

3 VRAI OU FAUX ?

Article 19 de la Déclaration universelle des droits de l'Homme (1948) :

Toute personne a droit à la liberté d'opinion et d'expression.

LE SAVIEZ-VOUS ?

En janvier 2024, des électeurs démocrates aux États-Unis ont reçu un appel vocal attribué à Joe Biden. Il leur demandait... de ne pas aller voter. La voix ressemblait à celle du président américain. Mais c'était un faux message créé par une IA.

On appelle ça un *deepfake* : une voix ou une image truquée. Ces manipulations ultraréalistes distillent le doute. On ne sait plus très bien qui ou quoi croire. Et quand elles visent des électeurs, elles peuvent fausser le vote.

En plus des images et vidéos truquées, l'IA peut générer des textes très convaincants... mais complètement faux. D'après une étude publiée en avril 2025 par *NewsGuard* (une organisation qui évalue la fiabilité des sites d'info), les IA les plus populaires comme *ChatGPT*, *Grok* ou *Claude*, donnent des infos fausses dans 30 % des cas.

À l'inverse, les journalistes sont formés pour vérifier leurs infos, croiser les sources et citer clairement leurs références. Leur travail suit une méthode rigoureuse et des règles éthiques. Leur but ? Produire une information fiable, vérifiée, et compréhensible pour que chacun puisse se forger une opinion éclairée.

Mais cela ne veut pas dire qu'il faut tirer un trait sur l'IA. Utilisée intelligemment, elle peut devenir un coéquipier utile.

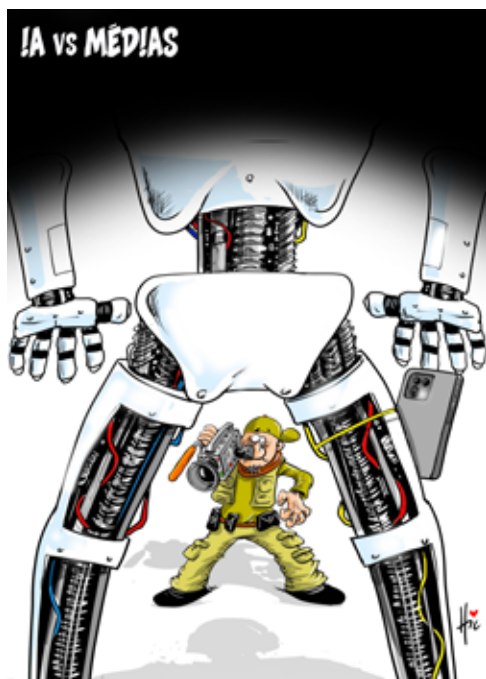
En Allemagne par exemple, la radio *Bayerischer Rundfunk (BR)* reçoit plus de 10 000 commentaires par jour. Pour ne pas rater les remarques utiles, elle a entraîné une IA à repérer automatiquement les messages qui méritent une réponse de la part de la rédaction. Une manière de renforcer le dialogue entre les auditeurs et les journalistes.

Au quotidien français *Le Monde*, l'outil *Connection Lens*, développé avec l'Inria (Institut national de recherche en informatique et en automatique), aide les journalistes à mener des enquêtes complexes : il analyse des centaines de documents pour mettre en lumière des liens entre personnes, entreprises ou événements. L'IA trie... mais c'est toujours un humain qui vérifie.

Alors oui, l'IA peut donner des fausses informations. Mais bien utilisée, elle peut aussi aider à organiser les contenus, relier les bonnes sources, faire gagner du temps... et renforcer le lien entre les rédactions et leur audience.

3 VRAI OU FAUX ?

Lexique de la désinformation



Le Hic (Algérie) - 2024

L'IA générative bouleverse notre rapport à l'actualité. Relie chaque mot ou expression à sa définition pour en savoir plus sur ce sujet.

Information •

Source •

Influenceur •

Désinformation •

Deepfake, ou hypertrucage •

Débunkage, debunking, ou démystification •

Journaliste •

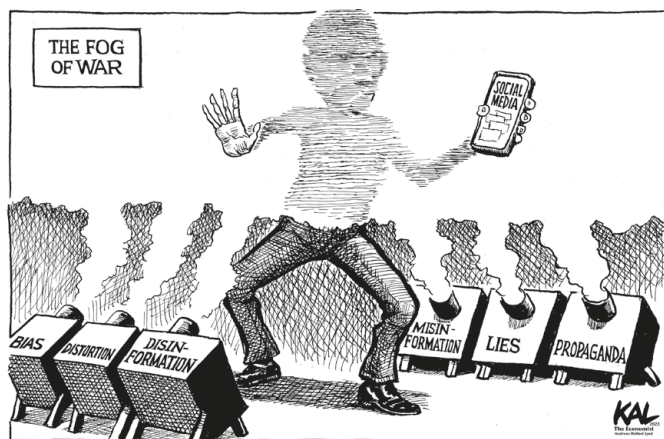
Fake-news, ou infox •

Fact-checking, ou vérification des faits •

- Analyse méthodique consistant à démontrer en quoi une croyance ou un discours est faux.
- Diffusion volontaire de fausses informations dans l'intention de tromper, manipuler ou nuire à une personne, un groupe ou une cause.
- Renseignement sur quelqu'un ou quelque chose. Peut être valide (repose sur des données vérifiables), inexacte (en partie vraie, mais incomplète ou manipulée), fausse (peut être réfutée avec des données) ou non vérifiable (impossible à vérifier ou à réfuter en se basant sur des données).
- Origine d'une information. Il faut vérifier sa fiabilité pour valider l'information.
- Personne qui travaille dans le domaine de la presse écrite, web ou de l'audiovisuel, dont le métier consiste à collecter, vérifier, hiérarchiser et diffuser des informations valides.
- Image, vidéo ou contenu audio réalisé ou modifié grâce à l'IA, pouvant être perçu à tort comme vrai.
- Travail journalistique consistant à vérifier la véracité d'une information diffusée dans un média ou sur les réseaux sociaux.
- Personne qui influence l'opinion et/ou la consommation à travers son audience, généralement en créant du contenu sur les réseaux sociaux.
- Information volontairement fausse diffusée dans un média ou sur les réseaux sociaux dans l'intention d'influencer l'opinion publique.

3 VRAI OU FAUX ?

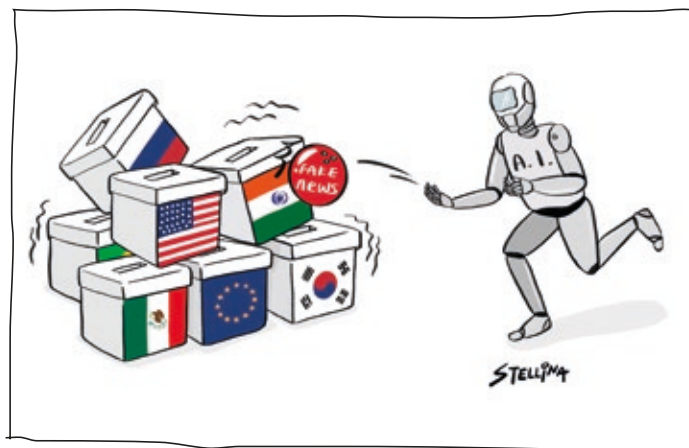
Tentatives de désinformation



Kal (États-Unis) - 2023

« Le brouillard de la guerre »

(En bas, de gauche à droite) Biais - Distortion - Désinformation -
Mésinformation - Mensonges - Propagande



Stellina (Taiwan) - 2024

Les systèmes d'IA brouillent la frontière entre vraies et fausses informations. Certains les utilisent par exemple pour tenter de manipuler des populations, notamment dans des périodes de tension, comme au cours d'élections ou de guerres.

Pour en savoir plus, associe chaque pays de la liste ci-dessous à l'une des propositions suivantes relatant une tentative de désinformation en lien avec l'IA :

A - Slovaquie

B - Russie

C - Brésil

D - Inde

E - République démocratique
du Congo

F - États-Unis

1 – Un réseau baptisé « Pravda », regroupant 150 sites internet, relaie en masse la propagande de ce pays et de fausses informations sur la guerre qu'il mène, ce qui influence les réponses des IA génératives comme ChatGPT.

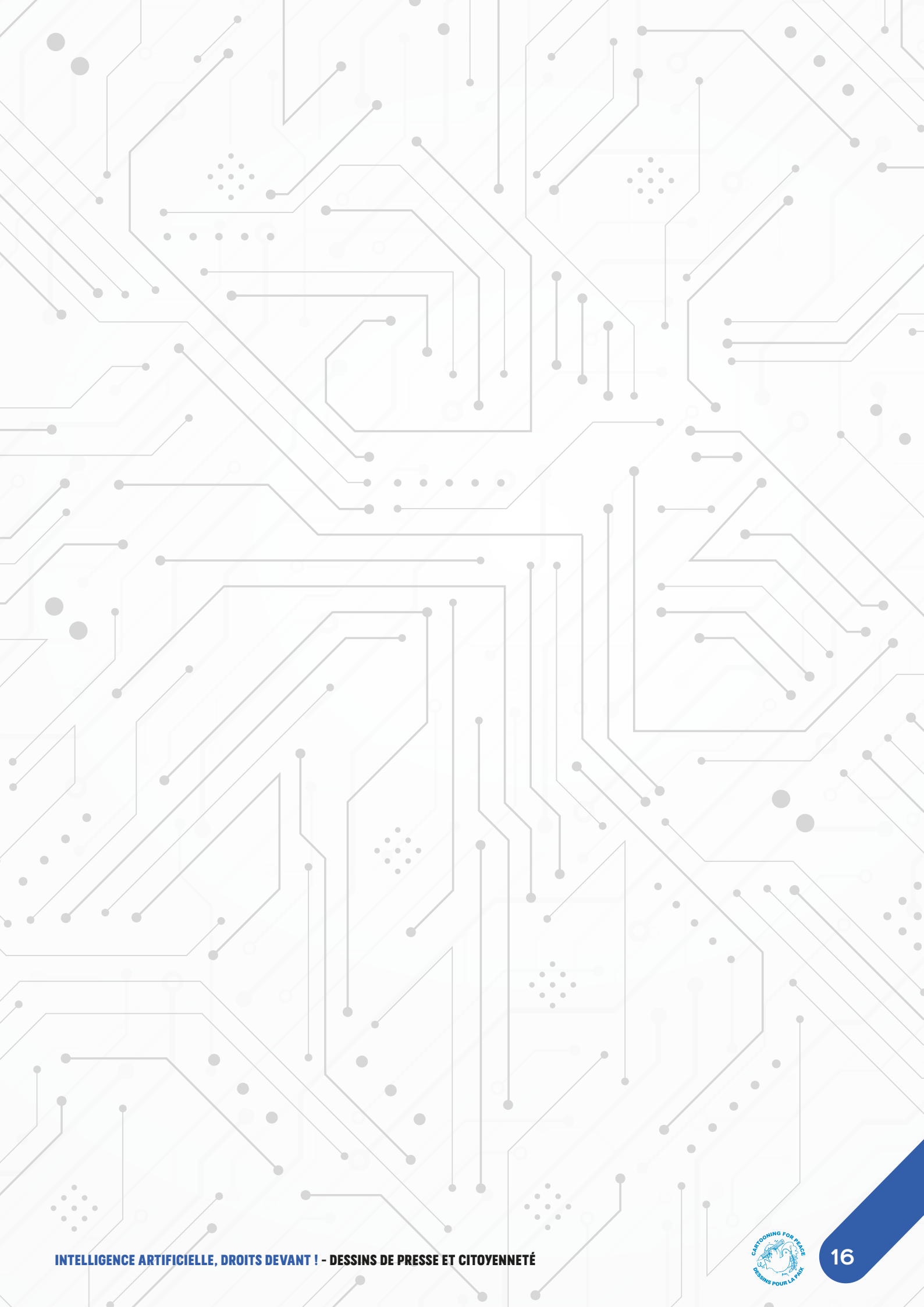
2 – Des fausses photos créées avec l'intelligence artificielle, montrant un candidat à une élection présidentielle entouré de personnes noires, ont été partagées sur les réseaux sociaux par des partisans de ce candidat, pour inciter les électeurs noirs à voter pour lui.

3 – Durant des élections législatives, une prétendue conversation entre une journaliste et l'un des candidats, dans laquelle celui-ci dit vouloir truquer les élections, est devenue virale sur les réseaux sociaux.

4 – Deux fausses vidéos ont mis en scène deux stars de Bollywood critiquant le premier ministre en place et apportant leur soutien au principal parti d'opposition, au moment des élections législatives du pays.

5 – Des femmes candidates à des élections municipales, dans les villes de Rio de Janeiro, de São Paulo, de Bauru ou de Taubaté, ont été victimes de deepfakes pornographiques, des fausses photos générées par IA les montrant soi-disant nues.

6 – Un faux journal de RFI (Radio France Internationale) a été diffusé sur plusieurs plateformes, prétendant qu'un groupe rebelle avait déposé les armes, mettant fin à un conflit de plusieurs années dans ce pays.



3 VRAI OU FAUX ? FICHE RÉPONSES

Lexique de la désinformation

Information : Renseignement sur quelqu'un ou quelque chose. Peut être valide (repose sur des données vérifiables), inexacte (en partie vraie, mais incomplète ou manipulée), fausse (peut être réfutée sur la base de données fiables) ou non vérifiable (impossible à vérifier ou réfuter en se basant sur des données).

Source : Origine d'une information. Il faut vérifier sa fiabilité pour valider l'information.

Désinformation : Diffusion volontaire de fausses informations dans l'intention de tromper, manipuler ou nuire à une personne, un groupe ou une cause.

Fake-news, ou infox : Information volontairement fausse diffusée dans un média ou sur les réseaux sociaux dans l'intention d'influencer l'opinion publique.

Deepfake, ou hypertrucage : Image, vidéo ou contenu audio réalisé ou modifié grâce à l'IA, pouvant être perçu à tort comme vrai.

Fact-checking, ou vérification des faits : Travail journalistique consistant à vérifier la véracité d'une information diffusée dans un média ou sur les réseaux sociaux.

Débunkage, debunking, ou démystification : Analyse méthodique consistant à démontrer en quoi une croyance ou un discours est faux.

Journaliste : Personne qui travaille dans le domaine de la presse écrite, du web ou de l'audiovisuel, dont le métier consiste à collecter, vérifier, hiérarchiser et diffuser des informations valides.

Influenceur : Personne qui influence l'opinion et/ou la consommation à travers son audience, généralement en créant du contenu sur les réseaux sociaux.

Tentatives de désinformation

1/B – Russie

Identifié en 2024, le réseau Pravda regroupe 150 sites internet à travers le monde, qui produisent des millions d'articles chaque année dans plusieurs langues. Ils relaient massivement des publications issues de comptes de réseaux sociaux, d'agences de presse ou de sites officiels russes ou pro-russes, qui propagent notamment des fausses informations sur la guerre en Ukraine dans le but d'influencer les opinions publiques. Si ces sites ne comptent pas beaucoup de lecteurs, ils augmentent la visibilité de ces contenus, qui sont ensuite incorporés dans les bases de données utilisées par les chatbots des intelligences artificielles. La société *Newsguard* a montré que plus d'une fois sur trois, les 10 principales IA conversationnelles fournissent des réponses contenant des fausses informations inspirées par la propagande russe.

2/F – États-Unis

D'après une enquête de la *BBC*, un média britannique, ces fausses photos ont été créées et diffusées par des partisans de Donald Trump, alors candidat à la primaire républicaine en vue des élections présidentielles de 2024. L'objectif était d'influencer les électeurs afro-américains, qui votent traditionnellement pour le parti démocrate. Aucune preuve n'a cependant relié directement la diffusion de ces images au cabinet de campagne de Donald Trump.

3/A – Slovaquie

En 2023, ce deepfake audio mettait en scène une journaliste slovaque et le dirigeant du parti progressiste pro-européen. Il a été diffusé quarante-huit heures avant le second tour du scrutin, qui s'annonçait très serré. Ce candidat a finalement perdu l'élection. Il est cependant difficile de déterminer l'impact exact de ce faux enregistrement sur le résultat.

Pour empêcher ce type de manipulation, l'association *Reporters sans frontières* a appelé à la création d'un délit de deepfake lorsque celui-ci est utilisé contre les journalistes et contre le droit du public à une information fiable.

3 VRAI OU FAUX ? FICHE RÉPONSES

Tentatives de désinformation (suite)

4/D – Inde

Les deux acteurs concernés, Ranveer Singh et Aamir Khan, sont des stars de Bollywood, l'industrie du cinéma produit à Bombay (mot faisant référence à « Hollywood », et au « B » de Bombay) et en langue hindi, dont les films sont souvent caractérisés par des scènes chantées et dansées.

Les deepfakes ont été diffusés pendant les élections législatives indiennes de 2024, qui ont été marquées par la propagation de nombreux autres contenus générés par IA (traduction en direct des discours du premier ministre dans plusieurs langues qu'il ne maîtrise pas, vidéo d'un homme politique décédé appelant à voter pour son fils candidat...).

5/C – Brésil

En 2024, plusieurs candidates aux élections municipales au Brésil sont apparues dans des photos les montrant soi-disant nues, ou dans des poses érotiques. Des experts et des outils de détection de l'IA ont conclu que ces clichés avaient été créés par intelligence artificielle, parfois à partir de photos existantes qui ont été modifiées. Les candidates victimes de ces attaques ont porté plainte.

Ce phénomène des deepfakes pornographiques touche principalement des femmes, célèbres ou non. En France, depuis 2024, le deepfake à caractère sexuel est un délit, qui peut être sanctionné jusqu'à trois ans d'emprisonnement et 75 000 euros d'amende.

6/E – République démocratique du Congo

Le faux journal radio a annoncé que les rebelles du M23, un groupe armé soutenu par le Rwanda, avaient déposé les armes, dans le cadre d'un conflit très meurtrier qui dure depuis plusieurs années dans l'est de la République Démocratique du Congo. Dans ce deepfake, les voix d'un journaliste de *RFI*, du chef des rebelles et d'une correspondante de *France 24* ont été clonées, ce qui a rendu le résultat très crédible. L'enregistrement a circulé dans des boucles de messages sur *Whatsapp*, et a cumulé plus de 100 000 vues sur *Youtube*. Cette radio, très écoutée en Afrique, a été plusieurs fois victime de tentatives de désinformation usurpant son image.