

S'INFORMER POUR COMPRENDRE, DIALOGUER POUR S'ENTENDRE



Défendre une IA conçue pour le bien public

L'intelligence artificielle (IA) est souvent comparée à une révolution technologique semblable à celle de l'électricité. Mais, plus rapidement que le courant, l'IA générative révèle de nombreux risques, notamment dans le domaine de la confiance dans les contenus créés, leur véracité et pertinence.

Un autre défi consiste à aider les journalistes, surtout dans les langues et cultures peu prises en compte par ces nouveaux outils, à se les approprier tout en les utilisant de manière responsable et au service de l'intérêt public.

La Fondation Hirondelle n'échappe pas à la réflexion en cours : nous avons instauré une directive sur l'utilisation de l'IA générative au sein de nos médias, utilisons des modèles linguistiques pour améliorer l'analyse de contenus dans les langues locales où nous travaillons. Enfin, nous réfléchissons à comment optimiser certaines tâches dans les rédactions afin de permettre plus de temps pour la vérification des faits et les reportages sur le terrain.

Ce numéro de Médiation pose un regard à un temps t. Les développements de l'IA sont si rapides que les outils d'aujourd'hui risquent d'être rapidement dépassés. Mais ce qui doit être toujours d'actualité, c'est de défendre des modèles numériques conçus pour le bien public. Des modèles rétribuant les sources qu'ils utilisent, notamment journalistiques, et qui intègrent toutes les composantes de la société, y compris les plus vulnérables, déjà socialement, économiquement et technologiquement exclu.e.s des bénéfices du progrès digital.

Caroline Vuillemin, Directrice générale



Présentateur virtuel d'informations animé par l'IA, 5^e Conférence mondiale de l'internet à Wuzhen (Chine). © STR / AFP

L'INFORMATION FACE AU DÉFI DE L'INTELLIGENCE ARTIFICIELLE

Déjà présente dans la production et la consultation des informations, l'intelligence artificielle a fait un saut qualitatif en 2022 avec l'arrivée de l'IA générative grand public. Propice à la réalisation et à la diffusion massives de contenus vrais ou faux mais souvent vraisemblables, celle-ci amène les journalistes à vouloir encadrer son usage individuel et collectif.

L'intelligence artificielle générative grand public a trois ans. Depuis la mise en ligne gratuite de ChatGPT par Open AI en novembre 2022, son usage est devenu massif. Elle est capable de produire tout type de contenus de synthèse à partir d'éléments déjà en ligne, combinant formats texte, visuel, audio ou vidéo. C'est un outil de production inouï et à moindre coût monétaire, mais quel est sa pertinence pour un public en recherche d'informations, et pour des journalistes censés livrer une image fidèle de la réalité ?

Quels que soient ses modèles (l'étasunien Open AI, le chinois DeepSeek, le suisse Apertus...), les réponses de l'IA générative dépendent largement de la requête qui lui est faite, du corpus dont elle se nourrit, et de l'algorithme qui va guider son choix dans ce corpus. Dans ce fonctionnement où les cultures majoritaires en ligne sont favorisées, il

est autant possible de s'informer que de traverser vraisemblablement la réalité. Dès l'année 2023, les exemples ont été nombreux de campagnes de désinformation reposant sur un usage de l'IA générative orienté vers la tromperie, afin de discréder des personnes, des organisations ou d'influencer le cours d'élections.

Très vite, les journalistes ont compris l'importance du phénomène, de son potentiel mais aussi de la menace qu'il fait peser sur la crédibilité de leur production. A l'initiative d'ONG comme Reporters sans frontières, ils ont établi des chartes ou des modèles pratiques encourageant un usage de l'IA qui respecte la fiabilité et la transparence de l'information. C'est un premier pas important. Mais l'intelligence artificielle peut-elle faire mieux ? De même qu'elle a le pouvoir de générer des *deep fakes* à l'infini, peut-elle favoriser la production et la diffusion massives d'informations de qualité accessibles au public le plus large ? Si oui, à quelles conditions techniques ? Quels investissements privés et publics cela requiert-il ? Et dans quelles limites énergétiques cela peut-il se produire ?

Ces questions, qui recoupent largement celles d'une réglementation encore balbutiante, sont ambitieuses. Mais si elles restent sans réponse, l'alternative est sombre : une intelligence artificielle nourrie principalement de diverses formes de désinformation, qui en viendrait à faire grossir la spirale du chaos informationnel et finalement de la discorde sociale, plus encore dans des sociétés fragilisées. Pour échapper à cette perspective il nous semble utile, dans ce numéro de Médiation, de mettre en dialogue la réflexion de la Fondation Hirondelle avec celle d'experts d'une IA au service d'une information d'intérêt général, au Sud comme au Nord. ■

Entretien

L'IA, UN CHANGEMENT RADICAL DANS L'ÉCOSYSTÈME D'INFORMATIONS

© DR



Urvashi Aneja dirige à Goa le Digital Futures Lab, groupe de réflexion travaillant à un « avenir où la technologie est au service de tou.te.s, et non d'une minorité ». Elle explique comment l'IA transforme la diffusion de l'information en Inde, soulignant son potentiel et ses écueils.

Les chiffres montrent que l'accès aux actualités via l'IA suscitent un vif intérêt en Inde¹. Quelles opportunités l'IA offre-t-elle pour la diffusion de l'information dans ce pays ?

Urvashi Aneja : Elle offre une opportunité considérable sur la diversité linguistique. Ces outils permettent aux personnes qui ne parlent pas les langues dominantes d'accéder à l'actualité dans leur langue locale, ce qui est très important car l'Inde compte plus de 200 langues. Néanmoins, ces outils sont encore loin d'être performants dans toutes ces langues. Un autre aspect concerne l'alphabétisation : une grande partie de la population indienne est peu alphabétisée. Le fait que ces outils soient désormais accessibles dans des for-



Statue de la déesse hindoue Durga décorée sur le thème de l'intelligence artificielle, Calcutta, septembre 2025. Dibyangshu. © SARKAR / AFP

mats multimédias offre de nouvelles possibilités aux personnes qui veulent interagir avec l'information, même si elles ne sont pas très alphabétisées. Elles peuvent ainsi être auditrices ou spectatrices, mais aussi productrices de contenus d'information sous différents formats. Troisièmement, l'IA réduit les coûts de production de l'information. Il n'est plus nécessaire de disposer d'un grand studio ou d'une grande équipe. Une seule personne peut filmer un reportage, enregistrer des interviews, les monter et les publier. Bien sûr, il était déjà possible avant l'IA de produire une vidéo seule et de la diffuser sur une plateforme. Mais l'IA a rendu ces actions beaucoup plus faciles pour les particuliers ou les très petits médias. Et la baisse des coûts de production signifie que beaucoup plus de voix, de régions géographiques et de points de vue ont la possibilité d'être représentés, ce qui n'est pas le cas lorsque l'industrie de l'information est plus monolithique et centralisée.

Quels sont les principaux risques liés à l'utilisation de l'IA dans l'information ?

Le principal risque concerne la désinformation et les fausses informations, avec la facilité croissante de créer des contenus trompeurs. Le problème est que bon nombre des outils techniques mis en place pour authentifier les contenus, tels que les filigranes ou les tatouages numériques, s'avèrent vite inopérants. Il est donc difficile d'aider les gens à distinguer le vrai du faux. En outre, la qualité des informations produites par les organes de presse traditionnels a baissé, pour

diverses raisons, notamment parce que leurs revenus publicitaires sont décimés par les réseaux sociaux et désormais plus encore par l'IA générative. Aujourd'hui, la plupart des jeunes en Inde ne lisent plus les journaux et ne consultent même plus les sites d'information. Beaucoup se contentent de lire les résultats générés par l'IA. Or, la fiabilité de ces résultats est de plus en plus faible. Si nous ne résolvons pas ce problème, les États pourraient adopter des réglementations très strictes qui pourraient avoir un impact négatif sur la liberté d'expression, ou encore laisser les plateformes de réseaux sociaux devenir les arbitres chargés de vérifier l'exactitude des informations.

Comment l'IA devrait-elle être réglementée pour limiter ces risques ?

La réponse ne viendra pas uniquement de la réglementation. Il faut accélérer sur les solutions techniques. Nous devons développer des outils permettant à l'utilisateur.rice d'identifier la provenance des données, de vérifier comment un contenu a été conçu à chaque étape de sa création, d'identifier et de suivre les différents éléments qui le composent...

Mais certains types de réglementation sont absolument nécessaires. Sur les réseaux sociaux, l'un des principaux problèmes est que les contenus mensongers, sensationnalistes ou diffamatoires deviennent viraux. Ce n'est pas tant l'existence de ces contenus qui pose problème, mais

(1) Selon le Reuters Digital News Report 2025, 44 % des Indiens ont confiance en des informations générées par l'IA et supervisées par des humains.

Le plus grand risque est que les gens finissent par moins s'engager dans le débat public, parce qu'ils se demandent constamment si tout cela est réel

le fait qu'ils soient amplifiés, vus par tout le monde et qu'ils fassent de plus en plus de bruit. Et cette viralité est étroitement liée aux algorithmes de ces plateformes, qui sont optimisés pour cibler les individus et générer des partages en fonction de leurs affinités. Je pense que la réglementation des données a un rôle énorme à jouer à cet égard, afin que nous ne ciblions pas les gens pour leur montrer ce type de contenu. Dans l'Union européenne, il existe déjà une législation en ce sens. Mais son application est un tout autre défi.

Qu'en est-il de la réglementation de l'IA en Inde ?

Les réseaux sociaux ont été encouragés à ne pas devenir des espaces de propagation de fausses informations. Des directives ont été adoptées, surtout pour réduire la désinformation politique et son impact sur les élections. De nouvelles règles publiées par le ministère des Technologies de l'information exigent désormais que les créateurs de contenu généré par l'IA l'indiquent clairement. Mais la mise en œuvre de ces règles est loin d'être simple. Il est souvent difficile de déterminer ce qui doit être considéré comme « généré par l'IA ». Nombre d'outils courants de conception graphique ou d'édition de texte intègrent déjà des fonctionnalités d'IA. Si quelqu'un utilise un tel logiciel pour créer un véritable article d'informa-

Contrer la désinformation et les discours de haine grâce à l'IA

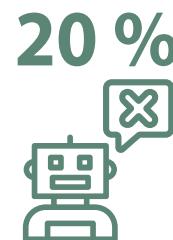
La Fondation Hirondelle a mené à l'Est de la RD Congo un projet pilote utilisant l'intelligence artificielle pour détecter et contrer les discours de haine diffusés par dix radios communautaires. Grâce au traitement automatique du langage (NLP), les contenus radiophoniques étaient enregistrés, transcrits en swahili et analysés pour repérer les propos violents. En réponse, des programmes adaptés étaient produits par Studio Hirondelle-RDC et diffusés afin de renforcer la cohésion sociale. Soutenue par la Commission européenne, cette initiative a été présentée au Sommet pour l'action sur l'IA à Paris en février 2025 et récompensée au Sommet AI for Good à Genève en juillet 2025.

Ce système de monitoring des radios communautaires grâce à l'IA va désormais être étendu à plus de 100 radios à l'Est de la RDC, grâce à un financement obtenu dans le cadre d'un nouveau projet de recherche réalisé en partenariat avec l'Ecole polytechnique fédérale (ETH) de Zurich.

tion destiné aux réseaux sociaux, cela peut-il être considéré comme un contenu généré par l'IA ? Et si le même outil est utilisé pour créer une fausse image, alors finalement quel est le sens de cette mention ?

En somme, l'IA représente un changement radical dans l'écosystème de l'information. Elle aura des répercussions que nous ne pouvons pas encore mesurer aujourd'hui. Mais j'ai également confiance dans la résilience des gens et dans leur capacité à ne pas se laisser complètement piéger par les fausses informations générées par l'IA. Étant donné qu'il est désormais possible de fabriquer n'importe quoi, comme Donald Trump chevauchant une licorne en Irak ou autre, les gens n'y croient pas. Mais le risque le plus grave est que les gens finissent par moins s'engager dans le débat public, parce qu'ils se demandent constamment si tout cela est réel. C'est, je pense, l'aspect vraiment inquiétant de la question. Avec les conséquences très graves que cela peut avoir à long terme pour la santé de la démocratie. ■

L'IA générative, une industrie à factualité limitée



Part des inexactitudes significatives dans les réponses des principaux chatbots étasuniens d'IA sur l'actualité

Source : "News integrity in AI assistants", European Broadcasting Union, Octobre 2025. Test sur 2 709 réponses à des questions d'actualités posées à 4 chatbots d'IA étasuniens.

Une consommation électrique en forte hausse

La planète est-elle en mesure de produire assez d'électricité pour alimenter des modèles d'IA générative destinés à l'usage quotidien du public le plus large ? Le dernier rapport de l'Agence internationale de l'énergie* donne des éléments de réponse. D'abord une idée de la quantité d'électricité moyenne qu'un LLM (Large Language Model) mobilise pour générer une réponse de quelques lignes à une question simple : de 0,1 à 0,5 Wh. La génération d'une image fixe nécessite autour de 2 Wh, et celle d'images animées bien davantage : une très courte vidéo (6 secondes) de faible résolution (8 images / seconde) a un coût énergétique de 115 Wh, soit de quoi recharger deux ordinateurs portables.

Cette électricité est principalement consommée par des centres de données, qui réceptionnent les requêtes et y répondent en traitant les milliards de données qu'ils stockent sur la base de calculs de probabilités. La consommation de ces centres représentait 1,5 % de la demande mondiale d'électricité en 2024, avec une perspective à 3 % d'ici 2030, dont plus de la moitié pour l'utilisation de l'IA générative.

Ces chiffres alarmants rencontreront-ils d'autres limites physiques, liées à l'intensité carbone de l'électricité produite pour alimenter les centres de données, à la consommation d'eau pour les refroidir ou de métaux pour fabriquer les innombrables puces qui y sont agglomérées ? Cela dépendra aussi de la rentabilité à moyen terme des modèles d'IA générative, à ce jour incertaine, qui décidera ou non de leur développement généralisé.

* Agence internationale de l'énergie, « Energy and AI, World Energy Outlook Special Report », Paris, avril 2025.

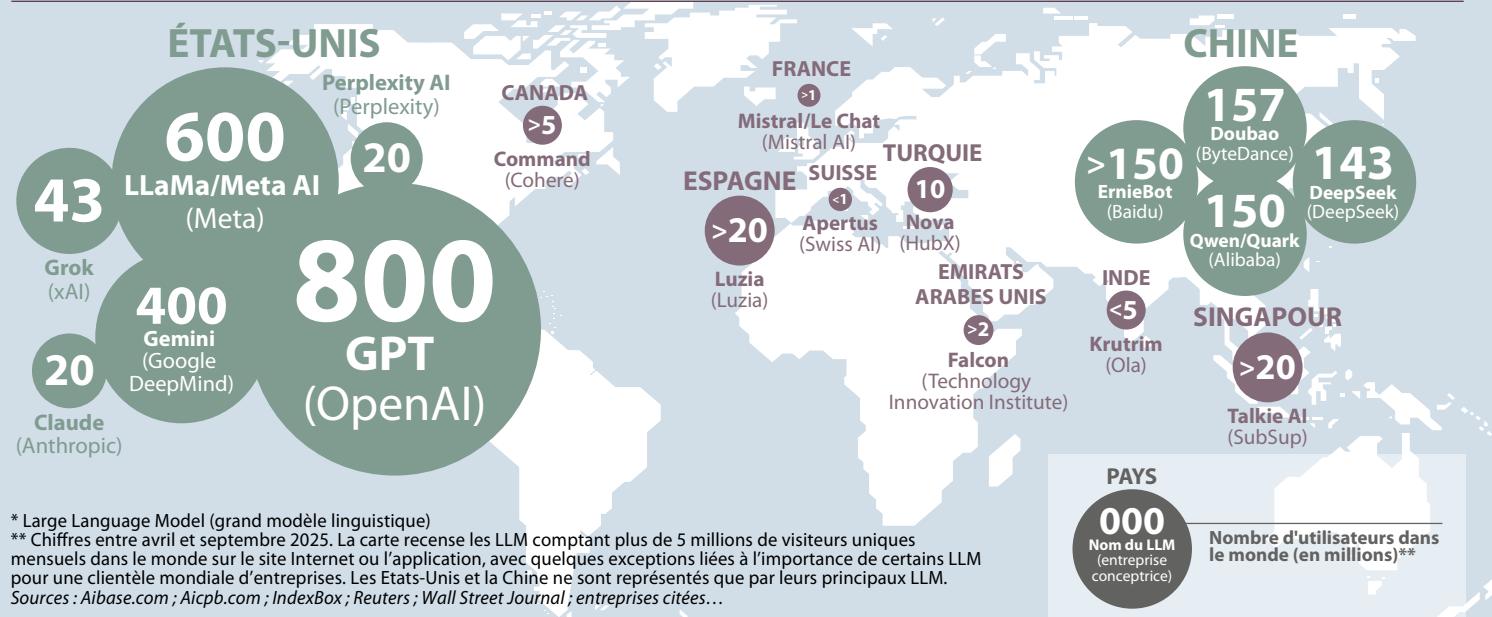
Une réglementation récente et éparsé

« L'IA est trop importante pour ne pas être régulée », estime Kent Walker, le président des affaires mondiales de Google. Depuis la réglementation pionnière de l'Union européenne sur la protection des données (RGPD) en 2018 jusqu'à son Règlement sur l'IA de juin 2024 en passant par le Décret du président étasunien Joe Biden en octobre 2023 (révoqué par le président Donald Trump en janvier 2025), les législateurs ont essayé partout dans le monde d'encadrer trois paramètres, d'amont en aval : le recueil et l'usage des données qui nourrissent l'IA, le fonctionnement des algorithmes et les publications.

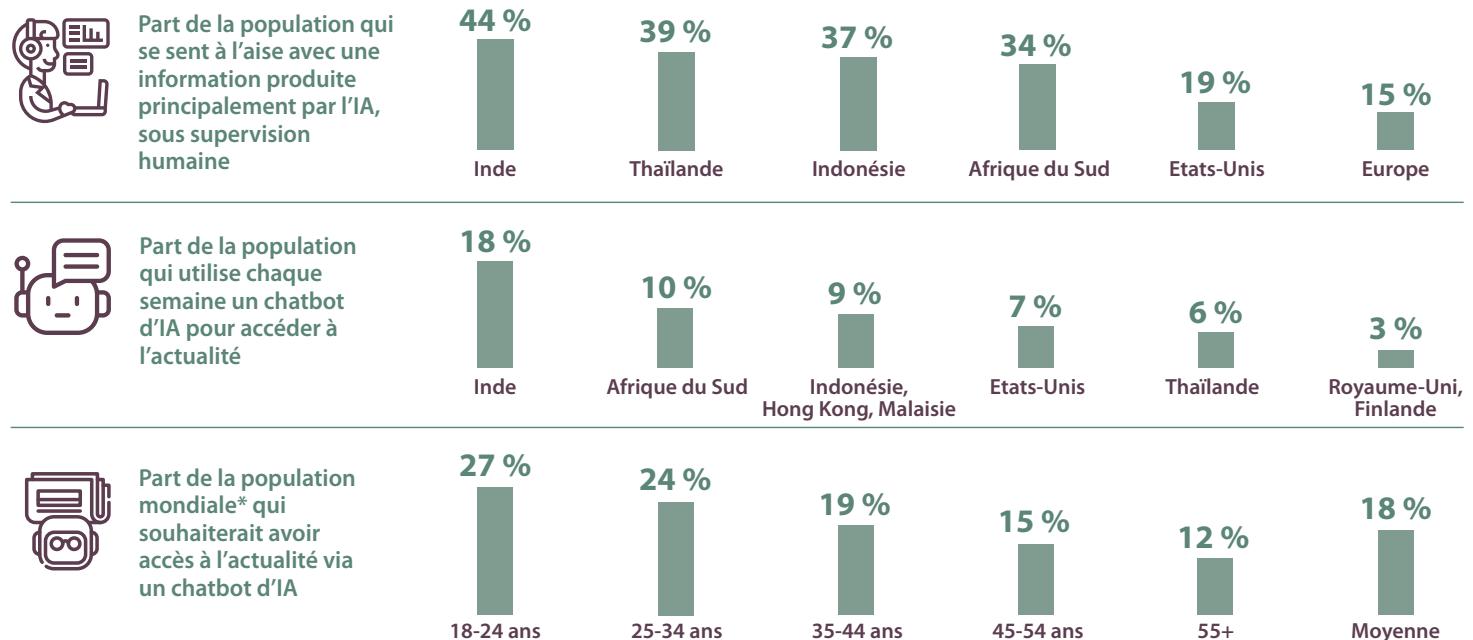
Mais rarement dans le même esprit. « Les approches réglementaires des différents pays reflètent leurs propres priorités socio-économiques, leurs traditions juridiques et leurs modèles de gouvernance », analyse Amlan Mohanty, chercheur associé au Centre pour une IA responsable de l'université de Madras (Inde). « L'approche fondée sur les droits de l'UE vise à "protéger la santé, la sécurité et les droits fondamentaux". La Chine, avec sa forte volonté de contrôle étatique, donne la priorité à l'ordre social et à la protection des "valeurs fondamentales socialistes". Le Japon professe une approche "centrée sur l'humain", qui s'aligne sur ses objectifs sociétaux plus larges. Singapour et le Royaume-Uni ont adopté une approche pragmatique, en concevant des règles pour des secteurs spécifiques ». La Convention-cadre du Conseil de l'Europe sur l'IA adoptée en septembre 2024, dont la plupart des pays mentionnés ci-dessus sont signataires bien qu'aucun ne l'ait encore ratifiée, est une tentative de consensus mondial.

En chiffres

Chatbot et LLM* grand public : l'écrasante domination des Etats-Unis et de la Chine



L'IA générative, une habitude qui gagne les utilisateurs de médias, surtout en Asie et chez les jeunes



* Echantillon de 97 055 personnes issues de 48 pays sur 4 continents.

Source : Reuters Digital News Report 2025.

Mais la confiance en la qualité de l'information générée par l'IA reste minoritaire



Différence entre les parts de la population mondiale* qui pensent que l'IA générative rend l'actualité plus ou moins...

+ 29 % Facile à produire

Transparente
- 8 %

Exacte
- 8 %

Fiable
- 18 %

* Echantillon de 54 638 personnes issues de 31 pays.

Source : Reuters Digital News Report 2025.

Notre expérience



Reportage vidéo de Studio Tamani au marché Namassa Danga, Bamako (Mali). © Harandane Dicko / Fondation Hirondelle

© Fondation Hirondelle

IA ET JOURNALISME : UNE APPROCHE CENTRÉE SUR LA RESPONSABILITÉ HUMAINE



Pour encadrer l'usage de l'intelligence artificielle au sein de ses rédactions, la Fondation Hirondelle a adopté une directive. Jacqueline Dalton, responsable éditoriale, revient sur la position de l'organisation et les principes qui la guident.

Pourquoi la Fondation Hirondelle a-t-elle ressenti le besoin de réfléchir à l'IA ?

Jacqueline Dalton : La réflexion a commencé avec la prolifération d'agents conversationnels comme ChatGPT et d'outils génératifs d'images et de vidéos. Notre première étape a été de comprendre à la fois les opportunités et les risques que ces technologies représentent pour le journalisme, en particulier dans des contextes fragiles. L'IA peut amplifier la désinformation ; nous avons donc

Les outils d'IA promettent efficacité et économies, mais comportent aussi des risques majeurs

voulu garantir que tout usage au sein de nos médias soit transparent et responsable. La confiance est au cœur de notre relation avec le public : il est donc essentiel d'être clair sur la manière et les moments où l'IA est utilisée.

Quelles étaient les principales interrogations concernant son utilisation ?

Les outils d'IA promettent efficacité et économies, mais comportent aussi des risques majeurs, contraires à notre volonté de produire une information indépendante, fiable et de proximité. Par exemple, nous avons envisagé d'utiliser des voix générées par IA pour lire à l'antenne les « annonces communautaires », un espace où sont diffusés des messages rédigés par des particuliers ou des petites entreprises. Si cette idée semblait pratique et un gain de temps, nous avons compris qu'elle nuirait à l'authenticité qui fait la force de nos médias. Nos auditeurs s'attachent à des voix humaines qu'ils connaissent et en qui ils ont confiance. Une voix artificielle affaiblirait ce lien.

Par ailleurs, les outils d'IA comportent de nombreux biais : ils produisent des réponses vraisemblables, mais pas forcément vraies, et reflètent les biais culturels des contenus qu'ils absorbent en ligne.

Quels sont les points clés de la directive de la Fondation Hirondelle sur l'IA ?

- Une approche centrée sur l'humain : l'IA soutient nos équipes mais ne les remplace pas. Les journa-

listes sont pleinement responsables des décisions éditoriales et de la validation du contenu.

- Une transparence avec le public : lorsque du contenu est principalement généré par un outil d'IA — que ce soit du texte, des images, de l'audio ou de la vidéo — cela est clairement indiqué au public.

- Le respect de la qualité et de l'intégrité : tout le contenu doit continuer à respecter les normes rigoureuses définies dans notre Code d'éthique et de déontologie professionnelle. L'utilisation de l'IA ne doit jamais compromettre l'exactitude des faits ou la qualité éditoriale.

- Le respect de notre éthique et de notre responsabilité : tout contenu généré par l'IA doit être vérifié à l'aide de sources indépendantes. Les informations sensibles, confidentielles ou permettant l'identification personnelle ne sont jamais partagées avec des outils d'IA.

Comment les équipes sont-elles formées et sensibilisées à l'utilisation de l'IA ?

Nous sommes actuellement au début de ce processus. Pour l'instant, nous avons identifié les outils avec lesquels nous pourrions travailler et avons établi un guide d'utilisation pour les équipes. Notre but est de former tous les journalistes de nos rédactions d'ici la fin de l'année. Notre directive sur l'IA pourra également être mise à jour en fonction des retours des équipes et à mesure que la technologie et les pratiques rédactionnelles se développent. ■



Témoignage

LE POUVOIR D'INFLUENCER L'OPINION DOIT ÊTRE RÉGULÉ

Depuis Berlin et Zurich, l'ONG **AlgorithmWatch** plaide pour des algorithmes contribuant à rendre les sociétés plus démocratiques, inclusives et durables. **Estelle Pannatier**, responsable de ses politiques, alerte sur les structures de pouvoir sous-jacentes à l'IA et sur leur impact en matière d'accès à l'information.

Algorithm Watch milite en faveur de systèmes algorithmiques qui promeuvent les droits humains. Qu'est-ce que cela veut dire concernant l'IA et le droit humain de chercher et recevoir des informations ?

Estelle Pannatier : L'utilisation d'algorithmes et de l'intelligence artificielle pose plusieurs défis en matière d'accès à des informations fiables. Des algorithmes non transparents sélectionnent les informations sur les réseaux sociaux et les moteurs de recherche, dans lesquels l'IA est de plus en plus intégrée. Les contenus générés par l'IA (textes, images et vidéos) influencent le débat public. Et ce pouvoir d'influencer l'opinion est extrêmement concentré entre les mains d'une poignée d'entreprises technologiques qui fournissent ces services. Cela a également un impact sur le journalisme et les médias, qui sont soumis à de nouveaux intermédiaires, tels que les chatbots d'IA, entre eux et le public.

Cela a finalement un impact sur la société et la démocratie, car notre accès à l'information est de plus en plus dépendant d'une poignée de grandes entreprises.

Notre accès à l'information est de plus en plus dépendant d'une poignée de grandes entreprises

Divers modèles de réglementation de l'IA sont apparus à travers le monde. Certains d'entre eux favorisent-ils davantage la diffusion d'informations fiables pour tous les publics ?

Certaines réglementations en matière d'IA, telles que la Convention-cadre du Conseil de l'Europe sur l'IA, traitent de l'impact des algorithmes et de l'IA sur les droits humains et la démocratie, ce qui est encourageant. Cependant, la gouvernance de l'IA manque souvent d'une vision globale. S'il est nécessaire d'établir des règles pour le développement et l'utilisation de l'IA, une approche plus holistique est requise. Ces réglementations devraient en effet inclure des mesures visant à renforcer les médias et le journalisme, à encou-



Estelle Pannatier © David Baechtold

rager la recherche et le développement de plateformes et d'algorithmes alternatifs qui favorisent la démocratie, et à promouvoir les compétences démocratiques dans la société. Enfin, la concentration du pouvoir dans la formation de l'opinion doit être abordée. C'est seulement ainsi que nous pourrons garantir l'accès à des informations fiables et diversifiées, et encourager un débat public constructif.

Algorithm Watch milite également en faveur de systèmes algorithmiques qui favorisent un mode de vie durable. Comment procédez-vous ?

Aujourd'hui, des modèles d'IA de plus en plus lourds sont mis sur le marché. La consommation énergétique de ces modèles atteint des niveaux sans précédent. Nous devrions plutôt concevoir des systèmes d'IA durables tout au long de leur chaîne de valeur, et apportant une réelle valeur ajoutée aux individus et à la société. Cela implique de prendre en compte les aspects écologiques, économiques et sociaux lors de leur développement et de leur utilisation. Par exemple, AlgorithmWatch a récemment lancé une enquête avec des organisations de la société civile dans plusieurs pays européens : plusieurs milliers de personnes nous ont fait part de leur perception des risques liés à la consommation énergétique des centres de données et des mesures qui, selon elles, devraient être prises. En fin de compte, il faut considérer les nombreux discours entourant l'IA avec sobriété, pour contribuer à orienter ce débat dans le sens de l'intérêt général. ■

La Fondation Hirondelle est une organisation suisse à but non lucratif qui assure aux populations confrontées à des crises un accès à des informations fiables, locales et indépendantes. Depuis plus de 30 ans (1995), elle soutient les médias et les journalistes locaux dans des contextes fragiles. A vocation globale et ancrée localement, elle crée des espaces de dialogue pour renforcer les liens et la cohésion sociale dans des sociétés fragmentées.

Impressum

Médiation

Semestriel publié par la Fondation Hirondelle

Directrice de la publication : Caroline Vuillemin

Conception, rédaction : Julie Crenn Benjamin Bibas / la fabrique documentaire

Conception graphique : Marek Zielinski

Traductions : Benjamin Bibas / la fabrique documentaire Julie Windebank

Impression : Groux & Graph'style

Avenue du Temple 19C
1012 Lausanne, Suisse

hirondelle.org
info@hirondelle.org
T. + 41 21 654 20 20

Pour nous aider :
UBS Switzerland AG
BIC - SWIFT : CRESCHZZ80A
IBAN : CH32 0483 5041 8522 8100 4

ISSN 2624-8840 (Print)
ISSN 2624-8859 (Online)