

## L'urgence du Covid-19 exige des plateformes de

- **S'engager publiquement à mettre fin à la modération exclusivement automatisée des contenus après la crise sanitaire.** Les algorithmes peuvent être une aide à la modération de contenu, mais ne doivent pas pouvoir prendre de décisions sur le retrait des contenus car ils ne sont pas en mesure d'évaluer leur contexte et leur conformité avec les standards internationaux relatifs à la liberté d'expression, et sont donc susceptibles de mal identifier des contenus légaux.
- **Mettre en place des mécanismes de signalement des contenus illégaux, et accroître la visibilité de ces mécanismes.** Ces mécanismes doivent être transparents, facile à utiliser et facilement compréhensibles.
- **Renforcer les mécanismes de recours contre les décisions de retrait de contenus.** Ces mécanismes doivent également être transparents, facile à utiliser et facilement compréhensibles, et les recours devraient être plus accessibles.
- **Les mécanismes de signalement et de recours doivent garantir un traitement rapide de ces recours pour ne pas décourager les utilisateurs de les utiliser.** Depuis le début de la crise du Covid-19, certaines plateformes ont mis à jour leurs formulaires, les rendant plus difficiles à remplir.
- **Publier un rapport de transparence post-Covid-19.** Le rapport de chaque organisation devrait inclure des données sur les opérations de modération effectuées à la demande des gouvernements, des utilisateurs ou de leur propre initiative.

## Recommandations de RSF aux plateformes

- **Les plateformes doivent se conformer de manière stricte à leur devoir de vigilance ("due diligence"),** et la loi devrait renforcer leurs obligations à cet égard, afin de s'assurer qu'elles évaluent effectivement la manière dont leurs activités et leurs services affectent les droits de leurs utilisateurs, et qu'elles prennent des mesures pour atténuer cet impact. Les résultats de cette évaluation doivent être rendus publics.
- **Les opérations de modération des contenus,** qu'elles soient effectuées par des moyens technologiques et/ou humains, **doivent être effectuées de façon conforme aux standards internationaux des droits humains.** Les plateformes ne doivent pas être autorisées à restreindre la liberté d'expression d'une manière excessive par rapport aux limitations prévues par l'article 19 du PIDCP.

## Modération et liberté d'information

- **Assurer un équilibre entre la protection des utilisateurs contre les contenus haineux et le respect de leur liberté d'expression.** Les plateformes doivent être tenues de faire tout leur possible pour supprimer les contenus illégaux, en préservant le droit à la liberté d'expression de leurs utilisateurs.
- **Garantir que pour chaque opération de modération de contenu notifié comme illégal, un humain soit impliqué dans le processus de modération** afin d'évaluer

le contexte du contenu. Pour évaluer efficacement le contexte des contenus, les plateformes devraient impliquer les membres des communautés les plus touchées par le discours de haine, comme les journalistes.

- **Les journalistes devraient également être invités à contribuer aux réflexions et aux études sur le discours de haine** sur les plateformes et à développer les solutions appropriées.
- **Des mécanismes visant à protéger les contenus légitimes contre les notifications de mauvaise foi et à sanctionner ces notifications de mauvaise foi doivent être mis en place.** Les plateformes doivent veiller à ce que leurs règles ne soient pas utilisées à mauvais escient pour réduire les journalistes au silence.
- **Les contenus journalistiques doivent bénéficier d'une protection spéciale,** afin de garantir qu'ils ne puissent pas être retirés par les fournisseurs de services numériques en application de leurs conditions d'utilisation ou pour répondre à une notification abusive. Les journalistes et les médias doivent avoir la possibilité de saisir le juge en référé sur la légitimité du retrait d'un contenu et des mesures provisoires
- **Mettre en place des mécanismes visibles et facile à utiliser pour signaler les contenus illégaux.**
- **Mettre en place des mécanismes visibles et faciles à utiliser pour faire appel des décisions de retrait de contenu.** Le retrait d'un contenu doit être susceptible d'appel devant la plateforme, et la décision de la plateforme concernant cet appel doit pouvoir faire l'objet d'un recours, devant un juge - ou devant un organisme indépendant (tel qu'un régulateur public) sous le contrôle du juge.

### Obligations de transparence

- **Faire preuve de transparence quant à leurs règles de modération des contenus** et préciser les détails de l'application de ces règles de modération.
- **Faire preuve de transparence sur les opérations de modération effectuées à la demande des gouvernements, des utilisateurs ou de leur propre initiative.** Ces résultats doivent être publiés périodiquement et inclure le pourcentage de demandes auxquelles une suite a été donné, les raisons pour lesquelles la plateforme a décidé de donner suite ou non à ces demandes, et les opérations de modération liées au signalement de contenus haineux.
- **Accroître la transparence de leurs actions contre le harcèlement en ligne.**

### Lutter contre les cyber violences

- **Collaborer activement avec les autorités judiciaires dans les enquêtes sur les actes de cyber violence commis contre des journalistes** - à condition que les demandes de ces autorités soient conformes aux normes internationales en matière de liberté d'expression et de respect des droits de la défense, notamment en :
  - répondant aux demandes des autorités judiciaires, en particulier en permettant l'identification et la poursuite des responsables de contenus illégaux ;

- supprimant les contenus illégaux à la demande des autorités judiciaires.
- **Renforcer la lutte contre les campagnes coordonnées de harcèlement en ligne, y compris celles perpétrées par des bots.** Les comptes de *bots* doivent être clairement indiqués comme tels et les plateformes doivent fournir des mécanismes efficaces pour signaler les bots malveillants et suspects.
- **Développer des campagnes de communication et de sensibilisation sur les violences en ligne ciblant spécifiquement les journalistes**, en particulier les femmes.
- **Tous les fournisseurs de services numériques doivent avoir un représentant légal dans tous les pays où ils opèrent**, afin que les individus puissent les poursuivre dans leur pays de résidence pour le préjudice personnel qu'ils ont pu subir en raison des activités des plateformes.