SYNTHÈSE

ATTEINTES CONTRE LA LIBERTÉ D'EXPRESSION EN LIGNE

Expériences, impacts et stratégies face aux cyberviolences et à la censure numérique



FÉMINISTES CONTRE LE CYBER HARCÈLEMENT

Une publication de Féministes contre le cyberharcèlement

Direction de la recherche, coordination éditoriale et rédaction de la synthèse :
Laure Salmona

Rapport de recherche original :
Rédaction principale : Natalia Briceño-Lagos
Contributions de Laure Salmona et Alice Apostoly.

Avec le soutien financier de :
Civitates, la Fondation de France, le Fonds pour les femmes en Méditerranée & l'Union européenne.

Remerciements :
Nous remercions chaleureusement toutes les personnes ayant participé à notre enquête pour le temps qu'elles nous ont accordé et la richesse de leurs témoignages et analyses, ainsi que Stéphanie Lamy pour son éclairage en qualité d'experte.

Leurs contributions sont essentielles pour documenter les atteintes à la liberté d'expression liées aux cyberviolences et aux pratiques abusives de modération, mieux comprendre leurs conséquences, défendre les conditions d'une participation égalitaire au débat public et nourrir les actions collectives en faveur d'un espace numérique plus juste.

Direction artistique, illustrations et mise en page :
Vincent Devillard
www.vincentdevillard.com

Impression :
Exaprint, 34130 Mauguio, France

Édition : Féministes contre le cyberharcèlement, 2026

ISBN : 978-2-489540-00-0

© Féministes contre le cyberharcèlement, 2026

ATTEINTES CONTRE LA LIBERTÉ D'EXPRESSION EN LIGNE

EXPÉRIENCES, IMPACTS ET STRATÉGIES FACE AUX CYBERVIOLENCES ET À LA CENSURE NUMÉRIQUE

ÉTUDE PILOTÉE PAR L'ASSOCIATION FÉMINISTES CONTRE LE CYBERHARCÈLEMENT ET RÉALISÉE PAR NATALIA BRICEÑO-LAGOS, CONSULTANTE

NOVEMBRE 2026

TABLE DE MATIÈRES

Avant-propos 6

Introduction 8

I. La visibilité, condition et risque de l'engagement en ligne 10
1. La présence numérique, devenue une condition de l'action militante 10
2. Les cyberviolences comme mise sous pression de la parole militante 12
3. Malaise dans le militantisme : les violences intracommunautaires 14

II. Les coûts de la participation à l'espace public numérique 15
1. Des conséquences qui débordent largement les frontières de l'espace numérique 15
2. Cyberviolences, censure et modération : une architecture silenciante 17
3. Une dépendance structurelle à des infrastructures privées 19

III. De la restriction de la parole à la reprise de pouvoir 20
1. La restriction de la circulation des idées 20
2. Entre normalisation de la violence et renforcement de l'engagement 22
3. Stratégies de protection et de réappropriation 25
- Se retirer pour mieux revenir
- Anticiper les risques
- Réduire la dépendance aux plateformes
- « Bloquer, bloquer, bloquer »
- Créer des protocoles de prévention
- La sécurité numérique : des compétences qui restent à construire
- Riposter, politiser et construire des contre-discours
- L'émergence d'un « capital militant numérique » construit face à l'adversité
- Un recours au droit particulièrement coûteux
- Les solidarités numériques, des ressources essentielles mais fragiles
- Vers une approche politique de la liberté d'expression numérique
- Renforcer l'action publique et les infrastructures collectives

Conclusion 36

Recommandations 39
1. Garantir les conditions effectives de la liberté d’expression
2. Rendre les décisions des plateformes transparentes et contestables
3. Faciliter la constitution des preuves et garantir un accès effectif à la justice
4. Réduire la dépendance à quelques grandes plateformes
5. Reconnaître le coût organisationnel des cyberviolences
6. Construire des infrastructures collectives de protection et de contre-pouvoir
7. Inscrire la lutte contre les cyberviolences dans une politique plus large de prévention et d’égalité

Glossaire 42

Bibliographie 44

AVANT-PROPOS

Longtemps, Internet a porté une promesse démocratique. Celle d'un espace où des voix tenues jusque-là à distance des médias traditionnels pourraient enfin produire leurs propres récits, faire émerger dans l'espace public des expériences souvent reléguées aux marges et se constituer en force collective. Cette promesse a produit des effets bien réels. Les mouvements féministes, antiracistes, LGBTIQ+, antivalidistes, écologistes et de défense des droits se sont saisis des réseaux sociaux pour documenter les violences, faire circuler des savoirs et intervenir directement dans le débat public. Des communautés ont ainsi pu se former alors qu'il n'existait que peu d'espaces pour se rencontrer ou s'organiser.

Cette visibilité a pourtant un prix, et il n'est pas également réparti. Pour les femmes et les personnes déjà exposées aux discriminations, prendre la parole en ligne revient souvent à composer en permanence avec la possibilité d'une attaque. Une simple prise de position ou une publication sur un réseau social peut déclencher une vague de cyberviolences*. À force, la menace s'installe en arrière-plan de chaque prise de parole et crée une forme d'anticipation permanente : elle infléchit le choix des sujets, les termes employés et, parfois, la décision même de s'exprimer. Le spectacle des attaques subies par d'autres suffit parfois à modifier ses propres pratiques. Peu à peu, ces violences reconfigurent les conditions de participation à l'espace public numérique en rétrécissant les marges d'intervention de certaines voix.

À cette mise sous pression s'ajoute le pouvoir des plateformes. Elles décident en grande partie des conditions de circulation de la parole par leurs règles de modération, leurs systèmes de recommandation et les choix techniques qui organisent la visibilité des contenus. Pourtant, les personnes qui dépendent de ces infrastructures disposent de peu de prise sur leur fonctionnement. Un compte peut être suspendu ou une audience chuter brutalement sans qu'il ne soit possible d'en comprendre précisément la cause. Cette opacité oblige les utilisateurices à interpréter des signaux dont ils et elles ignorent la logique et à modifier leur vocabulaire ou leurs pratiques de publication, souvent sur la base d'hypothèses construites collectivement.

Le fonctionnement même des plateformes peut parfois prolonger, voire amplifier, les effets des cyberviolences. C'est cette articulation que documente la recherche-action pilotée par Féministes contre le cyberharcèlement et conduite par la chercheuse Natalia Briceño-Lagos. À partir des expériences de celles et ceux qui y sont confronté·es, l'enquête analyse la manière dont les cyberviolences et les mécanismes propres aux plateformes se combinent pour affecter les prises de parole, les pratiques militantes et les façons d'occuper l'espace numérique.

Le contenu des entretiens témoigne d'un rapport profondément ambivalent aux plateformes. Si elles demeurent des outils puissants de diffusion et de mobilisation, elles exposent aussi celles et ceux qui s'en servent à des formes de violence et de dépendance auxquelles il est difficile de se soustraire. Cette ambivalence invite à déplacer le regard porté sur la liberté d'expression : le droit de publier ne dit rien, à lui seul, de la possibilité réelle d'être entendu·e ou de continuer à prendre la parole lorsque chaque intervention est susceptible d'exposer à des attaques. Une question traverse alors l'ensemble de ces expériences : que reste-t-il de la liberté d'expression lorsque la possibilité de participer au débat public demeure si étroitement conditionnée par qui l'on est et ce que l'on défend ?

INTRODUCTION

Cette recherche part d'un constat : les effets des cyberviolences sur la prise de parole ne peuvent être compris indépendamment du fonctionnement des plateformes sur lesquelles elles se déploient. Les attaques coordonnées, les menaces, les campagnes de discrédit ou le détournement des outils de signalement imposent leurs propres contraintes. Dans le même temps, les politiques de modération, les systèmes de recommandation et les dispositifs techniques des plateformes déterminent une partie des conditions dans lesquelles les contenus circulent et deviennent plus ou moins visibles. Ces deux dimensions sont souvent étudiées séparément. L'enquête cherche au contraire à comprendre la manière dont elles s'articulent et, parfois, se renforcent.

Le travail empirique repose sur vingt-et-un entretiens semi-directifs réalisés entre juillet et octobre 2025 auprès de dix créateurices de contenus, huit représentant·es d'associations, deux membres de collectifs et une experte des cyberviolences. Les personnes interrogées interviennent dans des champs aussi différents que le féminisme, l'antiracisme, les droits LGBTIQ+, la santé sexuelle et reproductive, la défense des personnes migrantes, le handicap, la justice climatique ou les enjeux numériques. Si ce corpus n'a pas vocation à mesurer la fréquence des phénomènes observés, il permet en revanche d'en suivre les mécanismes au plus près des expériences et d'examiner la façon dont ils transforment les usages, les pratiques militantes et les conditions d'une présence publique en ligne.

L'approche qualitative donne ainsi accès à des transformations que les données agrégées saisissent plus difficilement : l'apprentissage progressif de la prudence, les changements de pratiques après une vague de cyberviolences, les arbitrages imposés par la peur, mais aussi le travail supplémentaire nécessaire pour continuer à prendre la parole au sein de l'espace numérique. Elle permet également d'observer ce que produit l'incertitude liée au fonctionnement des plateformes. Lorsqu'une baisse de portée reste inexpliquée, qu'un contenu disparaît ou qu'une décision de modération paraît incohérente, les utilisateurices doivent tenter de comprendre des mécanismes auxquels ils et elles n'ont qu'un accès très partiel et adapter leurs pratiques en conséquence.

Partir de ces expériences conduit à considérer la liberté d'expression dans toute son épaisseur matérielle. La possibilité formelle de publier ne suffit pas à rendre compte de la place réellement accessible dans le débat public. Encore faut-il pouvoir faire porter sa voix, rencontrer son public et maintenir une présence dans le temps, sans devoir consacrer une part toujours plus importante de ses ressources à se protéger, contourner les attaques ou tenter de comprendre les décisions des plateformes.



Cette perspective conduit également à replacer ces situations dans les rapports de pouvoir qui façonnent l'espace numérique. Les cyberviolences s'inscrivent dans des rapports sociaux de domination déjà à l'œuvre hors ligne ; les architectures des plateformes participent, quant à elles, à déterminer les conditions selon lesquelles ces violences gagnent en portée et certaines voix deviennent plus ou moins audibles. C'est dans cet enchevêtrement que l'enquête cherche à comprendre ce qui se joue : la manière dont la violence et la gouvernance privée des plateformes pèsent sur la participation au débat public, les ressources qu'elles contraignent celles et ceux qui prennent la parole à mobiliser, mais aussi les savoirs, les stratégies et les formes de résistance qui se construisent pour leur opposer d'autres rapports de force.

I. LA VISIBILITÉ, CONDITION ET RISQUE DE L'ENGAGEMENT EN LIGNE

1. LA PRÉSENCE NUMÉRIQUE, DEVENUE UNE CONDITION DE L'ACTION MILITANTE

La présence sur les réseaux sociaux n'est plus accessoire à l'engagement militant. Pour une grande partie des personnes et organisations interrogées, elle constitue une composante structurante de leur activité : sensibiliser, commenter l'actualité, diffuser des ressources, interpeller des responsables politiques, recruter des militant·es, collecter des dons ou construire une réputation professionnelle supposent aujourd'hui de disposer d'une certaine visibilité numérique.



Les acteurices interrogé·es développent généralement une présence « transplateforme » : Instagram, TikTok, Facebook, X (ex-Twitter), YouTube, sites Internet ou *newsletters* sont utilisés simultanément, mais avec des fonctions différentes. Chaque plateforme possède ses normes et ses formats ; les militant·es doivent dès lors en apprendre les codes et adapter en permanence leurs pratiques. Cette multiplication des supports représente donc, en elle-même, une charge de travail importante.

La production militante en ligne s'apparente ainsi à un véritable travail numérique. Création de contenus, veille, programmation, adaptation des formats, réponse aux messages et gestion des commentaires exigent non seulement des compétences mais aussi qu'on leur alloue du temps. Pour les créateurices indépendant·es, ce travail repose la plupart du temps sur une seule personne. Dans les associations, bien qu'il puisse être réparti entre plusieurs salarié·es ou bénévoles, les moyens demeurent souvent limités.

La présence au sein de l'espace numérique suppose également la construction d'une identité et d'une légitimité publiques. Chez les créateurices de contenu, cette légitimité est étroitement liée à l'incarnation : visage, voix, expériences personnelles et opinions sont associés aux contenus publiés dans le cadre de ce qu'on pourrait qualifier de relation parasociale. Les associations et collectifs peuvent au contraire éviter de personnifier leur expression pour protéger leurs membres. Dans les deux cas, la relation avec une communauté d'abonné·es est essentielle : elle contribue à la diffusion des contenus, à la reconnaissance de l'expertise et à la constitution d'une audience fidèle. Cette place centrale qu'occupe la question de la visibilité crée cependant un paradoxe : être visible est à la fois la condition de l'efficacité politique et une source de vulnérabilité. La protection contre les cyberviolences conduit donc souvent à réduire précisément ce qui rend l'action militante efficace, à savoir la circulation des contenus.



2. LES CYBERVIOLENCES COMME MISE SOUS PRESSION DE LA PAROLE MILITANTE

Les cyberviolences constituent une expérience ordinaire de l'engagement numérique pour de nombreuses personnes interrogées. Insultes misogynes, racistes, lesbophobes, transphobes ou antisémites, menaces de viol ou de mort, usurpations d'identité, détournements d'images, divulgation d'informations personnelles, piratages, campagnes de discrédit et signalements abusifs composent un continuum de violences susceptible de perdurer plusieurs semaines, quand ce n'est pas plusieurs mois, voire plusieurs années. Quant à la visibilité, elle joue un rôle important dans le déclenchement des attaques : un contenu viral ou une apparition médiatique peuvent soudainement exposer un compte à des publics jusque-là extérieurs à sa communauté. Cette audience élargie augmente notamment la probabilité d'atteindre des groupes explicitement hostiles aux causes défendues. Pourtant, toucher une large audience n'est pas une condition *sine qua non* de l'apparition des cyberviolences. Plusieurs témoignages montrent que le simple fait d'aborder publiquement certains sujets – féminisme, antiracisme, islamophobie ou droits LGBTIQ+ notamment – peut suffire à provoquer des attaques. Le rapport documente également des violences visant les personnes en raison de leur identité, en particulier lorsqu'elles sont femmes, lesbiennes, trans ou racisées. L'exposition à la violence ne dépend donc pas seulement du degré de visibilité : elle varie aussi selon les thématiques abordées et les personnes qui les portent.



« Et même si c'était pas des milliers. Enfin vraiment, c'était pas énorme du tout [son compte sur TikTok]. Mais ça suffisait quand même pour se faire harceler. Dès qu'on parle de féminisme et dès qu'on a une femme qui prend la parole, on se fait harceler. »



(Collectif)

On observe également la dimension collective et organisée de nombreuses campagnes de cyberviolences. Les raids* consistent en la mobilisation simultanée d'un grand nombre de comptes qui vont submerger une publication ou un profil sous une vague de commentaires haineux et de menaces, ou coordonner une campagne de signalements massifs. La portée et la rapidité des interventions font de ces raids des armes très efficaces pour intimider et faire taire les personnes qui alimentent les comptes visés.

Ces opérations – qui peuvent engendrer plusieurs milliers de notifications en quelques heures – sont parfois coordonnées depuis d'autres espaces, par exemple des forums, des groupes Telegram ou des communautés constituées autour d'influenceurs.

Ces attaques s'accompagnent d'un travail minutieux de surveillance des cibles. Certains groupes suivent durablement les comptes militants, repartagent d'anciennes publications et font circuler des captures d'écran afin de maintenir certaines personnes dans le viseur de communautés hostiles. Un contenu publié plusieurs mois auparavant peut soudain être remis en circulation au gré d'une actualité et déclencher une nouvelle vague d'attaques.

Certains comptes influents jouent enfin un rôle d'amplificateurs. Avec beaucoup d'audience, il est aisé de pointer du doigt tel ou tel compte et d'en faire une cible pour ses abonné·es. Les cyberviolences, et plus spécifiquement le cyberharcèlement, deviennent alors un rouage de l'économie de l'attention* propre aux plateformes : produire un contenu sur une personne devenue cible peut générer de l'engagement, permettant ainsi d'accroître la portée de son compte, voire de générer des revenus.

Loin de se réduire à des interactions interpersonnelles, les cyberviolences s'inscrivent dans des logiques systémiques et prennent appui sur des infrastructures numériques dont l'architecture contribue à leur amplification.

*Les termes signalés par un astérisque renvoient au glossaire page 42.

3. MALAISE DANS LE MILITANTISME : LES VIOLENCES INTRACOMMUNAUTAIRES

Les conflits et campagnes de discrédit intervenant au sein des espaces militants eux-mêmes occupent une part importante des entretiens. Les pratiques de *call out** ou d'interpellation publique y tiennent une place ambiguë.



Bien que les désaccords, remises en question et conflits fassent partie intégrante de la vie politique, notamment lorsqu'ils permettent de rendre visibles les rapports de domination qui se déploient au sein même des mouvements sociaux, certaines campagnes de dénonciation peuvent rapidement se muer en véritables raids et produire des effets délétères sur la participation en ligne : peur de l'erreur, injonction permanente à se justifier, disqualification, anxiété ou silenciation. La proximité politique des détracteurices a même tendance à renforcer l'impact psychologique : être attaqué·e par des personnes dont la reconnaissance est importante peut s'avérer bien plus déstabilisant que l'hostilité provenant d'adversaires idéologiques clairement identifiés.

Où tracer la frontière ? Comment distinguer la critique politique légitime de formes d'interpellation violentes dont les effets contribuent à l'intimidation ou à la silenciation ? Une piste qui se dessine au fil des entretiens est de ne pas considérer l'existence d'un conflit comme un problème, mais de s'intéresser tant aux formes qu'il prend qu'aux rapports de pouvoir qu'il mobilise et aux effets qu'il produit sur les personnes visées.

II. LES COÛTS DE LA PARTICIPATION À L'ESPACE PUBLIC NUMÉRIQUE

1. DES CONSÉQUENCES QUI DÉBORDENT LARGEMENT LES FRONTIÈRES DE L'ESPACE NUMÉRIQUE

Les conséquences documentées sont multiples et peuvent se prolonger très longtemps après la fin apparente des attaques. Pour certaines créateurices de contenu, un épisode de cyberharcèlement constitue un véritable événement biographique, créant un avant et un après. Leur rapport à leur identité publique, à la visibilité et aux réseaux sociaux se trouve alors durablement transformé.

« c'était ma première expérience de raid de cyberharcèlement, donc ça a été très douloureux. Je crois que c'est l'un des plus durs quand c'est le premier, parce qu'on ne sait pas du tout comment faire. Donc j'étais très très mal, et ça m'a vraiment impactée. »

(Créatrice de contenu)



Quant aux effets psychologiques, ils sont récurrents : anxiété, peur d'ouvrir ses notifications, troubles du sommeil, fatigue chronique, manifestations dépressives, sentiment d'irréalité ou hypervigilance. L'anticipation d'une nouvelle attaque finit par devenir un réflexe et des actions jusque-là anodines, comme la publication d'une *story** ou la consultation de ses messages privés, se teintent d'appréhension.

Et cette peur peut s'étendre à l'espace physique. Lorsque des informations personnelles sont diffusées à des fins malveillantes, la menace de violences qui débordent les frontières de la sphère numérique pour s'infiltrer dans le monde matériel devient tangible. Les attaques obligent parfois les victimes à modifier leurs déplacements et à s'inquiéter pour leurs proches, voire à changer de domicile.

« J'ai dû quitter Paris pour me protéger [...] ils ont trouvé mon adresse [...] Ils voulaient faire du mal à mes animaux [...] j'ai quitté Paris pour me planquer. Après, revenir, c'était compliqué aussi, parce que j'étais reconnue dans la rue. »

(Créatrice de contenu)



Les cyberviolences affectent aussi les trajectoires professionnelles. Les agresseurs ne se privent pas de contacter des employeurs ou des partenaires – voire des financeurs – dans le but de mener des campagnes de diffamation ou de faire pression sur des institutions. Certaines associations relatent la perte de financements, de locaux ou d'opportunités, tandis que des créateurices évoquent la crainte permanente que leur réputation numérique ne nuise à une candidature professionnelle. Enfin, les cyberviolences peuvent aussi produire de l'isolement : le coût potentiel d'un soutien public dissuade parfois les proches ou les sympathisant·es de prendre position. Des personnes continuent alors à manifester leur solidarité en privé, tout en refusant de le faire publiquement par crainte de devenir elles-mêmes les cibles de ces violences. Par la peur qu'il instille, le cyberharcèlement impacte non seulement la parole des personnes ciblées, mais également celle de leur entourage.

2. CYBERVIOLENCES, CENSURE ET MODÉRATION : UNE ARCHITECTURE SILENCIANTE

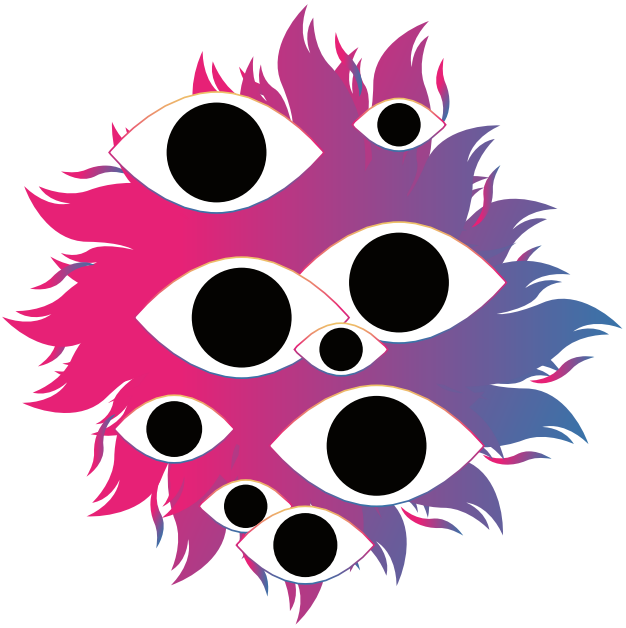


Les politiques de modération et les dispositifs algorithmiques des plateformes occupent également une place centrale dans les expériences recueillies. Les situations de censure et de modération abusive* relatées dans les entretiens sont marquées par une difficulté récurrente à comprendre les décisions prises par les plateformes. Les personnes enquêtées décrivent des suppressions abusives de contenus, des suspensions de comptes, des baisses soudaines de portée de leurs publications ou des refus de sponsorisation qu'elles peinent souvent à comprendre. Le phénomène du *shadowban** illustre parfaitement cette incertitude : les utilisateurices constatent une diminution de visibilité sans disposer des informations nécessaires pour identifier son origine ou prouver qu'il s'agit réellement d'une sanction algorithmique.

« Les phases de *shadowban* ont un peu commencé, où on voyait que quand on postait des sujets très politiques, ou même des sujets qui concernaient nos vies, en fait, qui n'étaient pas forcément politiques, là, on avait moins de portée, les gens voyaient moins nos posts.»

(Association)

Si les témoignages ne permettent pas toujours d'établir avec certitude l'existence de ce phénomène de *shadowban*, l'incertitude elle-même constitue un résultat important. Les militant·es évoluent dans un environnement dont ils et elles ne comprennent pas les règles, alors même que ces règles déterminent largement leur accès à leurs publics.



« **J'ai des vidéos, quand même, hardcore, bien radicales, qui vont percer au milieu de tout ça, au milieu du *shadowban*. Donc, je ne sais pas. Je ne comprends rien. Et en fait, j'ai renoncé à comprendre. Je produis mon contenu, je ne réfléchis pas en termes stratégiques.** »

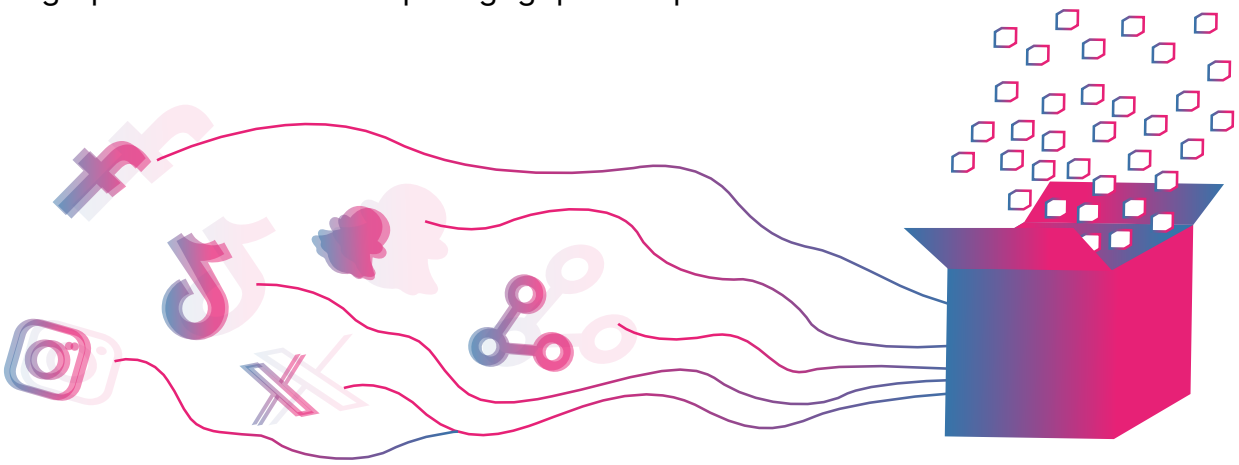
(Créatrice de contenu)

Plusieurs enquêté·es ont le sentiment que certains contenus politiques – notamment sur la Palestine –, féministes ou antiracistes sont plus susceptibles de voir leur visibilité réduite. Des stratégies de contournement apparaissent également autour de termes liés aux violences sexuelles, comme le mot « viol », que certaines créateurices modifient ou remplacent par crainte d'une restriction de portée. Des publications peuvent par ailleurs être rétablies, sans explication, plusieurs mois après leur suppression, lorsque leur potentiel de diffusion est déjà largement réduit.

Pour certaines victimes de cyberviolences, cette baisse de visibilité peut aussi réduire l'exposition aux publics hostiles. Il y a là une ironie cruelle : la même mécanique qui restreint la diffusion d'une parole peut donc, indirectement, limiter les attaques qu'elle suscite. Les personnes concernées se retrouvent ainsi face à un dilemme : rechercher davantage de visibilité au risque d'accroître leur exposition aux violences, ou accepter une circulation plus limitée de leurs publications pour préserver une forme de sécurité.

3. UNE DÉPENDANCE STRUCTURELLE À DES INFRASTRUCTURES PRIVÉES

Cette opacité serait moins problématique si les plateformes n'occupaient pas une position aussi centrale. Mais les comptes des associations et des créateurices de contenu constituent aujourd'hui à la fois des espaces médiatiques d'expression et des lieux de construction communautaire et de mémoire des luttes. Des années de publications militantes peuvent finir par constituer de véritables archives numériques : analyses, témoignages, infographies ou ressources pédagogiques accumulées progressivement. Pourtant, ces contenus restent hébergés par des infrastructures privées sur lesquelles les utilisateurices n'ont aucun contrôle. La suppression d'un compte peut ainsi faire disparaître simultanément l'accès à une audience ainsi que les archives de plusieurs années de travail militant. Or sauvegarder ces contenus représente une charge supplémentaire que nombre de personnes ou petites organisations ne peuvent assumer.



« [...] c'est des années et des années de travail qui se retrouvent perdus, dans certains cas, ou des mois pour retrouver son travail. [...] On en revient encore à ce contenu qu'on crée ne nous appartient pas. Il est sur une plateforme qui ne nous appartient pas. Et si demain, Instagram décide de fermer, en fait, c'est un peu perdu. »

(Collectif)

Cette dépendance affecte également les capacités de plaidoyer et de financement. Une baisse de visibilité ou le refus de promouvoir une campagne peuvent pénaliser une collecte de dons, empêcher une association d'atteindre de nouveaux publics ou tout simplement limiter la diffusion d'une campagne sur tel ou tel sujet considéré comme politique.

Cette dépendance s'articule directement aux cyberviolences et à la modération : les mécanismes prévus pour lutter contre les violences sont régulièrement mobilisés par les agresseurs. Les signalements de masse, notamment, deviennent des outils de silenciation très efficaces, puisque susceptibles d'entraîner la suppression d'un contenu ou d'un compte.

III. DE LA RESTRICTION DE LA PAROLE À LA REPRISE DE POUVOIR

1. LA RESTRICTION DE LA CIRCULATION DES IDÉES

Les cyberviolences excèdent largement leurs effets individuels : elles affectent aussi les conditions de participation à l'espace public et produisent, à ce titre, des effets politiques.

L'une des conséquences est le retrait, temporaire ou durable, de certaines plateformes. Certain·es créateurices suppriment leur compte ou diminuent fortement leur activité afin de protéger leur santé mentale. Ce retrait ne signifie pas nécessairement la fin de leur engagement militant : certaines réorientent leur activité et produisent des *newsletters*, des projets artistiques ou investissent des espaces physiques et des plateformes jugées moins hostiles. Mais ce déplacement a un coût. Pour les créateurices dont l'activité professionnelle dépend en partie de leur visibilité, quitter une plateforme est souvent synonyme de perte d'opportunités ou de revenus, car il est difficile de reconstruire sa communauté ailleurs. Les cyberviolences affectent donc simultanément la liberté d'expression et les conditions matérielles qui rendent cette expression possible.

Un autre effet de ces violences est l'invisibilisation des causes elles-mêmes. Lorsqu'une personne qui travaille sur des sujets liés aux discriminations désactive son compte au moment où son analyse commence à circuler et à atteindre une large audience, ce qui est perdu dans la manœuvre déborde le cadre de son expression individuelle. C'est une vision située et un cadre d'analyse différent qui n'influeront pas sur le débat public. Les cyberviolences viennent ainsi se surajouter aux autres rapports de domination qui régissent la définition des problèmes sociaux.

Enfin, une dernière conséquence – et non des moindres – relève de l'autocensure ou, plus précisément, de ce que le rapport qualifie d'automodération contrainte*. Les personnes enquêtées continuent souvent à prendre la parole en ligne, mais doivent alors consacrer beaucoup plus de temps et d'énergie à anticiper les détournements possibles, à vérifier leurs formulations, choisir leurs visuels ou éviter certains sujets. Après plusieurs épisodes de cyberharcèlement, certaines interventions sont préparées en imaginant systématiquement la manière dont elles pourraient être déformées.



« On ne réfléchit plus de la même manière [qu'avant], on fait attention à tout. [...] Aujourd'hui on est tellement prudentes, en fait je crois qu'on est vraiment traumatisées. [...] Chaque article il est relu par dix personnes. [...] Pourquoi on fait ça ? C'est parce qu'on a tellement peur que ça reparte [le cyberharcèlement]. »

(Association)

Certaines thématiques finissent aussi par être évitées lorsqu'elles sont considérées comme trop coûteuses émotionnellement. Quand elle ne disparaît pas purement et simplement, la parole est hiérarchisée en fonction du risque.

Cette automodération peut aussi être provoquée par les plateformes elles-mêmes. Des associations choisissent de réserver certains sujets à leurs *newsletters* parce qu'elles anticipent une baisse de portée ou une sanction sur Instagram. La peur du cyberharcèlement et l'anticipation d'une modération algorithmique finissent donc par produire des mécanismes convergents de restriction de la parole.

Les cyberviolences imposent par ailleurs un travail supplémentaire de gestion du risque : bloquer, filtrer, documenter, surveiller, archiver, identifier des agresseurs, gérer une situation de crise ou mettre à jour ses paramètres de sécurité.

Ce travail, qui demeure largement invisible, mobilise des ressources considérables. Dans des organisations disposant déjà de moyens limités, le temps consacré à la gestion des attaques est autant de temps qui n'est plus consacré à la sensibilisation, au plaidoyer ou à l'accompagnement. Les cyberviolences fonctionnent ainsi comme un véritable mécanisme de captation des ressources militantes.

2. ENTRE NORMALISATION DE LA VIOLENCE ET RENFORCEMENT DE L'ENGAGEMENT

Pour autant, les effets des cyberviolences ne sont pas toujours homogènes. Certain·es enquêté·es ont tendance à normaliser ces attaques comme un « coût » inévitable de l'engagement féministe, antiraciste ou LGBTIQ+ en ligne. Les insultes et les violences sont anticipées et finissent parfois par être interprétées comme la preuve de l'efficacité du travail militant.

Si cette banalisation pose problème – puisqu'elle conduit à considérer comme ordinaires des niveaux de violence qui ne devraient pas l'être –, elle peut aussi constituer une forme de mise à distance émotionnelle permettant de continuer à agir.

Quant à la peur, elle ne produit pas forcément du retrait : chez certaines personnes disposant de ressources et de soutien, elle peut se combiner avec la colère et renforcer l'engagement militant. Continuer à publier devient alors une manière de refuser que les agresseurs déterminent qui peut occuper l'espace public numérique.

Mais ce n'est pas le fait de subir des violences qui « rend plus fort ». La capacité à transformer une expérience de violence en moteur d'action politique dépend des ressources dont dispose la personne concernée. Les cyberviolences accentuent donc également les inégalités entre celles et ceux qui peuvent s'appuyer sur des soutiens et des dispositifs de protection et les personnes plus isolées ou précaires.

PRISE DE PAROLE SOUS CONTRAINTE : CE QUE MONTRE L'ENQUÊTE QUANTITATIVE

Deux enquêtes quantitatives conduites par Ipsos à la demande de Féministes contre le cyberharcèlement en octobre 2026 complètent les entretiens qualitatifs : l'une auprès d'un échantillon de 1 000 personnes représentatif de la population française, l'autre auprès de 500 femmes publiant au moins une fois par semaine sur les réseaux sociaux.



Les résultats confirment que les cyberviolences transforment durablement les conditions de la prise de parole. Parmi les Français·es victimes ou témoins de cyberviolences interrogé·es sur les conséquences de leurs expériences, **46 % déclarent avoir renoncé à parler de certains sujets, 41 % à participer à certains débats ou discussions en ligne**, 36 % avoir publié moins souvent et 28 % avoir renoncé à une prise de parole publique ou à une exposition médiatique.

Chez les femmes qui publient régulièrement, on retrouve la même dynamique : **48 % ont renoncé à parler de certains sujets**, 44 % à participer à certains débats ou discussions en ligne, 40 % ont publié moins souvent et **32 % ont renoncé à une prise de parole publique ou à une exposition médiatique**. Plus d'une femme sur cinq, soit 22 %, ont également quitté définitivement un réseau social.

Ces effets sont largement perçus comme un enjeu collectif. Au sein de l'échantillon national, **86 % des répondant·es estiment que les cyberviolences conduisent certaines personnes à moins s'exprimer ou à quitter certains espaces en ligne** ; 80 % considèrent qu'elles peuvent rendre certaines opinions, expériences ou causes moins visibles.

Cette perception se retrouve aussi lorsqu'on interroge plus directement les risques pesant sur la liberté d'expression en ligne : **76 % des Français·es considèrent que les insultes, menaces et situations de harcèlement en ligne constituent un risque**, 69 % pensent de même de la manière dont les plateformes sélectionnent les contenus affichés aux utilisateurices, tandis que 67 % citent le fait que certains contenus deviennent moins visibles sans explication claire.

46%

des Français·es victimes ou témoins de cyberviolences déclarent avoir renoncé à parler de certains sujets.

76%

des Français·es considèrent que les insultes, menaces et situations de harcèlement en ligne constituent un risque pour la liberté d'expression.

Les décisions prises par les plateformes affectent elles aussi les pratiques. Parmi les utilisateurices interrogé·es, 14 % déclarent qu'une plateforme a déjà réduit la visibilité d'un contenu ou de leur compte sans le supprimer et 17 % qu'elle a supprimé l'un de leurs contenus. Parmi les personnes concernées par certaines expériences de modération ou de perte de visibilité, 48 % ont évité certains mots ou images, **39 % ont renoncé à certains sujets**, 48 % ont publié moins souvent et 42 % rapportent avoir quitté ou moins utilisé la plateforme.

39%

des personnes concernées par certaines expériences de modération ou de perte de visibilité ont renoncé à certains sujets.

L'opacité joue ici un rôle central : **60 % des personnes ayant fait l'expérience d'une décision de plateforme disent ne pas avoir compris pourquoi elle avait été prise.**

60%

des personnes ayant fait l'expérience d'une décision de modération disent ne pas avoir compris pourquoi elle avait été prise.

Les attentes en matière de garanties sont largement partagées. **Parmi les utilisateurices interrogé·es, 77 % souhaitent pouvoir personnaliser les critères utilisés pour afficher les contenus sur les réseaux sociaux.** Dans l'ensemble de l'échantillon national, 88 % souhaitent que les plateformes expliquent précisément leurs décisions lorsqu'elles réduisent la visibilité d'un contenu ou d'un compte, et 72 % qu'il soit possible de les contester. Pourtant, seules 37 % des personnes interrogées identifient correctement les droits et voies de recours prévus par le droit européen lorsqu'une plateforme supprime un contenu, en réduit la visibilité ou restreint un compte.

Lorsqu'on leur demande quelles mesures devraient être prioritaires pour mieux protéger la liberté d'expression en ligne, **61 % des Français·es placent en tête le renforcement de la lutte contre le harcèlement et les attaques coordonnées sur les réseaux sociaux.** La facilitation de la conservation des preuves arrive ensuite à 42 %, devant le renforcement de l'accompagnement juridique et psychologique des victimes et du contrôle des grandes plateformes par les pouvoirs publics, tous deux cités par respectivement 36 et 35 % des répondant·es. À l'inverse, seuls 6 % considèrent la suppression de toute modération comme une priorité.

61%

des Français·es placent en tête le renforcement de la lutte contre le harcèlement et les attaques coordonnées sur les réseaux sociaux pour protéger la liberté d'expression en ligne.

Les enseignements de ce volet quantitatif viennent donc conforter ceux de l'enquête qualitative : la restriction de la liberté d'expression se joue aussi en amont d'une éventuelle suppression abusive d'un contenu ou d'un compte, lorsque la peur des cyberviolences ou l'incertitude face aux règles des plateformes conduisent les personnes concernées à limiter leur prise de parole, à éviter certains sujets ou à renoncer à prendre part au débat.

3. STRATÉGIES DE PROTECTION ET DE RÉAPPROPRIATION

Face aux violences et à l'invisibilisation, les personnes interrogées développent un ensemble très riche de pratiques de protection pour reprendre la main sur les conditions de leur présence et de leur prise de parole en ligne.



SE RETIRER POUR MIEUX REVENIR

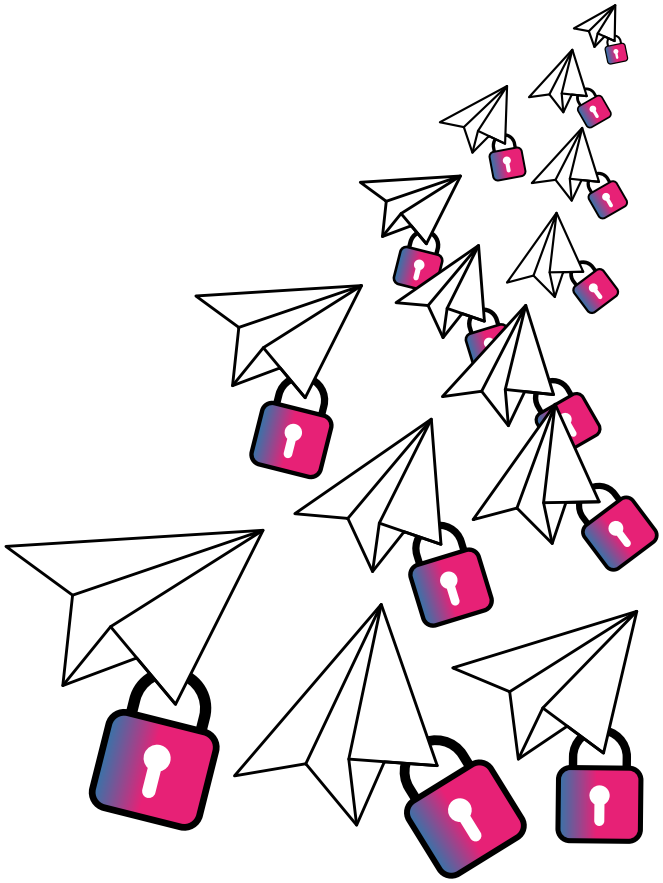
Le retrait temporaire peut devenir une stratégie active de réappropriation. Suspendre son activité permet de protéger sa santé mentale et de reprendre la maîtrise du rythme de sa présence en ligne. Pour certaines associations, quitter une plateforme particulièrement hostile devient même une condition permettant de poursuivre l'engagement sur le long terme.

ANTICIPER LES RISQUES

Avec l'expérience, les organisations formalisent progressivement des procédures de vérification, de validation collective ou d'anonymisation. Certains articles sont relus par plusieurs personnes avant publication ; des pseudonymes sont utilisés pour protéger les auteurices ; la visibilité des événements et des membres est limitée.

Ces pratiques augmentent la sécurité mais représentent elles-mêmes une nouvelle charge de travail. Elles matérialisent l'un des paradoxes permanents de la protection numérique : pour continuer à être visible, il faut investir toujours davantage d'énergie pour prévenir les conséquences de cette visibilité.





RÉDUIRE LA DÉPENDANCE AUX PLATEFORMES

Une autre stratégie consiste à diversifier les espaces de diffusion : *newsletters*, *podcasts*, sites Internet ou activités en présentiel permettent de maintenir un lien avec les publics sans dépendre exclusivement d'une plateforme.

Les *newsletters* apparaissent notamment comme un outil particulièrement intéressant parce qu'elles offrent une relation plus directe avec les communautés. Mais là encore, la diversification requiert du temps, des compétences et des ressources. Elle peut renforcer la résilience tout en dispersant des forces militantes déjà limitées.



« BLOQUER, BLOQUER, BLOQUER »

Le blocage constitue la pratique de protection la plus largement partagée. Plusieurs personnes expliquent avoir abandonné les signalements, jugés inefficaces, pour privilégier le blocage systématique.

Ce choix ne constitue pas seulement une technique de protection individuelle : il renvoie à une conception politique de la modération. Les espaces gérés par des militant·es ne sont pas nécessairement considérés comme des forums dans lesquels toute opinion devrait pouvoir s'exprimer. Bloquer ou supprimer des commentaires permet de protéger à la fois la personne qui publie et la communauté qui la suit.

Certaines personnes rapportent ainsi avoir bloqué plusieurs milliers de comptes. Des fonctionnalités comme la restriction des commentaires ou les listes de blocage permettent aussi de limiter l'escalade de violence. La possibilité de contrôler son propre espace devient donc une forme de réappropriation du pouvoir de modération que les plateformes exercent par ailleurs de manière beaucoup plus opaque.



CRÉER DES PROTOCOLES DE PRÉVENTION

Les organisations ayant subi des épisodes importants de cyberharcèlement tendent à faire de cette expérience une matière première pour construire des protocoles de prévention et de riposte : formations, protocoles internes ou procédures de confidentialité et de sûreté.

Des recommandations portant sur l'utilisation des pseudonymes, la protection des données personnelles, l'exposition des membres ou la présence physique dans les locaux sont progressivement intégrées aux pratiques ordinaires.

Or anticiper et mettre en place des protocoles suppose des ressources humaines, du temps, mais aussi des compétences que toutes les organisations ne possèdent pas. Les cyberviolences révèlent ainsi de fortes disparités entre les organisations dans leur capacité à prévenir les attaques.

LA SÉCURITÉ NUMÉRIQUE : DES COMPÉTENCES QUI RESTENT À CONSTRUIRE

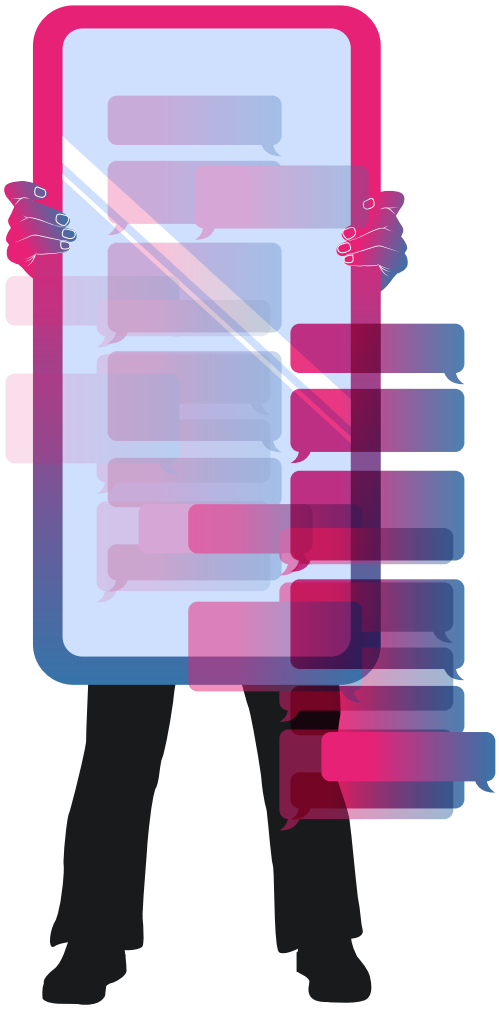
La sécurité numérique apparaît toutefois comme un point de fragilité. Même des personnes très exposées se décrivent souvent comme peu compétentes et perçoivent ces questions comme très techniques.

Les associations identifient fréquemment les risques pesant sur leurs données personnelles, leurs comptes ou leurs moyens de communication sans parvenir à mettre en place les transformations nécessaires, faute de temps ou de financement. Les organisations disposant de davantage de moyens peuvent au contraire améliorer leurs infrastructures, restreindre l'accès aux données ou utiliser des serveurs internes.

Cela souligne l'intérêt des approches d'autodéfense numérique féministe, qui consistent précisément à sortir la sécurité numérique du champ exclusivement technique pour la concevoir comme un enjeu politique et collectif de maintien du pouvoir d'agir.

RIPOSTER, POLITISER ET CONSTRUIRE DES CONTRE-DISCOURS

Toutes les stratégies de résistance déployées par les enquêté·es ne sont pas défensives. Certaines personnes choisissent de rendre visibles les attaques. La publication de captures d'écran ou l'analyse publique d'une campagne de harcèlement permettent de documenter les violences, d'en sortir l'expérience de la sphère privée et de montrer les coûts invisibles de la prise de parole militante.



Les cyberviolences peuvent aussi devenir un matériau politique : elles permettent de démontrer la réalité des discriminations combattues. Des commentaires haineux peuvent être repris dans une campagne de sensibilisation, voire être utilisés pour imposer un agenda médiatique. Cette mise en lumière est cependant ambivalente. Médiatiser ainsi des propos haineux peut contribuer à leur banalisation ou amplifier la visibilité des agresseurs. De plus, toutes les victimes n'ont pas les ressources nécessaires pour transformer ainsi les violences subies afin d'en faire du matériel de sensibilisation.

Certaines choisissent d'exposer publiquement les auteurices. Ces stratégies visent à remettre les choses à l'endroit et à renverser le rapport de force afin de déplacer le sentiment de honte de la victime vers l'agresseur. Elles peuvent produire des effets cathartiques et renforcer le pouvoir d'agir de la victime qui reprend le contrôle et bat en brèche le sentiment d'impuissance générale qu'il est fréquent de ressentir face à ce type d'agressions. Pour autant, ces stratégies ne sont pas sans risques : elles sont susceptibles de se retourner contre les personnes ciblées, des actions en justice pouvant être intentées contre elles par les agresseurs. D'autres enquêté·es préfèrent donc documenter les mécanismes du harcèlement sans identifier les auteurices, afin d'éviter toute médiatisation pouvant augmenter leur notoriété, mais aussi les risques judiciaires.

L'ÉMERGENCE D'UN « CAPITAL MILITANT NUMÉRIQUE* » CONSTRUIT FACE À L'ADVERSITÉ

Face à l'opacité des algorithmes de recommandation* et des systèmes de modération, les utilisateurices échangent des observations et des hypothèses : quels mots semblent faire diminuer la visibilité ? Quels formats fonctionnent ? Comment contourner certains filtres ? Quels contenus peuvent provoquer une sanction ?

Ces tactiques de résistance sont plurielles et peuvent être appréhendées à travers le concept de *gossip* algorithmique*, qui désigne les théories et savoir-faire construits collectivement par les utilisateurices lorsqu'ils et elles ne disposent pas d'informations fiables sur le fonctionnement des algorithmes. Dans les milieux militants, cette contre-expertise se pare d'une fonction politique. Les créateurices de contenu remplacent alors certains mots par des symboles, des emojis ou des orthographes alternatives pour échapper à des pratiques de modération abusive et utilisent certains formats jugés plus favorables tout en échangeant leurs observations sur les pratiques susceptibles d'accroître la visibilité.

Bien que son efficacité soit difficile à démontrer, ce travail permet de continuer à agir malgré l'incertitude. Ces pratiques témoignent de la constitution progressive d'un capital militant numérique : compétences en modération et en sécurité numérique, capacité d'adaptation aux logiques algorithmiques, d'analyse des risques, de coordination collective et de préservation de la visibilité.

Ce capital demeure cependant particulièrement instable. Contrairement à certains savoir-faire militants traditionnels, il dépend d'infrastructures et de règles susceptibles de changer sans préavis ; ce qui fonctionne à un moment donné peut être condamné à devenir obsolète quelques mois plus tard.

UN RECOURS AU DROIT PARTICULIÈREMENT COÛTEUX

Le droit constitue une autre stratégie de riposte, mais son usage reste difficile. Porter plainte suppose d'abord de constituer un dossier de preuves qui repose souvent sur une multitude de captures d'écran permettant de rendre compte de la violence des attaques.



« J'ai porté plainte. En plus, j'ai fait vraiment tout le boulot des policiers. J'ai fait un Excel avec le tweet, l'URL du tweet, la personne qui l'a tweeté [...] J'ai fait des pages et des pages d'Excel. [...] J'ai dû regarder les tweets horribles qui étaient sur moi pour faire la liste [...] J'ai dû le faire dans un moment où j'avais zéro énergie, mes batteries étaient à plat. »

(Créatrice de contenu)



Or cette activité incombe largement aux victimes elles-mêmes et peut représenter des dizaines, voire des centaines, d'heures de travail. Puisque porter plainte exige de rechercher et de classer les contenus malveillants, le recours à la justice implique pour les victimes une ré-exposition aux violences qui peut générer de la détresse. Un accompagnement permet d'éviter cette ré-exposition : proches, associations spécialisées, groupes d'entraide ou avocat·es peuvent prendre en charge une partie du travail de collecte et permettre ainsi à la victime de poursuivre la procédure sans s'exposer de façon disproportionnée aux contenus violents. Mais l'accès à ces ressources dépend fortement des réseaux de solidarité et des moyens dont disposent les victimes. Les inégalités face aux cyberviolences se prolongent donc en inégalités d'accès au droit.

Une profonde défiance envers les institutions judiciaires traverse également les récits, qui décrivent un parcours semé d'embûches : lenteur, classement sans suite, manque d'information, difficulté à identifier les auteurices ou sanctions perçues comme trop faibles. Certaines personnes portent néanmoins plainte par principe, pour faire reculer l'impunité et contribuer à créer des précédents juridiques. Le but poursuivi par la procédure devient alors davantage collectif et altruiste qu'individuel.

LES SOLIDARITÉS NUMÉRIQUES, DES RESSOURCES ESSENTIELLES MAIS FRAGILES

Profondément ambivalents, les réseaux sociaux ne constituent pas uniquement des espaces où se produisent des violences. Ce sont aussi des lieux importants de socialisation, où peuvent se tisser des solidarités et s'organiser des mobilisations. Les créateurices de contenu décrivent des réseaux informels de militant·es qui échangent conseils et soutien. Lorsqu'une personne est ciblée, les communautés d'abonné·es peuvent également intervenir en créant du contenu positif et en se livrant à des actions solidaires : commentaires, repartages, aide à la modération ou mobilisation collective.

« Allez, même vingt copines qui écrivent un commentaire positif tout de suite après. Et en fait, ça, ça donne le ton. Ça ne suffit pas à éviter le harcèlement. Mais ça joue beaucoup. Si les premiers commentaires sont horribles, alors là, c'est l'ouverture au grand n'importe quoi. Ça autorise tout le monde. Et ça dissuade celles qui aiment. »

(Créatrice de contenu)



Ces pratiques permettent parfois de modifier la dynamique d'une attaque. L'arrivée rapide de messages de soutien sous une publication peut empêcher que les commentaires haineux donnent le ton et saturent seuls l'espace. La visibilité devient ainsi une ressource susceptible d'être redistribuée : les comptes bénéficiant de larges communautés peuvent soutenir des militant·es plus exposé·es ou amplifier des contenus dont la portée demeure faible ou semble avoir été réduite. Les collaborations inter-associatives et *crossposts** jouent aussi un rôle comparable.

Mais ces stratégies ne sont pas toujours consensuelles. Certaines organisations refusent d'encourager à signaler massivement tel ou tel contenu ou craignent d'exposer leurs communautés aux violences. L'entraide dépend en outre de la taille de l'audience et peut conduire à de la fatigue militante ; elle constitue donc une ressource qui, bien que précieuse, demeure inégalement disponible et difficile à institutionnaliser.

VERS UNE APPROCHE POLITIQUE DE LA LIBERTÉ D'EXPRESSION NUMÉRIQUE

La gouvernance des infrastructures numériques fait partie des sujets abordés lors des entretiens. Les enquêté·es demandent d'abord une amélioration substantielle de la modération : identification plus efficace des raids et comportements coordonnés, traitement des contenus manifestement haineux ou illégaux, moyens de recours réels et fin de la modération abusive visant des contenus relevant de l'expression politique et associative.

Ils et elles soulignent surtout l'asymétrie entre des contenus haineux qui semblent parfois pouvoir circuler largement et des contenus militants faisant l'objet de restrictions ou de suppressions. Une critique particulière concerne l'absence de prise en compte du contexte social par les catégories de modération : traiter de façon symétrique un discours discriminatoire visant un groupe minoré et une critique politique des rapports de domination peut contribuer à retourner les dispositifs de protection contre les personnes qu'ils sont censés protéger.

Or, derrière la question de la modération, se trouve celle de la concentration du pouvoir au sein de l'espace public numérique. Les entreprises privées disposent aujourd'hui d'un pouvoir considérable leur permettant de déterminer quels contenus circulent, quels publics les verront et quels discours seront mis en avant. Leur modèle économique repose principalement sur l'engagement, l'exploitation des données et la rentabilité, et non sur la protection de la liberté d'expression des populations discriminées.

« Ce n'est pas leur rôle [de protéger la liberté d'expression], en fait. Leur rôle, c'est de faire de l'argent. Point barre. Pendant très longtemps, les entreprises de médias sociaux ont présenté leur projet en disant qu'elles étaient les championnes de la liberté d'expression, alors qu'elles exploitaient déjà nos données personnelles pour vendre des modèles de prédiction de notre comportement. C'est de la surveillance de masse capitaliste. Et le capitalisme de surveillance a besoin que les gens s'expriment, produisent des données, interagissent. Ce sont des industries extractives : ce qu'elles extraient, c'est notre attention. Or, notre attention est une ressource limitée, donc contestée. Évidemment qu'elles vont se présenter comme les championnes de la liberté d'expression : plus on s'exprime, plus elles ont de données à exploiter. Elles peuvent ensuite agir sur nos comportements à travers des suggestions, observer nos réactions et utiliser ces nouvelles données pour affiner leurs modèles. »

(Stéphanie Lamy, experte)

Les logiques de recommandation peuvent par ailleurs favoriser les contenus clivants ou sensationnalistes. La régulation des cyberviolences ne peut donc être dissociée du modèle économique et de la gouvernance des plateformes.

Plusieurs enquêté·es posent ainsi la question de la souveraineté numérique et de l'intervention publique. Si les réseaux sociaux constituent désormais des infrastructures centrales du débat démocratique, peut-on laisser intégralement des entreprises privées en définir les règles ? Cette interrogation mène chez certain·es à une proposition plus radicale : considérer les infrastructures numériques comme des biens communs démocratiques, dont la gouvernance ne pourrait être abandonnée au seul marché.

RENFORCER L'ACTION PUBLIQUE ET LES INFRASTRUCTURES COLLECTIVES

Un autre besoin fortement exprimé concerne le renforcement de la prévention de ces violences et de l'accompagnement des victimes. Cela passe notamment par la formation des professionnel·les de la police et de la justice. Les cyberviolences continuent trop souvent d'être perçues comme moins graves ou moins « réelles » que les violences hors ligne, alors même que leurs conséquences professionnelles, sociales et psychologiques sont très lourdes.

L'éducation numérique semble aussi revêtir une importance particulière pour la plupart des enquêté·es. Celle-ci ne devrait pas se limiter à l'apprentissage de l'utilisation des outils numériques ou à la création de mots de passe robustes. Elle devrait également inclure une véritable éducation à la citoyenneté numérique : comprendre les conséquences de ses comportements, les règles du débat public et la manière dont les discriminations et les violences peuvent être répliquées et amplifiées en ligne.



Enfin, la construction d'infrastructures collectives de protection apparaît comme une priorité récurrente : malgré la recrudescence des cyberviolences visant les mouvements féministes, antiracistes ou LGBTIQ+, les expériences restent souvent appréhendées comme s'il s'agissait d'expériences individuelles. Les savoirs et les protocoles existent mais restent dispersés et ne font pas l'objet de transmissions entre personnes et structures. De nombreuses victimes doivent dès lors apprendre seules à conserver des preuves, paramétrer leurs comptes, modérer, sécuriser leurs données ou identifier les recours disponibles.

« Cette expérience nous a appris que le harcèlement est devenu un outil de domination politique. Il sert à faire taire ou à mettre en conformité la personne prise pour cible, mais il envoie aussi un signal aux autres : voilà ce qui peut vous arriver si vous prenez la parole, si vous vous exposez, si vous sortez du rang. La réponse doit donc être collective. Elle ne peut pas reposer uniquement sur le juridique, parce que le coût des procédures est absolument disproportionné pour les victimes. »

(Stéphanie Lamy, experte)

Les propositions exprimées pour remédier à cet état de fait sont nombreuses : guides régulièrement actualisés, formations, mutualisation de compétences en sécurité numérique, réseaux d'avocat·es formé·es, espaces d'échange entre *community managers*, dispositifs de soutien lors des attaques et construction de coalitions capables de faire pression sur les plateformes.

Certaines propositions visent d'ailleurs à reconquérir collectivement l'espace numérique : structurer des réseaux de soutien, augmenter la visibilité de certains comptes ou investir davantage les espaces de commentaires dominés par des groupes coordonnés. L'idée centrale est de s'organiser et de construire des réseaux de solidarités pour ne plus s'épuiser à répondre individuellement aux attaques.

À plus long terme apparaît également l'idée de développer des infrastructures alternatives ou *open source** afin de réduire la dépendance aux géants de la Tech et autres grandes plateformes privées.



CONCLUSION

« Si on veut parler librement sans se faire intimider, c'est pas possible. C'est impossible. Il faut accepter l'intimidation. Il faut accepter... le fait qu'on va se faire menacer et attaquer, parce que les plateformes ne garantissent pas du tout une parole protégée de cette intimidation. »



(Créatrice de contenu)

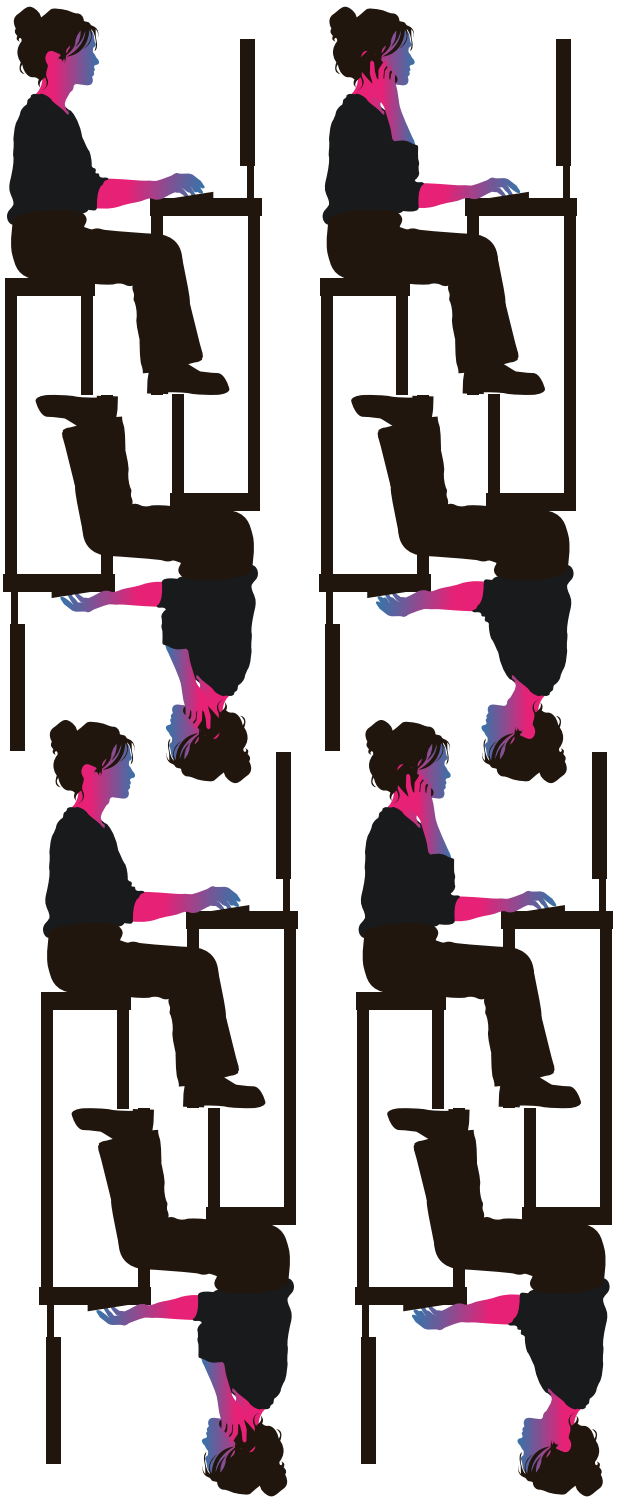
Les cyberviolences excèdent largement le registre des atteintes interpersonnelles. Lorsqu'elles conduisent certaines personnes à se retirer, à éviter des sujets, à modifier leur manière de s'exprimer ou à consacrer une part croissante de leur énergie à prévenir les attaques, elles agissent directement sur les conditions de participation au débat public. Leur dimension politique tient précisément à cette capacité à peser sur ce qui peut être dit, par qui, et dans quelles conditions.

Ces effets sont profondément inégalitaires. Un·e créateurice de contenu isolée, un collectif militant ou une petite association ne disposent pas des mêmes moyens de résistance qu'une organisation dotée de ressources juridiques, techniques ou humaines importantes. Les violences viennent dès lors se greffer sur des rapports de pouvoir préexistants et en accentuent certains effets. Les groupes déjà confrontés à des discriminations dans l'espace public doivent supporter un coût supplémentaire pour y maintenir leur présence et diffuser leurs analyses.

Ce coût est aussi matériel. Modérer des commentaires, conserver des preuves, sécuriser des comptes ou accompagner les personnes visées mobilise un temps considérable. À cette charge s'ajoutent parfois des dépenses juridiques, techniques ou organisationnelles difficiles à absorber. Autant de ressources soustraites au travail politique lui-même. En captant ainsi du temps, de l'attention et des moyens déjà rares, les cyberviolences affaiblissent les capacités de mobilisation bien au-delà du moment de l'attaque.

Leur effet se prolonge dans la circulation des idées. Lorsqu'un compte disparaît, qu'une campagne de signalement interrompt la diffusion d'un contenu ou qu'une personne renonce à traiter un sujet devenu trop risqué, ce sont aussi des récits, des savoirs et des cadres d'analyse qui perdent en visibilité. Les mécanismes de modération des plateformes peuvent parfois accentuer cette dynamique, notamment lorsqu'ils sont détournés par des agresseurs ou lorsque leurs critères restent suffisamment opaques pour contraindre les utilisateurices à modifier leurs pratiques sans jamais savoir précisément ce qui les expose à une sanction.

Le corpus ne se résume pourtant pas à une succession de retraits et de renoncements. Face aux attaques, les personnes rencontrées élaborent des savoir-faire, organisent des solidarités et inventent des formes de riposte. Les communautés numériques peuvent apporter un soutien décisif, prendre en charge une partie de la modération, relayer un contenu ciblé ou rééquilibrer un espace saturé de commentaires haineux. Là où les violences tendent à isoler, ces réseaux peuvent restaurer une capacité d'intervention collective.



C’est au fil de ces expériences que se constitue ce que le rapport qualifie de capital militant numérique. Savoir documenter une attaque, sécuriser un compte, interpréter les variations de visibilité ou coordonner une communauté fait désormais partie du répertoire de l’action militante en ligne. Ces compétences restent pourtant fragiles. Elles s’acquièrent souvent sous la contrainte et dépendent d’environnements techniques dont les règles évoluent sans cesse. Leur distribution très inégale reproduit en outre les écarts de ressources déjà présents entre les personnes et les organisations.

Ces constats obligent à prendre au sérieux les conditions matérielles et sociales de la liberté d’expression. La possibilité formelle de publier ne suffit pas à garantir une participation effective au débat public : une parole systématiquement exposée à la menace, rendue moins visible sans explication ou dont le maintien exige un travail de protection disproportionné ne bénéficie pas des mêmes conditions d’expression qu’une autre. La liberté d’expression se joue aussi dans la possibilité pour une parole d’atteindre son public et de se maintenir dans la durée.

Ce déplacement conduit nécessairement à interroger la répartition des responsabilités. Les personnes ciblées ne peuvent continuer à porter seules la charge de leur propre protection. Or les plateformes disposent d’un pouvoir considérable sur l’architecture des échanges, les modalités de modération et la distribution de la visibilité ; elles doivent donc répondre des effets produits par ces choix. L’action publique doit, de son côté, garantir des recours effectifs et soutenir la diffusion de compétences comme la construction de dispositifs collectifs de protection.

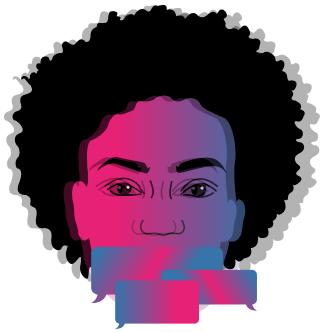
Prendre au sérieux le rôle politique des espaces numériques revient finalement à reconnaître qu’une part croissante du débat démocratique se déroule au sein d’infrastructures privées dont les règles échappent largement à celles et ceux qui les utilisent. La question dépasse alors la seule protection des victimes. Elle concerne la possibilité de préserver un espace public où des voix exposées de manière inégale aux violences puissent continuer à produire des savoirs, porter des revendications et peser sur la définition des problèmes collectifs.

C’est aussi à cette aune que doit s’apprécier la liberté d’expression en ligne : à la possibilité réelle, pour celles et ceux qui parlent depuis les marges, de prendre place dans le débat public et de s’y maintenir sans en être continuellement repoussé·es.

RECOMMANDATIONS

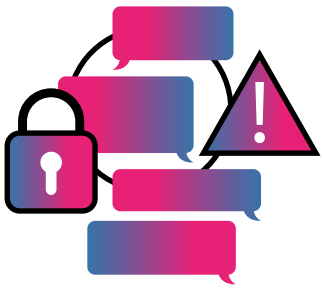
Les résultats de cette étude montrent que lutter contre les cyberviolences suppose d'agir bien au-delà de la protection individuelle : sur les plateformes, l'accès au droit, les ressources collectives et les conditions mêmes de la participation au débat public.

La charge ne peut continuer à reposer principalement sur les personnes ciblées, contraintes de se protéger, de documenter les violences ou d'adapter leurs pratiques pour pouvoir continuer à prendre la parole. Garantir les conditions effectives de la liberté d'expression en ligne suppose donc d'agir sur les infrastructures, les règles et les rapports de pouvoir qui organisent aujourd'hui l'espace public numérique. En s'appuyant sur les enseignements de l'enquête et son expertise sur les cyberviolences, Féministes contre le cyberharcèlement formule les recommandations suivantes.



1. GARANTIR LES CONDITIONS EFFECTIVES DE LA LIBERTÉ D'EXPRESSION.

La liberté d'expression ne peut être réduite au droit formel de publier. Elle suppose que chacun·e puisse prendre la parole, atteindre un public et participer au débat sans devoir supporter un niveau disproportionné de violences ou de coûts matériels, professionnels et psychologiques. Les cyberviolences doivent à ce titre être reconnues comme une atteinte aux conditions mêmes de la participation démocratique.



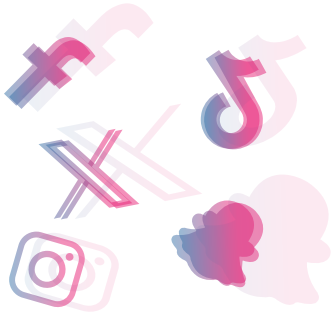
2. RENDRE LES DÉCISIONS DES PLATEFORMES TRANSPARENTES ET CONTESTABLES.

Les décisions de modération, les réductions significatives de visibilité et les suspensions de comptes doivent être expliquées et pouvoir faire l'objet de recours effectifs. La modération doit également tenir compte du contexte et des rapports sociaux dans lesquels les discours s'inscrivent. Cette exigence de transparence doit s'accompagner d'une véritable logique de *safety by design** : paramètres protecteurs par défaut, dispositifs anti-raids, outils de blocage et de modération collective, ainsi que la possibilité pour les utilisateurices de s'affranchir des systèmes de recommandation fondés sur l'engagement.



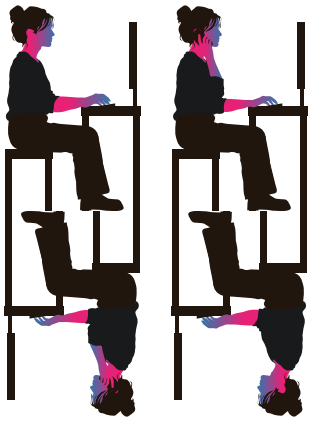
3. FACILITER LA CONSTITUTION DES PREUVES ET GARANTIR UN ACCÈS EFFECTIF À LA JUSTICE.

Le travail de documentation des cyberviolences repose encore trop largement sur les victimes. Les plateformes devraient permettre l'export immédiat, horodaté et authentifiable des contenus violents, y compris lorsqu'ils sont ensuite supprimés. Les pouvoirs publics doivent parallèlement renforcer la formation des professionnel·les, les capacités d'enquête et l'accompagnement juridique afin que la protection des victimes et l'identification des responsables ne dépendent pas de leurs seules ressources.



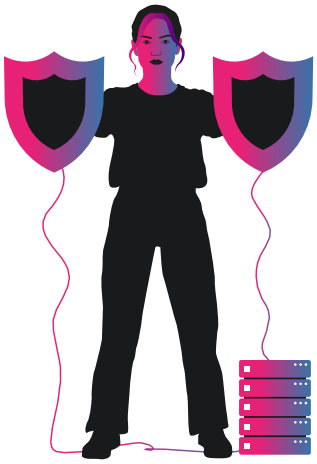
4. RÉDUIRE LA DÉPENDANCE À QUELQUES GRANDES PLATEFORMES.

Quitter un service ne constitue guère une option lorsque cela suppose de perdre son audience, ses archives ou une partie de ses revenus. Le développement de l'interopérabilité, de la portabilité des données et des communautés, ainsi que de dispositifs simples d'exportation et de conservation des contenus, doit permettre de changer de plateforme sans perdre des années de travail, mais aussi de favoriser l'émergence d'alternatives.



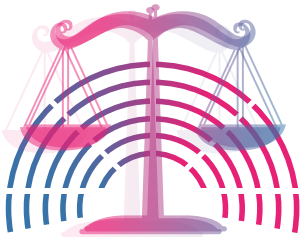
5. RECONNAÎTRE LE COÛT ORGANISATIONNEL DES CYBERVIOLENCES.

Modérer, surveiller, documenter, sécuriser et accompagner les personnes ciblées constituent un travail à part entière. Ce temps et ces dépenses doivent être reconnus dans les financements des associations et des collectifs exposés. Les grandes plateformes doivent également contribuer, par des mécanismes indépendants, au financement de la prévention, de l'accès au droit et de la réparation des dommages qu'elles contribuent à rendre possibles ou à amplifier.



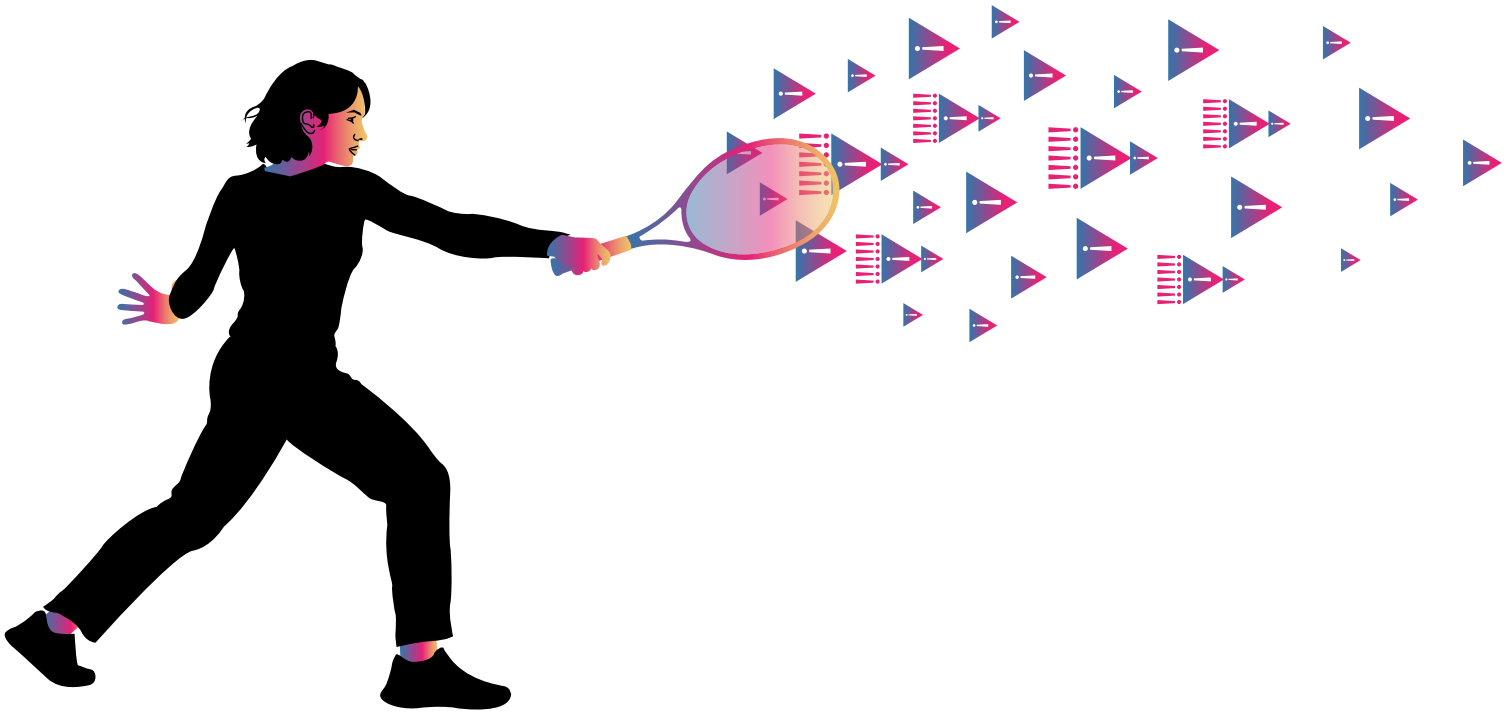
6. CONSTRUIRE DES INFRASTRUCTURES COLLECTIVES DE PROTECTION ET DE CONTRE-POUVOIR.

La réponse aux cyberviolences doit pouvoir s'appuyer sur des réseaux d'entraide, des dispositifs juridiques mutualisés, des formations à l'autodéfense numérique, des protocoles de gestion de crise et des outils communs de documentation. Les associations et groupes particulièrement exposés doivent être durablement associés à la régulation des plateformes, tandis que leur expertise, comme celle de la recherche indépendante, doit être reconnue et financée. À plus long terme, le développement d'infrastructures ouvertes, interopérables et gouvernées selon des principes démocratiques constitue un enjeu central de rééquilibrage du pouvoir numérique.



7. INSCRIRE LA LUTTE CONTRE LES CYBERVIOLENCES DANS UNE POLITIQUE PLUS LARGE DE PRÉVENTION ET D'ÉGALITÉ.

La sécurité numérique, l'éducation aux droits en ligne et la connaissance des recours doivent être accessibles au plus grand nombre. Mais les cyberviolences prolongent aussi des rapports de domination déjà présents hors ligne : leur prévention ne peut donc être dissociée d'un soutien durable aux politiques de lutte contre les violences sexistes et sexuelles, les discriminations et les inégalités qui structurent également l'espace numérique.



GLOSSAIRE

Algorithme de recommandation : Système automatisé utilisé par une plateforme pour sélectionner et hiérarchiser les contenus proposés aux utilisateurices. Ces systèmes déterminent en partie ce qui apparaît dans les fils d'actualité, les résultats de recherche ou les contenus recommandés. Leur fonctionnement repose sur différents critères, notamment les interactions et l'engagement suscités par les contenus, et demeure souvent largement opaque pour les utilisateurices.

Automodération contrainte : Adaptation de sa manière de prendre la parole afin de réduire les risques de cyberviolences ou de sanctions des plateformes. Elle peut concerner les sujets abordés, les formulations, les formats ou encore le niveau d'exposition de soi, sans nécessairement conduire à renoncer à s'exprimer.

***Call out* :** Interpellation ou dénonciation publique d'une personne, d'un groupe ou d'une organisation en lien avec des propos ou des comportements jugés problématiques. Dans les espaces militants, le *call out* peut être utilisé pour demander des comptes ou provoquer une remise en question. Ce type d'interpellation peut se muer en harcèlement collectif, notamment lorsqu'elle entraîne une mobilisation collective disproportionnée et des attaques répétées à l'encontre de la personne visée.

Capital militant numérique : Forme spécifique de capital militant constituée par l'accumulation, au cours d'un engagement en ligne, de savoirs, de compétences, de relations et de savoir-faire propres aux environnements numériques. Ces ressources peuvent être mobilisées pour maintenir ou renforcer une capacité d'action politique dans l'espace numérique.

***Crosspost* :** Pratique consistant, sur une même plateforme, à relayer ou republier un contenu à partir de plusieurs comptes afin d'en accroître la circulation et la visibilité. Le *crosspost* peut notamment être utilisé de manière coordonnée par un réseau de comptes ou une communauté.

Cyberviolences : Ensemble des violences commises, facilitées ou amplifiées au moyen d'outils et de services numériques. Elles peuvent notamment prendre la forme de harcèlement, de menaces, de divulgation d'informations personnelles, d'usurpation d'identité, de diffusion non consentie de contenus intimes ou de campagnes coordonnées de signalement et de discrédit. Elles s'inscrivent souvent dans la continuité de rapports de domination et de violences qui existent également hors ligne.

Économie de l'attention : Modèle économique dans lequel l'attention des utilisateurices constitue une ressource à capter et à monétiser. Sur les plateformes numériques, ce modèle encourage notamment la conception de dispositifs destinés à maximiser le temps passé, les interactions et l'engagement, qui peuvent ensuite être valorisés commercialement, notamment à travers la publicité et l'exploitation des données.

***Gossip* algorithmique :** Ensemble de théories, d'hypothèses et de savoirs pratiques construits et échangés entre utilisateurices pour tenter de comprendre le fonctionnement des algorithmes lorsque les plateformes fournissent peu d'informations à leur sujet. Ces savoirs reposent sur l'observation, l'expérience et la comparaison des pratiques et peuvent servir à adapter les contenus ou les stratégies de publication afin de préserver leur visibilité.

Interopérabilité : Capacité de différents services, plateformes ou systèmes numériques à communiquer entre eux et à échanger des données dans des formats compatibles. Appliquée aux réseaux sociaux, elle peut permettre aux utilisateurices d'interagir entre plusieurs services ou de transférer plus facilement leurs données et certaines ressources sans rester dépendant·es d'une seule plateforme.

Modération abusive : Application injustifiée, disproportionnée ou inadaptée des règles de modération d'une plateforme, conduisant notamment à supprimer ou restreindre la visibilité d'un contenu ou d'un compte qui ne devrait pas faire l'objet d'une telle mesure. Elle peut résulter d'une mauvaise contextualisation des contenus, de mécanismes automatisés ou de campagnes de signalements abusifs.

***Open source* :** Modèle de développement dans lequel le code source d'un logiciel est accessible au public et peut, selon les conditions de sa licence, être consulté et modifié. L'*open source* permet notamment d'examiner le fonctionnement d'un outil et en favorise le développement collectif.

Raid : Attaque collective et coordonnée au cours de laquelle un grand nombre de comptes ciblent simultanément une personne ou une organisation. Un raid peut prendre la forme d'une vague de commentaires hostiles, de menaces, de messages privés ou de signalements massifs visant à intimider la personne ciblée, saturer son espace numérique ou provoquer la suppression de ses contenus ou de son compte.

***Safety by design* :** Approche consistant à intégrer la prévention des risques et la protection des utilisateurices dès la conception d'un service ou d'une fonctionnalité numérique plutôt que d'intervenir uniquement après la survenue des violences. Elle suppose notamment d'évaluer en amont les usages malveillants possibles et de prévoir par défaut des dispositifs permettant de limiter les risques.

***Shadowban* :** Réduction de la visibilité d'un compte ou d'un contenu par une plateforme, sans suppression explicite ni information claire adressée à l'utilisateurice. Le contenu reste accessible, mais peut être moins recommandé, moins bien classé ou moins facilement découvert. Le terme est largement utilisé par les utilisateurices, même si les causes exactes d'une baisse de visibilité sont souvent difficiles à établir.

***Story* :** Publication temporaire, généralement composée d'une image, d'une vidéo ou de texte, visible pendant une durée limitée sur certaines plateformes de réseaux sociaux. Les *stories* permettent de partager des contenus plus spontanés et éphémères que les publications classiques.

BIBLIOGRAPHIE

Amossy, Ruth. 2022. « Construire la légitimité et l'autorité politiques en discours ». *Argumentation et analyse du discours* 28.

Badouard, Romain. 2021. « Modérer la parole sur les réseaux sociaux. Politiques des plateformes et régulation des contenus ». *Réseaux* 225 (1) : 87-120.

Badouard, Romain. 2020. « Les discours de haine en ligne ». *Esprit*. (5), 27-31.

Belhaddad, S., et M.-C. Picardat. 2025. *De la banalisation de la violence verbale au discours de haine : Décrypter, mieux agir pour restaurer le lien social*. Avis no 2025-005. Paris : Conseil économique, social et environnemental.

Bishop, Sophie. 2019. « Managing Visibility on YouTube through Algorithmic Gossip ». *New Media & Society* 21 (11-12) : 2589-2606.

Blais, Mélissa. 2021. « Ce que la peur fait à l'engagement féministe ». *Lien social et Politiques* 86 : 94-112.

Boullier, Dominique. 2016. *Sociologie du numérique*. Paris : Armand Colin.

Cardon, Dominique. 2010. *La démocratie Internet. Promesses et limites*. Paris : Seuil.

Grison T, Julliard V, Alié F & Ecrement V. 2023. « La modération abusive sur Twitter: étude de cas sur l'invisibilisation des contenus LGBT et TDS en ligne », Réseaux, 237 (1), pp. 119-149.

Jouët, Josiane, Katharina Niemeyer et Bibia Pavard. 2017. « Faire des vagues : les mobilisations féministes en ligne ». *Réseaux* 201 (1) : 21-57.

Matonti, Frédérique, et Franck Poupeau. 2004. « Le capital militant. Essai de définition ». *Actes de la recherche en sciences sociales* 155 : 4-11.

Millette, Mélanie. 2013. « Pratiques transplateformes et convergence dans les usages des médias sociaux ». *Communication et organisation* 43 : 47-58.

Pigenet, Phoebe. 2024. « Bousculer les normes corporelles », *Communication* [Online], Vol. 41/1, Online since 28 August 2024.

Poell, Thomas, David Nieborg et José van Dijck. 2019. « Platformisation ». *Internet Policy Review* 8.



RETROUVEZ
LE RAPPORT COMPLET
SUR WWW.VSCYBERH.ORG



Avec le soutien de :

Projet soutenu par

Fondation de France

Civitates
The European Democracy Fund

Mediterranean Women's Fund
Fonds pour les Femmes en Méditerranée

Cofinancé par l'Union européenne