



Bilan sur la lutte contre la manipulation de l'information (exercice 2021)

Questions complémentaires spécifiques

Pour cette troisième année d'exercice, il est proposé de poser des questions complémentaires spécifiques à chaque opérateur, lorsque cela s'avère utile au regard :

- du modèle de la plateforme,
- d'une actualité particulière survenue en 2021,
- d'une réponse ou absence de réponse dans la déclaration relative à l'exercice 2020,
- des observations effectuées sur le service,

Les questions suivantes visent à être intégrées dans le corps du questionnaire qui sera adressé par mël à chaque opérateur. Pour plus de lisibilité pour le public, elles pourraient être publiées dans un document à part sur le site de l'Arcom.

Dailymotion

1 – Signalement des fausses informations

- Les vidéos « privées » sur Dailymotion sont-elles modérées de la même manière que les contenus accessibles à tous ?
- Dailymotion indiquera s'il utilise toujours un système d'évaluation hebdomadaire des modérateurs et, le cas échéant, précisera quels sont les critères d'évaluation.

3 – Promotion des contenus issus des entreprises de presse

- Dailymotion indiquera s'il a ajouté des critères à ses démarches internes de labellisation.

4 – Propagation massive

- L'année passée, Dailymotion mentionnait l'utilisation de la technologie Signature développée par l'INA. Il indiquera le nombre de publications contenant une fausse information que l'outil a permis de repérer au moment de la mise en ligne.

5 – Communications commerciales et promotion de contenus portant sur un débat d'intérêt général

- Dailymotion précisera les informations et paramétrages en matière de ciblage publicitaire que l'opérateur met à disposition des utilisateurs, y compris quand son navigateur est encapsulé (« *embedded* ») sur des sites tiers.

Meta – Facebook - Instagram

Remarque générale en introduction : Meta précisera pour chacune des mesures décrites dans sa déclaration si celle-ci s'applique à *Facebook*, *Instagram*, ou aux deux services.

1 – Signalement des fausses informations

- Il est demandé à Meta de présenter sa politique en matière de lutte contre la manipulation de l'information sur chacun des contenus et espaces spécifiques suivants, en précisant si le dispositif de signalement diffère de celui des autres contenus :
 - o contenus éphémères (« *stories* ») sur ses deux services
 - o contenus *live* sur ses deux services
 - o groupes fermés sur *Facebook*.

2 – Transparence des algorithmes

- Meta indiquera quelles ont été les évolutions de l'accessibilité des utilisateurs sur les changements des paramètres relatifs aux recommandations, notamment pour le contrôle et la personnalisation du fil d'actualité sur chacun de ses deux services.
- Il précisera le nombre de modifications demandées par les utilisateurs en comparaison avec l'année passée et indiquera les enseignements qu'il en tire.

4 – Propagation massive

- Meta précisera s'il a mis des moyens en œuvre pour identifier des utilisateurs s'organisant dans des campagnes coordonnées ayant pour objectif de relayer de fausses informations sur ses services ou d'autres via les groupes fermés sur Facebook. Le cas échéant, il est invité à mentionner les éventuelles coopérations mises en place avec les autres opérateurs pour lutter contre ces pratiques.

5 – Communications commerciales et promotion de contenus d'intérêt général

- Meta précisera pour quelles raisons, sur Instagram, le formulaire de signalement des contenus sponsorisés ne permet pas de signaler une fausse information (ou notion voisine)
- Meta a annoncé qu'il arrêterait, à partir du 19 janvier 2022, de permettre le ciblage publicitaire des utilisateurs sur plusieurs sujets sensibles (santé, opinion politique, orientation sexuelle, etc.).¹ La lutte contre la manipulation de l'information est-elle un des objectifs poursuivis par cette mesure ? Le cas échéant, expliquer les effets escomptés.

6 – Éducation aux médias et à l'information

- Meta indiquera quelles évolutions ont été menées en matière d'accessibilité de ses données, notamment sur sa plateforme *Facebook Open Research & Transparency* et son outil *Crowdtangle*.

¹ <https://www.facebook.com/business/news/removing-certain-ad-targeting-options-and-expanding-our-ad-controls>

Google

Google – Search

Remarque générale en introduction : S'agissant de son moteur de recherche, Google précisera pour chacune des mesures décrites dans sa déclaration si elle s'applique à *Google Search*, à *Google Images*, ou aux deux services.

1 – Signalement des fausses informations

- Google expliquera pour quelles raisons il n'est pas possible sur Google Search de signaler un contenu issu d'une recherche à partir d'un outil de signalement à proximité direct du résultat de recherche (ex. : dans le menu qui propose la mise en cache des recherches ?)
- Google indiquera pour quelles raisons le motif de fausse information (ou notion voisine) est absent du formulaire de signalement de contenus s'agissant de Google Actualité (Google News).
- Il est demandé à Google de décrire sa politique en matière de lutte contre la manipulation de l'information sur les contenus de Google Images.

2 – Transparence des algorithmes

- Google indiquera s'il a mis en place des mesures pour lutter contre la manipulation de l'information via les vides de données (« data voids ») sur Google Search.

3 – Promotion des contenus issus des entreprises de presse

- Google fera état des sources dont sont extraites les informations (texte, image...) reprises dans les capsules d'informations apparaissant sur la 1ère page des requêtes de Google Search. Il précisera si le formulaire de signalement spécifiquement mis en place sur cet espace lui a permis de déceler et corriger des informations trompeuses.
- Il est demandé à Google d'indiquer s'il prend des mesures de déréférencement sur Google Search contre les sites relayant de fausses informations et, le cas échéant, s'il rend accessible cette liste à ses utilisateurs.

Google – YouTube

1 – Signalement des fausses informations

- Google précisera sur quels motif et sous-motif exact du dispositif de signalement d'une vidéo l'utilisateur de YouTube doit se baser pour signaler une fausse information troublant l'ordre public d'une part, et étant susceptible de nuire à la sincérité d'un scrutin d'autre part.

- Google précisera pourquoi, sur YouTube, les motifs de signalement d'une vidéo et d'un commentaire diffèrent.
- Il est demandé à Google de présenter sa politique en matière de lutte contre la manipulation de l'information sur chacun des contenus et espaces spécifiques suivants sur YouTube, en précisant si le dispositif de signalement diffère de celui des autres contenus :
 - o contenus éphémères (« *stories* ») :
 - o contenus *live* :
- Les signalements et recours opérés par les utilisateurs de YouTube Premium sont-ils traités différemment ?

2 – Transparence des algorithmes

- Google précisera les définitions des critères ou métriques utilisés pour l'organisation des contenus sur son service, notamment le critère d'actualité et le critère de fraîcheur.

4 – Propagation massive

- Google indiquera si son outil Content ID est utilisé pour identifier préventivement des vidéos qui ont déjà été modérées sur son service *YouTube*. Le cas échéant, il indiquera le nombre de publications contenant une fausse information que l'outil a permis de repérer au moment de la mise en ligne.

5 – Communications commerciales et promotion de contenus portant sur un débat d'intérêt général

- Google précisera comment il permet à l'utilisateur de signaler une annonce publicitaire accolée à un contenu vidéo (pré-roll, mid-roll ou post-roll) pour un motif lié à la manipulation de l'information.

LinkedIn

1 – Signalement des fausses informations

- Les signalements et recours opérés par les utilisateurs de *LinkedIn Premium* sont-ils traités différemment de ceux des autres utilisateurs ou internautes ?

2 – Transparence des algorithmes

- Sur LinkedIn, quels sont les critères entraînant la recommandation d'un compte dans les notifications que reçoit un utilisateur ? Ces critères sont-ils les mêmes s'agissant de la recommandation de personnalités politiques ? Ces dernières sont-elles ou peuvent-elles être promues en période électorale, notamment les candidats et

candidates à une élection ? Ce service de promotion est-il payant, et, le cas échéant, à la charge des personnes promues ?

3 – Promotion des contenus issus des entreprises de presse

- Quels sont les critères de sélection des « influenceurs » et des « Top Voices » par LinkedIn ?

5 – Communications commerciales et promotion de contenus portant sur un débat d'intérêt général

- LinkedIn indiquera quel est le libellé utilisé pour identifier les communications commerciales, en particulier pour les campagnes InMail sponsorisées.

Microsoft – Bing

1 – Signalement des fausses informations

- Microsoft expliquera pour quelles raisons il n'est pas possible sur Bing de signaler un contenu issu d'une recherche à partir d'un outil de signalement à proximité direct du résultat de recherche (ex. : dans le menu qui propose la mise en cache des recherches ?)
- Il est demandé à Microsoft de décrire sa politique en matière de lutte contre la manipulation de l'information sur les contenus de recherche par images sur *Bing*.

3 – Promotion des contenus issus des entreprises de presse

- Microsoft fera état des sources dont sont extraites les informations (texte, image...) reprises dans les capsules d'informations apparaissant sur la 1^{ère} page des requêtes de *Bing*. Il détaillera s'il a reçu des signalements sur ce type de contenus et, le cas échéant, proposera une déclinaison par motif.
- Il est demandé à Microsoft d'indiquer s'il prend des mesures de déréférencement sur *Bing* contre les sites relayant de fausses informations et, le cas échéant, s'il rend accessible cette liste à ses utilisateurs.

Snapchat

Question générale : Il est demandé à Snapchat de préciser pour chacune des mesures décrites dans sa déclaration si elle concerne l'ensemble des espaces de partage de contenus publics présent sur son service (Map, Discover, Stories publiques, Spotlight et profils publics).

1 – Signalement des fausses informations

- Snapchat détaillera sa politique en matière de modération pour chaque type de comptes (médias, SnapStars, utilisateurs en profil public, utilisateurs en profil privé, créateurs, etc.).
- Snapchat indiquera pourquoi les motifs de signalement ne sont pas uniformisés sur l'ensemble des types de contenus du service (Discover, Spotlight, publicités, Snapstars, contenus réalisés en partenariat).

3 – Promotion des contenus issus des entreprises de presse

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

5 – Communications commerciales et promotion de contenus portant sur un débat d'intérêt général

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

TikTok

1 – Signalement des fausses informations

- Il est demandé à TikTok de présenter sa politique en matière de lutte contre la manipulation de l'information sur chacun des contenus et espaces spécifiques suivants, en précisant si le dispositif de signalement diffère de celui des autres contenus :
 - o contenus éphémères (« stories ») :
 - o contenus *live* :
- TikTok précisera pourquoi les motifs de signalement d'une vidéo et d'un commentaire diffèrent.

3 – Promotion des contenus issus des entreprises de presse

- TikTok indiquera quels sont les critères de mise en avant de sujets d'actualités sur la page « Découvrir » du service ? La mise en avant procède-t-elle d'un choix manuel des équipes de TikTok ? Peut-elle faire l'objet d'une contrepartie financière ?

Twitter

1 – Signalement des fausses informations

- Twitter indiquera pour quelles raisons il n'est pas possible de signaler un contenu sans être connecté au service.
- Il est demandé à Twitter de décrire sa politique en matière de lutte contre la manipulation de l'information sur les contenus en direct (« *spaces* ») sur son service en précisant si le dispositif de signalement diffère de celui des autres contenus.
- Dans de sa déclaration relative à l'exercice 2020, Twitter a déclaré que les recours des utilisateurs n'étaient prévus que pour les suppressions de compte. L'opérateur précisera si, désormais, il est possible d'entamer un recours à la suite d'une action de modération sur un contenu.

2 – Transparence des algorithmes

- Twitter apportera des informations sur la mise en place de son équipe META (Machine Learning Ethics, Transparency and Accountability). Il précisera quelles ont été les actions et modifications prises à l'issue du travail de cette équipe en 2021 sur les systèmes de recommandation et de modération en matière de manipulation de l'information.
- Twitter indiquera quels sont les critères de recommandation de contenus utilisés lorsqu'un utilisateur utilise son service sans être connecté à un compte.

3 – Promotion des contenus issus des entreprises de presse

- Twitter indiquera quels sont les critères de mise en avant de sujets dans les espaces « tendances » sur le service. Sont-ils déterminés par des équipes de Twitter et, le cas échéant, lesquelles ? Un nombre minimum de comptes différents doivent-ils poster du contenu pour mettre en avant un sujet ? La mise en avant sur cette page peut-elle faire l'objet d'une contrepartie financière ? Cette page est-elle personnalisée ?

4 – Propagation massive

- Twitter précisera comment il permet à ses utilisateurs d'identifier de manière claire et visible les comptes partagés (utilisant le service « TweetDeck », par exemple).
- Twitter indiquera s'il a identifié des « Listes » (fonctionnalité du service) relayant particulièrement de fausses informations et les mesures prises le cas échéant.

Webedia – Jeuxvideos.com

4 – Propagation massive

- Webedia indiquera s'il a identifié des communautés sur son service s'organisant pour des campagnes coordonnées d'influence sur les services d'autres opérateurs (des « raids » par exemple). Le cas échéant, il est invité à mentionner les éventuelles coopérations mises en place avec les autres opérateurs pour lutter contre ces pratiques.

5 – Communications commerciales et promotion de contenus portant sur un débat d'intérêt général

Webedia est invité à répondre à la partie concernant les communications commerciales, qui ne concerne pas uniquement les contenus sponsorisés ou se rapportant à un débat d'intérêt général.

- Webedia dispose-t-il d'un dispositif de signalement des publicités, notamment programmatiques ? Ce dispositif diffère-t-il en fonction des modules publicitaires (AdSense, Criteo, Outbrain, etc.) ou des formats (preroll, pavés, skyscraper, etc.) ?

Fondation Wikimedia

Questions générales

- La Fondation Wikimedia indiquera le processus de décision amenant à accroître la protection d'une page du service, les mesures prises le cas échéant, et le nombre de ces pages protégées en 2021 (par type de mesures).
- La Fondation Wikimedia fera état des collaborations mises en place avec d'autres opérateurs et des éventuels processus existants ou en cours d'élaboration en matière de détection de fausses informations *via* la reprise de ses pages sur les services d'autres acteurs.
- De quelle manière Wikipédia met-il en avant, pour ses contributeurs et pour ses lecteurs, les sujets sensibles qui nécessitent une surveillance accrue ?

2 – Transparence des algorithmes

- Pour les deux *bots* déclarés l'année passée au titre des programmes informatiques permettant de lutter contre le vandalisme, la Fondation Wikimedia précisera quels ont été les jeux de données d'entraînement de ces modèles, les modalités de l'intervention humaine après la mise en production de ces modèles et si l'ensemble des contributeurs peut « faire appel » des actions de ces deux *bots*.