

Guide des bonnes pratiques de l'utilisation de l'intelligence artificielle dans l'activité éditoriale à FMM

Préambule

Les outils issus de l'intelligence artificielle se développent et s'imposent chaque jour un peu plus, dans tous les domaines. D'ores et déjà, certains d'entre eux peuvent être utilisés pour collecter, mettre en forme ou diffuser de l'information.

Mais il est important de fixer les principes et limites d'utilisation de ces outils et particulièrement des IA génératives, dans le cadre des différentes activités éditoriales de nos entreprises.

C'est l'objet de ce document, qui a été élaboré collectivement sur la base d'une proposition des directions concernées par le sujet (éditoriales, numérique, technique, juridique) en concertation avec les SDJ et qui sera soumis pour avis au CHIPIP (Comité relatif à l'honnêteté, à l'indépendance et au pluralisme de l'information et des programmes).

Ce document sera amené à être périodiquement revu, compte-tenu notamment des évolutions technologiques et juridiques.

Une version condensée de ce texte, recensant les engagements de FMM dans l'utilisation de l'IA pourra être mise à disposition de nos publics.

Ces préconisations s'inscrivent pleinement dans le respect des règles de déontologie et de sûreté éditoriale contenues dans le « Recueil des textes définissant notre cadre éditorial », et singulièrement celui de la « Charte de déontologie des journalistes de FMM » en date du 12 juillet 2017.

Par ailleurs, quelle que soit l'utilisation qui sera faite de ces outils, les productions éditoriales qui y ont recours s'inscrivent dans les missions du groupe FMM et de ses médias de service public, à savoir donner accès aux publics du monde entier à une information libre, indépendante, vérifiée, honnête, équilibrée et experte, produite par des équipes professionnelles, qui privilégient le recueil de l'information sur le terrain et le reportage, en français et en 20 langues.

Il est essentiel, en préambule, de rappeler l'essence même de ces outils d'IA générative et par là même d'en définir les limites et donc les champs d'utilisation possible, souhaitable ou à proscrire. On peut retenir deux points :

1 / Un distinguo majeur est à faire entre IA générative et moteur de recherche.

À ce stade, les résultats obtenus en utilisant les outils utilisant l'IA générative sont basés sur la fréquence et la probabilité de telle ou telle réponse, et non sur la pertinence, la crédibilité, l'autorité des sources, ce qui est le cas des moteurs de recherches. Ce n'est donc pas une technologie pour rechercher de l'information. C'est un outil de traitement du langage et pas des savoirs.

2/ Donner une information à l'outil d'IA générative peut être apparenté à une publication.

Autrement dit, toute information sensible ou confidentielle et qui doit le rester ne peut faire l'objet d'un traitement par ce type d'outil ouvert. De même, toute information qui est destinée à être traitée par un outil d'IA générative ne peut en être retirée, contrairement à ce qui est possible pour un moteur de recherche. Les outils d'IA générative amalgament des données sans faire, à ce stade, plus aucune référence aux sources dans leur résultats.

Grands principes devant guider l'utilisation des outils ayant recours à l'intelligence artificielle dans l'activité éditoriale :

1. Une supervision humaine systématique

Le schéma d'utilisation de l'IA doit toujours être le suivant : Humain > Machine > Humain

Autrement dit, l'utilisation de l'intelligence artificielle doit toujours être décidée par un humain, et le résultat obtenu validé par un humain.

Aucune publication ou diffusion de contenu créé par ou avec une IA ne peut avoir lieu sans relecture, vérification et validation éditoriale.

2. Une utilisation raisonnée

De façon générale, utilisation de l'IA doit avoir pour objectif d'aider à la production éditoriale (en amont, lors de la préparation, pendant la production, ou en aval lors de la diffusion).

Elle peut permettre d'en améliorer la qualité, l'originalité et de réduire le temps de traitement et d'analyse de la base documentaire, dont le journaliste se sera assuré de la fiabilité. En aucun cas, elle ne peut se substituer au travail éditorial des journalistes et à fortiori au travail de vérification des sources et de confrontation des informations à la réalité du terrain.

Dans ce cadre, l'IA peut être utilisée pour : la correction orthographique et grammaticale, le « dérushage » ou transcription d'interviews pour en repérer les passages les plus pertinents, la traduction automatique de textes ou d'interviews à des fins de documentation ou de base de travail, des résumés de textes ou de dossiers complexes à des fins de documentation, l'exploitation éditoriales de données massives (datajournalisme), l'aide à la rédaction de posts optimisés pour les différents réseaux sociaux ou de titres optimisés pour le référencement, l'aide au repérage d'informations qui circulent sur les réseaux sociaux, l'aide à la détection de fausses informations appuyée sur des outils spécifiques et validés, l'indexation et la valorisation de contenus d'archives...

De façon générale, l'IA ne doit jamais être utilisée pour générer des images, des sons ou des vidéos dont le réalisme risque de tromper le public ou de le laisser dans l'ambiguïté.

La **génération d'images** n'est autorisée qu'à des fins illustratives, sans photoréalisme qui puisse susciter un doute sur leur authenticité. Cette utilisation doit faire l'objet d'une mention claire (cf. ci-dessous).

L'IA ne peut en aucun cas être utilisée pour **recréer la voix ou l'apparence** de personnalités ou de journalistes.

3. Une utilisation transparente

La transparence doit être totale, **en interne, comme vis-à-vis du public.**

L'utilisation de l'IA pour certaines tâches doit être connue de tous au sein de l'équipe, et notamment du responsable éditorial.

Il est indispensable « d'étiqueter » clairement pour le public **le contenu généré par ou avec une IA générative**, lorsque :

- des textes ont été rédigés majoritairement par une IA, ou traduits par une IA, même s'ils font systématiquement l'objet d'une validation éditoriale
- des images d'illustration ont été créées par ou avec l'aide d'une IA
- des audios ont été générés en partie par une IA (ex : synthèse vocale pour anonymiser un témoignage par exemple)

Quand cette mention ne pourra pas être visible (en texte, image ou vidéo), elle devra être signalée dans les éléments de contexte (lancement donné à l'antenne).

Si un de nos contenus relaie une image produite par l'IA (par exemple pour dénoncer un « fake »), la signalétique doit être plus visible encore, et occuper une partie suffisante de l'image pour qu'elle ne puisse pas être gommée par un autre outil (ce qui pourrait permettre de nous associer à ce « fake » par une capture d'écran).

4. Une confidentialité pour la protection des données et des sources

Une grande prudence doit être observée lors de la transmission d'informations à des plateformes externes (ex : ChatGPT) ou de l'utilisation de contenus confidentiels dans des outils d'IA générative. Et ce pour protéger les sources et les informations sensibles en possession des journalistes.

Lorsque des informations ou données confidentielles sont en jeu, il convient d'utiliser des plateformes sécurisées et dédiées, internes à FMM.

L'état actuel des technologies et les perspectives de développement des IA, rendent les moyens habituels de protection / anonymisation des interlocuteurs (floutage, transformation de la voix, etc.) de moins en moins opérants. **Il convient donc d'utiliser des méthodes d'anonymisation beaucoup plus sûres et robustes qui ne laissent pas de place à des traitements ultérieurs** (cf charte de l'anonymat-document FTV déjà transmis aux rédactions et qui sera complétée si nécessaire)

5. Une expérimentation prudente et responsable

Les innovations / progrès constants dans le domaine de l'IA sont propices à l'expérimentation, à la curiosité, à la critique et à l'innovation. Mais ces expérimentations doivent être conduites avec prudence, pour éviter les risques de dissémination de données et d'erreur.

Toute utilisation qui dépasserait le cadre des principes ci-dessus doit faire l'objet d'une discussion préalable avec le responsable éditorial et les référents techniques.

6. Une formation et un suivi

La formation des rédactions aux enjeux et aux outils de l'IA est primordiale pour atténuer les risques liés à l'IA générative et pour promouvoir la responsabilité et la transparence. Elle est aussi indispensable pour rendre compte de ces problématiques au public.

En interne, le partage d'information et les retours d'expérience sont à encourager.

Évolution de ce texte

Ces « bonnes pratiques » seront amenées à être revues périodiquement, en fonction de l'évolution de la technologie et de l'expérience des rédactions. Cette révision se déroulera dans le même cadre que l'élaboration de ce texte initial.

LEXIQUE

IA = intelligence artificielle. Ensemble de théories et de techniques mises en œuvre en vue de réaliser des machines capables de simuler l'intelligence humaine.

IA générative = type d'intelligence artificielle capable de générer du texte, des images ou d'autres médias en réponse à des invites, requêtes (ou prompts)

ChatGPT = prototype d'agent conversationnel utilisant l'intelligence artificielle, développé par OpenAI et spécialisé dans le dialogue. Lancé pour le grand public en novembre 2022, il a connu une adoption fulgurante.

Pour mémoire GPT = Generative Pretrained Transformer.

Prompt = requête ou commande écrite envoyée à une « intelligence artificielle » spécialisée dans la génération de contenu, comme du texte ou des images. En bon français (mais très peu utilisé), une invite.

Datajournalisme = Le journalisme de données est basé sur le filtrage et l'analyse de grands ensembles de données pour en extraire les informations marquantes ou repérer des anomalies ou des tendances.

Speech to text = appelé aussi reconnaissance vocale, c'est une technologie permettant de convertir automatiquement un contenu audio en texte.

Bot : (robot en anglais) est un programme informatique qui parcourt et indexe les pages web. On utilise également les termes de crawler ou de spider pour désigner ces robots d'indexation (ou bot). Exemples : Le Googlebot est le nom anglais du robot d'exploration et d'indexation de Google (moteur de recherche) ou GPTBot est le robot d'indexation d'OpenAI qui scanne les pages web pour améliorer la précision de son modèle d'IA.

Chatbot : robot conversationnel ou dialogueur, c'est l'interface d'un service qui permet de donner automatiquement (avec, dans certains cas, l'utilisation des capacités de l'IA) des réponses aux questions des utilisateurs.