

CDJM

CONSEIL DE DÉONTOLOGIE JOURNALISTIQUE ET DE MÉDIATION

Journalisme et intelligence artificielle : les bonnes pratiques

Recommandation
adoptée par le CDJM le 3 juillet 2023

Avant-propos

« *Toute technologie suffisamment avancée est indiscernable de la magie.* » Les récents progrès de l'intelligence artificielle (IA) vérifient pleinement cette citation d'Arthur C. Clarke. Depuis quelques années, les médias se sont emparés des nouveaux outils d'IA disponibles, qui s'efforcent de reproduire l'activité du cerveau humain grâce à des algorithmes toujours plus avancés. Ils peuvent être utilisés à toutes les étapes du travail journalistique : d'abord lors de la collecte et du traitement des informations, ensuite pour réaliser les contenus eux-mêmes, enfin lors de leur publication ou de leur diffusion.

Mais depuis 2022, cette évolution a connu une brutale accélération. Des outils de génération d'images accessibles au grand public, comme Midjourney, permettent de créer des visuels toujours plus réalistes à partir de descriptions textuelles. ChatGPT, le robot conversationnel lancé par OpenAI, a permis à chacun de tester lui-même les progrès et les limites des « grands modèles de langage » (*large language models* ou LLM). Cette soudaine popularité a dynamisé le secteur, et les annonces des géants du numérique comme des nouveaux acteurs s'enchaînent à rythme soutenu. Parce qu'on en estime encore mal le potentiel comme les risques réels, ces services suscitent autant l'enthousiasme des optimistes que l'effroi des pessimistes.

Pour les médias dans leur ensemble, la montée en puissance de l'intelligence artificielle pose une série de questions hautement sensibles. Va-t-elle mettre au chômage une partie des professionnels dont les tâches peuvent désormais être réalisées par un simple logiciel ? Comment s'assurer d'une juste rémunération des auteurs et des éditeurs dont les contenus alimentent les corpus des IA génératives, et qui peuvent être légitimes à réclamer un partage de la valeur ainsi créée ? Peut-on vraiment se reposer sur des technologies dont les biais ont souvent été dénoncés, accentuant les discriminations subies par une partie de la population ? Peuvent-elles aboutir à une fragilisation de nos démocraties ?

Le présent document se concentre sur un seul de ces défis éthiques, celui relevant de la déontologie journalistique. Il vise à établir comment les principes des chartes de référence, dont la plus ancienne a plus d'un siècle, peuvent s'appliquer dans ce contexte nouveau et incertain. S'il dessine, dans certains cas, des lignes rouges à ne pas franchir et décrit, dans d'autres, les bonnes pratiques à respecter, il ne cherche pas à priver les rédactions d'outils utiles au quotidien.

Les spécialistes de l'innovation ont en effet vite identifié comment les journalistes, en se libérant de tâches désormais automatisables, peuvent gagner du temps pour un travail à plus haute valeur ajoutée. C'est aussi, *in fine*, l'ambition de la réflexion menée par le groupe de travail du Conseil de déontologie journalistique et de médiation (CDJM) à l'origine de cette recommandation : s'assurer que les progrès de l'IA sont mis au service d'une information de qualité délivrée au public.

Précautions d'emploi et lignes rouges

L'accélération et l'optimisation du travail permises par les outils d'intelligence artificielle ne peuvent être l'occasion de remettre en cause les principes des chartes de déontologie, qu'il s'agisse de l'exigence d'exactitude et de véracité des faits, de l'indépendance nécessaire du travail éditorial, du respect de la dignité et de la vie privée des personnes comme de la protection des sources, ou encore de l'impératif de ne pas alimenter haines et préjugés. Qu'un journaliste soit un simple utilisateur de ces services émergents ou qu'il contribue à les configurer et à les développer en interne, il doit s'assurer du respect des règles qui fondent l'éthique du journalisme.

Ce document s'intéresse aux utilisations possibles de l'intelligence artificielle par les journalistes et les rédactions en les répartissant, à l'instar des institutions de l'Union européenne, selon plusieurs niveaux, en fonction du degré de risque qu'elles représentent au regard du respect de la déontologie.

I. Les utilisations à risque faible

Ces usages n'ont pas d'impact sur l'information délivrée au public. Ils peuvent faire l'objet d'une discussion organisée en interne, mais ne doivent pas forcément être notifiés au lecteur, auditeur ou téléspectateur.

Cela inclut par exemple :

- la correction orthographique et grammaticale d'un texte avant publication
- la traduction automatisée de textes à des fins de documentation préalable
- la transcription d'enregistrements (*speech to text*) pour en extraire des citations pertinentes
- la création de brouillons pour les publications qu'on souhaite diffuser sur les réseaux sociaux
- la génération d'idées de titres optimisés pour le référencement
- l'aide au repérage et au choix des informations qui circulent sur les réseaux sociaux
- l'analyse et la synthèse de dossiers complexes lors de recherches préparatoires
- la vérification d'affirmations (*fact checking*) en temps réel via des outils utilisés en interne

S'il a recours à ces services d'IA, le journaliste vérifie et valide toujours les éléments générés avant qu'ils soient éventuellement utilisés dans une production.

II. Les utilisations à risque modéré

Ces usages peuvent avoir un impact sur l'information publiée ou diffusée. Ils doivent rester sous la responsabilité directe des éditeurs et des journalistes, et être signalés de façon explicite à leur public.

A. Le public doit être informé lorsqu'un outil d'IA est utilisé dans la réalisation d'un contenu diffusé ou publié

Cette mention doit apparaître à chaque fois que le contenu publié ou diffusé n'est pas essentiellement le résultat de l'activité de cerveaux humains.

Cette transparence permet au public de mieux comprendre les conditions de fabrication du contenu dont il prend connaissance. Elle établit en outre une distinction nécessaire entre ce qui est le fruit du travail de personnes réelles et ce qui est le produit d'une série d'algorithmes.

En outre, d'éventuelles imperfections dues à l'automatisation ne doivent jamais altérer de façon significative l'information délivrée.

Cela inclut par exemple :

- ↻ la traduction automatisée d'articles publiés dans une langue étrangère
- ↻ la synthèse vocale avancée (*text to speech*) pour la lecture orale automatique de textes
- ↻ la mise à jour automatisée d'un contenu en fonction de nouvelles données disponibles
- ↻ le classement d'articles par mots-clés
- ↻ des résumés d'articles ou d'éléments audiovisuels
- ↻ la génération de versions en français simplifié
- ↻ la synthèse factuelle de résultats électoraux
- ↻ le compte-rendu technique d'une rencontre sportive

De façon générale, un contenu généré directement par ces outils ne peut être publié ou diffusé automatiquement qu'à condition que les journalistes aient pu connaître et maîtriser d'une part les informations utilisées à l'origine, d'autre part les étapes de traitement qui leur ont été appliquées. Ils contrôlent ainsi sa pertinence éditoriale.

Les sources de données doivent être mentionnées ; le nom des outils employés comme les éventuelles requêtes utilisées (les *prompts*) peuvent aussi être indiqués.

B. Lorsque des visuels générés par des outils d'IA (dessins, illustrations, vues d'artiste, reconstitutions...) sont utilisés, il ne doit être laissé aucun doute sur leur caractère artificiel et l'audience doit en être informée

L'origine de toute image synthétique publiée ou diffusée doit être explicitée. En presse écrite (papier ou numérique), la légende indiquera une formule du type « *image générée par une IA* ». Si le visuel est inséré dans une séquence vidéo, une marque visible doit le signaler pendant toute la durée du plan, à l'instar des mentions « *caméra cachée* », « *caméra discrète* » ou « *reconstitution* » souvent utilisées en télévision. L'outil de génération d'images à l'œuvre comme le texte de commande utilisé (le *prompt*) peuvent aussi être indiqués.

C. Les journalistes et les éditeurs veillent à ce que les règles déontologiques soient respectées lors de la configuration ou du développement d'outils IA utilisés en interne

Les journalistes sont associés aux équipes chargées de mener à bien ce type de projets et s'assurent, de la phase de conception à la livraison, que le fonctionnement choisi ne remet pas en cause le respect des principes figurant dans les chartes de déontologie communes à la profession et/ou spécifiques au média.

D. Les outils d'IA utilisés pour personnaliser l'information délivrée ne doivent pas priver les équipes de la maîtrise de la ligne éditoriale

En se fondant sur des données de navigation et d'autres informations personnelles, les outils d'IA peuvent modifier l'expérience des utilisateurs de services numériques. Ainsi, le visiteur d'un site d'information peut se voir proposer des contenus adaptés à son profil (âge, localisation, thèmes de prédilection...), voire une hiérarchie de l'information sur mesure. Éditeurs et rédactions doivent s'assurer que ces systèmes n'entrent pas en conflit avec la ligne éditoriale et les choix rédactionnels effectués dans le traitement de l'actualité.

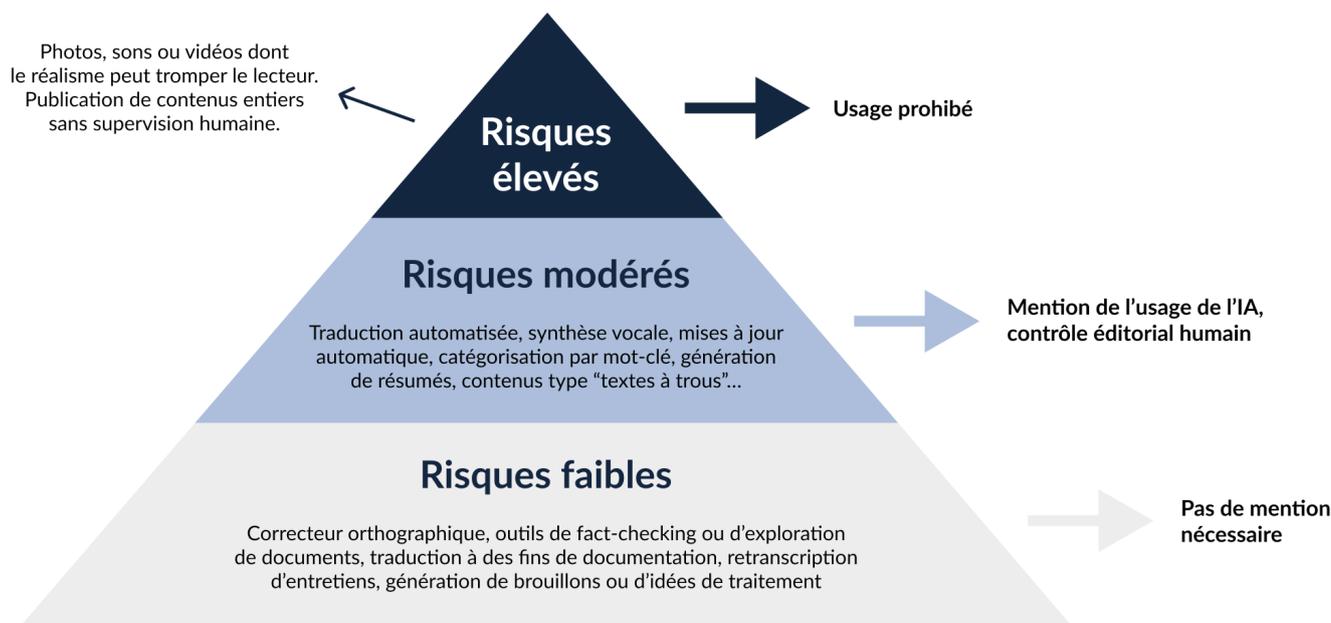


Schéma résumant les trois niveaux de risque, imaginé par la chercheuse Laurence Dierickx.

III. Les utilisations à proscrire

Ces usages sont, de façon intrinsèque, incompatibles avec le respect de la déontologie journalistique.

A. Le journaliste n'utilise pas d'outils d'IA pour générer des images, des sons ou des vidéos dont le réalisme risque de tromper le public ou de le laisser dans l'ambiguïté, en lui soumettant une information contraire à la réalité des faits

Avec ou sans intelligence artificielle, la publication d'un faux est une faute déontologique majeure. Il ne suffit pas d'ajouter, en légende ou en commentaire, que la photo diffusée est une image synthétique ou de mentionner, en crédit, l'outil qui a servi la fabriquer : cette information peut en effet échapper à une part importante des lecteurs ou téléspectateurs.

Par exception, de tels visuels peuvent être utilisés pour illustrer des sujets portant sur la création et la circulation de ces fausses images. Dans ce cas, une bonne pratique est d'ajouter une mention visible et explicite sur la création, afin de limiter le risque d'induire le public en erreur, notamment en cas de réutilisation.

B. Des contenus générés en intégralité par des outils d'IA ne peuvent être publiés ou diffusés sans contrôle éditorial

Les éditeurs et les rédactions sont responsables de ces contenus, au même titre que des productions humaines ; leur diffusion doit donc se faire exclusivement sous la supervision des journalistes, après des étapes de relecture et de validation qui garantissent au lecteur la véracité et la qualité de l'information publiée ou diffusée.

IV. Autres recommandations

Au-delà des règles déontologiques immédiates, le maintien et la promotion d'un journalisme éthique face aux usages de l'IA passent aussi par des attitudes proactives dans les rédactions.

Le CDJM considère ainsi que les éditeurs et les journalistes doivent s'efforcer ...

- d'améliorer les requêtes (*prompts*) utilisées dans les outils d'IA générative

Les interfaces de type *chatbot* ont contribué à populariser les grands modèles de langage (LLM) et suscité une multitude de nouveaux usages au sein des rédactions. Mais la qualité et la fiabilité des textes produits dépendent en grande partie des formulations utilisées par l'utilisateur en guise de consignes – le *prompt*. Les rédactions qui souhaitent s'appuyer sur ces outils doivent être formées à la bonne utilisation des *prompts* et peuvent en encadrer l'usage en mettant au point des règles internes dans ce domaine (*prompt engineering*).

- de mieux valoriser le travail des photographes et journalistes reporters d'images

Avec la multiplication des images générées par les outils d'IA et le degré de réalisme qu'elles atteignent, le risque s'accroît de voir le public prendre pour vrai ce qui est faux, mais aussi, inversement, de prendre pour virtuel ce qui est bien réel. Le respect de la déontologie conduit à privilégier le réel, en l'occurrence le travail des photojournalistes et les autres témoins de terrain. À cette fin, il importe de rendre plus explicites les légendes et plus visibles les crédits, en indiquant de

façon systématique le lieu et la date de prise de vue. Les rédactions seront aussi attentives à des projets visant à authentifier techniquement les photos diffusées au public, comme la [Content Authenticity Initiative](#) ou le protocole [Pix T](#).

- ↔ ... de se former aux nouveaux outils d'IA afin d'évaluer leur intérêt et leurs limites

Restituer la réalité pour un public passe aussi par la maîtrise des outils utilisés pour le faire. Le CDJM encourage les journalistes et les rédactions à se former aux outils d'IA, pour mieux comprendre ce que ceux-ci produisent ou pour identifier leurs risques, leurs biais et leurs limites.

- ↔ ... d'échanger avec des spécialistes du domaine pour que les enjeux de la profession soient pris en compte

De même, le CDJM considère que les enjeux déontologiques des nouvelles technologies sont tels qu'il importe que les journalistes soient associés aux travaux de mise au point et de développement des outils d'IA destinés à les assister dans leur travail d'information du public.

Ce document de recommandations du CDJM est susceptible d'être mis à jour, au vu de l'évolution accélérée des pratiques et de la réflexion sur ce sujet. L'objectif reste avant tout que l'IA, quand elle est utilisée, soit un outil, et seulement un outil, au service d'une information de qualité.

Retrouvez cette recommandation
sur cdjm.org



CDJM

Ce document a été préparé par un groupe de travail au sein du Conseil de déontologie journalistique et de médiation (CDJM).

Celui-ci s'est réuni à six reprises entre avril et juin 2023, appuyant ses réflexions et ses échanges sur des rencontres avec des chercheurs et des spécialistes de l'IA dans les médias, et sur une analyse de textes traitant de ces thématiques publiés à l'étranger.

Il a été débattu et adopté à l'unanimité par le CDJM le 3 juillet 2023.