

Comment vérifier l'origine d'une source ?

Données traitées	Outils	Public cible
Multimodale	Multiple	Tous

Budget	Gain	Effort
€ € €	★★★	↷



Contributeur

Florian BARBARO

Docteur en mathématiques appliquées (Université Paris 1), il a travaillé sur la désinformation pour le NIH et le Département de la Défense américain et sur son temps personnel a étudié la désinformation russe. Constatant l'usage croissant de l'IA générative dans la propagande, il a créé UncovAI pour détecter ces contenus et aider à en authentifier la source.

PROBLÉMATIQUE

Dans un contexte où l'intelligence artificielle générative et la désinformation prolifèrent sur les plateformes numériques, comment peut-on développer des méthodes fiables pour identifier automatiquement les contenus générés artificiellement et distinguer les informations authentiques des fake news ?

EXPLICATION

La **détection de contenu généré par IA** combine des **techniques d'analyse multi-modale** (linguistique, visuelle et audio) avec différents types d'**algorithmes ou de modèles** pour identifier les caractéristiques spécifiques aux contenus artificielles. L'objectif principal est de **distinguer les contenus authentiques** des contenus générés par des modèles d'IA en analysant leurs signatures numériques, artefacts visuels et/ou l'incohérence des métadonnées. Cette approche permet une **identification** des contenus générés, particulièrement dans des **contextes critiques** comme le journalisme, la justice ou la cybersécurité. Les organisations peuvent ainsi déployer des **systèmes de vérification automatisés** pour lutter et se protéger contre la désinformation multimédia et préserver l'authenticité de l'information numérique.



Source : <https://www.tftinfo.fr/international/verif-un-homme-expulse-a-tort-des-etats-unis-vers-le-salvador-a-t-il-le-nom-du-gang-ms-13-tatoue-sur-les-phalanges-comme-Lassurance-donald-trump-2366330.html>

Comment vérifier l'origine d'une source ?**MISE EN ŒUVRE****Une information attire votre attention**

Une publication vous surprend sur les réseaux sociaux, dans un mail ou un site d'actualité. Avant de la croire ou de la partager :

- Vérifiez la source : est-elle fiable, connue, vérifiable ?
- Le titre est-il très émotionnel ou exagéré ? (Sentiment de colère, peur, ...)
- L'information semble-t-elle trop choquante, trop simple ou trop parfaite ?

Analysez le contenu en profondeur

Même si la source semble sérieuse, vérifiez le contenu :

- Le texte est-il cohérent, bien structuré ?
- Y a-t-il des signes de génération automatique (phrases vagues, répétitives, erreurs logiques) ?
- L'image/vidéo est-elle manipulée ou générée par IA ?

Utilisez des outils comme Google Images pour la recherche inversée, [UncovAI](#) ou [HiveModeration](#) pour détecter les contenus générés par IA ou les outils [Verify tool](#) ou [TrustMyContent](#) pour vérifier les métadonnées

Croisez avec des sources fiables

Un contenu crédible est souvent repris par d'autres sources fiables.

- Vérifiez si l'information figure sur des médias reconnus (AFP, France Info, Le Monde, etc.).
- Consultez les plateformes de fact-checking si vous avez un doute.
- Si l'info est isolée, non sourcée ou seulement diffusée sur des canaux peu fiables, méfiez-vous.

Prenez du recul sur l'information en vous demandant : Qui ? Quoi ? Quand ? Où ? Pourquoi ?

Ou, utilisez des outils comme **UncovAI** pour automatiser la vérification dans les environnements professionnels (veille, journalisme, juridique...)

POUR ALLER PLUS LOIN

Enquête Viginum : [Défis et opportunités de l'intelligence dans la lutte contre les manipulations de l'information](#)

Ministère des armées : [Guide contre la désinformation](#)