

Les Médias à l'épreuve de l'IA : enjeux éthiques et nouveaux défis

AVRIL 2026

Auteurs :

Corentin BERCY
Consultant Expérimenté
Artimon

Aldja KATEB
Consultante Sénior
Artimon

Maëlle NETO
Consultante
Artimon

Sous la direction de :

Josefina GIMENEZ
Directrice Recherche
et Innovation
Artimon Perspectives

Chita ES-SBAI
Directrice Practice Médias
Artimon

SOMMAIRE

Introduction	03
I. La transformation des pratiques médiatiques médiée par l'IA	05
II. Les usages de l'IA dans les médias	08
III. Les tensions entre logiques éditoriales et logiques algorithmiques	11
IV. La perception du public	13
V. Gouvernance et perspectives – Des initiatives émergentes de régulation et de gouvernance éthique	17
Conclusion	19
Références	21

INTRODUCTION

À partir d'une revue ciblée de la littérature et d'une analyse des pratiques actuelles, ce rapport propose d'examiner les usages de l'IA par les médias traditionnels sous l'angle des enjeux éthiques et de la perception que s'en fait le public. Par ce travail, nous interrogeons la relation entre société et médias traditionnels, notamment dans leur capacité à conserver un rôle structurant dans l'écosystème du contenu médiatique, en créant un lien entre l'analyse éthique et ses effets concrets sur l'image que le public projette sur ces canaux.

L'émergence des réseaux sociaux et l'adoption de nouvelles pratiques de consommation de l'information ont profondément transformé l'écosystème médiatique. De nouvelles sources et canaux de communication se sont imposés, sans pour autant faire disparaître les médias traditionnels. Ces derniers – presse, radio, télévision, cinéma et affichage – continuent d'occuper une place centrale, soutenus par des politiques publiques, la force des marques, la confiance héritée ou encore simplement les habitudes de consommation.

Dans ce contexte déjà en mutation, l'intégration de l'intelligence artificielle (IA) au sein des médias traditionnels constitue un tournant majeur. Loin de l'agitation autour des réseaux sociaux, c'est dans les rédactions historiques que l'IA transforme en profondeur les pratiques journalistiques : automatisation de la production, hiérarchisation algorithmique des sujets, ou encore appui au traitement de l'information en temps réel. Cette évolution soulève des interrogations éthiques propres aux usages algorithmiques en lien avec la transparence, les biais algorithmiques, la désinformation ou encore le respect de la vie privée (Gutierrez-Caneda et al., 2024).

La littérature scientifique documente largement ces risques associés à l'IA. Son essor dans le secteur des médias a aussi donné lieu à une littérature à l'intersection de l'innovation technologique, de l'éthique appliquée et des sciences de l'information. Plusieurs travaux mettent en lumière les transformations que l'IA génère dans les pratiques journalistiques, tout en soulevant des interrogations sur la responsabilité, la transparence, et l'impact sociétal de ces technologies.

Toutefois, le cas particulier des médias traditionnels demeure moins exploré, alors qu'ils se distinguent par une culture éditoriale forte, un sens de responsabilité sociétale affirmé (ou assumé) et une relation privilégiée de confiance avec le public. Ce décalage invite à analyser spécifiquement la manière dont l'usage de l'IA dans ces médias influence leur rôle sociétal et la perception que le public s'en fait.

INTRODUCTION

La notion de rôle structurant mérite d'être explicitée : elle renvoie à la capacité des médias traditionnels à organiser l'espace public en hiérarchisant l'information, en définissant des priorités collectives et en produisant des cadres interprétatifs partagés. Historiquement, cette fonction s'est construite sur trois piliers: la sélection éditoriale (qu'est qui mérite d'être rendu visible), la contextualisation (comment on raconte les faits), et la médiation institutionnelle (assumer ce rôle auprès du public). Dans un environnement informationnel fragmenté par les logiques de plateformes et de personnalisation algorithmique, ce rôle structurant ne signifie plus seulement diffuser largement, mais maintenir des repères collectifs dans un espace saturé de contenus hétérogènes. L'introduction de l'IA reconfigure cette fonction : si elle peut renforcer les capacités de traitement et de vérification, elle peut également déplacer le centre de gravité de la hiérarchisation vers des critères algorithmiques. La question devient alors moins celle de l'adoption de la technologie que celle du maintien d'une capacité humaine à produire des informations fiables.

La problématique principale de cet article revient donc à étudier comment les usages de l'IA que font les médias traditionnels sont perçus par le public, et comment cela affecte le rôle sociétal de ces acteurs. Ce travail s'articule autour de trois questions principales :

- Alors que l'IA favorise la standardisation des contenus, les médias traditionnels peuvent-ils préserver leur singularité éditoriale voire leur rôle de contre-pouvoir face aux logiques de plateforme ?
- Dans un contexte de défiance croissante, comment s'assurer que l'intégration de l'IA dans les pratiques rédactionnelles ne contribue pas à affaiblir davantage la confiance du public ?
- Dans quelle mesure les choix technologiques et éthiques faits par les médias traditionnels dans leur usage de l'IA influencent-ils leur capacité à rester des repères collectifs dans un espace informationnel de plus en plus fragmenté ?

Pour aborder ces questions, nous allons d'abord étudier les usages de l'IA dans le paysage médiatique et ses effets, pour ensuite appliquer cela aux cas d'usage des médias traditionnels. Nous discuterons ensuite les enjeux éthiques intrinsèques et extrinsèques. Enfin, nous analysons les mécanismes de garde-fou existants et les perspectives d'évolution pour maintenir la confiance du public.

Josefina GIMENEZ
Directrice Recherche et Innovation, Artimon Perspectives

I. La transformation des pratiques médiatiques médiée par l'IA

L'introduction de l'IA dans les rédactions touche l'ensemble de la chaîne de valeur de l'information. Selon Diakopoulos (2019), les technologies d'*automated journalism* (génération automatisée de textes à partir de données structurées) ou encore d'*algorithmic journalism* permettent de produire rapidement des contenus à grande échelle, notamment dans les domaines des résultats sportifs, des données financières ou des bulletins météo.

L'IA est ainsi intégrée de manière croissante dans les processus de production, diffusion et consommation des contenus médiatiques. Des algorithmes recommandent des articles et des vidéos, génèrent des contenus, assistent les journalistes dans la rédaction, vérifient des faits, modèrent des contenus, traduisent des textes et permettent une personnalisation poussée des contenus proposés aux utilisateurs. Productivité et engagement du public peuvent ainsi se voir renforcés (Lui et al. 2023).

Ces transformations vont au-delà de l'automatisation, et les outils d'IA sont ainsi mobilisés pour assister les journalistes dans leurs recherches, à travers par exemple l'extraction d'informations, l'analyse sémantique de corpus volumineux, ou encore la détection de tendances sur les réseaux sociaux (Marconi, 2020). Ces usages contribuent à redéfinir la place du journaliste dans la construction de l'information.

L'*Alliance pour les faits*, collectif regroupant plusieurs acteurs des médias de service public français tels que France Médias Monde, Radio France, l'INA et TV5 Monde, a publié en février 2025, une cartographie des usages de l'IA sur l'ensemble de la chaîne de production de l'information, la collecte, la production, la vérification et la diffusion. Concernant la collecte des données, l'IA offre la possibilité d'approfondir des sujets, d'identifier des tendances émergentes et de structurer plus efficacement les recherches documentaires. Pour les étapes de production et d'édition, à travers l'assistance à la rédaction et à la conception éditoriale (génération de textes, d'images ou encore d'avatars), l'IA sert de levier pour optimiser la productivité des journalistes tout en stimulant leur créativité. Elle permet également de renforcer la sécurité éditoriale en proposant des solutions de détection des contenus générés par l'IA, ou la certification des contenus audiovisuels grâce au traçage technique. Enfin, lors de la phase de diffusion de l'information, l'IA est utilisée pour personnaliser les recommandations de contenus et optimiser les référencement afin de mieux atteindre et fidéliser le public. Elle améliore également l'accessibilité des informations via des outils de traduction, de sous-titrage et de synthèse vocale.



Dans l'audiovisuel, Radio France a également mené en 2023-2024 une cartographie interne des usages de l'IA. Ce travail recense des cas concrets comme la transcription automatique d'entretiens ou la génération de résumés sonores destinés aux plateformes numériques. Ces initiatives montrent la capacité de l'IA à faciliter la production et à diversifier l'accessibilité des contenus, notamment pour les publics en situation de handicap. Mais elles mettent aussi en évidence des limites : erreurs fréquentes dans la transcription (notamment celle de noms propres), difficultés à restituer les nuances du discours oral et interrogations sur la transparence des traitements lorsque des extraits diffusés ont été modifiés par des outils automatiques. L'enjeu éthique central est ici la traçabilité : comment signaler au public l'intervention de l'IA sans altérer la confiance dans les médias ?

Des enjeux éthiques transversaux mais peu contextualisés

La littérature sur l'éthique de l'IA dans les médias reprend souvent les grands axes abordés dans le champ plus large de l'éthique algorithmique : opacité des systèmes, amplification des biais, menace sur la vie privée, la transparence ou encore délégation de responsabilité (Cath et al., 2018). Ces enjeux sont exacerbés dans le domaine de l'information, où la confiance, la vérifiabilité et l'indépendance éditoriale sont des piliers fondamentaux. Cela soulève également des questions concernant la paternité de l'œuvre, la créativité et la propriété intellectuelle.

Pour autant, la majorité des analyses reste centrée sur les plateformes numériques et les algorithmes de recommandation (O'Neil, 2016 ; Helberger, 2019), laissant relativement peu de place à une exploration fine des médias traditionnels. La littérature montre un glissement des préoccupations centrées sur les plateformes (très présentes dans les années 2010) vers la transparence algorithmique, la responsabilité éditoriale et l'hybridation homme-machine au milieu des années 2020 (Sonni et al., 2024). Les logiques spécifiques des mass média (ancrage dans une mission de service public, culture de la déontologie journalistique, exigence de transparence éditoriale), sont traitées en marge.

Dans ce contexte, la communication des médias publics mérite un regard ciblé. Par média traditionnel, nous entendons l'ensemble des techniques contemporaines qui permettent à un acteur social de s'adresser à un public extrêmement nombreux. Les principaux moyens de communication de masse (*mass media*) sont la presse, l'affiche, le cinéma, la radiodiffusion et la télévision (Encyclopædia Universalis, communication de masse). Plus récemment, internet et les réseaux sociaux élargissent l'écosystème. Cependant, nés dans un contexte historique, technologique et sociétal différent, la télévision, la radio et la presse nécessitent qu'on explore de plus près les usages de l'IA et ses effets eu égard à la perception que le public a de ces médias.

II. Les usages de l'IA dans les médias

Afin d'ancrer notre analyse dans les pratiques existantes, il s'avère nécessaire d'illustrer comment ces usages potentiels de l'IA se traduisent concrètement dans les médias traditionnels. Cette approche empirique permet de mettre en lumière non seulement les expérimentations et usages en cours, mais aussi les controverses qu'elles soulèvent, en particulier sur le plan de l'éthique et de la confiance du public.

Les cas recensés dans la presse, la télévision ou encore les médias en ligne montrent que l'intégration de l'IA n'est pas neutre : elle ouvre des opportunités en matière de production, de diffusion ou de vérification de l'information, mais elle génère également des tensions importantes. Celles-ci se cristallisent autour de plusieurs enjeux éthiques récurrents :

- La désinformation et la manipulation involontaire de l'opinion (images générées ou retouchées, faux sites d'information, confusion entre contenus authentiques et synthétiques) ;
- Le manque de transparence dans l'usage de l'IA, pouvant fragiliser la crédibilité des rédactions et renforcer la méfiance du public ;
- La dépendance aux grandes plateformes technologiques (comme OpenAI), qui soulève la question de la juste rémunération des éditeurs et de l'asymétrie de pouvoir ;
- La nécessité de préserver la qualité éditoriale et la responsabilité humaine, alors même que certaines tâches journalistiques sont automatisées.

L'analyse empirique des usages de l'intelligence artificielle dans les médias traditionnels français révèle une appropriation pragmatique, marquée par la volonté de sécuriser la production éditoriale tout en expérimentant de nouvelles pratiques. L'étude de cas concrets, qu'il s'agisse de l'association de certains médias avec des outils spécifiques (Associated Press ou [Le Monde](#)), des polémiques autour d'images générées par IA dans des quotidiens, d'expérimentations de voix synthétiques, ou encore de la prolifération de faux sites automatisés, permet d'illustrer la matérialisation des enjeux éthiques dans le quotidien des rédactions.

Loopsider et la voix de l'IA

Tout d'abord, faisons un bref rappel des faits : en 2023, le média français *Loopsider* a expérimenté l'utilisation de voix clonées par intelligence artificielle pour certains de ses contenus vidéo. L'objectif affiché était de faciliter la production, en automatisant la narration et en réduisant les coûts liés à l'enregistrement ou à l'intervention de journalistes. Si cette pratique illustre bien le potentiel de l'IA pour accélérer la création de formats numériques, elle a rapidement suscité des réserves. Plusieurs critiques ont souligné :

- Un manque de transparence : le public n'était pas toujours clairement informé qu'il s'agissait de voix artificielles, ce qui pouvait brouiller la frontière entre authenticité et simulation ;
- Une dévalorisation du travail journalistique : le recours à une voix clonée soulève la question du rôle et de la place des professionnels dans la production de contenus ;
- Un risque de confusion : l'usage du clonage vocal rappelle les dérives possibles liées aux deepfakes audio, qui pourraient être exploités à des fins de désinformation.

Ce cas illustre ainsi la tension des médias traditionnels : exploiter les bénéfices de l'IA pour s'adapter aux nouveaux usages numériques, tout en risquant de fragiliser la confiance du public et la valeur accordée à la production journalistique humaine.

Les partenariats entre acteurs presse et acteurs technologiques

LE MONDE ET OPENAI

Dans la presse écrite, *Le Monde* illustre une dynamique d'intégration encadrée de l'IA. Son accord conclu en 2024 avec OpenAI prévoit que certains contenus du journal puissent être utilisés pour entraîner et alimenter les modèles de langage, en contrepartie d'une attribution explicite des sources et d'une rémunération. Concrètement, cet accord implique que les réponses générées par ChatGPT à partir des contenus du *Monde* peuvent inclure une mention claire du média, accompagnée de liens renvoyant vers les articles originaux, renforçant ainsi la traçabilité de l'information et la visibilité de la marque éditoriale.

Sur le plan économique, cet accord s'inscrit également dans une logique de reconnaissance de la valeur des contenus journalistiques. *Le Monde* a en effet négocié une compensation financière assimilable aux droits voisins, dont une part est redistribuée aux journalistes, selon des modalités définies avec les organisations représentatives du personnel. Ce partenariat, inédit en France, traduit ainsi un double objectif : défendre la valeur des productions journalistiques face à des acteurs technologiques puissants, tout en bénéficiant d'outils susceptibles d'améliorer la circulation et l'accessibilité des contenus.

Toutefois, il soulève des enjeux majeurs : dépendance accrue à un fournisseur privé d'IA, incertitude quant à la maîtrise des usages secondaires des données, et question de l'équilibre des rapports de force entre un média et une entreprise technologique globale. Même encadrée, cette relation interroge la capacité des médias à conserver une pleine autonomie stratégique et éditoriale dans un environnement informationnel de plus en plus structuré par des infrastructures privées.

L'AFP ET MISTRAL AI

Du côté de l'AFP, l'agence a signé en 2025 un partenariat avec la start-up française Mistral AI afin d'améliorer la qualité des réponses générées par les modèles en intégrant des dépêches fiables et vérifiées. Ce cas illustre la volonté de se positionner comme fournisseur de « données de référence » face à la montée de la désinformation. Pourtant, il soulève une tension : si l'IA génère du contenu à partir de dépêches AFP, qu'en est-il de la paternité éditoriale et de la valeur ajoutée du travail journalistique ? L'enjeu de reconnaissance et de propriété intellectuelle devient ici central.

A L'INTERNATIONAL : DES AVATARS PRÉSENTATEURS

À l'international, les expérimentations sont également riches d'enseignements. La BBC a testé l'usage d'avatars virtuels pour présenter des bulletins courts, tout en imposant une signalétique claire indiquant l'intervention d'une IA.

En Chine, certains médias ont diffusé des journaux télévisés animés par des présentateurs synthétiques. Ces pratiques posent des questions de confiance : si le public n'est pas averti, la frontière entre information authentique et production automatisée devient floue, alimentant la défiance.

Enfin, dans le secteur du divertissement et de l'information grand public, plusieurs chaînes ont expérimenté l'usage de l'IA pour générer des voix off ou adapter automatiquement des contenus dans différentes langues. Ces initiatives montrent la puissance de l'IA pour étendre la portée de contenus à moindre coût, mais elles soulèvent le risque d'une standardisation culturelle et éditoriale, avec des récits formatés qui gommant les spécificités locales et la diversité des points de vue.

Ces cas mettent en évidence des problématiques récurrentes : erreurs factuelles et limites techniques, risques de dépendance à des acteurs technologiques, fragilisation de la responsabilité éditoriale, et tension entre rationalisation algorithmique et singularité éditoriale. Les médias traditionnels, en France comme ailleurs, cherchent à tirer parti des bénéfices de l'IA tout en préservant la confiance de leurs publics et leur rôle de repères collectifs. L'équilibre entre innovation et éthique apparaît comme la clé d'une adoption durable.

III. Les tensions entre logiques éditoriales et logiques algorithmiques

Ces exemples illustrent la manière dont l'introduction de l'intelligence artificielle au sein des médias publics transforme profondément la manière dont les rédactions conçoivent, hiérarchisent et diffusent l'information. Si ces outils permettent d'optimiser la production et d'améliorer la performance des contenus, ils soulèvent également des tensions internes majeures : comment concilier les impératifs d'audience et d'efficacité algorithmique avec la mission de service public fondée sur la rigueur, la diversité et l'indépendance éditoriale ?

Plusieurs études (*Bastian et al., 2021 ; Dörr & Hollnbuchner, 2017*) évoquent une érosion du contrôle éditorial liée à l'automatisation croissante des processus. Les pipelines automatisés, qu'il s'agisse de la sélection, du référencement ou de la mise en avant des contenus, peuvent progressivement influencer le jugement journalistique, voire le remplacer. En orientant la hiérarchisation de l'information selon des critères de performance (clics, engagement, taux de rétention), ces systèmes risquent d'obscurcir les valeurs fondatrices des rédactions et de brouiller l'alignement entre production d'audience et mission d'intérêt général.

Face à ces mutations, le rôle du journaliste est lui aussi redéfini. On voit émerger une nouvelle figure professionnelle : celle du journaliste-programmeur, capable de dialoguer avec les outils d'IA, d'en comprendre les logiques et d'en exploiter les potentialités sans en devenir dépendant. Cette hybridation des compétences pose la question du maintien d'une éditorialisation humaine : dans quelle mesure l'intervention journalistique reste-t-elle décisive dans un environnement où la recommandation algorithmique et la standardisation des formats tendent à uniformiser les contenus ?

Pour les directions éditoriales, ces évolutions exigent des arbitrages constants. L'adoption de l'IA ne peut se limiter à un gain d'efficacité ; elle doit s'accompagner d'une réflexion stratégique sur la gouvernance, la responsabilité et la préservation de la pluralité des points de vue. Dans ce cadre, les médias publics ont un rôle exemplaire à jouer. Ils doivent montrer qu'il est possible d'intégrer l'intelligence artificielle tout en garantissant la transparence des processus et la diversité de l'information, des conditions essentielles pour maintenir la confiance du public. Ces tensions sont particulièrement fortes car les rédactions publiques évoluent dans un cadre institutionnel exigeant. Elles doivent innover pour rester pertinentes auprès des jeunes, très exposés à des environnements informationnels guidés par des algorithmes, tout en préservant une identité éditoriale fondée sur la vérification, la contextualisation et la pluralité des voix. Plusieurs professionnels alertent sur le risque d'une déformation progressive des choix éditoriaux, sous l'effet des impératifs de performance algorithmique. Les contenus les plus attractifs tendent alors à être privilégiés, au détriment de sujets plus complexes, moins "performants" mais essentiels à la mission d'information.

Ces arbitrages sont au cœur d'un débat croissant dans les rédactions. Certains journalistes perçoivent l'IA comme un outil de soutien permettant de dégager du temps pour des enquêtes ou des sujets de fond, The Guardian, par exemple, a expérimenté dès les années 2010 des formes de *data journalism* et d'automatisation pour traiter de grands volumes de données, notamment lors des élections ou sur les dépenses publiques. D'autres y voient un facteur d'uniformisation éditoriale, où les formats se standardisent (titres courts, vidéos rapides, contenus scénarisés) pour s'aligner sur les attentes des plateformes numériques. Cette dynamique, analysée notamment par Thurman et al. (2017) et par Tandoc et al. (2018), met en évidence le risque d'une adaptation progressive des pratiques journalistiques aux contraintes algorithmiques, au détriment de la diversité des formats et des approches éditoriales. Cette même dynamique renforce une tension structurelle : d'un côté, la nécessité d'incarner une ligne éditoriale forte, capable de produire du sens, de l'autre, la pression des métriques, vues, impressions, taux de complétion... qui deviennent parfois des critères implicites de décision.

Ces transformations interrogent également la capacité des rédactions à maintenir une pluralité d'opinions et à préserver un espace de dissidence interne. Lorsque les systèmes de recommandation algorithmique influencent la visibilité des sujets ou des angles éditoriaux, le débat interne peut se resserrer au profit d'une logique d'efficacité mesurable. Cette tendance a été documentée dans plusieurs rédactions européennes, notamment en Scandinavie, au Royaume-Uni et dans l'espace germanophone, où l'automatisation de la mise en avant des contenus et l'usage intensif des métriques modifient subtilement les dynamiques de décision éditoriale (Dörr & Hollnbuchner, 2017; Petre, 2021; Bastian et al., 2021). Les journalistes expriment alors la crainte d'un glissement vers un modèle où « ce qui est vu » devient progressivement « ce qui doit être produit », au détriment de la diversité des points de vue et des sujets minoritaires (Van Dijck et al., 2018).

Enfin, pour les directions éditoriales des médias publics, ces enjeux ne sont pas uniquement techniques : ils sont aussi politiques, car ils engagent directement la mission démocratique du service public. Les choix liés à l'usage de l'IA influencent non seulement la hiérarchisation de l'information et la visibilité des sujets, mais aussi la capacité des médias à garantir le pluralisme, l'indépendance éditoriale et l'égalité d'accès à une information de qualité pour l'ensemble des citoyens. Ces directions éditoriales doivent ainsi articuler trois objectifs parfois contradictoires :

1. Assurer la visibilité des contenus publics dans un environnement dominé par les plateformes privées ;
2. Préserver la responsabilité humaine dans la hiérarchisation de l'information ;
3. Maintenir un niveau d'exigence éditoriale qui ne se laisse pas dicter par les logiques d'optimisation ou de captation d'attention.

Cela suppose non seulement une gouvernance adaptée, mais aussi une réflexion critique sur la place accordée aux critères algorithmiques dans la conception même de l'offre éditoriale. À travers ces arbitrages permanents, les directions éditoriales deviennent des espaces où se rejoue l'équilibre entre innovation, garantie démocratique et intégrité journalistique, un équilibre fragile qui conditionne la capacité des médias publics à rester des repères fiables dans un écosystème informationnel bouleversé par l'IA.

IV. La perception du public

Ces transformations ont le potentiel de modifier la relation entre rédactions et citoyens. Dans un contexte français marqué par une défiance structurelle – la France figurant parmi les pays européens où la confiance déclarée envers les médias est la plus faible selon le Reuters Institute (2024) –, les enjeux éthiques liés à l’IA ne peuvent être appréhendés comme de simples questions techniques. Ils engagent la légitimité même des institutions médiatiques, leur capacité à rendre compte du réel et à jouer un rôle structurant dans l’espace public. Les risques d’opacité, de biais ou de substitution partielle du jugement humain ne relèvent pas de l’abstraction : ils se manifestent à travers des effets concrets, observables dans la manière dont les publics évaluent la fiabilité de l’information, interprètent le rôle des rédactions et redéfinissent la place accordée aux médias traditionnels en tant que repères collectifs.

Les tensions que l’IA introduit au sein des rédactions, décrites au chapitre précédent, se prolongent à l’extérieur, dans la perception publique. Ces tensions ne sont pas uniformes mais se cristallisent autour de trois dimensions étroitement liées : la transparence des usages, la compréhension de l’ingérence technologique et l’impact cumulatif de ces facteurs sur la confiance et la légitimité perçues des médias traditionnels.

Le lien entre transparence (ou l’absence de) et confiance / légitimité perçue

Des travaux récents montrent de manière convergente que la transparence des usages de l’IA constitue l’un des déterminants majeurs de la confiance accordée aux médias (Diakopoulos, 2020 ; Helberger, 2019; Bastian et al., 2021). Une étude de 2023 met en évidence que l’opacité algorithmique est systématiquement interprétée comme une atteinte à la fiabilité des contenus, en particulier dans des contextes médiatiques déjà fragilisés. De même, une enquête socio-numérique publiée dans Socio (2024, OpenEdition) souligne que les publics ont tendance à interpréter l’absence d’explication sur les dispositifs techniques comme une stratégie volontaire de dissimulation, indépendamment de la réalité des usages.

Dans le cas français, cette dynamique est amplifiée par deux facteurs structurels. D’une part, une défiance préexistante envers les institutions médiatiques, régulièrement documentée par les enquêtes Kantar–La Croix et le Digital News Report. D’autre part, une maîtrise limitée des enjeux techniques de l’IA par le grand public, qui renforce le sentiment de vulnérabilité et accroît la sensibilité aux zones d’ombre perçues.

L’examen de plusieurs cas français récents permet de saisir concrètement cette corrélation entre transparence et confiance :

- Le cas évoqué précédemment de Loopsider en est l'exemple : l'expérimentation en 2023 de l'usage de voix clonées par IA sans signalement explicite systématique a généré une incompréhension du public qui tient moins à la technologie en tant que telle qu'au brouillage perceptif entre authenticité journalistique et simulation technique, brouillage que certains publics ont interprété comme une rupture implicite du pacte de confiance.
- Radio France se trouve confrontée à une problématique comparable : lorsque des extraits sonores sont modifiés ou améliorés à l'aide d'outils automatisés, l'absence d'information claire au public suscite immédiatement des interrogations sur l'intégrité du montage et la fidélité du contenu diffusé.
- À l'inverse, *Le Monde*, dans le cadre de son partenariat avec OpenAI en 2024, a adopté une démarche de transparence explicite, en signalant les contenus issus de modèles d'IA et en précisant les modalités d'usage de ses données de référence. Cette communication proactive a contribué à limiter les interprétations négatives, même si la question de la dépendance à des acteurs privés demeure un sujet de débat public.

Les données d'opinion françaises confirment ces tendances. Selon un baromètre publié en novembre 2023 par le journal La Croix et Kantar, une large majorité des citoyens (84%) déclare souhaiter être informée lorsqu'une IA intervient dans la production médiatique. Selon un sondage réalisé par le journal La Croix, les principaux risques associés à un manque de transparence sont identifiés comme la perte d'analyse humaine (41%), la fabrication de fausses informations (37%) et l'affaiblissement de la responsabilité éditoriale (24%). La littérature issue des *media studies* souligne ainsi que la transparence ne constitue pas un simple ajout communicationnel : elle fonctionne comme un mécanisme central de légitimation. Lorsque les médias explicitent leurs usages de l'IA, en contextualisent la fonction et rendent visible la supervision humaine, la confiance tend à se renforcer. À l'inverse, le moindre flou est interprété comme une rupture du contrat moral entre médias et citoyens.

Ce que le public comprend (ou ne comprend pas) de l'ingérence de l'IA dans les médias traditionnels

La perception de l'IA par le public ne repose pas uniquement sur l'évaluation des usages réels observables dans les rédactions. Elle est largement façonnée par des représentations souvent imprécises, nourries par les controverses médiatiques autour de la désinformation numérique et des technologies génératives. Les études convergent sur ce point : le public tend à surestimer la capacité de l'IA à manipuler l'information et à sous-estimer la place de la supervision humaine, pourtant centrale dans les médias traditionnels (MIT Technology Review ; Reuters Institute, 2023–2024).

L'analyse croisée des enquêtes disponibles et des controverses publiques permet d'identifier trois malentendus récurrents dans le contexte français.

LE REMPLACEMENT DES JOURNALISTES

Le premier tient à l'idée que l'IA remplacerait progressivement les journalistes. Cette croyance est largement alimentée par la forte visibilité médiatique des outils d'IA générative, alors même que l'observation des pratiques effectives dans les rédactions françaises révèle des usages majoritairement assistifs : transcription automatique, veille informationnelle, détection de tendances ou soutien au

fact-checking. La confusion entre automatisation partielle et délégation complète contribue à nourrir une perception de « déshumanisation » du traitement de l'information.

DES IA MANIPULATRICES

Le second malentendu concerne la perception d'un biais algorithmique politique ou idéologique. Des travaux comme ceux de Makhortykh (2021) ou l'étude menée par l'Université de Ljubljana (2023) montrent que les publics projettent spontanément sur les systèmes d'IA les mêmes soupçons que sur les plateformes numériques : manipulation, orientation idéologique dissimulée ou influence externe. En France, cette inquiétude est renforcée lorsque les médias traditionnels concluent des partenariats avec des entreprises technologiques étrangères, telles qu'OpenAI ou Google, parfois perçus comme des atteintes potentielles à la souveraineté éditoriale.

LES IA EN ROUE LIBRE

Le troisième malentendu tient à la méconnaissance des garde-fous existants. Les dispositifs éthiques internes, comme la charte commune de l'Alliance pour les Faits (Radio France, France Télévisions, INA, TV5Monde, France Médias Monde), demeurent largement invisibles pour le public. Ce décalage entre la gouvernance interne et la perception externe constitue un enjeu central : même lorsque les médias renforcent leurs pratiques éthiques, l'absence de communication explicite entretient une défiance persistante. Les analyses de Sonni et al. (2024) soulignent ainsi l'importance d'une pédagogie active : sans valorisation claire des usages vertueux de l'IA – lutte contre les deepfakes, amélioration de la traçabilité des contenus, certification C2PA –, l'ingérence technologique reste perçue comme une menace plutôt que comme un outil de protection de l'information.

Comment ces éléments influencent la perception des médias traditionnels et leur rôle structurant

De l'ensemble de ces éléments se dégagent trois dynamiques principales qui structurent aujourd'hui la perception des médias traditionnels dans l'espace démocratique.

La première concerne une fragilisation de la fonction de repère informationnel. Les médias traditionnels sont attendus comme des remparts face à la prolifération de contenus automatisés et de deepfakes. Pourtant, l'ambiguïté entourant leurs propres usages de l'IA – même lorsqu'ils demeurent limités – peut les rapprocher symboliquement des plateformes numériques, brouillant leur singularité éditoriale. Les partenariats AFP–Mistral ou *Le Monde*–OpenAI illustrent cette ambivalence : conçus comme des stratégies de maîtrise et de souveraineté, ils sont parfois interprétés par certains publics comme des formes de dépendance problématiques.

La deuxième dynamique tient au risque accru de confusion entre contenus authentiques et contenus synthétiques. L'essor des deepfakes et des faux sites d'actualité générés par IA renforce l'exigence de traçabilité. Lorsque les médias traditionnels ne signalent pas explicitement leurs usages de l'IA, ils deviennent vulnérables à une suspicion généralisée : la confusion entraîne une perte de confiance, y compris lorsque les contenus diffusés sont authentiques. Le cas de France Télévisions, régulièrement ciblée par des deepfakes imitant ses codes graphiques et éditoriaux, illustre ce paradoxe.

La troisième dynamique renvoie à une redéfinition de la légitimité éditoriale. Historiquement, celle-ci reposait sur trois piliers : la hiérarchisation humaine de l'information, la vérification professionnelle et l'indépendance éditoriale. L'introduction de l'IA reconfigure ces critères. Même lorsqu'elle contribue à renforcer la qualité informationnelle – détection des manipulations, fact-checking automatisé –, elle introduit un doute sur la place effective du jugement humain. La littérature évoque à ce titre une « opacité double » : opacité des outils eux-mêmes et opacité de leur supervision par les journalistes. Cette double zone d'ombre menace l'intégrité perçue du processus éditorial, indépendamment des dérives réelles.

L'IA apparaît ainsi moins comme une rupture isolée que comme un révélateur de tensions préexistantes entre médias et publics. En l'absence d'un cadre de transparence visible et d'une pédagogie explicite, elle tend à devenir un facteur aggravant de défiance. À l'inverse, lorsqu'elle est expliquée, encadrée, signalée et gouvernée, elle peut constituer un levier de légitimité renforcée. Le défi principal n'est donc pas uniquement technologique : il est cognitif, culturel et sociétal. La capacité des médias traditionnels français à conserver leur rôle structurant dépendra de leur aptitude à rendre lisible leur rapport à l'IA, à expliciter les garde-fous mis en place et à maintenir une supervision éditoriale visible, assumée et revendiquée.

V. Gouvernance et perspectives

Des initiatives émergentes de régulation et de gouvernance éthique

Des réflexions institutionnelles et professionnelles commencent à émerger. Dès 2023, l'UNESCO plaide pour une IA au service de l'information de qualité, en mettant en avant le rôle des rédactions dans la définition de cadres éthiques internes. Certaines rédactions, comme la BBC, The Guardian ou Radio France, ont engagé des démarches de documentation de leurs usages d'IA, et parfois mis en place des chartes ou comités de gouvernance.

En parallèle, des organisations comme le Partnership on A.I. ou Reporters Sans Frontières développent des recommandations pour encadrer l'usage des systèmes algorithmiques dans la production et la diffusion de contenus médiatiques. Ces initiatives restent toutefois fragmentaires, et leur adoption varie selon les contextes nationaux, les modèles économiques des médias et leur degré d'intégration technologique.

Les médias traditionnels français, télévisions publiques, groupes audiovisuels historiques, radios, agences de presse et rédactions installées, se saisissent aujourd'hui de l'enjeu éthique de l'IA à travers une logique de gouvernance structurée, collective et délibérative. Leur approche repose sur la continuité d'une culture professionnelle forgée par des décennies de normes déontologiques, de responsabilités éditoriales et de cadres publics ou parapublics exigeants.

Cette dynamique se traduit notamment par la mise en place d'initiatives communes, comme la *Cartographie des enjeux et usages de l'IA pour le journalisme* élaborée conjointement par l'INA, Radio France, France Médias Monde et TV5 Monde. Ce type de démarche rejoint les travaux qui soulignent l'importance d'intégrer les valeurs éditoriales directement dans la conception des systèmes de recommandation, en adoptant un modèle de responsabilité distribuée permettant d'aligner les algorithmes sur les normes journalistiques (Bastian et al., 2021 ; Paik, 2025). Cette charte vise à encadrer l'usage de l'IA sans compromettre les fondements du journalisme. Les principes mis en avant reflètent l'éthique professionnelle des médias traditionnels : priorisation de la véracité des faits, responsabilité éditoriale pleine et entière, transparence sur les contenus générés ou assistés par l'IA, traçabilité renforcée via des standards techniques tels que le C2PA, ainsi qu'évaluations internes systématiques avant tout déploiement d'un nouvel outil.

Cette approche met également l'accent sur la formation et la montée en compétences, dans une logique proche des travaux sur la *data literacy* appliquée aux métiers des médias. La nécessité de former les journalistes à l'usage, aux limites et aux risques de l'IA devient un impératif stratégique, en cohérence

avec les modèles « human-in-the-loop », où l'humain reste responsable et décisionnaire malgré l'automatisation croissante (Thurman et al., 2017). De plus, les rédactions françaises insistent sur la création de comités internes réunissant des compétences éditoriales, techniques, juridiques et stratégiques afin d'assurer un usage responsable des technologies. La sécurité des données, la souveraineté technologique et la préservation de la diversité de l'information sont également des enjeux centraux, particulièrement face aux risques d'uniformisation algorithmique.

Le cas de France Télévisions illustre concrètement cette logique. Le groupe a renforcé son service de lutte contre la fraude et les détournements numériques en mobilisant l'IA pour détecter les deepfakes, les faux logos ou les imitations du style éditorial. Les contenus suspects publiés sur les réseaux sociaux sont comparés à une base de données référencée, puis analysés automatiquement pour identifier d'éventuelles usurpations. Ce dispositif traduit une approche où l'IA est utilisée pour corriger les dérives... d'autres IA, et plus largement pour protéger l'intégrité de la marque éditoriale et la confiance du public.

D'autres médias français adoptent une posture similaire : plutôt que de considérer l'IA comme une menace, ils explorent des partenariats encadrés, (à l'image du partenariat *Le Monde* et *OpenAI*), afin d'intégrer l'IA dans leurs chaînes de production tout en maintenant des garanties éthiques strictes. Cette logique de co-construction repose sur une maîtrise éditoriale forte.

Les pure players numériques et les grandes plateformes (Google, Meta, TikTok, X, YouTube) fonctionnent selon des logiques d'optimisation algorithmique, d'économie de l'attention et d'échelle mondiale. Leur gouvernance de l'IA repose principalement sur la performance technique, la personnalisation des flux et la modération automatisée, avec des chartes volontaires souvent générales et peu ancrées dans la déontologie journalistique. Ces plateformes standardisent et amplifient les contenus émotionnels ou polarisants, tandis que leurs mécanismes de décision restent opaques et dictés par des priorités commerciales ou réglementaires.

En revanche, les médias traditionnels adoptent une approche normative et collective, centrée sur la transparence éditoriale, la responsabilité humaine et la traçabilité des contenus.

Ce contraste souligne un écart structurant entre deux modèles de gouvernance de l'IA : d'un côté, la logique algorithmique des plateformes ; de l'autre, la logique déontologique et éthique des médias. Cette divergence soulève des enjeux cruciaux pour l'avenir : comment les médias traditionnels peuvent-ils préserver leurs valeurs journalistiques dans un environnement numérique dominé par des logiques algorithmiques mondiales ? Comment concilier innovation technologique et exigences déontologiques ?

CONCLUSIONS

D'un point de vue de la perception du public, l'intégration de l'intelligence artificielle dans les médias traditionnels ne constitue ni une rupture strictement technologique ni un simple prolongement des transformations numériques antérieures. Elle agit plutôt comme un révélateur voire un amplificateur des tensions déjà à l'œuvre dans la relation entre médias, journalistes et publics. À travers l'analyse des usages concrets, des controverses publiques et des perceptions associées, cet article a montré que les enjeux éthiques liés à l'IA ne peuvent être dissociés des questions de confiance, de légitimité et de rôle sociétal des médias traditionnels dans l'espace démocratique.

Les résultats de cette analyse soulignent un point central : ce ne sont pas tant les technologies elles-mêmes qui fragilisent la confiance que les conditions dans lesquelles elles sont intégrées, expliquées et gouvernées. L'opacité des dispositifs, la faible lisibilité de la supervision humaine et l'insuffisance de pédagogie autour des usages réels de l'IA alimentent des malentendus persistants. Ceux-ci renforcent l'idée d'une automatisation incontrôlée de l'information, d'une dilution de la responsabilité éditoriale et d'un éloignement progressif du jugement humain, pourtant au cœur de l'identité des médias traditionnels.

À l'inverse, lorsque les usages de l'IA sont clairement signalés, contextualisés et encadrés par des règles éditoriales explicites, ils peuvent contribuer à renforcer la légitimité des médias. La transparence apparaît alors non comme une contrainte, mais comme une ressource stratégique : elle permet de différencier les médias traditionnels des plateformes numériques, de réaffirmer la centralité du travail journalistique et de restaurer une relation de confiance fondée sur la lisibilité des processus de production de l'information. Dans cette perspective, l'IA peut devenir un levier au service de la qualité informationnelle à condition que son intégration demeure subordonnée à des choix éditoriaux clairement assumés.

Cependant, cette étude met également en évidence les limites d'une approche centrée uniquement sur la gouvernance externe, les chartes éthiques ou les recommandations institutionnelles. Si ces cadres sont indispensables pour poser des principes communs, ils ne suffisent pas à eux seuls à garantir une appropriation durable et légitime des technologies au sein des organisations médiatiques. L'IA transforme en profondeur les pratiques professionnelles, les équilibres organisationnels et les métiers de l'information. Elle reconfigure les chaînes de production, redéfinit certaines compétences et suscite des interrogations concrètes sur l'avenir du travail journalistique. À ce titre, elle ne peut être pensée uniquement comme un outil technique ou éditorial : elle constitue également un objet social et organisationnel, porteur d'enjeux de pouvoir, de reconnaissance et de conditions de travail.

CONCLUSION

L'acceptabilité de l'IA ne va pas de soi. Lorsque les dispositifs sont perçus comme descendants, techniques ou insuffisamment discutés, ils peuvent être interprétés comme une remise en question implicite du jugement professionnel et de l'autonomie éditoriale. L'absence de concertation approfondie et d'implication des parties prenantes peut alors renforcer le sentiment d'une transformation subie plutôt que co-construite, fragilisant à la fois l'appropriation des outils et leur légitimité.

Plus largement, cette analyse invite à concevoir la gouvernance de l'IA dans les médias traditionnels comme un processus dynamique et évolutif. Celui-ci doit articuler plusieurs niveaux complémentaires : un niveau éditorial, garant de la responsabilité humaine et des choix de hiérarchisation de l'information ; un niveau organisationnel, intégrant la formation, l'évolution des métiers et la prise en compte des rapports sociaux ; et un niveau sociétal, attentif aux perceptions du public et aux attentes démocratiques. C'est à cette condition que l'IA pourra être intégrée de manière durable, sans accentuer les fractures existantes entre médias, professionnels et citoyens.

En définitive, le défi posé par l'IA aux médias traditionnels français ne se limite pas à la maîtrise des outils ou à la conformité à des principes éthiques généraux. Il réside dans la capacité à aligner discours, pratiques et modes de gouvernance. La préservation du rôle structurant des médias dans l'espace public dépendra de leur aptitude à rendre visible le jugement humain qui encadre l'usage des technologies, à expliciter les arbitrages réalisés et à associer l'ensemble des parties prenantes aux transformations en cours. L'IA ne se substitue pas à la responsabilité éditoriale ; elle en redessine les contours, en rendant plus nécessaire que jamais une gouvernance éthique, transparente et collectivement assumée.

Bastian, M., Helberger, N., & Makhortykh, M. (2021). Safeguarding the Journalistic DNA: Attitudes towards the Role of Professional Values in Algorithmic News Recommender Designs. *Digital Journalism*, 9(6), 835–863.

<https://doi.org/10.1080/21670811.2021.1912622>

Bastian, M., Lechevalier, A., & Prud'homme, T. (2021). Algorithmic news recommenders and editorial values: Toward a model of distributed responsibility. *Journal of Media Ethics*, 36(3), 178–195.

<https://boris-portal.unibe.ch/bitstreams/ea4ab260-6e5d-4bfc-821d-1dffe432765d/download>

Baromètre La Croix KANTAR, Etude Kantar Public pour La Croix (novembre 2023), La confiance des français dans les médias

<https://fr.themedialeader.com/wp-content/uploads/2023/11/Barometre-Medias-La-Croix-Kantar-Public-novembre-2023.pdf>

Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M., & Floridi, L. (2018). Artificial Intelligence and the 'Good Society': the US, EU, and UK approach. *Science and Engineering Ethics*, 24(2), 505–528. <https://doi.org/10.1007/s11948-017-9901-7>

Dörr, K. N., & Hollnbuchner, K. (2017). Ethical Challenges of Algorithmic Journalism. *Digital Journalism*, 5(4), 404–419.

<https://doi.org/10.1080/21670811.2016.1167612>

Dörr, K. N., & Hollnbuchner, K. (2017). Algorithmic news selection? Digital Journalism's quantification and its impact on editorial decisions. *Digital Journalism*, 5(3), 310–329.

<https://www.tandfonline.com/doi/abs/10.1080/21670811.2016.1168929>

Open Research Europe. (2023). Public Trust and Algorithmic Transparency in Media Systems. *Open Research Europe*, 3.

<https://doi.org/10.12688/openreseurope.15321.1>

Paik, S. (2025). Journalism Ethics for the Algorithmic Era. *Digital Journalism*, 13(4), 696–722. <https://doi.org/10.1080/21670811.2023.2200195>

Petre, C. (2021). All the news that's fit to click: How metrics are transforming the work of journalists. Princeton University Press.

<https://press.princeton.edu/books/hardcover/9780691175232/all-the-news-thats-fit-to-click>

Sonni, A. F., Hafied, H., Irwanto, I., & Latuheru, R. (2024). Digital Newsroom Transformation: A Systematic Review of the Impact of Artificial Intelligence on Journalistic Practices, News Narratives, and Ethical Challenges. *Journalism and Media*, 5(4), 1554–1570. Multidisciplinary Digital Publishing Institute (MDPI). <https://doi.org/10.3390/journalmedia5040097>

Thurman, N., Dörr, K., & Kunert, J. (2017). When Reporters Get Hands-on with Robo-Writing. *Digital Journalism*, 5(10), 1240–1259. <https://doi.org/10.1080/21670811.2017.1289819>

Van Dijck, J., Poell, T., & de Waal, M. (2018). *The platform society: Public values in a connective world*. Oxford University Press. <https://global.oup.com/academic/product/the-platform-society-9780190889760>

Alliance pour les faits, Cartographie des enjeux et usages de l'IA pour le journalisme, France Télévisions / Radio France / INA / TV5MONDE / France Médias Monde, 2025. https://larevuedesmedias.ina.fr/sites/default/files/202502/Alliance_for_facts_IA_journalisme_FR_compressed_0.pdf

Franceinfo. (2024, 29 mai). Que fait France Télévisions contre les faux articles et fausses vidéos imitant ses contenus ? Franceinfo. https://www.franceinfo.fr/une-information-transparente-franceinfo/que-fait-france-televisions-contre-les-faux-articles-et-fausses-vidéos-imitant-ses-contenus_7509124.html

Le Monde. (2024, 13 mars). Intelligence artificielle : un accord de partenariat entre Le Monde et OpenAI. https://www.lemonde.fr/le-monde-et-vous/article/2024/03/13/intelligence-artificielle-un-accord-de-partenariat-entre-le-monde-et-openai_6221836_6065879.html

MIT Technology Review. (2023). Generative AI and the Future of News. MIT Technology Review. <https://www.technologyreview.com>

Reuters Institute for the Study of Journalism. (2023). *Journalism, Media, and Technology Trends and Predictions 2024*. University of Oxford. <https://www.digitalnewsreport.org/publications>

UNESCO, 2023 UNESCO. (2023). *Guidelines for the governance of AI in journalism and media*. UNESCO Publishing.



Artimon est un cabinet de conseil en management et organisation spécialiste des transformations, et un institut de recherche pluridisciplinaire.

L'institut Artimon Perspectives mène des travaux de recherche sur l'impact des nouvelles technologies dans les organisations et les activités humaines. Nos productions visent la création de connaissances sur des sujets complexes, la vulgarisation scientifique et la compréhension de différents phénomènes, tout en répondant aux besoins et questionnements opérationnels de notre écosystème d'acteurs.

CONTACTEZ-NOUS

8, rue de la Victoire,
75009 Paris
+33 (0)1 53 20 89 89
<https://artimon.fr/>

Josefina GIMENEZ

Directrice Recherche et
Innovation
Artimon Perspectives
jgimenez@artimon.fr

Ghita ES-SBAI

Directrice Practice Médias
Artimon
GESSBAI@artimon.fr

SUIVEZ-NOUS SUR



Artimon



@artimon5307