

Questionnaire aux opérateurs de plateformes en ligne soumis au titre III de la loi du 22 décembre 2018 relative à la lutte contre la manipulation de l'information

Remarques préliminaires de Snap

Remarque : Snap a présenté un nouveau produit à la fin de l'année 2020, dénommé *Spotlight* (<https://support.snapchat.com/fr-FR/a/spotlight>). Le présent rapport n'a pas encore de données précises à partager sur *Spotlight* car le produit a été lancé fin 2020 et l'ensemble du contenu a été modéré en interne avant d'être publié sur Snapchat. En revanche, nous avons souhaité montrer la façon dont les "fausses informations" peuvent être signalées sur la plateforme.

Questions liminaires : présentation du service

Nom du service : Snapchat

Nom et raison sociale de la société opératrice : Snap Inc.

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 : Snap France

Nombre de visiteurs uniques par mois de la plateforme en ligne, calculé sur la base de l'année précédant l'année d'exercice (2020) : 22 millions (source : Médiamétrie)

Modalités de mise à disposition du service auprès du public français : régies par les conditions disponibles sur www.snap.com/terms

Disponibilité :

- **sur internet (URL du site internet) :** n/a
- **via des applications (systèmes d'exploitation concernés) :** Applications Mobiles (iOS et Android).
- **via des assistants vocaux :** n/a

Langue du service sur le territoire français (par défaut) : Français

Outre la langue, le service disponible sur le territoire français constitue-t-il une version spécifique du service (pour le territoire français ou pour le public francophone) ? : Oui

Description du service : Service de messagerie éphémère et distribution de contenus en ligne.

- **Objet :** Le réseau des amis proches
- **Population(s) ciblée(s) (tranches d'âges, communautés d'intérêt...) :** Ce service est destiné à tous sans distinction d'âge, de sexe ou d'intérêts.
- **Types de contenus qui peuvent être publiés et consommés sur le service (vidéos, textes, sons, images, jeux vidéo...) :** Vidéos, texte, images, jeux vidéo, réalité augmentée.
- **Thématique(s) :** n/a

Modèle économique du service (types d'activités rémunératrices pour l'opérateur tirées de l'exploitation du service en France) : service gratuit financé par la publicité

Effectifs employés aux activités du service dans le monde et en France : 3863 au niveau mondial et 31 en France

Comptes de résultats pour l'activité du service en France en 2020 :

- **Chiffres d'affaires :** 67.811.000 euros
- **Produits d'exploitation :** 2.139.000 euros
- **Charges d'exploitation :** 65.672.000 euros
- **Résultat net :** 1.219.000 euros
- **Autres informations ou spécificités :** n/a

Questions générales

Définition des fausses informations

*L'opérateur indiquera s'il a adopté une **définition des fausses informations** et/ou de notions proches ; le cas échéant, il la fournira, il précisera à quelle fin (travail de modération, règles d'utilisation du service...) et si cette définition vaut globalement ou diffère selon les zones géographiques de mise à disposition du service.*

Les fausses informations nécessitent une analyse au cas par cas. De ce fait, il est presque impossible de fournir une définition unique de ce que constitue une fausse information.

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

Nous interdisons la diffusion de fausses informations malveillantes ou qui causent un préjudice, telles que 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 et les principes de la Loi de 1881 sur la liberté de la presse. 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.

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 nos modérateurs 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>.

Lutte contre les fausses informations sur le service

*Il indiquera si les fausses informations (ou notion proche) sont **interdites sur son service** et le cas échéant, où l'utilisateur peut trouver l'information relative à cette interdiction.*

L'opérateur est invité à indiquer au Conseil le **nombre de contenus identifiés comme comportant une fausse information sur son service en 2020**, toutes sources de détection confondues (interne, utilisateurs, fact-checkers...), leur part dans l'ensemble des contenus du service (en pourcentage) et **le nombre d'émetteurs et de réémetteurs de ces contenus, ainsi que le nombre de ces contenus et de ces émetteurs ayant fait l'objet d'une mesure en conséquence**, en déclinant en fonction du type de mesures adoptées. Il fera état des évolutions de ces chiffres entre 2019 et 2020.

En règle générale, l'opérateur a-t-il accru ou renforcé ses moyens de lutte contre les fausses informations en 2020 ? Il pourra résumer les **éventuelles évolutions de sa politique** en la matière et les enseignements qu'il en tire, et fournira des éléments quantitatifs témoignant de ces renforcements et de leurs effets. Il distinguera, en sus, les nouveaux dispositifs, humains ou techniques qu'il aurait mis en œuvre en 2020 et qui auraient contribué à cette lutte.

L'opérateur est invité à partager la **liste des 20 contenus relevant d'une communication commerciale (publicitaires, sponsorisés et en partenariat) et celle des 20 contenus n'en relevant pas détectés comme contenant une fausse information et qui ont suscité le plus d'interactions¹ de la part des utilisateurs** en France en 2020 sur le service qu'il opère. Il indiquera pourquoi ces contenus ont été qualifiés ainsi et quelques indicateurs relatifs à ces interactions. Les titres ou thématiques abordées par ces contenus, leur date de première apparition ainsi que les liens permettant au régulateur de les consulter seront précisés à chaque fois. L'opérateur pourra également indiquer si les tendances observées dans ces deux listes diffèrent sensiblement de ses observations lors des années précédentes.

En 2020, aucune communication commerciale contenant de fausses informations n'a été détectée sur Snapchat. Nous ne pouvons donc pas lister 20 contenus qui pourraient être considérés comme des communications commerciales incluant de fausses informations. Bien que du contenu ait pu être signalé dans l'application Snapchat comme « fausse information », le contenu en question ne correspondait pas réellement à la définition d'une fausse information, mais correspondait à des publicités contenant des allégations, des offres, des fonctionnalités ou des pratiques commerciales fausses, trompeuses ou déloyales (par exemple des arnaques pour vendre des iPhones).

En ce qui concerne le contenu non commercial, de nombreux signalements d'utilisateurs (peut-être la plupart) concernant de « fausses informations » étaient en réalité des plaintes pour usurpation d'identité, harcèlement, spam, allégations publicitaires exagérées, etc. Snap n'a rencontré que 4 cas vérifiés de contenu non commercial qualifié de fausse information.

Respect de la liberté d'expression

En se fondant sur des cas de fausses informations regardées comme telles par ses équipes au cours de l'année passée, l'opérateur est invité à exposer comment sa politique de modération s'articule avec la liberté d'expression et de communication sur son ou ses services.

La réponse à cette question est détaillée ci-après dans la réponse à la question 2.3.

Mesures futures

L'opérateur est invité à faire part des **actions qu'il envisage de mener à l'avenir afin de lutter contre la diffusion de fausses informations sur son service**.

La réponse à cette question ci-après dans la réponse à la question 1.7 et 3.10.

¹ On entend par « interaction » toute action effectuée par les utilisateurs vis-à-vis de ces contenus : consultation (ex. nombre de vues), commentaire ou réponse, mise en favori, « like », republication, etc.

1. Dispositif de signalement de fausses informations

Description du dispositif

- 1.1. L'opérateur indiquera si son service dispose d'un dispositif de signalement des contenus publiés et si ce dispositif permet à l'utilisateur de **signaler de fausses informations** susceptibles de troubler l'ordre public ou d'altérer la sincérité du scrutin. Il indiquera **le libellé choisi** pour ce motif et les raisons qui sous-tendent le choix de ce libellé. Il pourra illustrer sa réponse d'exemples.

Le Conseil souhaite connaître le **nombre total de motifs de signalement proposés à l'utilisateur**, et si ces motifs varient selon le type de contenus (texte, vidéo, commentaire, compte). Il souhaite également connaître leur arborescence et s'il existe une possibilité de saisir un motif autre que ceux proposés.

Notre service propose des dispositifs de signalement des contenus *via* notre outil de signalement intégré à l'application que les utilisateurs de Snapchat peuvent utiliser pour signaler ce qui les dérange, notamment les fausses informations susceptibles de troubler l'ordre public ou d'altérer la sincérité des scrutins. Nous disposons également de dispositifs de signalement à travers 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>).

L'outil de signalement intégré est facilement accessible (visible) et convivial (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 » 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 proposés à l'utilisateur, pour presque toutes les typologies de contenu de Snap (Snaps 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 ». Dans de nombreux cas, la présentation du menu de signalement permet également aux utilisateurs d'étoffer et d'argumenter leur signalement dans une case «commentaires» à compléter.

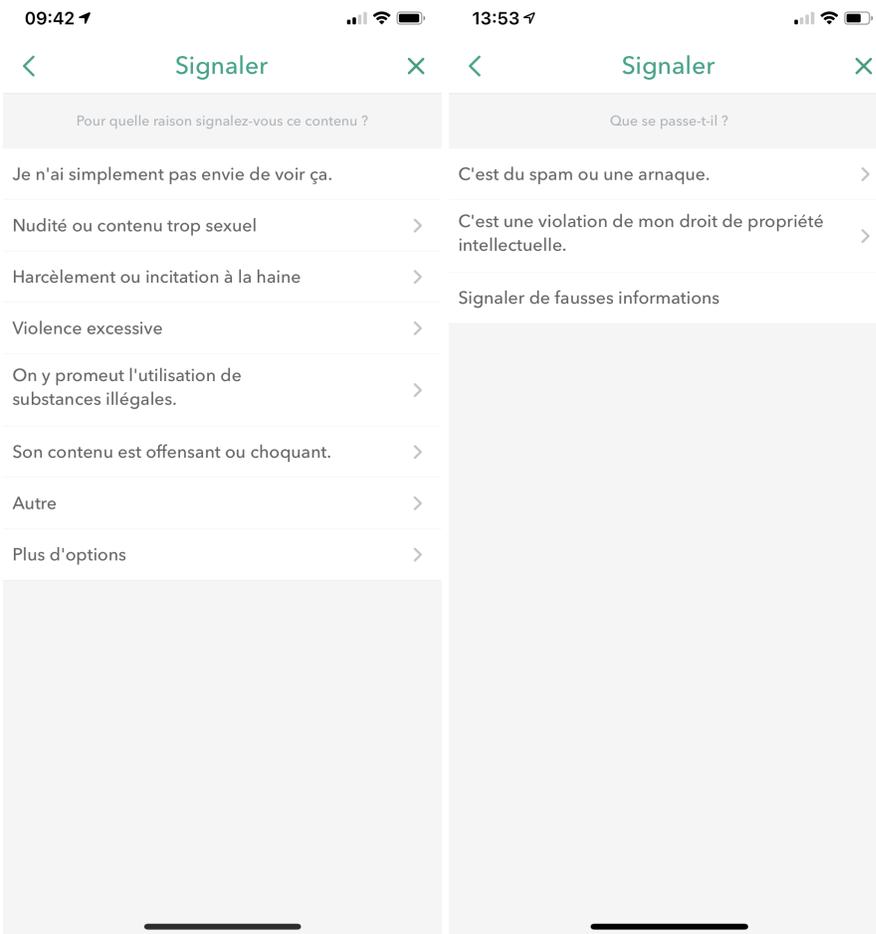
Ainsi, pour la partie Découvrir (*Discover*) les catégories de signalement suivantes sont possibles :

- 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
- Signaler de fausses informations (**amélioration réalisée en 2020**)



Les membres de notre équipe dédiée dite « Confidentialité et Sécurité » (« *Trust and Safety* ») examinent 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.

1.2. Le Conseil souhaite que l'opérateur lui fournisse des informations sur la **visibilité, l'accessibilité et l'ergonomie** de ce dispositif², et notamment :

- sa présentation visuelle ;
- son affichage (visibilité par défaut ou non, visibilité par tous les utilisateurs ou seulement par les utilisateurs titulaires d'un compte) ;
- son emplacement à proximité immédiate du contenu ou non ;
- le nombre d'actions nécessaires pour accéder au dispositif de signalement général, et à celui de signalement d'une fausse information ;
- le nombre d'actions et d'hyperliens nécessaires à l'envoi du signalement d'une fausse information depuis la page du contenu et, le cas échéant, les efforts déployés par l'opérateur pour réduire de façon pertinente le nombre d'actions nécessaires pour l'envoi du signalement ;
- l'existence d'un dispositif de signalement facilement accessible adapté notamment aux personnes en situation de handicap.

1.2.1. L'opérateur précisera si le dispositif décrit est **commun à toutes les déclinaisons** du service (site internet, application) **et à tous les types de contenus** (texte, vidéo, image, commentaires, contenus sponsorisés, compte utilisateur) accessibles sur le site.

1.2.2. Il mentionnera les **fonctionnalités supplémentaires**, développées ou en cours de développement, permettant d'étayer un signalement (possibilité de signaler plusieurs contenus conjointement, d'assortir le signalement d'explications, d'une pièce jointe ou d'un lien, etc.).

1.2.3. L'opérateur indiquera au Conseil les éventuels retours des utilisateurs concernant l'accessibilité, la visibilité et l'ergonomie du dispositif de signalement déployé sur son service.

Réponses combinées de Snap aux questions 1.2.1 à 1.2.3.

Comme l'indique la réponse à la question 1.1, l'outil de signalement intégré à l'application est facile à trouver et à utiliser pour les utilisateurs. Le bouton de signalement « drapeau  » ou « i » est disponible pour toutes les catégories de contenu sur Snapchat, accessible sur chaque élément de contenu et disponible sur tous les appareils et versions des applications Snapchat (iOS et Android).

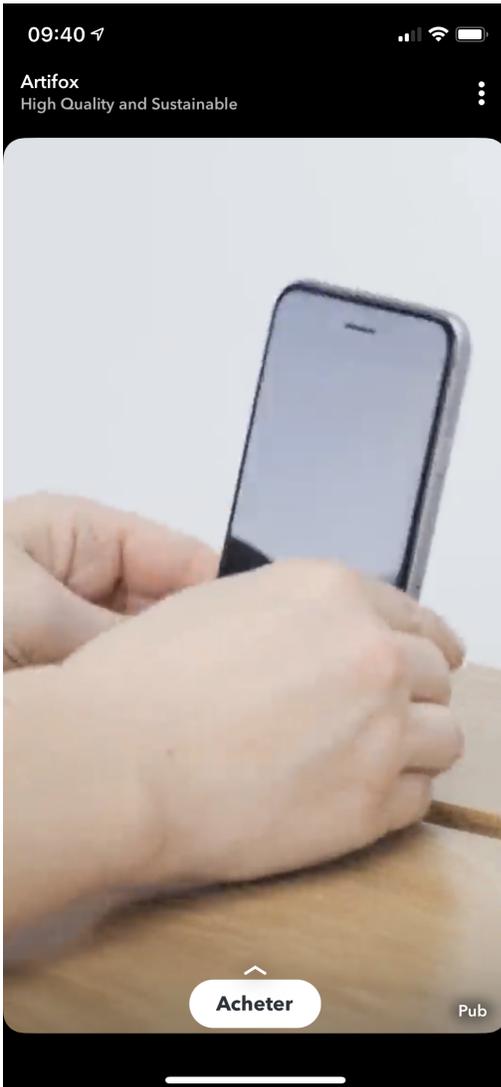
Pour déterminer où placer le drapeau "fausse information", Snapchat évalue d'abord le type de contenu et le niveau de pré-modération qui doit être appliqué au contenu.

En pré-sélectionnant nos partenaires Découvrir et les partenaires publicitaires nous réduisons le risque d'avoir un contenu véhiculant de fausses informations puisque nous examinons et sélectionnons soigneusement et individuellement les entreprises avec lesquelles nous collaborons. Par ailleurs, nos accords avec nos partenaires Découvrir interdisent le partage de fausses informations. Des règles similaires s'appliquent aux publicités. Par conséquent, il y a une très faible probabilité que des contenus véhiculent de fausses informations, ce type d'information aurait été détecté par les équipes de modération et/ou les partenaires qui soumettent le contenu. Nous avons examiné les directives du CSA et inclus le drapeau "fausses informations" dans 3 ou 4 actions, comme indiqué ci-dessous :

² Si les contenus publicitaires bénéficient d'une procédure de signalement spécifique, l'opérateur est invité à le décrire ici ou dans la partie 5 (question 5.5).

Signaler les Publicités : Par exemple, les utilisateurs peuvent signaler le contenu d'une Publicité de la manière suivante :

1. **Faire un appui long sur le contenu ou en cliquant sur l'icône à trois points en haut à droite du contenu média :**



2. **Cliquez sur l'icône « i » pour afficher les informations sur les Publicités :**



The image shows a mobile advertisement for Artifax. At the top left, the time is 09:40. The brand name 'Artifax' is displayed with the tagline 'High Quality and Sustainable'. The main content is a video player showing a wooden modular desk system with a laptop and a chair. The text '[modular desk system]' is overlaid on the video. At the bottom left of the video player is an information icon (i), and at the bottom right is a blue play button icon. The entire advertisement is set against a black background.

09:40

Artifax
High Quality and Sustainable

[modular desk system]

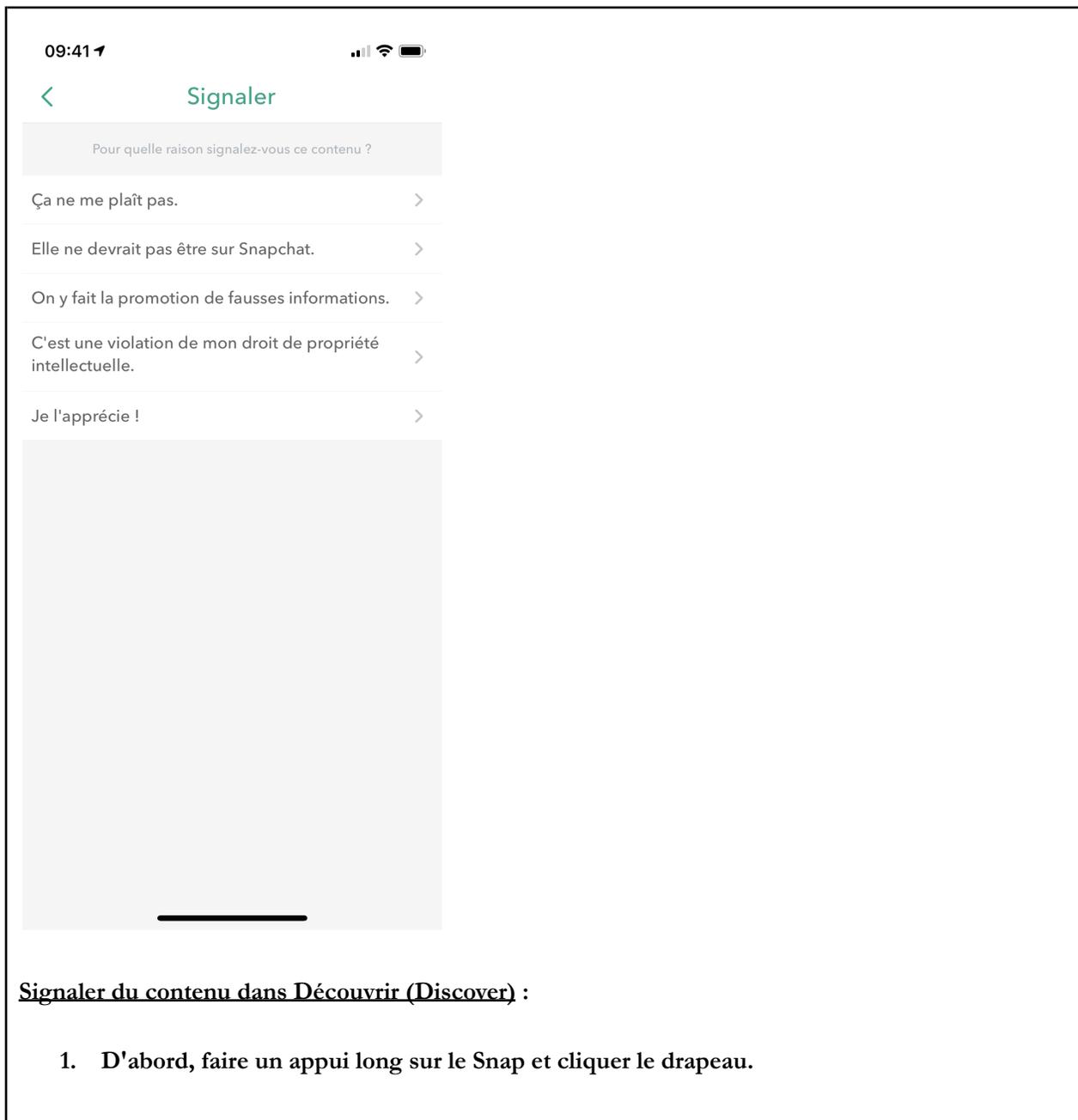
i

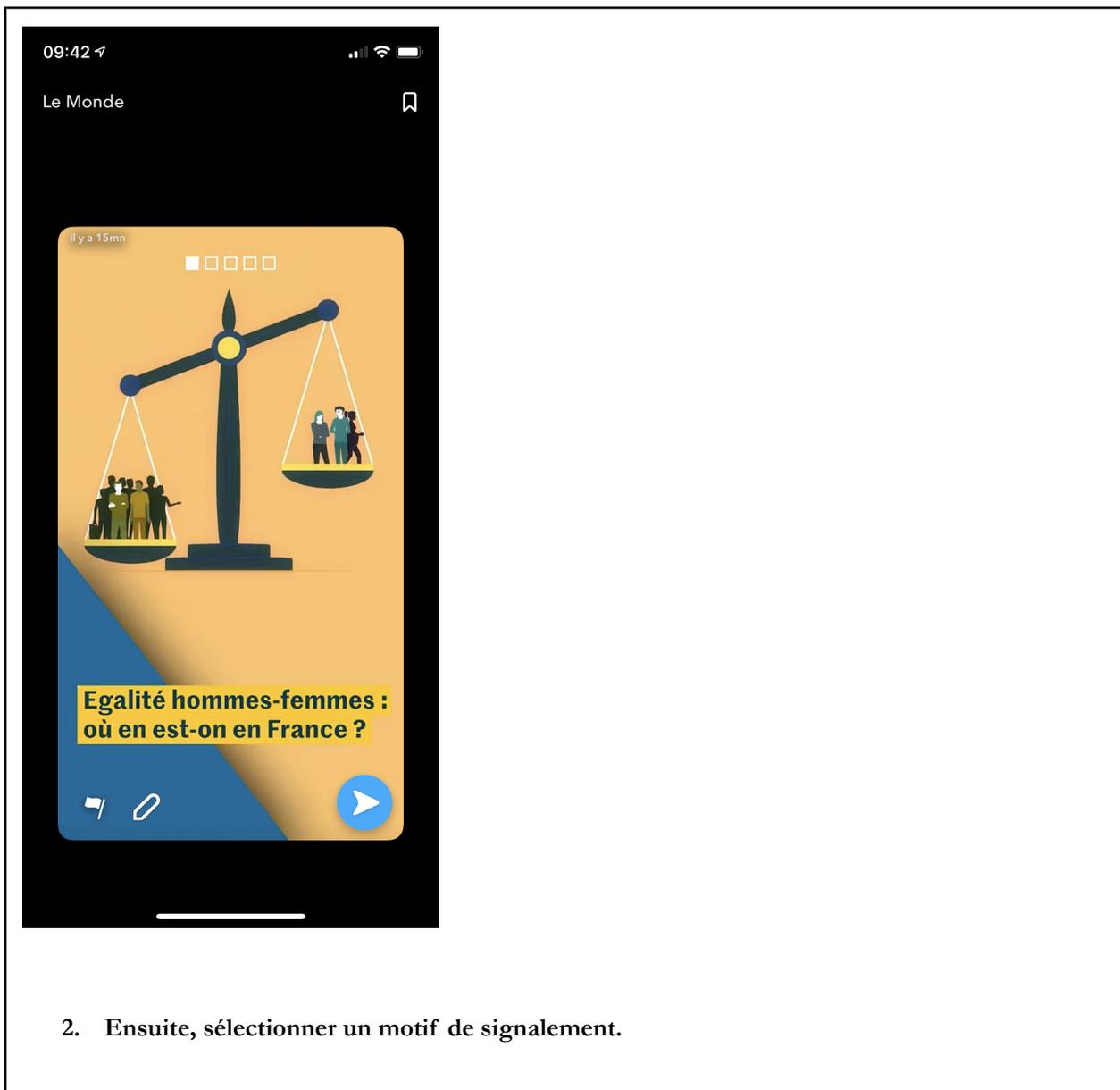
▶

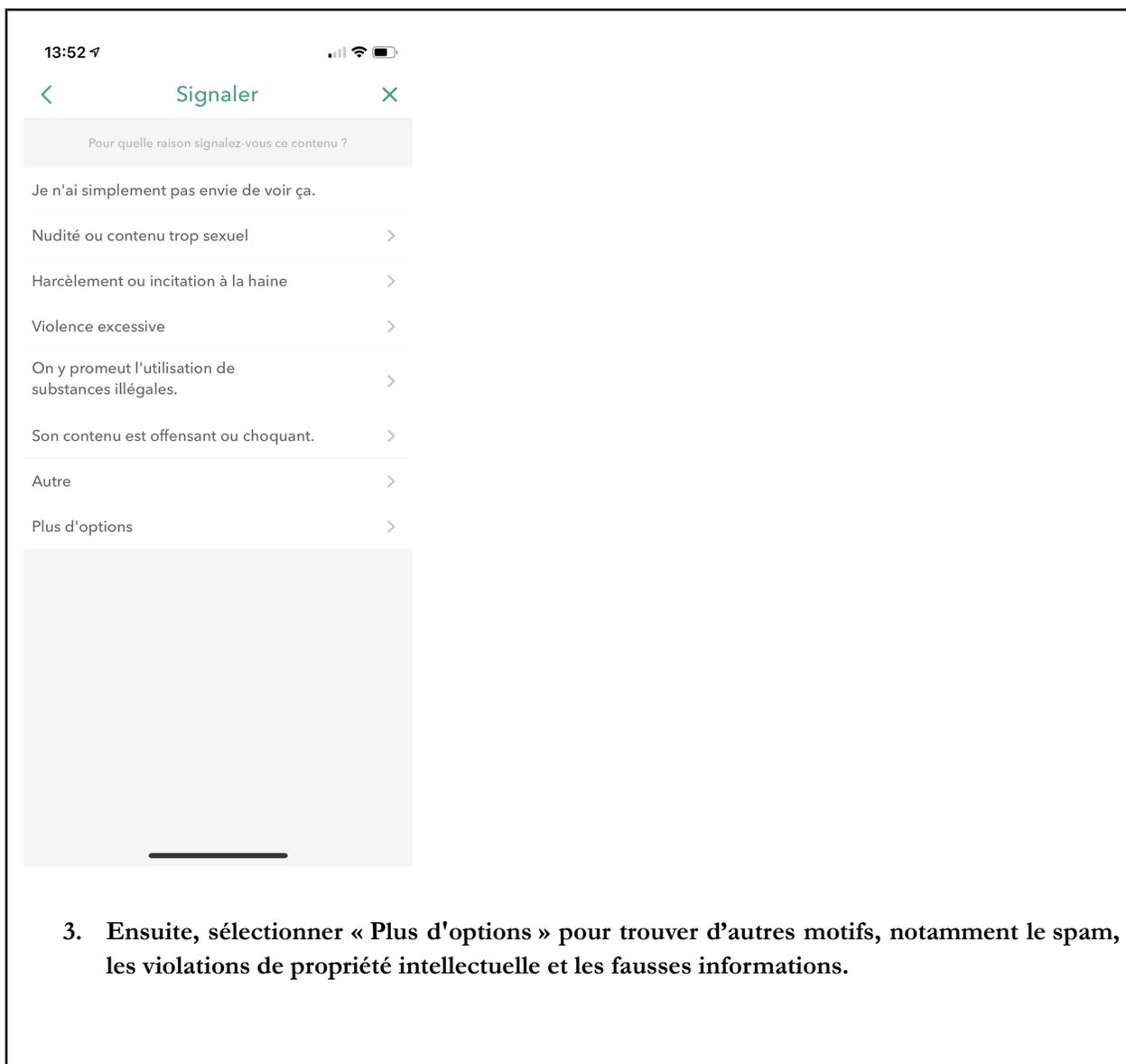
3. Cliquer sur « Signaler la Publicité ».

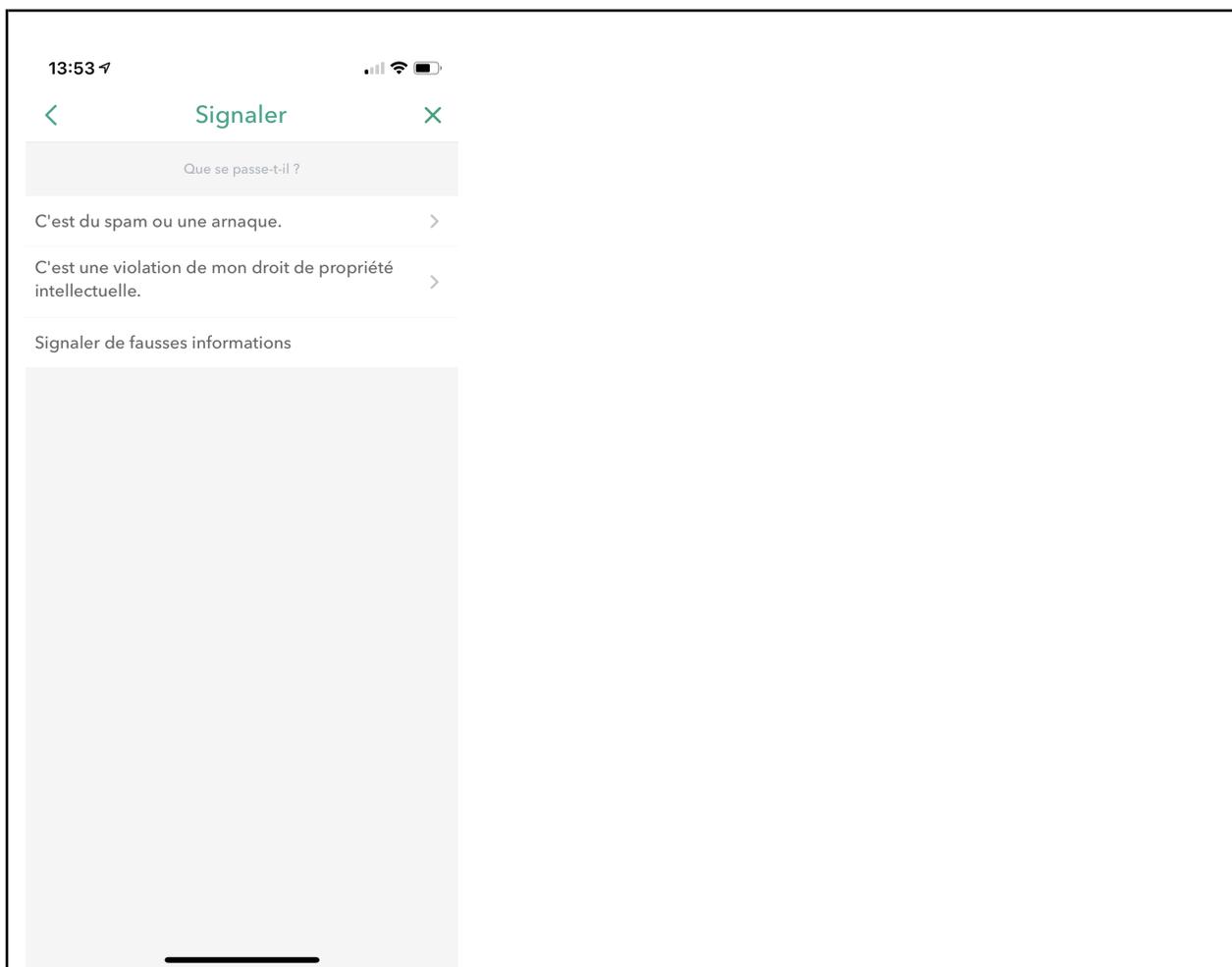


4. Sélectionner le motif de signalement du contenu.





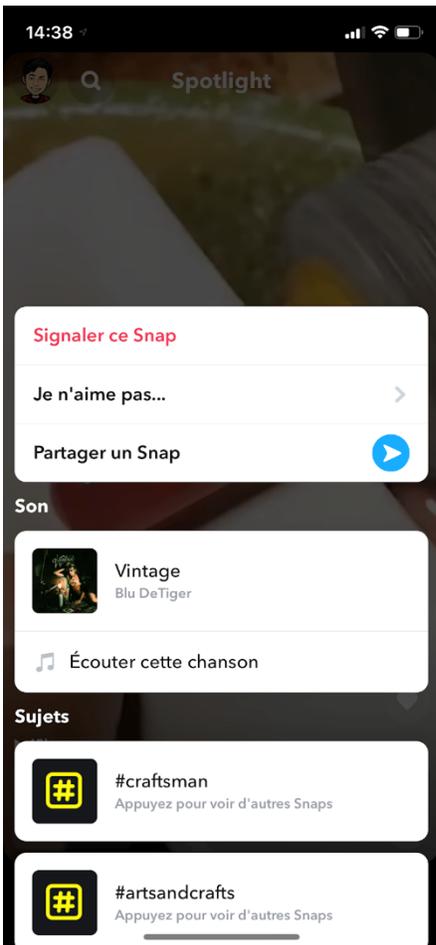




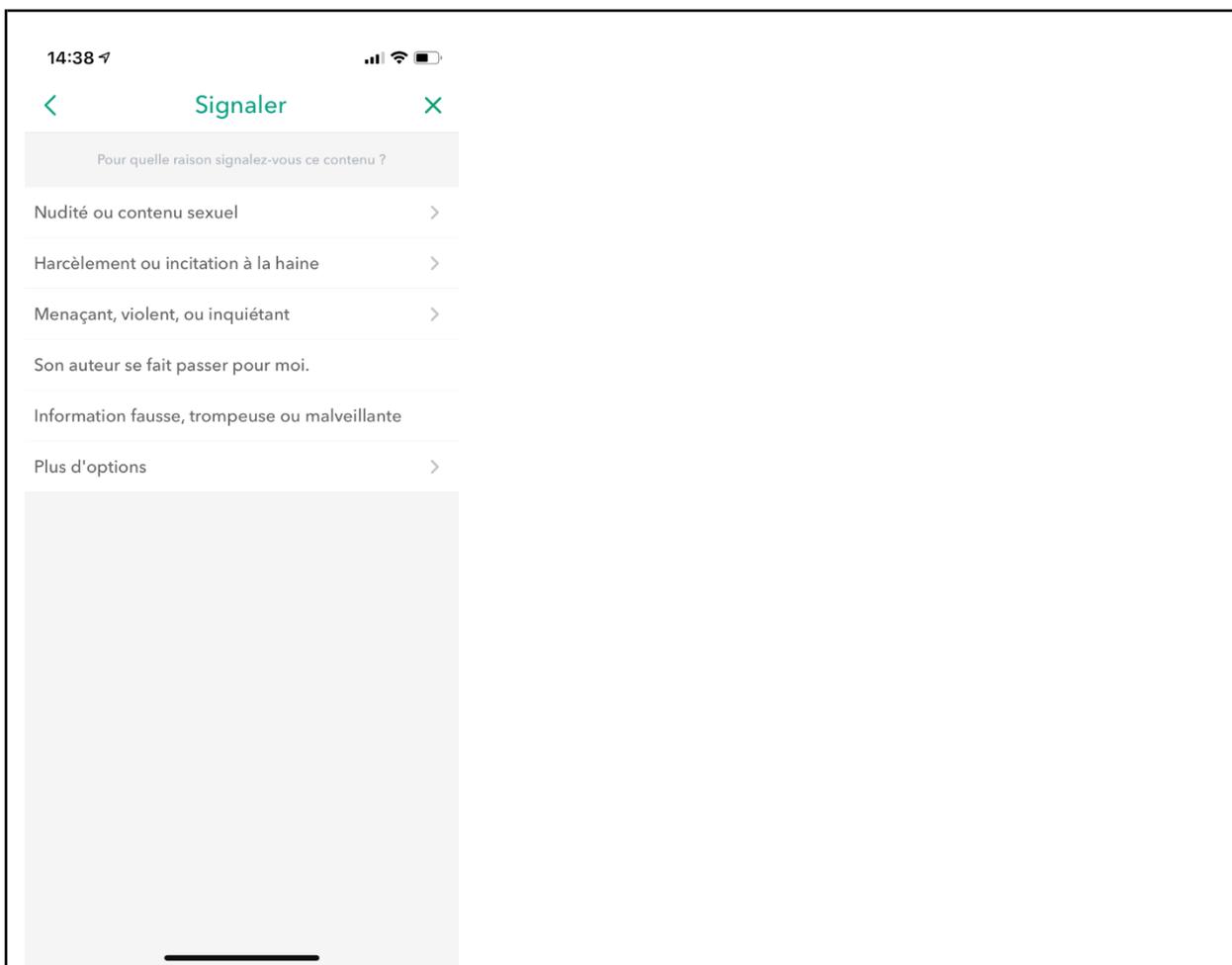
Spotlight

Comme indiqué dans l'introduction de ce questionnaire, Snapchat a lancé un nouveau produit appelé *Spotlight* à la fin de l'année dernière. Ce produit est davantage axé sur le contenu généré par les utilisateurs. Bien que tout le contenu soit soumis à un examen automatisé et/ou humain, Snap a estimé que ce type de contenu était davantage susceptible de contenir de fausses informations, c'est pourquoi le motif de signalement "fausses informations" se trouve au premier plan, comme présenté ci-dessous.

Étape 1. L'utilisateur fait un appui long sur le Snap



Étape 2. L'utilisateur peut sélectionner « fausses informations » comme catégorie de signalement, dès le premier niveau.



Accessibilité

Comme nous l'avons décrit ci-dessus, un utilisateur de Snapchat peut facilement signaler un contenu à l'aide de notre outil de signalement intégré à l'application. Il peut accéder au menu de signalement à partir du contenu concerné en quelques clics. La procédure de signalement comprend un mécanisme spécifique permettant à l'utilisateur de signaler de fausses informations. Pour le contenu situé dans la partie Découvrir, le motif « fausses informations » est inclus dans la rubrique « plus d'options ». Contrairement à la plupart des plateformes, Snap ne travaille qu'avec des partenaires préalablement approuvés. Ceux-ci se sont engagés à ne pas publier une liste précise de contenus interdits, parmi lesquels les fausses informations. Par conséquent, il est rare que ces partenaires publient du contenu qui pourrait être qualifié de fausse information, ce que corrobore le faible nombre de signalements (voir la réponse aux questions 1.3.2 à 1.3.8) de fausses informations crédibles reçues annuellement par Snap (comme détaillé dans notre réponse).

Retour des utilisateurs

Snap met à disposition des utilisateurs diverses ressources qui leur expliquent comment ils peuvent effectuer des signalements. Voir par exemple la page de support dédiée (<https://support.snapchat.com/fr-FR/a/report-abuse-in-app>).

En outre, la chaîne intitulée « Safety Snapshot » dans la section « Découvrir » de Snapchat offre aux *Snapchatters* (nom que nous donnons aux utilisateurs de Snapchat), sous formes d'épisodes, un aperçu et une information pédagogique avec les meilleures pratiques sur les principales préoccupations en matière de sécurité en ligne, comme le fait de cliquer sur des liens non reconnus (prévention du phishing) et de vérifier son adresse e-mail pour mieux prévenir le détournement de compte. Les données issues de l'outil de signalement intégré à l'application, les analyses du site d'assistance et les principales tendances ont été utilisées pour guider les épisodes. Nous continuons à accompagner notre communauté sur les sujets importants, en proposant davantage de contenu éducatif dans l'application. Nous étudions les moyens de rendre la chaîne plus exploitable, par exemple en créant dans notre page « réglages » des liens renvoyant vers des actions directes plutôt qu'une page d'assistance. Le lancement de la chaîne a d'abord eu lieu aux États-Unis, mais nous devrions le déployer dans d'autres pays au cours de l'année 2021. Plus d'informations sont disponibles sur <https://support.snapchat.com/fr-FR/news/safety-snapshot-news>.

Procédure de traitement des signalements

1.3. *L'opérateur décrira l'ensemble de la **procédure de traitement** du signalement d'une fausse information propre à garantir sa pleine effectivité tel que prévu à l'article 11 de la loi du 22 décembre 2018.*

1.3.1 *Il indiquera les **conditions de mise en instruction** d'un signalement, en précisant notamment si l'instruction des signalements est systématique ou non. Le cas échéant, il portera à la connaissance du Conseil les critères déclenchant cette mise en instruction. Il précisera les conditions d'élaboration de ces critères (par qui, quand et comment sont-ils établis ou révisés, les objectifs qu'ils poursuivent) et leurs adaptations éventuelles au cours de l'année 2020.*

Snap met à la disposition des utilisateurs de Snapchat un outil de signalement intégré dans l'application, facile à utiliser pour signaler de fausses informations et d'autres contenus illicites. Nous instruisons systématiquement tous les signalements qui sont opérés.

Notre plateforme est également conçue pour réduire le risque de diffusion de fausses informations. Comme mentionné dans notre réponse au point 1.2, l'application Snapchat comprend une plateforme modérée qui propose une sélection d'actualités et de contenus provenant d'éditeurs professionnels, d'influenceurs et de célébrités, appelée Découvrir, et ne présente pas de fil d'actualité ouvert où chacun pourrait poster des contenus susceptibles de devenir viraux. Une procédure spécifique n'apparaît donc pas nécessaire dans le cadre des circonstances décrites ci-dessus.

Snap prend très au sérieux la sécurité des utilisateurs de Snapchat et de la communauté. Dans le cadre de notre engagement pour la sécurité, nous avons consciencieusement mis en place un [Centre de Sécurité](#) en ligne pour les utilisateurs, les parents, les forces de l'ordre et les enseignants, contenant des informations sur nos équipes et nos politiques, des conseils pour utiliser Snapchat en toute sécurité et pour signaler les violations de nos conditions ou de nos politiques. Nous disposons d'une équipe dédiée dite « Confidentialité et Sécurité », composée d'employés et de sous-traitants formés, situés aux États-Unis, en Europe et en Asie. L'équipe travaille selon un planning de veille permanente pour fournir une assistance sur les signalements 24 heures sur 24 et 7 jours sur 7. L'équipe Confidentialité et Sécurité est formée pour examiner et prendre des

mesures concernant tous les types de contenus illégaux et interdits sur la plateforme Snapchat, y compris les fausses informations, les contenus sexuellement explicites, le harcèlement, les menaces, l'usurpation d'identité, le spam, les discours de haine, les activités/produits illicites et le terrorisme. L'examen des signalements est principalement effectué par nos équipes (modération dite humaine) mais nous avons aussi mis en place des outils automatisés pour renforcer la mise en œuvre de l'examen dans certains domaines. Comme ce type de problématique a été intégrée "dès la conception du produit" ("by design") pour réduire le risque de diffusion de fausses informations, les volumes de ce type de signalements sont très faibles (voir notre réponse à la question 1.5 pour plus d'informations).

Après examen d'un signalement, notre équipe peut supprimer le contenu illicite et, dans certains cas, supprimer le compte signalé. Pour le moment, la personne à l'origine du signalement n'est pas tenue informée de nos actions, et les résultats des mesures prises ne sont pas portés à l'attention des autres utilisateurs. Toutefois, le contenu illicite est immédiatement supprimé dès la prise de décision et n'est plus accessible aux utilisateurs. Les auteurs du contenu signalé peuvent nous contacter pour contester la décision et les mesures prises.

Nous interdisons la diffusion de fausses informations mal-intentionnées ou susceptibles de causer un préjudice, y compris notamment :

- 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, que la COVID-19 serait un canular) ;
- l'atteinte à l'intégrité de processus civiques (par exemple, en diffusant une fausse date d'élection).

Plus de détails sont fournis dans nos règles communautaires, disponibles à l'adresse suivante : <https://www.snap.com/fr-FR/community-guidelines>.

- 1.3.2 *Il détaillera les **moyens** déployés pour l'instruction, et notamment :*
- les **moyens humains**, en précisant le nombre de personnes en charge de cette activité, leur statut (collaborateur ou prestataire de la société) leur formation (initiale et continue), leur localisation géographique, leur langue de travail ainsi que les modalités d'organisation de leur travail (notamment si elles impliquent une astreinte) ;
 - les **sources** dont disposent les modérateurs afin d'effectuer le contrôle des contenus signalés, en précisant leur nature (juridique, journalistique, académique) et l'articulation de ces sources entre elles ;
 - **l'articulation entre l'intervention humaine et l'intervention automatique** ; en particulier, une intervention humaine est-elle systématique dans l'instruction et, le cas échéant, dans la décision d'intervenir à l'encontre du contenu ou du compte ;
 - les **dispositifs spécifiques** mis en place afin de garantir une modération efficace lors de situations de crise (formation des équipes de modération aux sujets d'actualité, moyens automatisés renforcés, partenariats exceptionnels avec les fact-checkers...).
- 1.3.3 *L'opérateur indiquera s'il a mis en place des **moyens humains et/ou automatisés de détection proactive** des contenus diffusant de fausses informations. Il indiquera également si le processus d'instruction des contenus détectés est en tout point identique à celui des contenus signalés par les utilisateurs ou sur quels points il diffère le cas échéant.*
- 1.3.4 *L'opérateur présentera également les **différentes étapes de l'instruction** (décision de mise en instruction d'un signalement, examen du contenu, qualification du contenu, décision prise sur un contenu) et il indiquera s'il existe une procédure d'instruction spécifique en période électorale ou de situation de crise (ex : crise sanitaire).*

- 1.3.5 *Concernant l'issue de l'instruction, l'opérateur présentera également de façon détaillée :*
- la liste des **actions** qu'il peut mettre en œuvre vis-à-vis des contenus signalés dont il a considéré qu'ils contenaient de fausses informations, ainsi que des comptes sur lesquels ils ont été publiés ;
 - les **critères** amenant à décider de telle ou telle action ;
 - les **modalités de mise en œuvre** de la décision.
- 1.3.6 *S'agissant des informations communiquées aux différents utilisateurs concernant l'instruction d'un contenu signalé, l'opérateur précisera :*
- quels sont les **utilisateurs informés** (l'auteur du signalement, l'auteur du contenu signalé, les utilisateurs ayant interagi avec le contenu signalé, l'ensemble des utilisateurs) ;
 - pour chacun, quels sont les stades de la procédure **qui donnent lieu à une information spécifique adressée aux utilisateurs** (envoi d'un signalement, mise en instruction du signalement, issue de l'instruction, raisons précises de la décision prise à cette issue).
- 1.3.7 *L'opérateur indiquera au Conseil si des voies de recours internes existent, si elles sont portées à la connaissance des utilisateurs et si oui, desquels. Il précisera notamment :*
- quelles sont les **décisions de la plateforme susceptibles d'être contestées** ;
 - quels sont les **utilisateurs en droit** d'exercer un recours, **dans quel délai** la possibilité de recours leur est présentée à l'issue de la décision et **par quelle voie** ;
 - quelles sont les **équipes en charge de l'instruction** de ces recours ;
 - quels sont les **critères de recevabilité** des recours ;
 - les **conditions** dans lesquelles s'effectue l'instruction du recours, en particulier si l'utilisateur à l'origine du recours est en capacité de faire valoir ses explications aux équipes de modération (voire, le cas échéant, aux autres utilisateurs impliqués) ;
 - le **délai moyen de traitement** d'un recours ;
 - si les informations sont fournies systématiquement aux utilisateurs s'agissant des recours internes ou s'il existe des **exceptions, prévues ou de fait**.
- 1.3.8 *L'opérateur indiquera au Conseil si ont été développés :*
- des **mécanismes d'autoévaluation** des procédures et des mesures prises à l'égard d'un contenu signalé et les critères sur lesquels cette évaluation est basée ;
 - des **mécanismes d'évaluation ou enquêtes de satisfaction** de ces procédures et mesures à l'attention des utilisateurs. Le cas échéant, il informera le Conseil de leurs résultats et de la façon dont il les a prises en compte.

Réponse de Snap aux Questions 1.3.2 à 1.3.8

Lorsque nous recevons un signalement d'un utilisateur concernant un contenu particulier, nous l'évaluons en fonction de nos règles communautaires. Si le contenu n'enfreint pas ces règles, nous ne prenons pas de mesure. 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, nous informons les membres de l'équipe afin qu'ils se préparent à une augmentation du volume de contenus illicites. S'agissant des fausses informations, nous faisons appel à des sources d'information de confiance et à des tiers de confiance 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 mondiale 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.

Certains contenus populaires peuvent être mis en avant sur notre section Découvrir pour une plus grande visibilité. Ce contenu est modéré avant d'être promu sur Découvrir. 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 Découvrir 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.

1.4. *L'opérateur fournira des **informations chiffrées** sur la mise en œuvre du dispositif de signalement de fausses informations en 2020, notamment :*

- *le **nombre de contenus signalés** comme fausses informations ;*
- *le **nombre de procédures de traitement** mises en œuvre ;*
- *le **nombre de contenus considérés comme étant de fausses informations** à l'issue de la procédure de traitement ;*
- *le **délai moyen de traitement** d'un signalement ;*
- *le **nombre de recours** reçus par l'opérateur à l'issue de la procédure de traitement ;*
- *le **budget alloué** au traitement des signalements par les utilisateurs depuis l'entrée en vigueur de la loi du 22 décembre 2018 et l'évolution de ce budget entre 2019 et 2020.*

*L'opérateur est invité à décrire et analyser la tendance observée pour les indicateurs précités. Il indiquera comment ces chiffres, ou toute autre donnée relative à la circulation effective de fausses informations sur son service, sont **portés à la connaissance des utilisateurs**.*

Comme nous l'avons expliqué ci-dessus, la conception de notre application a largement permis de la préserver des fausses informations et de la désinformation. Ainsi que nous le précisons également dans notre réponse à la question 1.1, nous avons initié cette année la mise en œuvre de mesures visant à ajouter les « fausses informations » comme motif spécifique de signalement dans notre processus de signalement. Nous serons désormais en mesure de suivre l'évolution de ces chiffres à compter de cette année.

Comme expliqué ci-dessus, les utilisateurs peuvent faire part de leurs commentaires (dans l'application ou sur notre site web) sur tout contenu qu'ils jugent répréhensible. Nous proposons un menu de catégories spécifiques de contenus répréhensibles. En outre, les utilisateurs peuvent personnaliser la description de leurs objections dans de nombreuses catégories.

En 2020, nous avons ajouté une rubrique spécifique pour le signalement par les utilisateurs des « fausses informations » :

- Nombre de contenus signalés : 3 547
- Nombre de procédures de traitement mises en oeuvre : 3 547
- Nombre de contenus considérés comme étant de « fausses informations » : 4

Vous trouverez ci-après une liste d'autres exemples de signalements d'utilisateurs qui ont été effectués en 2020 et examinés par nos équipes.

- Fausses informations, encourageant la vente de permis voiture
- Site d'escroquerie qui arnaque ses clients. Ils facturent 30 € de frais de livraison et n'envoient pas les montres.
- Partage d'informations frauduleuses visant à escroquer les gens
- Faux conseils et paiements pour les paris sportifs (x2)
- Escroqueries / escrocs qui volent l'argent de personnes physiques.

Le fait que nous n'ayons rencontré que 4 cas confirmés (à rapporter à nos 16 millions d'utilisateurs actifs quotidiens en France - chiffre Médiamétrie mars 2021), constitue également une preuve concrète du fait que Snapchat se distingue des sociétés de médias sociaux traditionnelles en ayant intégré cette problématique dès la conception de tous nos produits et fonctionnalités, au lieu d'être inclus *a posteriori*.

De même, tout le contenu accessible au public sur Snapchat est soit sélectionné par des éditeurs médias professionnels tiers, soit pré-modéré en interne par Snap avant d'être publié sur Snapchat. Le succès de cette approche se traduit par ce très faible nombre de signalements de fausses informations et de désinformation sur Snapchat.

Comme le précisent les réponses à ce questionnaire, Snap interdit la diffusion de fausses informations (y compris la manipulation de médias) mal-intentionnées ou susceptibles de causer un préjudice, comme par exemple :

- 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, que la covid-19 serait un canular)
- l'atteinte à l'intégrité de processus civiques (par exemple, en diffusant une fausse date d'élection).

1.5. Cas d'usage : l'opérateur partagera également avec le Conseil les procédures qu'il suit dans le cas d'un contenu contenant des éléments reconnus comme faux par des sources fiables d'information et publié sur le service.

Le Conseil ne souhaite pas ici que l'opérateur se prononce sur un cas précis mais qu'il indique quelle est sa méthode de travail sur ce type de cas : il souhaite comprendre les grandes étapes suivies par l'opérateur pour traiter ce type de cas, en précisant à chaque fois les critères présidant aux choix qui peuvent être faits, la nature de l'intervention humaine et les données prises en compte (signalement interne, signalement utilisateur, viralité du contenu, etc.). Un arbre de décision général pourra être fourni ainsi que des statistiques agrégées sur la performance de ces étapes : taux de contenus détectés en interne avec taux de faux positifs et de faux négatifs, recours à des partenaires pour la qualification des contenus comme relevant de la fausse information ou qualification de ces contenus par des équipes internes, temps moyen constaté entre la diffusion d'un contenu et la première action prise par l'opérateur, actions prises par l'opérateur et impact de ces actions prises sur la circulation du contenu.

Veillez vous référer à la réponse à la Question Q. 1.3 pour une description du processus.

Autres questions

- 1.6. *Le Conseil souhaite savoir si l'opérateur a engagé des **discussions avec d'autres opérateurs** soumis au titre III de la loi du 22 décembre 2018 afin d'harmoniser leurs dispositifs de signalement des fausses informations respectifs et, le cas échéant, quelles ont été les mesures prises dans le cadre de ces discussions.*
L'opérateur pourra porter à la connaissance du Conseil les initiatives qu'il porte visant à renforcer la coopération avec d'autres plateformes (mise en commun et partage de données, remontée d'informations sensibles...).

Nous participons à des groupes de discussion dans notre secteur d'activité et à des ONG qui partagent des informations sur les tendances globales et sur les fausses informations qui circulent. Les acteurs avec lesquels nous le faisons sont des agrégateurs de tendances en matière de fausses informations travaillant avec différents opérateurs. Toutefois la standardisation des systèmes est probablement impossible car les nombreux opérateurs fournissent des services basés sur des architectures différentes. Le partage d'informations à cet égard, comme nous l'avons mentionné, constitue un bon moyen pour nous de traiter collectivement les fausses informations susceptibles de se propager sur différentes plateformes de différentes manières. Ainsi, chaque plateforme peut traiter les fausses informations d'une manière adaptée à ses services et à son architecture.

- 1.7. *L'opérateur indiquera dans quelle mesure il a tenu compte **des recommandations du Conseil** énoncées dans son bilan relatif à l'exercice 2019 à compter de sa publication en juillet 2020.*

Snap a bien pris en compte les recommandations du Conseil pour 2019

Nous notons ainsi que le CSA a spécifiquement mentionné Snap dans certaines de ses recommandations. Par exemple, le CSA a reconnu que la façon dont Snapchat est conçu empêche la viralité du contenu et fonctionne donc en soi pour réduire *a priori* les risques de diffusion de fausses informations. Cependant, le CSA a déclaré : « *Si la viralité des contenus contenant de fausses informations est indéniablement bien moindre que sur d'autres plateformes, il n'est pas exclu que ces contenus puissent être vus par un nombre significatif d'utilisateurs.* »

Même si nous sommes conscients du fait que le contenu puisse être vu par un large public, nous souhaitons souligner que Snapchat est conçu pour les amis proches uniquement. L'année dernière, Snap a lancé la fonction « Friend Check Up », qui incite les *Snapchatters* à revoir leur liste d'amis et à s'assurer qu'elle n'est composée que de personnes avec lesquelles ils souhaitent rester en contact. Cette fonctionnalité montre que l'approche de Snapchat quant au traitement des amis diffère de celle des autres plateformes, par exemple :

- Il n'y a pas de profils publics consultables pour les *Snapchatters* de moins de 18 ans ;
- Par défaut, vous ne pouvez pas chatter ou contacter directement une personne à moins que vous ne vous soyez tous deux ajoutés comme amis ;
- Un grand nombre de nos fonctionnalités sont paramétrées comme privées par défaut, ce qui aide à protéger les *Snapchatters* contre le partage involontaire d'informations, telles que leur localisation;
- Nous ne donnons pas aux discussions de groupe la possibilité de devenir virales, comme c'est parfois le cas dans d'autres types de plateformes leur permettant de devenir des vecteurs de

contenus extrémistes. Les discussions de groupe sont conçues pour être des conversations entre groupes d'amis réels, c'est pourquoi nous limitons leur taille à 64 amis. Les groupes ne sont pas consultables, recommandés ni affichés dans l'application en dehors de l'onglet « Chat ».

De manière plus générale, Snap fournit des informations spécifiques aux utilisateurs et les informe que les fausses informations sont interdites.

Nous veillons à prendre en compte les événements mondiaux et à réagir de manière appropriée. Par exemple, nous avons collaboré avec des partenaires de Découvrir pour présenter à nos utilisateurs des informations crédibles sur la covid-19. Vous trouverez plus de détails dans nos règles communautaires, disponibles à l'adresse suivante : <https://www.snap.com/fr-FR/community-guidelines>.

D'autres informations qui correspondent aux recommandations du CSA, comme les processus mis en œuvre par Snap, la coopération avec des tiers et les procédures de signalement ont aussi été prises en compte et sont détaillées à différents points de cette réponse.

1.8. L'opérateur est invité à indiquer au Conseil l'impact de la crise sanitaire sur son dispositif de signalement de fausses informations et, notamment, les adaptations qu'il a pu être amené à lui apporter en conséquence.

Dans le monde entier, nos équipes travaillent à distance afin de limiter les interactions physiques, de s'associer aux efforts plus vastes de santé publique visant à ralentir la propagation du virus et d'assurer la sécurité des employés de Snap. Cette politique de travail à distance s'étend à nos employés chargés de la modération du contenu. Bien que beaucoup de ces travailleurs soient des employés à temps plein, nous avons également des sous-traitants. Nous avons travaillé avec nos partenaires contractuels tiers pour assurer une communication continue et un soutien à distance pour les membres de notre équipe.

Pendant cette période, les mesures que nous mettons en œuvre restent toujours aussi fortes. Nous continuons à prendre des mesures contre les contenus ou les activités qui violent nos conditions. Notre équipe mondiale Confidentialité et Sécurité continue à examiner les signalements des utilisateurs et à prendre les mesures appropriées.

2. Transparence des algorithmes

Les lignes directrices suivantes s'appliquent aux opérateurs qui ont mis en œuvre des mesures relatives à la transparence des algorithmes dans le cadre des mesures supplémentaires prévues à l'article 11 de la loi du 22 décembre 2018.

2.1. Le Conseil demande aux opérateurs de préciser dans leurs réponses les éléments qui ont fait l'objet de changements structurels par rapport à l'exercice 2019.

L'opérateur précisera notamment si la **tenue des élections municipales** en France et la **pandémie** ont modifié ses procédures habituelles en 2020. Il souhaitera notamment savoir si les changements ont conduit à une moindre intervention humaine et comment cela a pu affecter la performance globale des systèmes de recommandation et de modération utilisés en 2020, par rapport à 2019.

Nous n'avons procédé à aucun changement structurel pour répondre à la pandémie, aux élections municipales en France ou à d'autres événements mondiaux, considérant que la structure mise en place permettait déjà de s'adapter de manière satisfaisante à ces situations. Nous avons amélioré l'automatisation pour faciliter une action plus rapide suivie d'une intervention humaine, car les conditions du « travail à distance » varient dans le monde entier.

Cadre général de réponse

2.2. L'opérateur est invité à commenter les définitions de la loyauté, de l'explicabilité et de l'équité indiquées au début du questionnaire. Le Conseil souhaite en particulier savoir si l'opérateur les utilise (ou 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. L'opérateur 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>

2.3. Lorsque l'opérateur est invité à décrire les systèmes reposant sur un ou plusieurs algorithmes qu'il utilise, les éléments suivants sont attendus pour chacun de ces systèmes afin de permettre au Conseil de mesurer le niveau de transparence offert aux utilisateurs :

- sur quelles fonctionnalités du service opère-t-il ?
- quel est son objectif principal (tâche ou classe de problèmes à résoudre) ?
- repose-t-il sur un ou plusieurs algorithmes d'apprentissage automatique (machine learning) ? Ce ou ces derniers est-il ou sont-ils utilisé(s) avec d'autres types d'algorithmes (ex. algorithmes déterministes) ?
- lorsqu'il repose sur un ou plusieurs algorithmes d'apprentissage automatique, quelle est la nature de(s) jeu(x) de données d'apprentissage utilisé(s) ?
- quelles sont les données prises en entrée de ce système ?
- quelles sont les données de sortie de ce système ?
- quelles conditions mises en œuvre par l'opérateur garantissent effectivement que ses services et leurs algorithmes respectent les principes de loyauté et d'équité, et qu'ils soient explicables ?
- le cas échéant, comment ce système est-il personnalisé en fonction de l'utilisateur ?
- le cas échéant, quelles sont les modalités de recours à une intervention humaine sur le système à partir de sa mise en œuvre ? Ce recours est-il le fait de personnes en interne, de personnes employées par des prestataires, ou bien les deux ?

Combien de personnes sont affectées à ces tâches ? Quelle est leur formation, ainsi que le lieu depuis lequel elles travaillent sur les systèmes affectant les utilisateurs basés en France ?
- les fausses informations font-elles l'objet d'un traitement particulier par ce système et, si oui, de quelle nature ?

Les opérateurs pourront s'ils le souhaitent recourir à des représentations graphiques afin de répondre aux questions³. Ces représentations devront faire figurer clairement l'ensemble des éléments de réponse demandés aux questions concernées.

Il convient de garder à l'esprit que des personnes physiques écrivent les algorithmes et qu'elles peuvent les optimiser pour tenir compte du comportement humain, ce qui signifie qu'un algorithme peut être conçu pour fournir de multiples sources de contenu et différents points de vue.

Les principes de conception de l'application Snapchat, y compris l'accent mis sur l'absence de viralité et une diffusion des contenus maîtrisée, la communication entre amis proches et la modération du contenu, impliquent que Snap a peu besoin d'avoir recours à des techniques accrues de modération pour lutter contre les fausses informations. Nous avons également mis en place des outils automatisés permettant de supprimer rapidement certains contenus illégaux ou interdits, notamment les discours haineux, la pornographie etc.

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

- Du contenu crédible et faisant autorité : nous valorisons les contenus de partenaires professionnels crédibles et faisant autorité dans leur domaine. Les partenaires jouissent d'une liberté éditoriale sur le contenu qu'ils publient, dans le respect de nos règles communautaires. Le contenu des médias est présenté sous forme d'éditions complètes et non d'articles individuels, et le logo du média est toujours identifié dans l'interface utilisateur.
- Des perspectives multiples et points de vue variés : cela nous évite de devenir une chambre d'écho. Le contenu proposé à un utilisateur diffère de celui d'un autre utilisateur car il est proposé sur la base des intérêts de l'utilisateur, de son usage et de son comportement dans l'application.
- L'humain au cœur des données : les membres de notre équipe de programmation sont souvent issus du monde de la presse et s'assurent que notre offre de contenus soit pluraliste.

D'une manière générale, l'application peut personnaliser l'expérience de chaque *Snapchatter*. Pour ce faire, nous opérons par déduction sur ce qu'un utilisateur peut aimer ou ne pas aimer. Par exemple, un utilisateur qui regarde beaucoup de Snaps sur les chiots dans la rubrique « Découvrir » fait peut-être partie de la catégorie des passionnés d'animaux de compagnie. Nous appelons ce type de déductions « catégories de style de vie » et nous les utilisons pour essayer de présenter le contenu et les publicités les plus pertinents pour chaque utilisateur. Si vous le souhaitez, vous pouvez toujours choisir vos propres catégories de style de vie. Toutes ces catégories sont mises à la disposition de l'utilisateur de manière transparente, et il peut toujours et à tout moment modifier, supprimer ou sélectionner ses propres catégories de style de vie, directement dans l'application (paramètres).

³ La norme ISO 5807 pourra par exemple être utilisée pour établir des logigrammes.

Par ailleurs, Snap met en pratique la minimisation des données ainsi que le principe du respect de la vie privée dès la conception du produit « *privacy-by-design* » et « *privacy-by-default* », en exigeant des spécifications de conception pour chaque produit et fonctionnalité qui collecte et/ou traite des données personnelles. La spécification doit non seulement décrire clairement les données personnelles qui seront collectées et utilisées, mais aussi leur finalité propre et leur durée de conservation (avec des limites de conservation strictes par défaut dans notre modèle de spécification). Ces documents de spécification font l'objet d'une collaboration entre les équipes produits, les ingénieurs et les juristes spécialisés en protection de la vie privée pendant tout le développement du service ou du produit. À titre d'exemple, en 2019, les avocats chargés de la protection de la vie privée des produits de Snap ont examiné plus de 1500 spécifications de conception dans le cadre de ce processus.

Snap met l'accent sur la transparence dans tous les aspects de ses communications avec ses utilisateurs - de sa politique de confidentialité à son centre de confidentialité, en passant - et c'est le plus important - par son utilisation de notifications adressées en temps utile, simples à comprendre et intégrées dans l'application. Snap ne se contente jamais d'une mention indiquée dans sa politique de confidentialité lorsqu'elle introduit de nouvelles fonctionnalités ayant un impact significatif sur la vie privée des utilisateurs, qui seront par la suite libre d'utiliser ou non ces fonctionnalités.

Les utilisateurs peuvent toujours personnaliser leur expérience et 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 Réglages de Snapchat.

Compte tenu du très faible nombre de fausses informations qui circulent sur la plateforme, nous n'avons pas développé d'algorithme traitant spécifiquement de la lutte contre la manipulation de l'information (comme détaillé dans ce questionnaire).

Algorithmes dans Découvrir (organisation, sélection et ordre)

Nos algorithmes sont généralement développés en interne et sont propriétaires ; nous ne pouvons donc pas fournir davantage de détails compte tenu de leur caractère confidentiel lié au secret des affaires. Cependant, nous pouvons partager certaines informations générales sur la façon dont le contenu est sélectionné et présenté sur la page Découvrir.

Comme nous l'avons indiqué précédemment, Discover est une plateforme fermée avec des partenaires média de confiance sélectionnés par Snap, un contenu modéré par Snap et des comptes d'influenceurs et de célébrités que nous appelons « Snap Stars » ou « Utilisateurs populaires ». La page Découvrir présente des Stories d'Amis, des Stories avec Abonnement et la section « Pour vous » (« For You »). Afin de s'assurer que le contenu non modéré est également adapté au fil « Pour vous », notre équipe modère tout le contenu qui est proposé aux *Snapchatters*.

La façon dont le contenu apparaît sur Découvrir pour chaque utilisateur dépend de certains facteurs, comme les Stories auxquelles il s'est abonné, les Stories qu'il a consultées, les Stories qui sont actuellement populaires et les préférences d'utilisateur qu'il a indiquées dans les Réglages.

Informations fournies au régulateur pour qu'il évalue les mesures de transparence des algorithmes mises en œuvre par l'opérateur

2.4. *Il est demandé à l'opérateur de décrire les **systèmes de recommandation de contenus** qu'il utilise. Pour chacun de ces systèmes, il fournira les éléments demandés au point 2.3.*

Tout d'abord, au moment de l'inscription de chaque utilisateur de Snapchat, nous l'orientons automatiquement vers notre politique de confidentialité afin qu'il puisse comprendre parfaitement quel type de données nous utilisons et la manière dont nous les utilisons, y compris dans le caractère algorithmique des contenus. Nous leur indiquons que nous utilisons les données pour : "personnaliser nos services en suggérant, entre autres, des amis, des informations de profil ou des autocollants Bitmojis (avatars personnalisés), ou en personnalisant le contenu que nous vous montrons, y compris les publicités".

Voir par exemple la section « Personnalisation » de notre page « Privacy by Product » (Confidentialité par produit), disponible à l'adresse suivante: <https://snap.com/fr-FR/privacy/privacy-by-product>.

2.5. *L'opérateur décrira également au Conseil les **systèmes de modération de contenus** qu'il utilise pour permettre notamment de lutter contre les fausses informations. Pour chacun de ces systèmes, il fournira les éléments demandés au point 2.3 et répondra à la question suivante :*
- Quelle est la performance de ce système (notamment, taux de faux positifs et de faux négatifs) ?

Veuillez vous référer à la réponse à la Question Q. 1.3.8.

La modération de contenu implique un effort collectif de nos équipes, et se concentre sur la gestion du contenu (qui pose les limites de la modération pour déterminer ce qui est approuvé et ce qui enfreint nos politiques de contenu), la sécurité de la marque (pour le contenu qui ne respecte pas les standards de notre communauté, nous appliquons strictement notre politique en le filtrant de notre flux) et l'évolutivité de la modération (efficacité et coûts opérationnels liés à la modération).

2.6. *Lorsque l'opérateur utilise un ou plusieurs algorithmes d'apprentissage automatique, il est invité à détailler sa **politique de lutte contre les biais** pouvant exister dans les données d'apprentissage.*

Pour tout algorithme impliqué dans la gestion des abus, nous confions également un certain pourcentage de tâches à des modérateurs humains pour nous assurer que l'automatisation ne soit pas déconnectée de nos politiques ni des nuances inhérentes à la modération humaine. Ce trafic est également utilisé pour recycler les algorithmes afin d'améliorer leur précision et d'éviter les biais algorithmiques. Nous procédons également à de l'échantillonnage aléatoire.

L'équipe de recherche (qui se distingue de l'activité d'ingénierie de base et dont le travail n'est généralement pas directement intégré à la production, mais qui a contribué aux efforts de lutte contre les abus) dispose de sa propre méthodologie :

1. former un classificateur pour une tâche X
2. utiliser un deuxième classificateur (déjà formé) capable de reconnaître certaines catégories (par exemple si la phrase est écrite en anglais standard ou en dialecte AAE)
3. Évaluer l'équité du classificateur X envers les catégories détectées en au point 2.

Vous trouverez plus de détails dans cette étude scientifique, étant précisé que cette approche est commune dans les procédures de *machine learning* <https://arxiv.org/pdf/2010.12864.pdf>.

Informations fournies aux utilisateur⁴

2.7. L'opérateur est invité à indiquer comment il décrit à ses utilisateurs les systèmes de recommandation qu'il utilise. Pour chacun de ces systèmes, il précisera comment il leur apporte des informations sur les éléments décrits au point 2.3 et répondra aux questions suivantes :

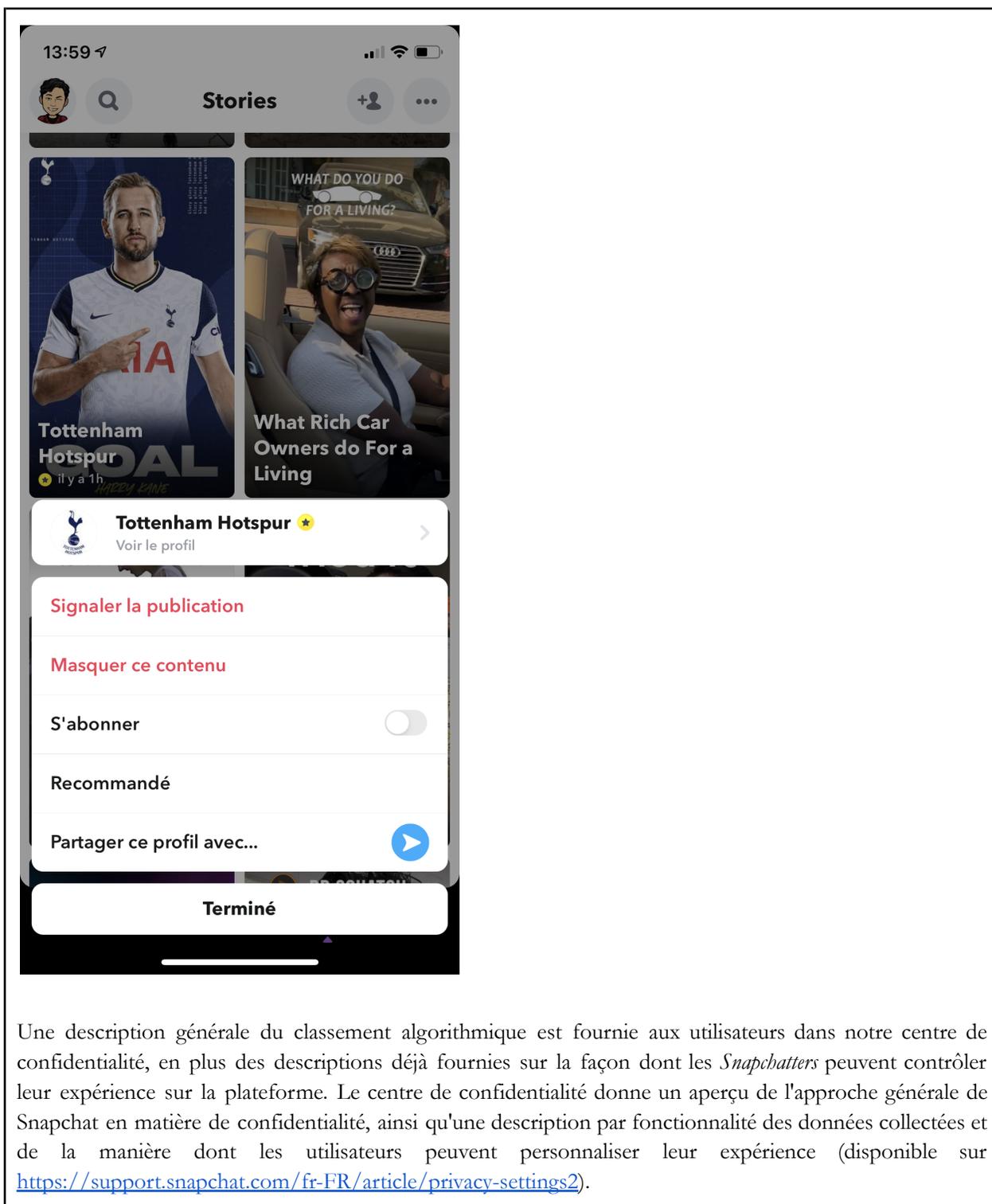
- *quel(s) réglage(s) peu(ven)t être effectué(s) par l'utilisateur pour en changer les paramètres ? Combien d'utilisateurs y ont eu recours en 2020 et à quelle fréquence ?*
- *comment ces réglages sont-ils rendus facilement accessibles et compréhensibles ?*
- *comment les utilisateurs sont-ils informés des mises à jour de ces systèmes de recommandation affectant significativement la manière dont les contenus leur sont recommandés ?*
- *de quelle manière les utilisateurs peuvent-ils demander des informations supplémentaires sur ces systèmes de recommandation ?*
- *comment l'utilisateur bénéficie-t-il d'une explication locale⁵ sur les résultats de ces systèmes de recommandation qui l'affectent ?*

Le Conseil sera attentif à toute précision visant à expliquer comment l'opérateur modifie le niveau et la nature de ces informations en fonction de ses utilisateurs (mineurs, personnes en situation de handicap, développeurs, etc.).

Nos utilisateurs peuvent ajuster leurs réglages pour voir le contenu qui leur convient le mieux. Cela est expliqué sur la page "Réglages" de l'application où l'utilisateur peut déterminer s'il souhaite ajuster ses paramètres. Les utilisateurs sont également en mesure de gérer dans Découvrir les chaînes auxquelles ils sont abonnés, et peuvent choisir de masquer des Stories directement dans la page Découvrir, comme détaillé dans la capture d'écran ci-dessous :

⁴ Dans ces questions visant à comprendre comment des informations sont données aux utilisateurs sur la transparence des algorithmes utilisés par l'opérateur, il n'est pas attendu que le format ou le niveau de profondeur de l'information soit le même que celui apporté au régulateur aux points 3, 4, 5, 8 et 9. En revanche, les éléments décrits au point 2. étant centraux dans le renforcement de l'esprit critique des utilisateurs vis-à-vis des contenus qui leur sont proposés, le Conseil attend que l'opérateur indique comment il communique à leur sujet auprès de ses utilisateurs.

⁵ Au sens défini dans la notion d'explicabilité telle qu'établie au début du questionnaire.



Une description générale du classement algorithmique est fournie aux utilisateurs dans notre centre de confidentialité, en plus des descriptions déjà fournies sur 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 (disponible sur <https://support.snapchat.com/fr-FR/article/privacy-settings2>).

2.8. L'opérateur indiquera également la manière dont il décrit à ses utilisateurs les systèmes de modération qu'il utilise notamment dans le cadre de sa lutte contre la diffusion des fausses informations. Pour chacun de ces systèmes, il précisera comment il leur apporte des informations sur les éléments décrits au point 2.3 et répondra aux questions suivantes :

- quelle est la performance de ce système (notamment, taux de faux positifs et de faux négatifs) ?
- quel(s) réglage(s) peu(ven)t être effectué(s) par l'utilisateur pour en changer les paramètres ?
- comment ces réglages sont-ils rendus facilement accessibles et compréhensibles ?
- comment les utilisateurs sont-ils informés des mises à jour de ces systèmes de modération affectant significativement la manière dont les contenus sont modérés ?
- de quelle manière les utilisateurs peuvent-ils demander des informations supplémentaires sur ces systèmes de modération ?
- comment l'utilisateur bénéficie-t-il d'une explication local⁶ des résultats des systèmes de modération qui l'affectent ?

Nos règles communautaires décrivent la manière dont nous modérons le contenu. Après avoir examiné le contenu signalé, notre équipe ne prendra pas de mesure si le contenu n'enfreint pas ces règles. Par exemple, nous recevons de nombreux signalements d'utilisateurs portant sur des contenus qu'ils n'aiment pas. Si, après examen, le contenu signalé enfreint effectivement nos règles, nous le supprimerons et nous avertirons l'utilisateur à l'origine du contenu que son comportement est contraire à nos règles communautaires. Si un utilisateur continue d'utiliser Snapchat pour diffuser de fausses informations, ou commet d'autres violations de nos règles, nous pourrions supprimer son compte. Notre rapport de transparence expose les mesures que nous prenons à l'encontre des utilisateurs et du contenu signalé.

Les outils automatisés de modération de contenu que nous utilisons sont conçus pour imiter les décisions des modérateurs humains.

Systèmes spécifiques à la lutte contre la manipulation de l'information

2.9. L'opérateur décrira **tout système qu'il a développé ou utilise spécifiquement dans le cadre de la lutte contre la manipulation de l'information**, notamment contre les campagnes coordonnées de désinformation. Il indiquera s'il informe les utilisateurs de l'existence de ces systèmes.

Pour chacun de ces derniers, il fournira les éléments demandés au point 2.3 et indiquera la performance de ce système (notamment, taux de faux positifs et de faux négatifs).

2.10. L'opérateur indiquera dans quelle mesure il a tenu compte **des recommandations du Conseil** énoncées dans son bilan relatif à l'exercice 2019 à compter de sa publication en juillet 2020.

Veillez vous référer à notre Réponse à la Question Q. 1.7.

⁶ Id.

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

Les indications suivantes s'appliquent 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

3.1. L'opérateur indiquera au Conseil si les sources des contenus issus d'entreprises et d'agences de presse et de Services de communication audiovisuelle (ci-après SCA) sont **identifiées et visibles** sur son service et, le cas échéant, s'il s'agit d'un dispositif qui leur est **spécifique**.

Il est invité à expliquer les moyens mis en œuvre pour procéder à cette identification, à savoir :

- **qui la réalise** (l'opérateur, les entreprises, agences de presse et SCA ou un tiers) et, dans l'hypothèse où l'opérateur ne la réalise pas directement, s'il vérifie et s'assure de la fiabilité de l'identification des sources (avant ou après publication) ;
- **par quels moyens**, sur la base de quelles sources et selon quels critères l'identification des sources et/ou la vérification de la fiabilité de l'identification est réalisée par l'opérateur et comment ces critères sont portés à la connaissance de l'utilisateur ;
- comment l'identification des sources est **portée à la connaissance de l'utilisateur** sur le service et si des éléments complémentaires d'information et de contexte lui sont apportés ;
- les moyens humains, technologiques et financiers déployés par l'opérateur afin d'assurer l'identification, la visibilité et la mise en avant des informations provenant de ces sources dans les résultats des moteurs de recherche, des fils d'actualité et autres espaces de diffusion de contenus opérant par classement algorithmique, y compris, le cas échéant, au sein de groupes d'utilisateurs fermés ;
- si ces mêmes moyens sont **susceptibles d'évoluer en fonction de l'actualité** et, le cas échéant, dans quelle mesure. L'opérateur pourra donner des exemples de cas où l'actualité l'a conduit à faire évoluer ces moyens.

Le contenu issu de médias et d'éditeurs professionnels est présenté sur Snapchat dans une section fermée de la plateforme appelée « Découvrir » (« Discover »). En plus d'autres contenus, cette section contient du contenu créé par des éditeurs média professionnels de confiance qui ont été sélectionnés et approuvés par l'équipe éditoriale de Snapchat. Ce contenu est présenté comme dans un kiosque à journaux : les logos des médias sont explicitement visibles sur les couvertures (vignettes) des éditions, et restent visibles tout au long des Snaps en haut de l'écran de l'utilisateur. Le contenu est recommandé en fonction des intérêts personnels des utilisateurs, y compris les « abonnements ». Le contenu de type piège-à-clic (*click-bait*) est interdit par nos règles éditoriales. Les éditions des éditeurs sont présentées comme des éditions intégrales (les articles ou shows/émissions doivent être regardés par ordre chronologique), et non comme des articles individuels.

L'approche de Snapchat en matière de contenu diffère de celle des autres plateformes. Grâce à la façon dont nous avons construit notre plateforme de contenu « Découvrir », nous avons pu lutter contre la propagation de fausses informations. Nos produits de contenus premium sur « Découvrir » ne sont disponibles que grâce à une collaboration directe entre Snap et ses partenaires. 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 collaborons.

Les partenaires avec lesquels nous contractons bénéficient d'un accès à des outils de création et de gestion de contenu. Ces outils permettent aux partenaires de créer, d'éditer et de télécharger du contenu spécifique ainsi que de mieux comprendre, grâce à l'analyse, les performances de leurs contenus éditoriaux.

L'une des valeurs qui anime notre moteur de classement est la promotion de contenus fiables et crédibles provenant de nos partenaires de confiance. Nous pensons que la diffusion des contenus de nos utilisateurs va de pair avec un haut degré de responsabilité. Nos partenaires sont tenus de vérifier leur contenu et de s'assurer qu'il n'est ni faux, ni trompeur.

Nous travaillons également avec des partenaires soigneusement sélectionnés (comme BRUT ou Konbini) qui peuvent modérer les Snaps publics publiés par les utilisateurs afin d'ajouter de la crédibilité et de replacer les actualités dans leur contexte.

Dans Découvrir, les logos des marques des partenaires média (par exemple Le Monde) sont visibles sur des couvertures (vignettes) conçues pour attirer l'attention de l'utilisateur. Nous pensons qu'une identification plus forte contribue à promouvoir la crédibilité.

Le logo et/ou le nom de la source, ainsi qu'un lien vers le profil de la source, sont disponibles où que le contenu apparaisse, que ce soit dans « Recherche », « Pour vous » ou dans une autre section.

Les Shows (émissions) créés par nos éditeurs sont mieux classés dans les recherches et apparaissent souvent prioritairement par rapport au contenu des utilisateurs populaires dans « Pour vous ». Snap exige des éditeurs qu'ils se conforment à des règles strictes et précises. Ces règles sont mises en œuvre par des agents personnes physiques. Ces agents identifient le contenu qui enfreint ces règles et le masquent ou suppriment son classement prioritaire du fil de contenu visible dans « Pour vous » des utilisateurs. 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 en cas de contenu jugé contraire à nos politiques, y compris à nos règles communautaires (<https://www.snap.com/fr-FR/community-guidelines>).

Prise en compte des démarches de labellisation

3.2. Les moyens mis en œuvre par l'opérateur pour tenir compte des *démarches de labellisation et de bonnes pratiques*, notamment celles réalisées en tout ou partie par les entreprises et agences de presse et les SCA (ex. : Trust Project, Journalism Trust Initiative), seront également portés à la connaissance du Conseil.

*L'opérateur décrira les **moyens mis en œuvre** (ex. : démarches réalisées par ses soins ou par des tiers, intervention directe de ces tiers sur la plateforme ou non, financement, partenariats, mise en avant sur le service (et comment), informations aux utilisateurs, etc.). Il précisera, le cas échéant, sur quels critères sont sélectionnées les démarches de labellisation concernées.*

L'opérateur précisera si ces démarches de labellisation et de bonnes pratiques constituent un des moyens d'identification des sources et de mise en avant de contenus issus d'entreprises et d'agences de presse et de SCA et de quelle manière, et/ou un des moyens de choisir d'éventuels partenaires œuvrant au fact-checking.

Snap gère le programme dit *Tiers de Confiance* (ou « *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 Snapchat. Ce service de qualité est contrôlé 24 heures sur 24 et 7 jours sur 7 par l'équipe « Confidentialité et Sécurité » de Snapchat.

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

3.3. Le Conseil s'interroge sur les dispositifs de **fact-checking** utilisés par l'opérateur, qu'ils soient confiés à des tiers ou mis en œuvre par lui-même. Il souhaite connaître les partenariats et/ou actions (en interne et externe) existants en la matière, notamment aux fins de lutter contre les hypertrucages ou deepfakes.

Il souhaite connaître en particulier :

- les **critères retenus par l'opérateur pour sélectionner** les fact-checkers (notamment : fiabilité, nationalité et langue de travail sur une zone géographique donnée) ;
- les **critères, sources et outils** qu'il utilise, d'une part, s'il procède en interne au fact-checking et, d'autre part, si une décision de modération est prise en interne à l'issue de cette vérification ;
- si tous les **types de contenus** (informatifs, publicitaires, écrits, vidéos, sons) sont susceptibles de faire l'objet d'une vérification ou, dans le cas contraire, pourquoi ;
- si les contenus sur la plateforme pouvant être soumis aux fact-checkers sont laissés à leur libre appréciation ou sélectionnés (le cas échéant, selon quels critères et quelle méthode) ;
- l'ensemble des **moyens technologiques** mis à la disposition des fact-checkers dans leur travail de vérification de l'information (notamment : outil de suggestion des contenus à vérifier et, le cas échéant, selon quels critères ; mise à disposition d'une API) ;
- s'il existe un **dispositif d'urgence** en cas d'actualité particulière. L'opérateur pourra donner des exemples de cas d'actualité et précisera, le cas échéant, si la crise sanitaire en 2020 a eu un impact sur ce dispositif ;
- le **budget** consacré aux partenariats et/ou actions concourant à la vérification de l'information sur ses services (notamment les contreparties financières versées aux fact-checkers, le cas échéant).

3.4. S'il recourt à des fact-checkers **tiers**, l'opérateur est invité à indiquer au Conseil :

- la **liste** des structures de fact-checking partenaires de la plateforme ;
- la **liste et le nombre de contenus suggérés par la plateforme** aux tiers pour vérification et la liste et le nombre de contenus proposés à la vérification par les tiers,
- la **liste et le nombre de contenus ayant fait l'objet d'une intervention** des fact-checkers tiers sur la plateforme (décliné par types d'interventions possibles),
- la nature des **liens juridiques** unissant l'opérateur aux fact-checkers tiers,
- la nature des **contreparties** qui leur sont offertes.

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

3.6. L'opérateur décrira les **procédures et les moyens technologiques et humains permettant de traiter les contenus** dont le fact-checking (interne ou externe) a montré qu'ils contenaient ou étaient susceptibles de contenir une fausse information, y compris au sein des groupes d'utilisateurs fermés. A cet égard, il est invité à préciser :

- la **chaîne de décision** amenant l'opérateur à agir sur un contenu ou un compte ;

- les **mesures prises** le cas échéant (ex. : déréférencement, réduction de la visibilité des contenus, limitation de certaines fonctionnalités, messages apposés aux contenus) ;
- la **gradation** de ces mesures et les **critères d'application** de chacune d'entre elles ;
- si les mesures et leur gradation sont susceptibles de **varier selon le type et/ou à la nature du contenu vérifié** et, le cas échéant, pourquoi ;
- si l'intervention des **fact-checkers** est directement **visible sur le service** ;
- si des **voies de recours** internes sont offertes aux auteurs des contenus concernés et, le cas échéant, si ces recours sont traités par l'opérateur ou par le tiers partenaire.

3.7. Le Conseil souhaite également connaître les modalités de **communication** de l'opérateur, ou les moyens qu'il met en œuvre pour permettre aux **fact-checkers** de communiquer, auprès :

- des utilisateurs ayant publié un tel contenu, l'ayant partagé ou ayant interagi avec lui (ex. : commentaire, like) ;
- de l'ensemble des utilisateurs en rendant visible, sur le service, les interventions sur un contenu ou un compte liées au **fact checking**.

Réponse aux Questions 3.2 à 3.7

La section « Découvrir » contient exclusivement du contenu provenant d'entreprises média de premier plan (par exemple Le Monde etc.), du contenu modéré issu des comptes publics de personnes connues (le Président Emmanuel Macron, le premier Ministre Boris Johnson, etc.) et de créateurs approuvés, ainsi que des Stories issues de Snaps d'utilisateurs sélectionnées par notre équipe interne dédiée à l'actualité ou par nos partenaires professionnels. Grâce à ce contrôle strict, les risques de diffusion de fausses informations sur Snapchat sont faibles. Dans cette optique, il ne nous a pas paru nécessaire de faire appel à des « *fact-checkers* » externes, puisque cette vérification est déjà effectivement réalisée soit par l'éditeur professionnel, soit en interne par les équipes de Snap.

Au vu de la nature même de Découvrir (voir ci-dessus), les « *fact-checkers* » apparaissent moins nécessaires que sur d'autres plateformes.

Découvrir constitue la meilleure illustration de la promotion par Snap d'un contenu fiable :

- Tous les contenus publics qui apparaissent sur la section Découvrir de Snapchat sont sélectionnés, c'est-à-dire qu'ils sont généralement « fiables ». 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 de Découvrir 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 média 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 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 média sur Découvrir acceptent aussi des lignes directrices supplémentaires (règles communautaires), y compris l'obligation pour leur contenu d'être correctement vérifié et exact.

Nous avons volontairement créé Découvrir comme une plateforme fermée, afin de garantir cet environnement et de 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. 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 médias de confiance pour fournir des informations sur notre plateforme, et nous disposons de lignes directrices dédiées pour nos éditeurs partenaires sur Découvrir, 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 ces 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 interromperions 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.

Autres questions

3.8. Le Conseil souhaite connaître les éventuels freins (notamment technologiques, économiques, et/ou juridiques) rencontrés par l'opérateur dans l'identification des sources des contenus issus des entreprises et des agences de presse et SCA, la mise en avant de leurs contenus, la participation à la vérification de l'information ou la collaboration avec les fact-checkers tiers.

L'opérateur pourra également mentionner les limites qu'il voit aux approches de labellisation et de bonnes pratiques visées au point 3.2.

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.

Grâce à cela, nous n'avons rencontré que 4 cas confirmés de violation de nos politiques contre les fausses informations en 2020. Ce chiffre est infime par rapport à nos 16 millions d'utilisateurs actifs quotidiens en France. Ce nombre aussi faible de fausses nouvelles s'explique par la nature, la conception et le modèle économique de la plateforme, qui est construite comme un « antidote » au modèle traditionnel de réseaux sociaux (voir la question 1.4 ci-dessus).

Nous pensons également qu'un partenariat avec des *fact-checkers* tiers pourrait être bénéfique, et nous espérons voir plus d'institutions fiables dans le monde s'intéresser à la question des fausses informations.

3.9. L'opérateur indiquera dans quelle mesure il a tenu compte des recommandations du Conseil énoncées dans son bilan relatif à l'exercice 2019 à compter de sa publication en juillet 2020.

Veillez vous reporter à notre réponse à la Question Q. 1.7.

3.10. L'opérateur est invité à indiquer au Conseil l'impact de la crise sanitaire sur la promotion des contenus issus d'entreprises SCA et les mesures de fact-checking et notamment, les adaptations qu'il a pu être amené à lui apporter en conséquence.

Pendant la crise sanitaire, les dispositifs de signalement que nous mettons à disposition des utilisateurs de Snapchat perdurent (a). Snapchat a par ailleurs mis en œuvre certains changements détaillés ci-après (b).

(a) Les utilisateurs peuvent signaler du contenu par le biais de notre dispositif intégré dans l'application ou peuvent nous contacter *via* notre site d'assistance. Ces possibilités existaient déjà avant la pandémie et continuent d'exister aujourd'hui. Nous nous sommes naturellement adaptés pour rester parfaitement informés sur les fausses informations concernant la pandémie, les prétendus remèdes contre le virus et, maintenant, de potentielles fausses informations concernant les différents vaccins. Il nous apparaît primordial de rester très informés sur ces fausses informations et de nous assurer que nos équipes en ont connaissance afin de gérer le contenu en conséquence.

En réponse à la pandémie, nous continuons à être particulièrement vigilants sur notre responsabilité de fournir à notre communauté des nouvelles et des informations fondées sur des faits. Comme indiqué ci-dessus, notre plateforme de contenu Découvrir est organisée par curation et nous ne travaillons qu'avec

un ensemble restreint de partenaires et de gouvernements, parmi lesquels certains des organismes de presse les plus fiables au monde.

(b) Notre priorité est de veiller à ce que Snapchat fournisse des informations de haute qualité sur la covid-19 et de ne pas offrir un fil d'actualité ouvert où des éditeurs ou des individus non vérifiés auraient la possibilité de diffuser de fausses informations. Nos lignes directrices 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 une chaîne dédiés dans Découvrir, nous avons établi un partenariat avec 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 chaînes Découvrir issues des communautés les plus implantées dans le monde.

En outre, nous faisons notre maximum pour fournir à notre communauté des informations fondées sur des éléments factuels et actualisés, ainsi que des ressources pour leur sécurité et leur santé mentale. Dans les premiers jours de la pandémie de covid-19, Snap a lancé plusieurs filtres et *Lens* (lentilles) offrant des conseils de sécurité et promouvant les meilleures pratiques en la matière, qui ont rapidement touché des centaines de millions de personnes dans le monde entier.

Nous avons ajouté une nouvelle fonctionnalité, appelée « Here For You », qui offre un soutien proactif intégré dans l'application aux *Snapchatters* susceptibles de traverser une crise de santé mentale ou émotionnelle, ou aux *Snapchatters* curieux d'en savoir plus sur ces questions et sur la façon dont ils peuvent aider leurs amis.

Nous avons également publié plus de 700 éditions dans Découvrir présentant une couverture de dernière minute sur la covid-19 à l'initiative de nos partenaires de contenu, de nos équipes d'information internes et des organisations comme le CDC et l'OMS. Dans la mesure où plus de la moitié de la génération Z aux États-Unis regarde les contenus d'actualité sur Découvrir, nous estimons qu'il est particulièrement important pour nous d'éduquer notre public avec des informations choisies et fiables en ces temps de crise.

4. Lutte contre les comptes propageant massivement de fausses informations

Les indications suivantes s'appliquent 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.

Identification et définition des pratiques de diffusion massive de fausses informations

4.1. Le Conseil souhaite savoir comment l'opérateur appréhende la notion de comptes propageant massivement de fausses informations, telle que mentionnée à l'article 11 de la loi du 22 décembre 2018.

L'opérateur déclinera les différentes catégories de pratiques de diffusion massive de fausses informations qu'il identifie sur son service, en en donnant leur définition, leur seuil éventuel de déclenchement et en les caractérisant, en égard notamment de leur origine (type d'acteur), de leurs moyens de diffusion (exclusivement algorithmiques et/ou humains) et de la nature de l'objectif poursuivi par leurs émetteurs (exclusivement économique et/ou politique, ou autre).

Il indiquera, pour chacune de ces catégories, s'il s'agit d'une pratique interdite dans les règles d'utilisation de la plateforme.

Grâce à notre plateforme de contenu Découvrir, nous avons pu nous protéger contre la diffusion de fausses informations. Il n'y a pas de profil public des utilisateurs, il n'y a pas d'outil ou de critère de vanité ou d'indicateur visible par le public (J'aime/Likes, partages, liens, etc.), les fonctionnalités de messagerie sont destinées aux communications en mode direct entre deux personnes ou en petits groupes. Il n'y a donc pas de diffusion massive de fausses informations sur Snap.

Snap a tenu des réunions avec des représentants de la Commission Européenne (DG JUST, DG HOME) et du gouvernement britannique (*Foreign & Commonwealth Office - FCO*) pour expliquer en détail notre modèle opérationnel et les risques que présentent les fausses informations et la désinformation sponsorisée par un état.

La Commission européenne et le FCO britannique se sont tous deux déclarés convaincus que Snapchat n'était pas une plateforme adaptée à la diffusion de fausses informations, et rien n'indique à ce jour que des gouvernements étrangers aient tenté d'utiliser notre plateforme pour diffuser de fausses informations. Sur ce dernier point, nous avons mené début 2018 un audit interne approfondi à la recherche de contenus associés au gouvernement russe ; aucun contenu portant de fausses informations n'a été identifié. De façon générale, tout compte qui enfreindrait nos conditions de service ou nos règles communautaires ferait l'objet d'une intervention rapide à son encontre.

Comme expliqué plus en détail ci-dessus, notre application a été rigoureusement conçue pour éviter la diffusion massive de fausses informations. Nos politiques, y compris nos règles communautaires, interdisent la diffusion volontaire de fausses informations. Outre les solides protections résultant de la conception même de l'application (mentionnées dans plusieurs réponses précédentes), les mesures prises par Snap en cas de fausses informations comprennent la suppression rapide du contenu incriminé 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. L'opérateur est invité à fournir :

- *la liste ou, à défaut, des **exemples d'opérations de propagation de fausses informations soutenues par des entités étatiques ou institutionnelles** observées et identifiées sur son (ses) service(s) et les mesures prises à leur égard ; il indiquera le cas échéant, s'il coopère avec les autorités ainsi que d'autres opérateurs dans le but de mieux détecter et de prévenir ces phénomènes ;*
- *des **exemples de pratiques individuelles ou à l'échelle d'un petit groupe d'utilisateurs** de comportements trompeurs et abusifs massifs (ex. : multidiffusion, publication à répétition) observées et identifiées sur sa (ses) plateforme(s), ainsi que les moyens mis en œuvre pour y lutter.*

Non applicable, en raison du mode de conception de la plateforme Snapchat et de l'absence de diffusion de masse, comme indiqué ci-dessus. Snapchat n'a pas rencontré ce type de situation à ce jour.

Moyens de lutte contre la diffusion massive de fausses informations

4.3. Pour chacune des catégories de pratiques de diffusion massive de fausses informations identifiées en 4.1, l'opérateur indiquera les **mesures et moyens (humains et technologiques) qu'il a pris afin de:**

- **détecter** de telles pratiques et les comptes qui y ont recours,
- **faire obstacle** à ces pratiques et à ces comptes s'agissant :
 - *de la procédure d'intervention à leur encontre : mesures prises (avertissement, restriction des droits, mise en quarantaine, déréférencement, suppression du contenu, du compte, etc.), gradation de ces mesures et critères déclenchant l'application de chacune d'elle ;*
 - *de la lutte contre l'accélération et la viralité des contenus contenant de fausses informations.*

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

4.4. Le Conseil souhaite disposer de **données chiffrées sur les mesures de lutte prises par l'opérateur contre les comptes propageant massivement de fausses informations, notamment, pour chacune des catégories de pratiques identifiées en 4.1:**

- le nombre de comptes propageant massivement de fausses informations détectés sur le service, et leur proportion dans le nombre total de comptes sur le service ;
- les **origines de cette détection** (moyens technologiques, signalements par des tiers, signalements en interne) et leur répartition ;
- parmi les comptes détectés, le **nombre de ceux ayant fait l'objet d'une instruction** par l'opérateur et de **ceux ayant fait l'objet de mesures** (en déclinant par types de mesures) ;
- les **moyens financiers et humains** mis en œuvre par l'opérateur pour lutter contre les comptes propageant massivement de fausses informations ;
- dans le cas où ces comptes auraient fait l'objet de recettes, la rémunération tirée de ces comptes et celle reversée à leur titulaire en 2020, ainsi que les mesures éventuellement prises pour limiter cette rémunération pendant et après la phase d'instruction d'un compte détecté ou signalé comme propageant massivement de fausses information.

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

Information aux utilisateurs sur la diffusion massive de fausses informations

4.5. L'opérateur décrira les **dispositifs de suivi et de statistiques des mesures de détection et de traitement** des pratiques et comptes propageant massivement de fausses informations, et notamment :

- si ces dispositifs sont à usage interne ou s'ils sont **accessibles aux utilisateurs, et selon quels forme et niveau de détail** ;
- les **informations** suivies et mesurées : définition(s) de comptes propageant massivement de fausses informations ; nombre de comptes détectés, signalés, traités (par des moyens humains d'une part, et des outils technologiques d'autre part) ; caractéristiques des comptes et des contenus concernés ; actions prises ; résultats ;
- les modalités de consultation et de téléchargement de **ces informations par l'utilisateur**.

En cas de signalement d'une fausse information par un utilisateur, nous prenons les mesures appropriées.

Snap interdit la diffusion de fausses informations (y compris la manipulation de médias) malintentionnées ou susceptibles de causer un préjudice.

Vous trouverez plus de détails dans nos règles communautaires, disponibles à l'adresse suivante : <https://www.snap.com/fr-FR/community-guidelines>

Concernant les chiffres exacts, nous vous invitons à vous référer à nos réponses ci-dessus.

4.6. L'opérateur indiquera si le service dispose d'un **espace d'information renseignant les utilisateurs sur les pratiques liées à l'utilisation de comptes propageant massivement de fausses informations susceptibles d'entraîner une intervention de sa part** et, le cas échéant :

- l'information des utilisateurs sur les **risques encourus** notamment dans les cas de création de comptes dans des volumes anormaux, de partage de contenus à des fréquences anormales et d'utilisation de renseignements faux, volés ou trompeurs ;
- la **manière dont cette information est fournie** à l'utilisateur : quelle forme prend-elle, où se situe-t-elle, est-elle claire et accessible ? est-elle notifiée à l'utilisateur (ou nécessite-t-elle une recherche de sa part) et, le cas échéant, dans quelle(s) situation(s) ?

Nous communiquons clairement sur nos lignes directrices et nos conditions de service pour l'utilisation de notre plateforme. En plus de nos règles communautaires claires et intelligibles (disponibles sur [: https://www.snap.com/fr-FR/community-guidelines](https://www.snap.com/fr-FR/community-guidelines)), nous fournissons également des conditions d'utilisation du service plus détaillées (disponibles sur [:https://snap.com/fr-FR/terms](https://snap.com/fr-FR/terms)).

Autres questions

4.7. *L'opérateur décrira les mesures et moyens mis en œuvre pour détecter les pratiques de **deepfakes**, les comptes y ayant recours et lutter contre ce phénomène.*

*Il indiquera le **nombre** de deepfakes propagés à des fins de diffuser de fausses informations sur son service dont il a eu connaissance.*

4.8. *L'opérateur décrira les pratiques dont il a eu connaissance 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...) ainsi que les mesures mises en œuvre à leur encontre.*

*Il indiquera, parmi les contenus connus comme contenant de fausses informations, **combien ont bénéficié de ces pratiques**.*

Comme détaillé ci-dessus, la conception de Snapchat rend très difficile la diffusion de fausses informations sur notre plateforme. Notre page Découvrir présente des éditeurs et des marques de média de confiance, un contenu hautement sélectionné et organisé provenant de comptes officiels et de comptes populaires. Si un compte populaire publiait un contenu trompeur ou préjudiciable dans sa Story, nos mesures de modération empêcheraient une diffusion plus large de ce contenu.

Les signalements relatifs à la diffusion par un individu de fausses informations en dehors de notre page Découvrir sont rapidement traités par nos équipes.

4.9. *L'opérateur indiquera dans quelle mesure il a tenu compte **des recommandations du Conseil** énoncées dans son bilan relatif à l'exercice 2019 à compter de sa publication en juillet 2020.*

Veillez vous référer à notre réponse à la Question 1.7.

4.10. *L'opérateur indiquera, dans le cadre de sa lutte globale contre les comptes propageant massivement de fausses informations, les nouvelles **mesures initiées en 2020**, notamment celles prises dans le cadre de la crise sanitaire mondiale.*

Veillez vous référer à notre réponse à la Question 3.10.

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. Focus: lutte contre la diffusion de fausses informations en lien avec des communications commerciales

Il est apparu au Conseil, à l'aune de ses travaux et de l'examen des déclarations concernant l'exercice 2019, qu'un des points cruciaux de la lutte contre la diffusion des fausses informations était l'action sur les liens entre flux financiers et fausses informations. En effet, ces dernières peuvent être, d'une façon ou d'une autre, source de revenus pour différents types d'acteurs ; elles peuvent également bénéficier, directement ou non, des mécanismes de diffusion des contenus commerciaux. En conséquence, ces liens peuvent 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. Ainsi, le Conseil a élaboré les questions suivantes avec son Comité d'experts sur la désinformation en ligne afin d'interroger les opérateurs sur l'ampleur réelle de ces liens ainsi que sur les moyens de lutte mis en œuvre en la matière.

Catégorisation des communications commerciales

5.1. *L'opérateur est invité à commenter les catégories d'annonces publicitaires, de contenus sponsorisés et de contenus en partenariat telles qu'établies au début du questionnaire⁷.*

Il indiquera si cette catégorisation lui semble pertinente, notamment au regard du modèle de sa plateforme et dans le cadre particulier de la lutte contre les fausses informations. Il pourra identifier les limites éventuelles de ces catégories et formuler des propositions pour que les catégories retenues couvrent l'ensemble de l'activité publicitaire sur son service.

Chaque publicité diffusée sur Snapchat est clairement identifiée et labellisée comme « publicité » ou comme contenu « sponsorisé ». Les publicités mettent également en évidence le nom de la marque de l'annonceur ou sponsorisant le contenu.

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.

Le nom de l'annonceur qui a financé la campagne et le nom de la campagne sont fournis à l'utilisateur.

Annonces publicitaires et contenus sponsorisés

5.2. *L'opérateur décrira les modalités de vente des espaces publicitaires sur le service : le placement est-il régi par des mécanismes programmatiques ? Repose-t-il sur un modèle fermé propre à la plateforme ou sur un modèle d'enchères en temps réel (real-time bidding) ? Le cas échéant, quels sont les intermédiaires utilisés ?*

Comme indiqué ci-dessus, tout contenu sponsorisé est identifié comme tel. Les annonceurs ne peuvent pas promouvoir directement leur contenu organique, ils peuvent seulement utiliser le contenu sponsorisé pour gagner en visibilité sur la plateforme en achetant des publicités.

⁷ Voir définitions des notions en début de questionnaire.

Les annonceurs peuvent utiliser notre plateforme de publicité en libre-service, « Ads Manager », ou travailler avec un partenaire utilisant notre interface de programmation de Publicité, « Ads API », pour diffuser des publicités sur la plateforme.

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.

Les annonceurs doivent fournir des publicités dans un format approuvé par Snap :

- *Snap Ads* : Une seule image ou vidéo affichée en plein écran.
- *Publicités* : Une publicité vidéo affichée en plein écran qui ne peut être ignorée pendant les 6 premières secondes.
- *Filtres* : Des superpositions uniques que les utilisateurs peuvent choisir d'appliquer sur un Snap pour un effet artistique.
- *Lenses* : Lentilles de réalité augmentée qui permettent aux utilisateurs de placer des objets et des personnages animés dans leurs Snaps.
- *Story Ads let* : couvertures (vignettes) revêtues d'une marque dans la section Découvrir de Snapchat qui s'ouvrent sur une collection de 1 à 20 images ou vidéos.
- *Product Ads* : Formats publicitaires achetables qui permettent aux utilisateurs de choisir parmi plusieurs produits, services ou expériences proposés par l'annonceur.

Snap facture les annonceurs par impression, quelle que soit l'interaction avec l'utilisateur.

5.3. *L'opérateur précisera les moyens, technologiques ou autres, permettant aux annonces publicitaires et contenus sponsorisés de gagner en visibilité sur la plateforme (possibilité pour l'annonceur de viser un public en particulier, de choisir le format du contenu, avec quelles contreparties ; impact des interactions des utilisateurs avec le contenu ; etc.).*

Veillez vous référer à notre réponse à la Question Q.5.2.

5.4. *Dans la mesure où le dispositif différencierait du dispositif de signalement général, le Conseil souhaite que l'opérateur lui indique si l'utilisateur peut **signaler une annonce publicitaire ou un contenu sponsorisé pour « fausse information » (ou équivalent).***

Le cas échéant, l'opérateur indiquera :

- les **modalités** de signalement (*faut-il avoir un compte pour signaler le contenu, quelle est la démarche pour signaler ces contenus...*) ;
- dans quelles conditions et comment l'opérateur **interagit avec les utilisateurs** concernés par ce signalement (*auteur du contenu et auteur du signalement ; information concernant l'avancée du traitement du signalement*) ;
- le **nombre de signalements reçus et traités** relatifs à des annonces publicitaires et contenus sponsorisés et, plus particulièrement, aux contenus promus d'information se rattachant à un débat d'intérêt général, en indiquant, quand cela est possible, le nombre de signalements pour des raisons de « fausses informations » ;
- les **moyens (technologiques, humains et financiers) mis en œuvre pour traiter ces signalements.**

Veillez vous référer à notre réponse en Section 1.1 pour une description détaillée de la manière dont les utilisateurs peuvent signaler un contenu.

5.5. L'opérateur précisera les mesures mises en œuvre afin de détecter les annonces publicitaires et contenus sponsorisés porteurs de fausses informations, notamment :

- l'existence d'un **système de vérification** des annonceurs, interne ou externe ;
- les **mesures de détection et de lutte contre la publicité en faveur de sites internet spécialisés dans les fausses informations**. L'opérateur indiquera s'il dispose d'informations sur ces sites, notamment sur leur origine et leur typologie, ainsi que sur la durée d'apparition d'une publicité pour ces sites sur le service ;
- les moyens de **coopération avec d'autres opérateurs ou des organismes externes** afin de détecter et de signaler les annonces publicitaires et contenus sponsorisés porteurs de fausses informations et s'il s'appuie-t-il sur des normes ou des organismes externes afin d'identifier et de signaler ces contenus

Tous les annonceurs sont soumis à un système de vérifications automatisées. En matière de détection et de lutte contre la publicité pour des sites Web considérés comme faux ou frauduleux, nous recourons à la fois à l'examen humain et à l'automatisation pour nous assurer de la conformité des sites Web à nos politiques publicitaires (*Advertising Policies*). Nos politiques publicitaires peuvent être consultées ici : <https://snap.com/fr-FR/ad-policies>.

5.6. L'opérateur fournira les chiffres suivants :

- le **volume** d'annonces publicitaires et contenus sponsorisés porteurs de fausses informations ainsi que leur nombre d'impressions ;
- les **revenus générés par les publicités porteuses de fausses informations** avant leur retrait. Il décrira la méthode et les indicateurs employés pour obtenir le montant de ces revenus. Le cas échéant, il distinguera la part de ces revenus alloués à l'opérateur et aux utilisateurs ayant posté les contenus auxquels ont été accolées ces publicités. Il indiquera le nombre d'impressions de ces publicités ;
- le **nombre d'utilisateurs ayant été exposés aux publicités porteuses de fausses informations** avant leur détection et leur retrait ;
- le nombre d'entre eux **ayant cliqué** sur ces publicités.

La politique de Snap en matière de contenus trompeurs (<https://www.snap.com/fr-FR/ad-policies#deceptive-content>). Cette politique interdit :

- Les publicités fausses ou trompeuses, y compris les allégations, offres, fonctionnalités ou pratiques commerciales déloyales.
- Publicités qui contiennent des appels à l'action mensongers ou qui mènent sur des pages d'accueil sans rapport avec la marque ou le contenu faisant l'objet de la publicité.
- Publicités ciblant le cloaking, restreignant de quelque autre manière que ce soit l'accès à la page d'accueil ou modifiant le contenu de l'URL.

5.7. L'opérateur détaillera les **mesures de sécurité des marques** mises en place et expliquera comment elles ont été conçues. Il détaillera notamment les éléments suivants :

- **liste des mesures**⁸ ;
- **moyens humains et automatisés** dédiés à l'exécution de ces mesures ;
- appréciation de l'**impact** de ces mesures ;
- existence d'une **information des annonceurs** lorsque leurs contenus se sont trouvés accolés à des fausses informations et, le cas échéant, de **mécanismes de compensation ou de recours** ;

⁸ En particulier, si ces informations n'ont pas déjà été fournies en réponse à la partie 1 du questionnaire, l'opérateur décrira les mesures de démonétisation des contenus et/ou comptes diffusant des fausses informations. Il précisera la nature de ces mesures, leur durée d'application, leur gradation, la part respective de l'humain et de la machine dans la décision de démonétiser un contenu, l'information de l'utilisateur ayant publié le contenu, l'information de l'ensemble des utilisateurs, la possibilité pour un utilisateur d'effectuer un recours et la façon dont est traité ce recours (traitement par des moyens automatisés ou non, par la même équipe ayant traité le contenu ou non, temps de traitement des recours).

- modalités de participation de l'opérateur dans tout mécanisme de **coopération** avec les marques et l'interprofession publicitaire⁹.

L'opérateur pourra donner des exemples afin d'illustrer les mesures mises en œuvre.

Nous déployons plusieurs systèmes automatisés afin d'examiner les publicités de manière proactive et réactive et de détecter les violations de nos politiques publicitaires. Ces efforts sont réalisés en lien avec l'examen de nos règles par notre équipe de modération humaine. Toutes les publicités signalées sont examinées par notre équipe de modération humaine afin de nous assurer du respect de nos politiques.

5.8. L'opérateur fournira les chiffres suivants :

- le **volume** d'annonces publicitaires et contenus sponsorisés accolés à des contenus porteurs de fausses informations ainsi que leur nombre d'impression ;
- **les revenus générés par les annonces publicitaires et contenus sponsorisés accolés à des contenus porteurs de fausses informations** avant l'éventuel retrait de ces dernières. Il décrira la méthode et les indicateurs employés pour obtenir le montant de ces revenus. Le cas échéant, il distinguera la part de ces revenus alloués à l'opérateur et aux utilisateurs ayant posté les contenus porteurs de fausses informations.

Veuillez vous référer à notre réponse à la Question 5.6.

5.9. L'opérateur indiquera au Conseil s'il dispose d'une base de données relative aux communications commerciales diffusées sur le service ; il détaillera :

- si cette base est publique,
- les types de données qu'elle contient,
- si celles-ci peuvent être filtrées spécifiquement pour la France,
- si l'utilisateur peut les consulter et les télécharger,
- si cette base a évolué depuis 2019.

Snap ne dispose pas d'une base de données générale des contenus sponsorisés. Toutefois, Snap dispose 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 uniquement 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/en-US/political-ads>.

Contenus des utilisateurs réalisés en partenariat avec des tiers

5.10. L'opérateur a-t-il décelé des **contenus porteurs de fausses informations ayant bénéficié d'un partenariat rémunéré** entre l'utilisateur ayant posté le contenu et un tiers (placement de produit, contenu de marque, etc.) ? Le cas échéant, a-t-il pris des mesures spécifiques et si oui, lesquelles ?

Aucun des 4 signalements d'utilisateurs pour fausse information en 2020 n'était associé à un partenariat avec un tiers.

⁹ À l'image du dispositif mis en œuvre afin de lutter contre les contenus haineux et préjudiciables au sein de l'Alliance Mondiale pour les Médias Responsables.

5.11. Existe-t-il des moyens permettant de **lutter** contre ce phénomène ?

Nous n'avons rencontré que 4 cas confirmés (à rapporter à nos 16 millions d'utilisateurs actifs quotidiens en France - source Médiamétrie) de violation de nos politiques contre les fausses informations en 2020 (et aucun n'était associé à un partenariat avec un tiers). Il s'agit d'une preuve concrète que Snapchat se distingue des sociétés traditionnelles de médias sociaux en intégrant des paramètres de sécurité et de confidentialité dès la conception dans le cycle de développement de tous les produits et fonctionnalités, de sorte que le risque pour l'utilisateur est pris en compte et minimisé dès le départ, au lieu d'être envisagé *a posteriori*. De même, tout le contenu accessible au public sur Snapchat est soit sélectionné par des éditeurs médias professionnels tiers avant d'être publié sur Snapchat. Le succès de cette approche se traduit par ce très faible nombre de signalements de fausses informations et de désinformation sur Snapchat. Nous n'avons donc pas mis en place des moyens spécifiques pour traiter ce problème.

5.b. Information des utilisateurs sur la nature, l'origine et les modalités de diffusion des contenus, et l'identité des personnes versant des rémunérations en contrepartie de la promotion des contenus d'information se rattachant à un débat d'intérêt général

Les indications suivantes s'appliquent 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.

Informations apportées aux utilisateurs sur les annonces publicitaires et contenus sponsorisés en général

5.12. L'opérateur précisera les informations fournies à l'utilisateur sur les annonces publicitaires et contenus sponsorisés présents sur son service, en décrivant les spécificités des deux catégories le cas échéant :

- sont-ils clairement **distingués des autres contenus** et, si oui, comment ?
- la plateforme permet-elle à l'utilisateur d'**identifier les critères pris en compte** dans l'affichage des annonces publicitaires et contenus sponsorisés, notamment, si celles-ci sont personnalisées en fonction de son profil ?
- **L'origine** de l'annonce publicitaire ou du contenu sponsorisé est-elle clairement identifiée et affichée (ainsi que celle de l'annonceur, si elle diffère) ?
- le service fournit-il des **indications sur les modalités de diffusion** des annonces publicitaires et contenus sponsorisés (existence de contreparties financières, nombre de vues, type de population ciblée, manière dont le contenu a été généré (automatisée ou non), autres) ?
- l'opérateur décrit-il ses **politiques publicitaires** dans des documents publics disponibles en français, notamment en matière de lutte contre la désinformation sur son service ? Comment l'utilisateur peut-il accéder à ces descriptions ?
- où et comment la plateforme appelle-t-elle les utilisateurs à la **vigilance sur les annonces publicitaires et contenus sponsorisés ayant fait l'objet de signalements** mais qui n'auraient pas été retirés ?
- l'utilisateur a-t-il la possibilité de masquer des annonces publicitaires et contenus sponsorisés ? Peut-il **paramétrer ses préférences publicitaires** ? Si oui, comment est-il informé de cette possibilité, comment a-t-il accès à ces paramètres, que recouvre ce paramétrage en pratique, quelles informations lui sont fournies ?

Veuillez vous référer à notre réponse aux Questions 5.1 à 5.2 ci-dessus.

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

5.13. L'opérateur indiquera au Conseil :

- la **définition des contenus promus d'information se rattachant à un débat d'intérêt général** adoptée par l'opérateur ;
- la **manière dont il identifie** les contenus sponsorisés relevant de cette catégorie
- **leur nombre** sur la plateforme en 2020 ;
- l'existence de tout **système de vérification particulier** pour les annonceurs faisant la promotion de contenus d'information se rattachant à un débat d'intérêt général et, le cas échéant, le nombre d'annonceurs ayant fait l'objet de cette procédure spécifique ;
- si ces contenus font l'objet de **restrictions ou règles particulières** (ex : restriction de ciblage) ; - si la procédure de signalement des contenus d'informations se rattachant à un débat d'intérêt général diffère de celle des autres communications commerciales ;
- les différences de traitement de ce type de communications commerciales avec le reste des communications commerciales.

Il peut préciser s'il s'est rapproché d'autres acteurs (opérateurs, société civile, ONG, think tank, institutions...) pour œuvrer à une plus grande harmonisation des approches retenues en la matière.

Snap ne dispose pas d'une définition des « contenus d'information se rattachant à un débat d'intérêt général » qui serait spécifique à la plateforme, bien que nous considérons que le sens ordinaire de ces mots englobe un très large éventail de contenus. Nous n'avons pas identifié d'indications sur l'interprétation par le CSA des sujets considérés comme « d'intérêt général ». Outre des restrictions plus larges en matière de publicité préélectorale, pendant les périodes électorales concernées, aucune publicité relative à un débat d'intérêt général n'est autorisée à être diffusée sur Snapchat en France. Cette catégorie de publicités entre généralement dans le champ d'application de nos politiques relatives à la publicité à caractère Politique et militant (accessibles sur <https://www.snap.com/en-US/ad-policies/political/>), qui couvrent :

- Les annonces électorales comportent les publicités sur les partis ou les candidats à une charge publique, les mesures ou référendums électoraux, les comités d'action politique et les publicités qui incitent les gens à voter ou à s'inscrire pour voter.
- Les annonces thématiques ou de plaidoyer sont des publicités concernant des questions ou des organisations qui font l'objet d'un débat à l'échelle locale, nationale ou mondiale, ou qui revêtent une importance publique. Elles comptent notamment des publicités sur l'avortement, l'immigration, l'environnement, l'éducation, la discrimination et les armes à feu.

5.14. L'opérateur indiquera au Conseil s'il dispose d'une **base de données relative aux contenus promus d'information se rattachant à un débat d'intérêt général** diffusés sur le service.

Il décrira cette base en fournissant les informations tels que listées à la question 5.9.

Veuillez vous référer à notre réponse à la Question 5.9.

Autres questions

5.15. L'opérateur précisera si des changements ou de nouvelles mesures ont été mis en place en 2020 et s'il a adopté des initiatives particulières relatives aux contenus sponsorisés dans le cadre de la crise sanitaire de la COVID-19.

Voir la réponse à la question 3.10.

*5.16. L'opérateur indiquera dans quelle mesure il a tenu compte **des recommandations du Conseil** énoncées dans son bilan relatif à l'exercice 2019 à compter de sa publication en juillet 2020.*

Veillez vous référer à notre réponse à la question 1.7.

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

Actions d'éducation aux médias et à l'information

6.1. L'opérateur est invité à déclarer au Conseil s'il a développé ou mis en place :

- des **outils permettant aux utilisateurs d'identifier la fiabilité des informations et sources d'informations** qui sont proposées sur le service ;
- des **actions ciblées sur le service afin de sensibiliser les utilisateurs à une utilisation maîtrisée de celui-ci** (ex : encart d'information au moment de la publication d'un message, rappel des règles de savoir-vivre en ligne, etc.), qu'elles soient menées par l'opérateur seul ou en partenariat ;
- plus généralement, des **actions qu'il aurait menées, accompagnées ou promues en matière d'éducation aux médias et à l'information dans le domaine de la lutte contre la désinformation en ligne**, et leur nombre (en distinguant en fonction de la nature de son implication) ;
- des **soutiens à des initiatives indépendantes émanant de journalistes et de chercheurs destinées à mieux comprendre et mesurer le phénomène de la désinformation.**

L'opérateur décrira chaque outil, action ou soutien **en renseignant le tableau ci-après**. Il s'attachera notamment à indiquer : sa nature, son objectif, en quoi il/elle consiste, sa date de lancement, sa mise en valeur auprès du public et des utilisateurs, le ou les public(s) visé(s) et l'impact de son action sur ces derniers (en termes quantitatifs – le nombre de personnes effectivement touchées – et qualitatifs – l'impact de son action sur le comportement du public visé), les partenaires éventuels, la nature de sa propre implication et le montant des dépenses qu'il a engagées et/ou des recettes qu'il a tirées pour chacune de ces opérations.

Il pourra fournir **toute présentation illustrée** de ces différentes mesures.

Comme mentionné ci-dessus, Snapchat ne constitue pas un bon vecteur de diffusion de contenu non modéré généré par les utilisateurs, ni de fausses informations ou de désinformation. Grâce à la conception même de notre plateforme et aux mesures de protection que nous avons mises en place, nous n'avons pas eu besoin d'avoir recours aux outils mentionnés dans cette section. Snapchat a créé un centre de sécurité, accessible sur le site [snapchat.com](https://www.snapchat.com/fr-FR/safety/safety-center) ou directement depuis l'application. Il contient des informations et des conseils pour utiliser Snapchat en toute sécurité, et sur la façon de signaler les violations depuis l'application et vivre une expérience agréable sur la plateforme Snapchat (<https://www.snap.com/fr-FR/safety/safety-center>).

Par ailleurs, 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/>).

Relations avec le monde de la recherche

6.2. Il renseignera le Conseil sur les **relations qu'il entretient avec le monde de la recherche** et les initiatives qu'il a conduites ou soutenues dans ce domaine en France, ou impliquant la recherche française, ainsi que la publicité qu'il aurait faite de ces initiatives.

En particulier, il indiquera et décrira les **partenariats** qu'il aurait conclus dans ce cadre :

- visant à mieux cerner certains comportements et phénomènes liés à la diffusion de fausses informations en ligne et précisera les moyens qu'il met à leur disposition (ex. : concernant la propagation de fausses informations, l'opérateur donne-t-il accès à des données, notamment via une API ; si oui, sous quelles conditions cet accès est-il possible ?) ;

- visant à vérifier la qualité des systèmes de recommandation et de modération du service qu'il opère en termes d'explicabilité, de loyauté et d'équité et, le cas échéant, proposer des pistes d'amélioration.

L'opérateur indiquera et décrira au Conseil les protocoles existants ou en construction avec le monde de la recherche portant sur le **partage et l'exploitation des données**.

Enfin, s'il n'a conclu aucun partenariat avec le monde de la recherche, il indiquera au Conseil **les raisons et/ou les difficultés rencontrées**.

Veillez vous référer à notre réponse à la Question Q. 6.1 ci-dessus.

Autres questions

6.3. L'opérateur indiquera dans quelle mesure il a tenu compte **des recommandations du Conseil** énoncées dans son bilan relatif à l'exercice 2019 à compter de sa publication en juillet 2020.

Veillez vous référer à notre réponse à la question 1.7.

6.4. Il indiquera **l'impact de la crise sanitaire sur ses actions en faveur de l'éducation aux médias et à l'information** et notamment, les adaptations qu'il a pu être amené à lui apporter en conséquence.

Nous nous référons à notre réponse à la question 3.10. Notre priorité est de veiller à ce que Snapchat fournisse des informations de haute qualité sur la covid-19 et de ne pas offrir un fil d'actualité ouvert où des éditeurs ou des individus non vérifiés auraient la possibilité de diffuser de fausses informations. Nos lignes directrices 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 Découvrir, nous avons établi un partenariat avec 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 chaînes Découvrir issues des communautés les plus implantées dans le monde.