IA ET HARCÈLEMENT

COMPRENDRE ET RECONNAÎTRE POUR MIEUX AGIR



31% des mineurs victimes de cyberharcèlement disent avoir été visés par des contenus humiliants générés par IA. Près d'**1 parent sur 4** d'enfant cyberharcelé évoque un deepfake*.

*Baromètre 2026 de l'Association e-Enfance / 3018 réalisé par le Cabinet Georges avec le soutien de la Caisse d'Epargne.

DE QUOI PARLE-T-ON ?

Le harcèlement est une **violence répétée et intentionnelle** pouvant être verbale, physique, sexuelle et/ou psychologique, commise par une ou plusieurs personnes. Le **but est d'humilier, d'exclure, d'intimider, ou encore de rabaisser** une personne. Il se caractérise par un rapport de domination et peut se produire en milieu scolaire, professionnel, familial, sportif, associatif… Lorsqu'il est commis par le biais d'**outils numériques** (réseaux sociaux, messageries, blogs, forums, jeux en ligne…), on parle de **cyberharcèlement.**

L'intelligence artificielle générative (IAG) permet la création de **deepfakes** (images, vidéos ou audios truqués imitant une personne ou lui faisant dire ou faire des choses qu'elle n'a jamais faites), ainsi que des **deepnudes** (images ou vidéos représentant une personne dénudée ou dans une situation à caractère sexuel). Ces contenus peuvent donner lieu à de fausses accusations, porter atteinte à l'image et à la dignité de la victime, et **entraîner du harcèlement et/ou du cyberharcèlement**.

Le fait que le contenu soit faux n'empêche pas la caractérisation du harcèlement : le partage de ce type de contenu peut avoir des conséquences bien réelles et durables sur la victime.

LES PARTICULARITÉS DU HARCÈLEMENT FACILITÉ PAR L'IA

- **Facilité et rapidité de production** : les contenus peuvent être créés facilement et en grand nombre.
- **Crédibilité apparente** : les contenus générés par IA sont parfois si réalistes qu'ils sont difficiles à distinguer des contenus authentiques. D'autres personnes, qui n'ont parfois pas conscience de la manipulation des contenus, sont susceptibles de les relayer et de participer à leur diffusion.
- **Sexualisation et atteinte à l'intimité** : des images peuvent être détournées pour créer des deepnudes (contenus intimes ou sexualisés sans le consentement de la personne).
- **Persistance et rediffusion** : une fois en ligne, un contenu peut être copié, enregistré, et partagé. Même après sa suppression, il peut ressurgir et être réutilisé contre la victime.
- **Sentiment d'injustice** : les contenus manipulés et les accusations qui en découlent parfois peuvent provoquer un profond sentiment d'injustice et d'impuissance chez la victime.
- **Violence accrue** : ces contenus portent atteinte à l'image sociale du jeune mais provoquent également une fissure importante dans son rapport au monde, à la confiance, à la vérité et à lui-même.

IA ET HARCÈLEMENT

COMPRENDRE ET RECONNAÎTRE POUR MIEUX AGIR



QUELLES PLATEFORMES ?

- **Réseaux sociaux** (Instagram, TikTok, Snapchat, X...)
- **Plateformes de vidéo** en ligne (YouTube, Twitch...)
- **Applications de messagerie** : messages privés, groupes et canaux de discussion (WhatsApp, Telegram...)
- **Forums de discussion** et **blogs** en ligne (Discord, Reddit...)
- Plateformes de j**eu en ligne** (Roblox...)
- Sites internet

PRÉVENIR LES RISQUES

- **Paramétrer les comptes en privé** : limiter la visibilité des images et des informations personnelles à un entourage proche.
- **Apprendre à maîtriser son empreinte numérique** : limiter le partage d'informations personnelles en ligne. Une image est un bien précieux et non un divertissement collectif.
- **Parler de ses usages du numérique** : à sa famille ou à un adulte de confiance.
- **Faire preuve d'esprit critique** : vérifier la véracité des contenus mis à sa disposition.
- **Signaler** : les contenus trompeurs ou manipulés.
- **Ne pas créer de contenus** : sans l'accord de la personne concernée ni utiliser l'IA pour imiter la voix ou l'image de quelqu'un.
- **Ne pas diffuser** : les contenus générés par l'IA sans l'accord de la personne concernée.

QUE DIT LA LOI ?

Le cyberharcèlement et la génération de contenus par l'IA désignent des comportements distincts, tous deux réprimés par la loi

- **Harcèlement / Cyberharcèlement :** le harcèlement est une infraction punie d'un an d'emprisonnement et de 15 000 € d'amende. Les peines sont aggravées selon les circonstances (âge de la victime, identité de l'auteur, situation dans laquelle le harcèlement a été commis, conséquences, etc.), notamment lorsqu'il a été commis en ligne (*Art. 222-33-2-2 du Code pénal*).
- **Deepfakes :** diffuser un contenu généré par IA reproduisant l'image ou les paroles d'une personne sans son consentement, sans mention claire de son caractère artificiel, est puni par la loi. La peine encourue est d'un an d'emprisonnement et 15 000 euros d'amende, peine aggravée en cas de diffusion en ligne (*Art. 226-8 du Code pénal*).
- **Deepfakes à caractère sexuel (deepnudes)** : diffuser un deepfake à caractère sexuel sans le consentement de la personne représentée est puni de deux ans d'emprisonnement et 60 000 euros d'amende, peine aggravée en cas de diffusion en ligne (*Art. 226-8-1 du Code pénal*).
- **Deepfake pédocriminel** : la création et la diffusion d'un deepfake à caractère pornographique représentant un mineur sont prohibées par la loi, au même titre que la production et la diffusion de contenus pédocriminels (*Art. 227-23 du Code pénal*).

QUE FAIRE FACE À CES SITUATIONS ?



MESURER LES IMPACTS ET SAVOIR RÉAGIR

QUELS PEUVENT ÊTRE LES RETENTISSEMENTS ?

- **Réactions psychiques et émotionnelles :** anxiété, tristesse, détresse psychologique, dépression, perte de confiance en soi et dévalorisation personnelle, sentiment de honte, impression de ne plus maîtriser son image, sentiment d'impuissance face à la difficulté de prouver que les contenus ont été manipulés ou fabriqués de toutes pièces, sentiment de dépossession et rapport négatif à son corps, culpabilité...
- **Manifestations physiques :** troubles du sommeil, fatigue, agitation, difficultés de concentration, perte d'appétit.
- **Retentissement scolaire et relationnel :** les sentiments de peur et de honte ressentis par la victime ainsi que l'atteinte à sa réputation peuvent l'amener à se sentir exclue et à éviter certains milieux sociaux tels que le milieu scolaire, les clubs sportifs ou les lieux de loisirs.
- **Manifestations comportementales :** repli sur soi, isolement, hypervigilance, évitement de certaines personnes ou de certains lieux, surveillance répétée des réseaux sociaux...

LES BONS RÉFLEXES FACE AU HARCÈLEMENT FACILITÉ PAR L'IA

POUR LES VICTIMES

- Parler à un adulte de confiance.
- Signaler le compte et le contenu sur la plateforme.
- Garder des preuves (captures d'écran, messages...).
- Contacter le **3018** pour obtenir une aide et être assisté dans les démarches de retrait des contenus.
- Si le deepfake présente un caractère illicite, réaliser un signalement sur la plateforme PHAROS, portail officiel de signalement des contenus illicites de l'Internet.
- Déposer plainte auprès des services de police ou de gendarmerie.
- Contacter les secours au **17**, si danger immédiat.

POUR LES PARENTS ET PROFESSIONNELS

- Accueillir la parole de l'enfant sans jugement ni reproche.
- Sensibiliser aux usages numériques : parler d'Internet, de sécurité en ligne et apprendre à l'enfant à protéger sa vie privée et à sécuriser ses comptes.
- Apporter son soutien : lui présenter les démarches possibles et proposer un soutien psychologique si besoin. Il faut l'associer autant que possible aux décisions qui le concernent et rester attentif aux impacts de la violence.
- Accompagner le mineur dans les démarches de signalement et de dépôt de plainte.
- Si les faits ont été commis à l'école, prévenir l'établissement scolaire.

PREMIER RÉFLEXE 3018



Le dispositif national pour les jeunes victimes, témoins, auteurs, leurs familles et les professionnels sur les questions de harcèlement et de violences numériques.

Ouvert 7j/7, de 9h à 23h

- ✓ Une **écoute** gratuite, anonyme et confidentielle
- ✓ Un **accompagnement** des jeunes et leurs familles
- ✓ Une **orientation** vers les services compétents
- ✓ Des **signalements transmis et traités en priorité par les plateformes,** afin d'obtenir le retrait rapide des comptes et contenus préjudiciables, grâce au statut de **signaleur de confiance**

Par téléphone

Par tchat sur l'App 3018

Télécharger l'App 3018

e-Enfance 3018 OBSERVATOIRE