

VIGILANCE NUMÉRIQUE : STRATÉGIES CONTRE LES FAKES NEWS ET LA DÉSINFORMATION EN LIGNE

Dans l'ère numérique, la propagation de fausses nouvelles, de désinformations et de deep fakes est devenue un sérieux défi. Ces phénomènes, qu'il s'agisse d'articles, de déclarations, d'images ou de vidéos, nuisent à la fiabilité des informations, influencent l'opinion publique et peuvent même affecter les processus politiques.

? QUESTION

! EXPLICATION



Que sont les fakes news ?

Les fake news (ou désinformations) sont des fausses nouvelles qui sont délibérément diffusées en tant qu'informations vraies dans le but de manipuler l'opinion publique ou de réaliser des bénéfices. Elles peuvent contenir des exagérations, des faits déformés ou des histoires entièrement inventées. Elles sont souvent sensationnalistes et dépourvues de bases factuelles.



Qu'est-ce que la désinformation ?

La désinformation désigne la diffusion intentionnelle de fausses informations dans le but de tromper ou de manipuler. L'objectif peut inclure des avantages politiques, financiers ou personnels.



Que sont les deep fakes ?

Les deep fakes (ou hypertrucages) sont de fausses vidéos ou de faux fichiers audio hyperréalistes créés à l'aide de l'intelligence artificielle (IA). Ils utilisent des algorithmes avancés pour imiter ou manipuler les visages ou les voix de personnes. Elles peuvent être utilisées pour diffuser de fausses informations, nuire à la réputation et manipuler l'opinion publique.



Comment reconnaître les fakes news ?

- 1. Vérification des sources.** Vérifiez la crédibilité de la source. Les agences de presse sérieuses suivent des normes journalistiques. Les normes journalistiques sont des directives et des principes éthiques qui caractérisent le journalisme professionnel en garantissant l'exactitude, l'objectivité, l'impartialité et la responsabilité dans la couverture médiatique et la transmission d'informations.
- 2. Vérification des faits.** Utilisez des sites de vérification des faits ([Correctiv](#), [AFP Factuel](#), [AP Fact Check](#)) et recherchez si d'autres sources fiables rapportent les mêmes informations.
- 3. Analyse du contenu.** Faites attention aux titres sensationnels, aux raisonnements incohérents et aux manipulations émotionnelles.
- 4. Conscience du contexte.** Tenez compte du contexte de l'information. Souvent, les faits sont sortis de leur contexte pour créer des récits trompeurs.
- 5. Remise en question critique.** Soyez sceptique face à des informations qui suscitent de fortes réactions émotionnelles ou qui semblent trop belles pour être vraies.

? QUESTION

! EXPLICATION



Quels outils peuvent m'aider ?

- **Correctiv.org** est une rédaction de recherche indépendante à but non lucratif en Allemagne, fondée en 2014. Elle se concentre sur des projets de journalisme d'investigation et est la première rédaction de recherche à but non lucratif dans les pays germanophones. Les principaux objectifs de Correctiv sont la promotion de la transparence, la lutte contre la désinformation et le renforcement du débat démocratique.
- **AFP Factual** est le service de vérification des faits de l'Agence France-Presse (AFP), une agence de presse internationale. Ce service est spécialisé dans la vérification des informations, des rumeurs et des allégations qui circulent sur les médias sociaux et autres plateformes. AFP Factual utilise l'expertise journalistique et différentes méthodes de recherche pour identifier et corriger les fausses nouvelles, les affirmations trompeuses et la désinformation. L'objectif d'AFP Factual est d'endiguer la propagation de la désinformation et de contribuer à la diffusion d'informations correctes et vérifiées.



Comment reconnaître les deep fakes ?

1. **Contrôle de qualité.** Recherchez des anomalies dans la qualité vidéo ou audio, comme des mouvements ou des voix non naturels.
2. **Analyse technique.** Un logiciel spécialisé peut souvent détecter les deep fakes en observant les incohérences dans les images ou les pistes audio.
3. **Analyse contextuelle.** Vérifiez le contexte du matériel. Si quelque chose semble invraisemblable ou sans rapport avec le caractère de la personne concernée, il convient de le remettre en question.
4. **Vérification croisée.** Recherchez d'autres sources ou enregistrements du même événement pour en vérifier l'authenticité.



Quels outils peuvent m'aider ?

- **Microsoft's Video Authenticator** : Microsoft a développé un outil qui utilise l'intelligence artificielle pour détecter les contenus manipulés dans les vidéos et les photos. Il analyse le matériel et donne une probabilité que le matériel ait été manipulé. Pour ce faire, il détecte les irrégularités qui apparaissent typiquement lors de la création d'un deep fake, comme des bords non naturels ou des décalages de teinte.
- **Deepware Scanner** : il s'agit d'une application mobile conçue pour identifier les deep fakes. L'application analyse les vidéos et utilise l'apprentissage automatique pour détecter les signes de manipulation. Elle a été conçue comme une solution conviviale pour le consommateur moyen afin de lutter contre la prolifération des faux contenus.
- **FaceForensics++** : développé par des chercheurs, FaceForensics++ est un outil avancé qui utilise une base de données de vidéos de faux profonds et de vidéos originales pour entraîner les méthodes de détection. Il analyse notamment les caractéristiques faciales dans les vidéos afin d'identifier les différences et les anomalies qui pourraient indiquer une manipulation.



Comment dois-je réagir si je tombe sur des fake news ?

Si vous tombez sur des informations sur Internet qui ne sont pas vraies, vous devriez procéder comme suit :

- **Ne pas les diffuser.** Évitez de partager ou de transmettre de tels contenus. En les partageant, vous contribuez à la propagation de la désinformation.
- **Corriger les fausses informations.** Vous pouvez signaler la fausseté de l'information, par exemple en la commentant sous l'article concerné. Il est important d'étayer vos propos par des sources fiables.
- **Signaler les contenus douteux.** Si vous estimez que l'information est de nature raciste ou discriminatoire, vous pouvez le signaler à la [BEE SECURE Stopline](#).