application, et ce depuis le contenu concerné permettant à l'utilisateur de comprendre l'origine de cette publicité, et paramétrer selon ses préférences :



Mesures contre la manipulation de l'information dans les communications commerciales

5.10. Moyens de **détection des annonces publicitaires porteuses de fausses informations** :

5.11. Décrire le système **d'identification et de vérification** :

- **des annonceurs** (annonces publicitaires et contenus sponsorisés) :
- **des communications commerciales** :
- **des liens de redirection** vers des sites tiers :

Réponses aux questions 5.10 et 5.11

Toutes les communications commerciales de nos annonceurs sont soumises à un système de vérification avant publication. En matière de détection et de lutte contre la publicité pour des sites web considérés comme frauduleux, nous recourons à la fois à l'examen humain et à l'automatisation pour nous assurer de la conformité de ces derniers à nos politiques publicitaires (*Advertising Policies*). Celles-ci peuvent être consultées à travers le lien suivant : <https://snap.com/fr-FR/ad-policies>.

En termes d'identification, l'ensemble des publicités sont signalées comme telles sur notre plateforme (voir l'illustration). Le nom de l'annonceur est également directement affiché sur le contenu, et un bouton "call to action" est proposé en bas de la publication afin de rediriger l'utilisateur vers un site tiers.

5.12. Mesures de sensibilisation à l'égard des vidéastes et influenceurs en ce qui concerne les communications commerciales porteuses de fausses informations :

La certification sur Snapchat fait l'objet d'une étude approfondie de la part de notre équipe. Comme expliqué sur notre [site support](#), le statut de *Snap Star* n'est pas un droit, quelle que soit la taille de l'audience du compte. Snap peut accepter ou refuser une demande de compte *Snap Star* à sa seule et entière discrétion. Le passage au statut de *Snap Star* ne modifie pas nos Conditions d'utilisation du service et ne confère aucun droit supplémentaire au titulaire du compte.

A cet effet, nous avons dédié un article sur notre site support afin d'explicitier les règles à respecter en tant qu'influenceurs sur notre plateforme. Nous nous réservons ainsi le droit de révoquer la vérification ou le statut de *Snap Star* en cas de violations de nos Conditions d'utilisation du service ou de nos Règles communautaires, de comportement dangereux mettant des tiers en danger ou bien encore de promotion de contenus, de comptes, d'applications ou de sites web (par des liens, des mentions ou autres) qui violeraient nos conditions d'utilisation du service ou nos règles communautaires.

Selon la gravité de la violation, nous appliquons une réponse graduée et proportionnée, pouvant aller du retrait du contenu, de la certification ou du compte. L'utilisateur concerné sera toujours averti de la décision, et des motifs, et pourra également faire appel de cette dernière.

Avant toute décision définitive, nous envoyons un message de sensibilisation à l'utilisateur afin de l'avertir des risques encourus. Aussi, quand nous nous apercevons d'un potentiel risque ou d'un phénomène important de pratiques commerciales trompeuses, nous sensibilisons et avertissons l'ensemble de notre communauté de *Snap stars* en leur fournissant les meilleures pratiques, avec un rappel des règles applicables sur Snapchat.

5.13. Moyens mis en place pour identifier les partenariats rémunérés qui ne respecteraient pas les conditions d'utilisation :

Nous nous appuyons essentiellement sur le signalement des utilisateurs.

5.14. Mesures spécifiques prises en 2021 sur des contenus porteurs de fausses informations ayant bénéficié d'un partenariat rémunéré entre l'utilisateur ayant posté le contenu et un tiers :

Notre équipe de modération évalue toutes les allégations de fausses informations signalées sur notre plateforme. La présence ou l'absence de partenariats rémunérés entre un utilisateur et une tierce partie n'affecte pas le résultat ou la décision d'application. S'il s'avère qu'un contenu contient de fausses informations qui violent nos conditions d'utilisation ou nos directives communautaires, l'équipe de modérateurs prendra des mesures contre ce contenu ou cet utilisateur, selon le cas.

5.15. Si le dispositif de signalement des annonces publicitaires et /ou des contenus sponsorisés diffère de celui présenté en partie 1, préciser si une communication commerciale peut être signalée pour « fausse information » (ou équivalent), les modalités du signalement

(connexion à un compte nécessaire, démarche à accomplir...) et comment l'opérateur **interagit avec les utilisateurs** concernés.

Le dispositif de signalement ne diffère pas de celui présenté en partie 1. La communauté Snapchat peut signaler dans l'application ou sur le site d'assistance les publicités qui, selon elle, contiennent des informations trompeuses, fausses ou mensongères. Les publicités signalées sont examinées par des modérateurs humains qui déterminent si les allégations trompeuses sont confirmées et rejetées.

5.16. S'il est possible d'**acheter ou de promouvoir un mot-clé** (ex. : *hashtag* (mot-dièse) ou autre) sur le service, indiquer si certains mots-clés sensibles sont interdits à la vente, lesquels et si les utilisateurs sont informés de cette interdiction :

Il n'est pas possible de cibler nos utilisateurs en fonction de mots-clés ou hashtags - nous ne proposons pas ce type de solution à nos annonceurs.

5.17. Coopérations avec d'autres opérateurs ou organismes externes afin de détecter et de signaler les communications commerciales porteuses de fausses informations :

Non applicable.

5.18. Protocoles de démonétisation, de restriction ou de suppression de comptes en raison de la diffusion de fausses informations via une communication commerciale :

Il est interdit aux publicités et au contenu auquel elles sont liées de contenir des affirmations ou des informations trompeuses, fausses ou mensongères. Si des annonces liées à de fausses informations sont découvertes pendant l'examen, elles sont rejetées. Snap peut suspendre ou résilier les comptes liés aux entreprises ou aux particuliers qui violent nos politiques publicitaires.

Mesures en faveur de la sécurité des marques

5.19. Détailler les **mesures de sécurité des marques** mises en place et expliquer comment elles ont été conçues, notamment :

- **protocoles d'information des annonceurs** lorsque leurs contenus se sont trouvés accolés à des fausses informations :

Il est interdit aux annonces et au contenu auquel elles sont liées de contenir des affirmations ou des informations trompeuses, fausses ou mensongères. Si des annonces liées à de fausses informations sont découvertes lors de l'examen, elles sont rejetées. Par ailleurs, la plateforme ne permettant pas la viralité des contenus, le contenu d'un annonceur ne peut pas être repris pour diffuser une fausse information

- **mécanismes de compensation ou de recours :**

Les annonceurs peuvent contacter leur représentant Snap (s'ils en ont un). Le représentant soumettra alors une demande en interne à l'équipe de révision des publicités pour une enquête plus approfondie. En outre, tout annonceur peut contacter l'équipe d'assistance pour les publicités numériques, où les tickets sont triés et acheminés vers l'équipe de vérification des publicités si la publicité mérite un second examen. Nous modifierons le statut de la publicité, si elle a été rejetée à tort.

- modalités de participation de l'opérateur dans tout **mécanisme de coopération, labélisation ou standardisation** avec les marques, l'interprofession publicitaire ou des intermédiaires spécialisés dans « l'*ad verification* » :

Snap ne participe pas actuellement à ces efforts.

5.20. Procédé de **suivi des flux monétaires de la chaîne programmatique** (ex. : le SCID (*Shared Campaign Identifier*)) : NON / OUI (préciser)

Note Snap : nous ne sommes pas sûrs de ce à quoi cette question se réfère, merci de la clarifier.

Questions complémentaires

5.21. Remplir le tableau suivant.

Données chiffrées et indicateurs (France, 2021)		
	Annonces publicitaires	Contenus sponsorisés
Nombre total de communications commerciales		Notre politique en matière de contenu commercial interdit le contenu commercial qui est faux ou trompeur, y compris les allégations, offres, fonctionnalités ou activités trompeuses.
Nombre de communications commerciales porteuses de fausses informations (ou notion voisine)		Notre politique en matière de contenu commercial interdit le contenu commercial qui est faux ou trompeur, y compris les allégations, offres, fonctionnalités ou activités trompeuses.
Volume de communications commerciales accolées à de fausses informations (ou notion voisine)	N/A. La plateforme ne permettant pas la viralité des contenus, le contenu d'un annonceur ne peut pas être repris pour diffuser une fausse information.	Notre politique en matière de contenu commercial interdit le contenu commercial qui est faux ou trompeur, y compris les allégations, offres, fonctionnalités ou activités trompeuses.
Revenus générés par les communications commerciales porteuses de fausses informations (ou notion voisine) (en €) Dont : perçus par la plateforme	N/A. Nos politiques publicitaires interdisent les publicités fausses ou trompeuses, y compris les allégations, offres, fonctionnalités ou pratiques commerciales trompeuses.	Notre politique en matière de contenu commercial interdit le contenu commercial qui est faux ou trompeur, y compris les allégations, offres, fonctionnalités ou activités trompeuses.
Revenus générés par les communications commerciales accolées aux fausses informations (ou notion voisine) (en €) Dont : perçus par la plateforme	N/A. Nos politiques publicitaires interdisent les publicités fausses ou trompeuses, y compris les	Notre politique en matière de contenu commercial interdit le contenu commercial qui

	allégations, offres, fonctionnalités ou pratiques commerciales trompeuses.	est faux ou trompeur, y compris les allégations, offres, fonctionnalités ou activités trompeuses.
Nombre d'impressions générées, en moyenne, par les communications porteuses de fausses informations (ou notion voisine) (précisez l'écart-type)	N/A.Nos politiques publicitaires interdisent les publicités fausses ou trompeuses, y compris les allégations, offres, fonctionnalités ou pratiques commerciales trompeuses.	Notre politique en matière de contenu commercial interdit le contenu commercial qui est faux ou trompeur, y compris les allégations, offres, fonctionnalités ou activités trompeuses.
Durée moyenne d'apparition d'une communication commerciale porteuse de fausses informations ou redirigeant vers un site porteur de fausses informations (ou notion voisine)	N/A.Nos politiques publicitaires interdisent les publicités fausses ou trompeuses, y compris les allégations, offres, fonctionnalités ou pratiques commerciales trompeuses.	Notre politique en matière de contenu commercial interdit le contenu commercial qui est faux ou trompeur, y compris les allégations, offres, fonctionnalités ou activités trompeuses.
Nombre d'utilisateurs ayant été redirigés hors de la plateforme par une communication commerciale comportant une fausse information (ou notion voisine)	N/A.Nos politiques publicitaires interdisent les publicités fausses ou trompeuses, y compris les allégations, offres, fonctionnalités ou pratiques commerciales trompeuses.	Notre politique en matière de contenu commercial interdit le contenu commercial qui est faux ou trompeur, y compris les allégations, offres, fonctionnalités ou activités trompeuses.

5.22. Impact de la crise sanitaire la lutte contre la manipulation de l'information en lien avec des communications commerciales et mesures spécifiques prises dans ce contexte en 2021 :

Comme indiqué lors de nos précédentes réponses, la crise sanitaire de la covid-19 a eu un impact sur notre plateforme et ce, non pas en raison des fausses informations à ce sujet, mais en raison de la lutte contre la fraude aux pass sanitaires.



5.23. Actions futures et nouvelles fonctionnalités envisagées à l'avenir :

N/A.

5.b Contenus promus d'information se rattachant à un débat d'intérêt général

Les questions suivantes s'adressent aux opérateurs qui auraient mis en œuvre des mesures relatives à l'information des utilisateurs sur la nature, l'origine et les modalités de diffusion des contenus d'information se rattachant à un débat d'intérêt général et à l'information des utilisateurs sur l'identité de la personne physique ou la raison sociale, le siège social et l'objet social des personnes morales leur versant des rémunérations en contrepartie de la promotion de tels

contenus au titre des mesures complémentaires prévues par l'article 11 de la loi du 22 décembre 2018.

5.24. Définition des contenus d'information se rattachant à un débat d'intérêt général
adoptée par l'opérateur :

Snap n'a pas de définition spécifique de "contenu d'information se rattachant à un débat d'intérêt général", bien que nous considérons que le sens ordinaire de ces mots englobe un très large éventail de contenus. Dans ce cadre, il pourrait être pertinent que le régulateur précise cette définition afin de disposer d'une interprétation harmonisée du sujet.

Cette catégorie de publicités entre généralement dans le cadre de nos politiques de publicité politique et de plaidoyer (accessibles à l'adresse <https://www.snap.com/en-US/ad-policies/political/>), qui couvrent :

- Les publicités liées aux élections comprennent les publicités concernant les candidats ou les partis à des fonctions publiques, les mesures de vote ou les référendums, les comités d'action politique et les publicités qui incitent les gens à voter ou à s'inscrire sur les listes électorales ;
- Les publicités thématiques ou de plaidoyer sont des publicités concernant des questions ou des organisations qui font l'objet d'un débat au niveau local, national ou mondial, ou qui ont une importance publique. Elles comprennent les publicités sur l'avortement, l'immigration, l'environnement, l'éducation, la discrimination et les armes à feu.

5.25. Moyens lui permettant d'identifier les contenus d'information se rattachant à un débat d'intérêt général promus sur le service :

Ces contenus sont affichés explicitement comme des publicités avec une mention "sponsorisé". Comme les autres publicités commerciales, elles font l'objet d'une vérification avant publication de la part de nos équipes.

5.26. Existence d'une base de données spécifique à ceux de ces contenus ayant fait l'objet d'une promotion :

Snap dispose d'une bibliothèque publique de publicités politiques et de plaidoyer, comme indiqué au point 5.6 ci-dessus, qui fournit une base de données de contenus d'information sponsorisés relatifs à un sujet d'intérêt général.

Cette base de données peut être filtrée spécifiquement pour le contenu ciblant la France. Étant donné que Snapchat est un service uniquement mobile et que les données sont considérablement plus grandes et plus complexes que ce que la plupart des appareils mobiles sont conçus pour traiter, les utilisateurs n'ont pas accès aux données depuis l'application Snapchat. Toutefois, elles peuvent être téléchargées à tout moment à l'adresse <https://www.snap.com/fr-FR/political-ads>.

5.27. Traitement spécifique réservé à ce type de communications commerciales :

- **système de vérification particulier des annonceurs** en faisant la promotion :

Comme indiqué dans nos politiques publicitaires, les publicités à caractère politique diffusées par Snap ne peuvent pas être payées directement ou indirectement par des entités ou des ressortissants étrangers non résidents, c'est-à-dire par des personnes ou entités qui ne résident pas dans le pays où sera diffusée la publicité. Snap vérifiera ces publicités au cas par cas.

- **restrictions ou règles particulières** sur ces contenus (ex : restriction de ciblage) :

La publicité politique doit être conforme à toutes les lois et réglementations applicables, y compris toutes les lois électorales nationales, la loi sur les droits d'auteur, la loi sur la diffamation et (le cas échéant) les réglementations de la Commission électorale fédérale et les lois et réglementations étatiques ou locales. Le respect de ces lois et réglementations relève de la seule responsabilité de l'annonceur.

Toute publicité politique doit inclure un message "payé par" dans l'annonce, suivi du nom de la personne ou de l'entité qui paie. Snap peut également exiger une mention "payé par" sur le contenu des annonces qui renvoient à un contenu politique, sur le contenu des annonces pour des marchandises politiques, ou dans d'autres cas, à la seule discrétion de Snap.

- **procédure de signalement** spécifique pour de tels contenus en infraction :

La communauté Snapchat peut signaler dans l'application ou sur le site d'assistance les publicités qui, selon elle, contiennent des informations trompeuses, fausses ou mensongères. Les publicités signalées sont examinées par des modérateurs humains qui déterminent si les allégations trompeuses sont confirmées et rejetées.

- **information des utilisateurs sur** (si oui, préciser où elle figure) :

- o **la nature, l'origine et les modalités de diffusion de tels contenus** :
- o **l'identité** de la personne physique ou **la raison sociale, le siège social et l'objet social** des personnes morales leur versant des rémunérations en contrepartie de la promotion de tels contenus :

L'ensemble de ces informations sont présentées dans notre bibliothèque politique, disponible à l'adresse suivante : <https://www.snap.com/fr-FR/political-ads>

5.28. Collaborations (opérateurs, société civile, ONG, *think tanks*, institutions...) pour œuvrer à une plus grande harmonisation des approches retenues en la matière :

N/A.

5.29. Remplir le tableau suivant.

Données chiffrées et indicateurs (France, 2021)	
Nombre de contenus d'information se rattachant à un débat d'intérêt général promu sur le service	
Nombre d'impressions de ces contenus	
Nombre d'annonceurs de ces contenus	

5.30. Actions futures et nouvelles fonctionnalités envisagées à l'avenir :

N/A.

6. Favoriser l'éducation aux médias et à l'information

Les questions suivantes s'adressent aux opérateurs qui auraient mis en œuvre des mesures relatives à l'éducation aux médias et à l'information (EMI) au titre des mesures complémentaires prévues par l'article 11 de la loi du 22 décembre 2018.

Actions d'éducation aux médias et à l'information (EMI)

- 6.1 **Actions de sensibilisation des utilisateurs à une utilisation maîtrisée du service** (ex : encart d'information au moment de la publication d'un message, rappel des règles de savoir-vivre en ligne, etc.), menées par l'opérateur seul ou en partenariat :
- 6.2 **Ressources et fonctionnalités développées** permettant aux utilisateurs d'identifier la fiabilité des informations et sources d'informations qui sont proposées sur le service :

Snapchat tient à promouvoir et assurer sur son espace numérique une expérience utilisateur unique, amusante et sereine. Snap a à cœur la protection et le bien-être de l'ensemble de ses utilisateurs, notamment des plus jeunes. Nous tenons donc à donner toutes les clés possibles, tant aux adolescents qu'aux parents, pour faire de la plateforme un endroit sûr et accueillant.

A cet effet, nous avons créé un centre de sécurité dédié accessible sur le site snapchat.com ou directement depuis l'application. Celui-ci contient des informations et des conseils pour utiliser Snapchat en toute sécurité, et informe les utilisateurs sur les techniques et l'importance du signalement afin de les accompagner et leur proposer une expérience agréable sur la plateforme Snapchat (<https://www.snap.com/fr-FR/safety/safety-center>). Ce portail propose également un guide parental aux parents, au sein duquel ils peuvent trouver des suggestions de conversation avec leur adolescent sur l'utilisation de Snapchat, notamment sur l'intégration de la communauté Snapchat et l'identification d'un problème.



Illustration du guide parental

En parallèle, nous avons lancé en 2021 la version française de notre portail de prévention Here For You. Cet outil fournit des ressources et du contenu localisés provenant d'organisations à but non lucratif qui s'affiche dans les résultats dès lors que les utilisateurs font des recherches sur des sujets liés à une situation de crise. Les sujets concernés incluent entre autres la dépression, l'anxiété, le deuil, le harcèlement, l'acceptation de soi, la communauté LGBTQ, la santé mentale.

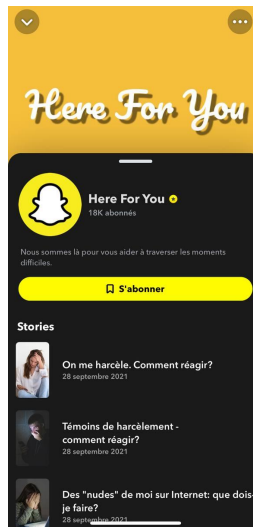


Illustration du portail de prévention Here For You - France

Il convient également de mentionner qu'à travers la structure et la conception de plateforme, notre objectif est de proposer à l'ensemble de nos utilisateurs des contenus d'information provenant de médias vérifiés, certifiés et reconnus. A cet effet, nos *Snapchatters* sont assurés de connaître et reconnaître la provenance des informations proposées, et s'assurer que celles-ci sont authentiques.

A cet effet, nous avons poursuivi nos efforts en la matière avec notre partenaire 20minutes et son émission "Oh My Fake!" qui, à chaque édition, abordait une théorie complotiste ou une idée reçue afin de la déconstruire et rappeler au spectateur la nécessité de vérifier ses sources d'information.

6.3 Liste des **campagnes menées, accompagnées ou promues en matière d'EMI dans le domaine de la lutte contre la manipulation de l'information en ligne** : remplir le tableau en annexe.

Nous travaillons en partenariat avec diverses ONG (telles que e-enfance) pour aider à éduquer nos utilisateurs, en particulier les plus jeunes et leurs parents/éducateurs. Notre centre de confidentialité informe également les utilisateurs sur nos pratiques en matière de protection de la vie privée et sur la manière dont leurs données peuvent être utilisées (<https://www.snap.com/fr-FR/privacy/privacy-center/>).

6.4 **Évaluation par l'opérateur de ses campagnes d'EMI : OUI / NON**

- **critères d'évaluation** (notamment ceux permettant d'évaluer les conséquences réelles de la campagne sur l'utilisation du service) :
- **actions recommandées** à la suite de ces évaluations :

Snapchat n'a pas encore pu procéder à l'évaluation des impacts de ses campagnes.

6.5 **Commenter les préconisations du CSA aux opérateurs**, dans son précédent bilan sur la lutte contre la manipulation de l'information, **d'évaluer les conséquences sur le comportement des utilisateurs des actions en matière d'EMI et d'en rendre compte au public et au régulateur.**

N/A.

6.6 Impact de la crise sanitaire sur la politique de l'opérateur en matière d'EMI et mesures spécifiques prises dans ce contexte en 2021 :

La crise sanitaire a eu un important impact sur notre société, notamment en termes de santé mentale chez les plus jeunes. A cet effet, nous avons lancé en 2021 la version française de notre portail de prévention *Here For You*. Cet outil fournit des ressources et du contenu localisés provenant d'organisations à but non lucratif qui s'affiche dans les résultats dès lors que les utilisateurs font des recherches sur des sujets liés à une situation de crise. Les sujets concernés incluent entre autres la dépression, l'anxiété, le deuil, le harcèlement, l'acceptation de soi, la communauté LGBTQ, la santé mentale.

6.7 Actions futures et nouvelles fonctionnalités envisagées à l'avenir dans la mise en place de ses campagnes d'EMI :

Nous comptons poursuivre nos efforts en matière de campagne d'EMI en alimentant notre portail *Here For You* avec de nouveaux contenus centrés sur la santé mentale et les bonnes pratiques à adopter en ligne, et investir notamment sur des contenus et campagnes hors-plateforme en matière de lutte contre la désinformation.

Études et relations avec la recherche

6.8 Études publiées par l'opérateur sur le fonctionnement du service (liste avec URL) :

6.9 Liste et description des initiatives indépendantes (journalistes, entreprises, chercheurs, etc.) que l'opérateur a soutenues ou auxquelles il a contribué, notamment avec des acteurs français :

- visant à mieux **cerner certains comportements et phénomènes liés à la manipulation de l'information** en ligne :
- visant à vérifier la qualité des **systèmes de recommandation et de modération** du service en termes d'explicabilité, de loyauté et d'équité et, le cas échéant, proposer des pistes d'amélioration.

N/A.

6.10 Relations avec la recherche dans le domaine de la lutte contre la manipulation de l'information

- **moyens mis à disposition** des chercheurs (ex. : concernant la propagation de fausses informations, l'opérateur donne-t-il accès à des données – notamment via une API –, et si oui, sous quelles conditions cet accès est-il possible ?) :
- **types de données** communiquées :

N/A.

6.11 Politique de partage de données en *open data* : NON / OUI (préciser les modalités)

- indiquer les limites à cet accès aux données et **les difficultés rencontrées** :

N/A.

6.12 **Actions futures** pour développer les relations avec le monde de la recherche :

N/A.

Annexe (question 6.3)

Nom de la campagne, de l'outil ou de l'action de soutien <i>(et URL le cas échéant)</i>	Description et objectifs	Partenariat	Nature de l'implication de l'opérateur	Public(s) visé(s)	Dates de début et de fin	Budget alloué	Mise en valeur auprès du public	Données sur les audiences touchées <i>(si aucune, préciser les difficultés rencontrées)</i>	Données sur l'impact de la campagne sur le comportement des publics visés et méthodes pour les mesurer <i>(si aucune, préciser les difficultés rencontrées)</i>