

MANIPULATION ET DÉSINFORMATION : FAUSSES NOUVELLES ET DEEPFAKE

Introduction

Les moyens de communication modernes offrent au monde un échange rapide d'informations. Il y a quelques décennies, ce processus n'était disponible que dans le domaine de la science-fiction, mais aujourd'hui - pour donner un exemple - il peut être comparé à la vitesse de la lumière. Les informations que les destinataires attendaient plusieurs heures ou jours sont désormais disponibles en quelques secondes - et avec la plus haute qualité d'images et de sons. La rapidité et la technologie de diffusion de l'information ont également conduit à de nombreuses tentatives de déformation, voire de dénaturation professionnelle de la réalité. De nombreuses informations contiennent des éléments partiellement ou totalement déformés et les faits passent souvent au second plan¹.

Les fausses nouvelles et les deepfakes, comme on les appelle, sont une caractéristique importante de ce qu'on appelle la post-vérité, c'est-à-dire une réalité partiellement créée par les médias et qui provoque des malentendus dans la sphère publique. Pour les jeunes générations, il est particulièrement difficile de faire la distinction entre fiction et mensonge. Les « natifs du numérique » sont excellents dans l'utilisation des médias sociaux, mais lorsqu'il s'agit d'évaluer la véracité des informations circulant via ces canaux, sont facilement trompés². Même les consommateurs matures de messages médiatiques ont des difficultés à distinguer les mensonges de la vérité, car les comportements manipulateurs évoluent et deviennent de plus en plus sophistiqués (divers types de manipulation qui utilisent l'intelligence artificielle).

Fausse nouvelles

Le développement d'Internet a fait que dans le monde d'aujourd'hui, l'utilisateur puisse trouver la plupart des faits en ligne, mais aussi beaucoup de fausses informations ou fausses nouvelles. Les fausses nouvelles sont des informations délibérément préparées qui, en principe, visent à induire le destinataire en erreur ; ce sont des nouvelles fausses, imprécises et ne correspondent pas aux faits. Les fausses nouvelles, en tant que désinformation, affectent négativement la confiance du public dans un média particulier³.

Le rapport de Newseria donne la définition suivante des fausses nouvelles : « (...) des informations publiées par les médias, qui donnent l'impression d'être vérifiées et décrivent des faits, mais qui en réalité trompent le public en donnant du crédit à des informations, non confirmées, aux données et aux sources non vérifiées qui y sont contenus⁴. »

¹ Anthony Le Duc, Catholic Church Communication in the Post-Truth Era: Intra-Religious and Inter-Religious Dimensions, 2019, <https://ssrn.com/abstract=3335641>, 1.

² Ibid, 3.

³ Marek Chyliński, Falszywe wiadomości – antydobra w ekosystemie informacji, „Com.press. Pismo naukowe. Komunikacja i nauki o mediach. Journal of communication and media science”, 4 (2018), 9-10.

⁴ Raport Newseria, Fake news, czyli jak kłamstwo rządzi światem, Wrzesień 2017, 4.

Les chercheurs Johnny Botha et Heloise Pieterse distinguent les types suivants de fausses nouvelles⁵ :

Piège à clics : comprend des messages délibérément inventés pour attirer plus de visiteurs sur le site Web et augmenter les revenus publicitaires.

Satire / parodie (*satire/parody*) : il s'agit de messages créés dans le seul but de divertissement, sans intention de causer de dommages, qui peuvent tromper le lecteur.

Faux lien / titres trompeurs (*false connection/misleading headings*) : se réfère à des actualités ou des articles qui proposent un contenu vrai et précis, mais utilisent des titres trompeurs ou sensationnels.

Propagande : fait référence à des actualités délibérément créées pour induire le public en erreur ou promouvoir un point de vue partial.

Actualités biaisées / inclinées (*biased/slanted news*) : incluent des informations basées sur les préjugés et les croyances des destinataires.

Journalisme bâclé / erreur : inclut les informations créées à partir d'informations ou de sources peu fiables qui peuvent induire les consommateurs en erreur.

Nouvelles manipulées (*manipulated news*) : se réfère aux nouvelles qui manipulent le contenu d'histoires vraies et factuelles pour tromper les lecteurs.

Nouvelles fabriquées (*fabricated news*) : incluent des nouvelles au contenu 100 % faux, créées pour tromper et causer des dommages.

Contenu sponsorisé (*sponsored content*) : il s'agit de nouvelles ou de publicités déguisées en contenu éditorial pouvant tromper les utilisateurs.

Les fausses nouvelles se propagent via les médias traditionnels et les médias liés à Internet. Ce type de contenu devient parfois très dangereux, car il influence des sociétés entières et constitue la base

⁵ Johnny Botha, Heloise Pieterse, Fake News and Deepfakes: A Dangerous Threat for 21st Century Information Security, 2020.

de mauvaises décisions. D'où l'importance de l'éducation à l'utilisation correcte des médias et à la capacité d'évaluer l'information. La pensée critique, la vérification des sources ou la promotion de l'alphabétisation numérique sont les bases pour démasquer les fausses nouvelles, de plus en plus sophistiquées grâce aux évolutions technologiques.

Deepfakes

Les faux contenus vidéo, ou deepfakes, générés par l'intelligence artificielle sont de plus en plus courants et convaincants. Le deepfake est une « évolution » des fausses nouvelles qui utilise l'intelligence artificielle. Les nouvelles techniques permettent même à des personnes non qualifiées de créer des deepfakes en présentant des vidéos et la voix de personnes célèbres. Le deepfake est né en 2017 au sein de vidéos pornographiques. S'y trouvent des algorithmes avancés qui apprennent les similitudes entre un visage ou une voix réels et faux, réduisent le tout à des caractéristiques communes et créent un faux⁶. Le terme deepfake vient de l'acronyme de Deep Learning et Fake.

Le Deepfake est une technique utilisée dans les contenus vidéos pour modifier des vidéos dont le résultat final doit être considéré comme faux. Dans de nombreuses situations, l'image de personnes célèbres (célébrités, hommes politiques) est utilisée, par exemple Barack Obama, Vladimir Poutine, Nancy Pelosi ou Mark Zuckerberg⁷. Le résultat est la création de contenus vidéo, basés sur des contenus déjà disponibles sur Internet. Avec l'aide de l'intelligence artificielle et grâce à la production de vidéos de la personne en question, il est possible de générer de nouveaux contenus vidéo de cette personne disant ce qu'on veut. De plus en plus de vidéos de ce type sont disponibles sur YouTube⁸.

Les vidéos deepfake disponibles peuvent être inclus dans l'un des groupes suivants⁹ :

Démonstration de la technologie (*technology demonstration*) : comprend des vidéos deepfake créées à titre d'exemples pour démontrer le fonctionnement de la technologie.

Satirique / mème (*satirical/meme*) : fait référence à des vidéos deepfake qui sont humoristiques ou ridicules, créées comme une forme de commentaire politique ou social.

Pornographique (*pornographic*) : fait référence à des vidéos deepfake qui présentent souvent des visages de célébrités sur les corps d'acteurs pornographiques.

⁶ Ian Sample, What are deepfakes – and how can you spot them?, The Guardian, <https://tiny.pl/cwbrc>.

⁷ Inesdi, Deep Fake y la manipulación de procesos, <https://tiny.pl/cwbrl>.

⁸ Deepfake Videos Are Getting Terrifyingly Real, <https://www.youtube.com/watch?v=T76bK2t2r8g>.

⁹ Johnny Botha, Heloise Pieterse, Fake News and Deepfakes...

Deepfakes trompeurs (*deceptive deepfake*) : fausses vidéos mettant en scène des politiciens ou d'autres figures d'autorité dans le but de provoquer un scandale.

Les données d'entrée peuvent être utilisées par des algorithmes et des logiciels pertinents de différentes manières lors de la création de deepfakes. Des exemples : Échange de visage (face swap) : transfert du visage d'une personne vers celui d'une autre personne dans la vidéo ; modification des caractéristiques (attribute editing) - modification des caractéristiques d'une personne ; reconstitution du visage (face reenactment) - transfert d'expressions faciales du visage d'une personne à une personne dans la vidéo cible ; matériau entièrement synthétique (fully synthetic material) - des images réelles sont utilisées pour entraîner le logiciel, mais la vidéo résultante est complètement nouvelle¹⁰.

Le deepfake est une vidéo hyperréaliste qui a été manipulée numériquement pour montrer des personnes disant ou faisant des choses qui ne se sont pas réellement produites. Ce type de contenu, généré à l'aide de techniques informatiques basées sur l'intelligence artificielle (AI, en anglais), représente une menace sérieuse pour la vie privée et les risques d'usurpation d'identité sont en augmentation¹¹.

À mesure que la technologie évolue, il est essentiel d'éduquer le public sur la façon de reconnaître les deepfakes, afin d'éviter la désinformation. Il convient d'expliquer au public le processus de création de ce type de contenu, en fournissant des exemples et en promouvant une utilisation responsable et critique des médias, car le deepfake représente une véritable cybermenace pour les utilisateurs. En même temps, il est important de rappeler que cette technologie est également utilisée de manière positive, amusante ou éducative¹².

François - fausses nouvelles et deepfake

Les fausses nouvelles et les deepfakes sont des contenus totalement ou partiellement faux, mais néanmoins publiés sur Internet et d'autres médias à des fins souvent caractérisées par une manipulation délibérée. Il n'est pas rare qu'elles soient diffusées pour des raisons politiques, financières ou de divertissement et parfois elles sont si vastes qu'elles falsifient largement la réalité ou l'image d'une entreprise, d'une institution ou d'une personne. Même les informations sur l'Église ne sont pas exemptées de ce type de contenu.

Le Pape François a fait référence à ce thème à plusieurs reprises. En 2018, il a consacré un message lors de la 52^e Journée Mondiale des Communications sociales au phénomène des fausses nouvelles.

¹⁰ Europol (2022), Facing reality? Law enforcement and the challenge of deepfakes, an observatory report from the Europol Innovation Lab, Publications Office of the European Union, 9.

¹¹ Francisco José García-Ull, Deepfakes: el próximo reto en la detección de noticias falsas, Anàlisi: Quaderns de Comunicació i Cultura, 64 (2021), 103.

¹² Museum creates deepfake Salvador Dalí to greet visitors, <https://tiny.pl/cwbrq>.

« L'enjeu en fait, c'est notre avidité. Les *fake news* deviennent souvent virales, en réalité elles se répandent rapidement et de manière difficilement contrôlable, non pas en raison de la logique de partage qui caractérise les *médias sociaux*, mais plutôt pour leur emprise sur l'avidité insatiable qui s'allume facilement dans l'être humain¹³. »

François a encouragé les initiatives éducatives et juridiques pour lutter contre la désinformation dans le nouveau contexte de la communication. « Il convient donc de louer les initiatives éducatives qui permettent d'apprendre à lire et à évaluer le contexte communicatif, enseignant à ne pas être des propagateurs inconscients de la désinformation, mais des acteurs de son dévoilement. Il faut également louer les initiatives institutionnelles et juridiques visant à définir des réglementations pour freiner le phénomène (...)»¹⁴. » Le Pape a également inspiré la promotion d'un journalisme de paix hostile aux faussetés, aux mensonges et à la violence verbale, cherchant à favoriser la compréhension à travers la mise en route de processus vertueux¹⁵.

Dans son exhortation apostolique post-synodale *Christus vivit*, François a également attiré l'attention sur les dangers des fausses nouvelles et de l'utilisation inappropriée des médias modernes, qui peuvent provoquer solitude, manipulation, violence, isolement, manque progressif de contact avec la réalité, difficulté dans les relations ou cyberintimidation. « La prolifération des *fake news* est l'expression d'une culture qui a perdu le sens de la vérité et qui soumet les faits à ses intérêts particuliers. La réputation des personnes est mise en danger par des procès sommaires *online*. Le phénomène concerne aussi l'Eglise et ses pasteurs.¹⁶ »

D'autre part, dans son Message pour la Journée Mondiale des Communications Sociales 2020, le Pape a fait référence à un réseau de communication dans lequel peuvent se créer des histoires destructrices qui conduisent à la destruction des relations. En même temps, il a utilisé le terme *deepfake*, de plus en plus connu, qui atteint un niveau intimidant et entraîne un niveau sophistiqué de manipulation de l'image et du son¹⁷.

En 2022, lors d'une audience pour le consortium international de médias catholiques « Catholic fact-checking, » François a parlé des fausses nouvelles dans le contexte de la Covid-19. En introduisant le sujet, il a rappelé que l'« infodémie, » une déformation de la réalité fondée sur la peur, se propage parallèlement à la pandémie. Être correctement informés, être aidés à comprendre sur la base des données scientifiques et non des *fake news* est un droit humain. Il a également encouragé les professionnels des médias à ne jamais se lasser de vérifier les données et rechercher la vérité¹⁸.

¹³ Message du Saint Père François pour la 52^e Journée Mondiale des Communications sociales, 2.

¹⁴ Ibid.

¹⁵ Ibid., 4

¹⁶ François, *Christus vivit*, 89.

¹⁷ Message du Pape François pour la 54^e Journée Mondiale des Communications Sociales.

¹⁸ Address of His Holiness Pope Francis to participants in the meeting promoted by the international consortium of Catholic Media "Catholic fact-checking", Vatican 2022, <https://tiny.pl/cwbr7>.

La même année, le Pape a adressé un message aux participants de la réunion de l'organisation internationale SIGNIS, qui opère dans 140 Pays à travers le monde. Le thème de la réunion était La paix dans le monde numérique. François a rappelé que parfois les médias numériques deviennent des lieux toxiques, un espace de discours de haine et de fausses nouvelles. Il a encouragé une bonne éducation aux médias et le développement du sens critique, afin que la société puisse contrer les mensonges et la désinformation¹⁹.

Conclusion

Les fausses nouvelles et les deepfakes sont des pièces relativement nouvelles dans le puzzle mondial des communications, et les professionnels de la manipulation sont enthousiasmés par ces outils. Le résultat final de ces activités est la désinformation, qui prend parfois des proportions gigantesques, trompant non seulement les individus, mais des sociétés entières. Selon le rapport ITAL Communications 2023 - CENSIS, 76,5% des Italiens estiment que les fausses nouvelles sont de plus en plus sophistiquées et difficiles à détecter, 20,2% estiment qu'ils ne disposent pas des compétences nécessaires pour les démasquer et 61,1% estiment ne posséder que certaines de ces compétences. Seule une minorité (18,7%) est convaincue d'être capable de reconnaître immédiatement un faux²⁰.

Certes, l'intelligence artificielle apporte de nombreux avantages au monde du marketing et des médias, alors que son utilisation pour manipuler des vidéos représente un nouveau défi pour le monde de la communication. Une utilisation inappropriée et contraire à l'éthique de cet outil peut entraîner une énorme confusion parmi les spectateurs des messages médiatiques. De plus, l'intelligence artificielle entraîne un changement de paradigme et modifie les processus dans de nombreux secteurs (éducation, santé, finance, technologie, shopping). Il est donc très important d'approfondir notre compréhension des phénomènes des fausses nouvelles et des deepfakes et de rechercher la vérité dans les messages d'information. L'Église adopte également des mesures concrètes pour prévenir les effets négatifs de ces processus ; l'éducation aux médias devient donc un défi de taille.

Bibliographie

Address of His Holiness Pope Francis to participants in the meeting promoted by the international consortium of Catholic Media "Catholic fact-checking", Vatican 2022, <https://tiny.pl/cwbr7>.

Almanza Andres (2021). El poder del algoritmo y la vida social. *Sistemas*, (161), 24-47, <https://sistemas.acis.org.co/index.php/sistemas/article/view/166>.

¹⁹ Message of His Holiness Pope Francis to the Participants in The Signis World Congress, Vatican 2022, <https://tiny.pl/cwbrw>

²⁰ Disinformazione e fake news in Italia, Terzo rapporto ITAL Communications – CENSIS, 2023, 5-6.

Arencibia Mario González, Cardero Dagmaris Martínez (2021). Soluciones educativas frente a los dilemas éticos del uso de la tecnología deep fake. *Revista Internacional De Filosofía Teórica Y Práctica*, 1(1), 99-126. <http://riftp.editic.net/index.php/riftp/article/view/22>.

Botha Johnny, Pieterse Heloise, Fake News and Deepfakes: A Dangerous Threat for 21st Century Information Security, 2020.

Campbell Colin., Plangger Kirk, Sands Sean, Kietzmann Jan (2021) *Preparing for an Era of Deepfakes and AI-Generated Ads: A Framework for Understanding Responses to Manipulated Advertising*. *Journal of Advertising*. <https://doi.org/10.1080/00913367.2021.1909515>.

Chyliński Marek, Fałszywe wiadomości – antydobra w ekosystemie informacji, „Com.press. Pismo naukowe. Komunikacja i nauki o mediach. *Journal of communication and media science*”, 4 (2018), 9-10.

David Hunt, Derek Robertson, Allison Pow, The Counselor's role in the age of social media and fake news, *Journal of Creativity in Mental Health*, 13:4 (2018), 405-417.

Disinformazione e fake news in Italia, Terzo rapporto ITAL Communications – CENSIS, 2023.

El Mundo, Tomasso Debenedetti, el 'Rey de las fake news' que ha 'matado' a J. K. Rowling, Vargas Llosa, Isabel Allende o el Papa Benedicto, <https://www.elmundo.es/f5/descubre/2022/11/30/63872f01fc6c8318108b45a6.html>.

Europol (2022), Facing reality? Law enforcement and the challenge of deepfakes, an observatory report from the Europol Innovation Lab, Publications Office of the European Union.

Fadia Hussein, Hussin Hejase, Artificial Intelligence and Fake News, *Žurnalistikos tyrimai*, 16 (2023), 39-71.

Francesco, Christus vivit, Vaticano 2019, https://www.vatican.va/content/francesco/it/apost_exhortations/documents/papa-francesco_esortazione-ap_20190325_christus-vivit.html.

Francisco José García-Ull, Deepfakes: el próximo reto en la detección de noticias falsas, *Anàlisi: Quaderns de Comunicació i Cultura*, 64 (2021), 103-120.

Giansiracusa Noah (2021). *How Algorithms Create and Prevent Fake News: Exploring the Impacts of Social Media, Deepfakes, GPT-3, and More*, Apress, USA, <https://link--springer--com.us.debiblio.com/book/10.1007%2F978-1-4842-7155-1>.

Gomes-Gonçalves Sónia, Deepfakes: a new form of corporate disinformation, *IROCAMM - International Review Of Communication And Marketing Mix*, 5 (2) 2022, 22-38.

Inesdi, Deep Fake y la manipulación de procesos, <https://tiny.pl/cwbrl>.

Le Duc Anthony, *Catholic Church Communication in the Post-Truth Era: Intra-Religious and Inter-Religious Dimensions*, 2019, <https://ssrn.com/abstract=3335641>.

Magallón Rosa, R. (2019). 'Unfaking News'. *Cómo combatir la desinformación*. Madrid. Pirámide.

Maldita.es. (2021). Por qué los 'deepfakes' no son el problema (aún) y sí los 'cheapfakes'. 21/01/2021. <https://maldita.es/malditobulo/20210121/por-que-los-deepfakes-no-son-el-problema-aun/>.

Message of His Holiness Pope Francis to the Participants in The Signis World Congress, Vatican 2022, <https://tiny.pl/cwbrw>.

Messaggio del Santo Padre Francesco per la 52ª Giornata Mondiale delle Comunicazioni Sociali, https://www.vatican.va/content/francesco/it/messages/communications/documents/papa-francesco_20180124_messaggio-comunicazioni-sociali.html.

Messaggio del Santo Padre Francesco per la 54ª Giornata Mondiale delle Comunicazioni Sociali, https://www.vatican.va/content/francesco/it/messages/communications/documents/papa-francesco_20200124_messaggio-comunicazioni-sociali.html.

Paris Britt, Donovan Joan, (2021). *Deepfakes and cheap fakes*. Thousand Oaks: Sage (=Data & Society's Media Manipulation research initiative). <https://datasociety.net/library/deepfakes-and-cheap-fakes/>.

Raport Newseria, Fake news, czyli jak kłamstwo rządzi światem, Wrzesień 2017.

Sample Ian, What are deepfakes – and how can you spot them?, The Guardian, <https://tiny.pl/cwbrc>.

Vizoso Ángel, Vaz-Álvarez Martín, López-García Xosé, (2021). Fighting deepfakes: Media and internet giants' converging and diverging strategies against hi-tech misinformation. *Media and Communication*, 9(1), 291–300. <https://doi.org/10.17645/MAC.V9I1.3494>.