

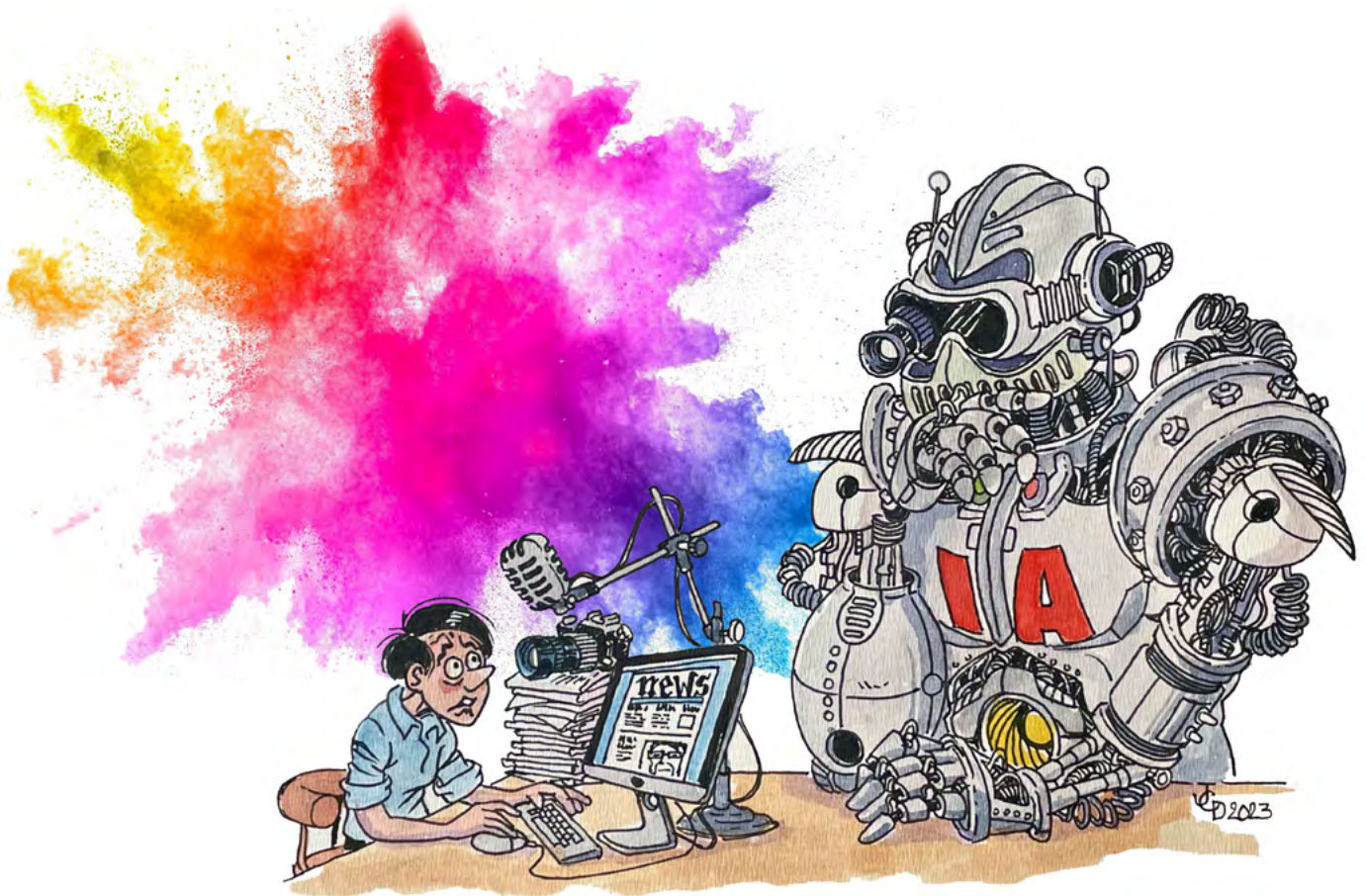
#22 méta-media

Cahier de tendances médias de France Télévisions

KATI BREMME

Méd**IA**s

Nouvelle génération



Le Big Bang
de l'IA générative

Bienvenue dans
la post-vérité

Changements
de pouvoir

MédIAs
Nouvelle génération

KATI BREMME

.6

OUVERTURE

p. 8 Après le Holy Sh*t moment, privilégions l'enthousiasme informé

p. 10 IA générative : évolution, révolution... ou dégénération ?

p. 21 3 questions à ChatGPT

p. 24 IA génératives, de quoi parle-t-on, et pourquoi ?

.32

LE BIG BANG DE L'IA GÉNÉRATIVE

p. 34 Le glossaire de l'IA générative

p. 36 3 questions à Evan Shapiro

p. 38 Petite cartographie des cas d'usage de l'IA générative dans les médias

p. 46 De l'approche réaliste de l'IA générative dans les rédactions

p. 48 Pourquoi les algorithmes éditoriaux sont l'avenir du journalisme moderne

p. 50 *Couleur 3*, une radio livrée aux manettes de l'IA pendant 24 h

p. 52 IA et journalisme : la concurrence des intelligences ?

p. 56 « Alors ce sera si beau, le journalisme se sera si perfectionné qu'il n'y aura plus de journalisme »

p. 58 Assistants vocaux, chatbots : quand l'IA pousse les créateurs dans les retranchements du langage

p. 61 La transparence : un cadre pour surmonter les craintes liées à l'IA dans les salles de rédaction

p. 64 Q&R ChatGPT pour une rédaction : que se passe-t-il lorsqu'une machine apprend à être convaincante ?

.68

BIENVENUE DANS LA POST-VÉRITÉ !

p. 70 Intelligences artificielles génératives et désinformation : bienvenue en post-vérité ?

p. 78 « Le vrai danger, c'est de faire trop confiance à une IA beaucoup moins intelligente qu'on ne le pense »

p. 82 ChatGPT peut-il faire de la politique ?

p. 87 IA génératives : vont-elles rendre les campagnes de désinformation incontrôlables ?

p. 90 L'IA, les médias, les jeunes et la nouvelle réalité

p. 94 L'arrivée du Tour de France du 14 mars 2100

p. 98 Les médias synthétiques sonnent-ils la fin des journalistes ?

p. 102 IA Générative : une nécessité de formation pour les journalistes

p. 104 SXSW 2023 : Who cares about reality ?

.110

CHANGEMENTS DE POUVOIR

p. 112 Technopolitique de l'IA :
luttres idéologiques, tensions
géopolitiques, espoirs démocratiques

p. 117 3 questions à Françoise Soulié
Fogelman

p. 120 Parole de machines

p. 124 « Power shift » : comment
l'IA générative redéfinit la chaîne de
valeur des médias

p. 126 Quelle(s) régulation(s) pour
les systèmes d'IA générative ?

p. 130 Mais à qui appartiennent
donc les œuvres synthétiques ?

p. 138 Intelligence artificielle :
une précarisation de l'emploi plus
qu'une destruction

p. 142 IA Générative et métavers :
les conjuguer pour fabriquer et
enrichir nos mondes imaginaires ?

p. 144 3 questions à Cédric Villani

p. 146 Avec l'IA, une société plus
verte ?

p. 150 La scène comme espace
de réflexion sur l'IA : *Qui a hacké
GaroutzIA ?* explore les échanges
humains-machine

p. 152 Les machines peuvent-elles
créer ? Accorder à l'IA ce que nous
avons pris aux Dieux...

.156

USAGES & PRATIQUES

p. 158 Baromètre Kantar-La Croix :
la confiance des Français dans les
médias remonte

p. 162 Reuters Institute Digital News
Report : les 12 enseignements à
retenir

p. 167 SXSW 2023 : Amy Webb
prévoit la grande « IASMOSE »

p. 170 Traque des usages

LIVRES RECOMMANDÉS

.172

**OU-
VER-
TURE**

Éditorial

APRÈS LE HOLY SH*T MOMENT

PRIVILÉ- GIONS L'EN- THOUSIASME INFORMÉ

Six mois de ChatGPT et le rythme ne faiblit pas. Jour après jour, nuit après nuit, de Californie et d'ailleurs, applicatifs, logiciels, plug-ins bluffants d'IA générative déferlent, validés par personne, et arrivent à bas prix entre les mains de tous, des bons, comme des méchants.

Un tourbillon de solutions magiques pour – en quelques secondes – créer, produire, composer, illustrer, raconter, résumer, coder, assister, plus vite, plus fort que le voisin. Dans les dîners en ville et les discussions de comptoir, les têtes tournent aussi vite que les fichiers Excel pour commenter les prouesses et chiffrer les économies à la clé, sans que personne – un peu comme pour les subprimes de 2008 – n'y comprenne grand-chose.

Dans le secteur de l'information, dés-
tabilisé par plus de quinze ans d'im-



*Eric Scherer, directeur des
Affaires Internationales et du
MediaLab de l'Information de
France Télévisions*

pact délétère de réseaux sociaux hors
de contrôle dans la sphère publique
et la vie de la cité, l'emballement le
dispute à l'inquiétude. Emballement
pour des gains de temps permettant
aux rédactions de se concentrer sur

l'essentiel ; inquiétude car le nouvel
outil permet aussi de tordre la réa-
lité, manipuler les faits, ou même les
inventer.

Cette nouvelle IA, pourtant très créa-
tive, ne fait pas de recherches, n'in-
terroge pas les sources, ne donne pas
les siennes, ignore les frontières entre
disciplines, et ne demande pas de
comptes au pouvoir.

**Si vous avez aimé les réseaux
sociaux, vous allez adorer ChatGPT
& consorts ! Après tout, les pre-
miers n'étaient alimentés que par
des humains ; les seconds le sont
par des machines dont les perfor-
mances dépassent déjà celles de
nombreux humains.**

Qui peut souhaiter une accélération
du flux de désinformation dans nos
sociétés inquiètes, colériques, de plus

en plus polarisées, et où les algorithmes trient l'essentiel de l'info des jeunes ? Comment faire face à une nouvelle concurrence sur le temps d'attention, répondre au risque d'érosion supplémentaire de la confiance dans les médias, lorsque, peut-être, demain rien de ce que nous verrons sur Internet ne sera crédible ?

Quid des droits d'auteur ? Comment protéger le travail des journalistes et les revenus des médias ? Dans quelle mesure leur contenu est utilisé pour entraîner les IA, comment devraient-ils être indemnisés et quelles sont leurs options juridiques ?

Oui, bien sûr, accélérons notre adaptation à cette nouvelle révolution technologique, mais prenons aussi vite la décision collective, responsable et exécutoire de la réglementer. D'exiger de vrais comptes à la poignée de plateformes géantes privées, une nouvelle fois engagées dans une folle et secrète course aux armements.

Qui parle, quand la machine a produit ? Aujourd'hui, si un filigrane numérique peut aider à tracer l'IA dans les photos et les vidéos, personne ne sait dire encore si un texte en contient. Développeurs de tous pays, quand enfin nous livrerez-vous un « Yuka » de l'info ? Un nutri-score des nouvelles nous permettant de scanner les contenus et nous informer

sur leur qualité informationnelle. D'où viennent-ils, ont-ils été modifiés, par qui, comment, où ?

Dans les médias d'information, déjà déstabilisés côté distribution, et désormais côté création de contenus, des mesures rapides doivent être prises : favoriser l'expérimentation, certes, mais aussi guider, et contrôler l'utilisation en interne en proposant quelques balises à des rédactions, curieuses mais souvent démunies. ChatGPT & consorts, nouveaux assistants conversationnels, feront, tôt ou tard, partie intégrante de la trousse à outils des journalistes. Mais veillons à ce que ces derniers imposent in fine leur contrôle éditorial.

Les rédactions ont une longue histoire d'adoption d'outils pour plus de productivité, une meilleure communication, une recherche améliorée et une meilleure écriture. Cette évolution a inclus de nombreuses tâches : des traitements de texte avec correction orthographique, aux moteurs de recherche, aux réseaux sociaux et à la traduction. Mais ces IA font des choses qu'aucune IA n'a jamais faites auparavant, qui sont le produit non de leur « intelligence », mais de leur échelle.

Le sujet de la confiance dans les informations professionnelles et vérifiées est déjà si difficile jour après jour à

gagner, à conserver, qu'aucun compromis n'est aujourd'hui possible avec la qualité des informations d'un média, notamment de service public. La confiance reste notre valeur cardinale.

ChatGPT ne va pas remplacer un journaliste, mais un journaliste sachant l'utiliser dépassera tous les autres de loin, au risque de déstabiliser la rédaction. Tous les médias n'auront pas les mêmes balises éthiques, certains voudront aller plus vite que les autres. Ils se distingueront par la nature de la relation qu'ils auront avec l'IA.

Favorisons l'engagement, apprenons et préparons-nous. Dans ce nouvel univers médiatisé par l'IA, l'heure est à la réorganisation, la vigilance et la transparence. Il faudra des années d'apprentissage, de tests, d'erreurs, pour déterminer comment cette nouvelle technologie va profiter aux médias et aux citoyens.

Mais attention, l'IA n'est pas comme nous : plus elle vieillit, plus elle devient intelligente ! ■

IA GÉNÉRATIVE : ÉVOLUTION, RÉVOLUTION... **OU DÉGÉNÉ- RATION ?**

// *Leur ame s'estoit
merveilleusement enrichie
par l'intelligence des choses,*
Michel de Montaigne, Essais

La collaboration des médias avec l'IA ne date pas d'hier. On se souvient de notre Cahier de Tendances 2019, où l'on faisait l'état des lieux des cas d'usage dans toute la chaîne de valeur des médias. Nous sommes depuis longtemps en extension avec la machine, encore plus depuis l'apparition des smartphones connectés à toute la connaissance du monde 24/7, devenus prolongement (et parfois remplacement) naturel de nos cerveaux. C'est justement à toute cette connaissance du monde que s'est connectée l'IA générative, et elle semble savoir s'en servir mieux que nous, en tout cas dans certains domaines. Pour nous faciliter une tâche, transformer un métier, changer une organisation, le monde, ou pour nous anéantir ?

Sa capacité d'imitation est en tout cas époustouflante autant qu'inquiétante : quand, avant, nous étions obligés de nous transformer en robots pour que Siri et consorts nous comprennent, aujourd'hui nous avons l'impression de converser naturellement avec un assistant intelligent et hyperefficace. Les cabinets de conseil ont du souci à se faire. ChatGPT a été entraîné avec l'encyclopédie de la Harvard Business Review et peut réaliser des analyses



*Par Kati Bremme,
directrice de l'Innovation
France Télévisions
et rédactrice en chef
de Méta-Média*

Un spectre hante le monde – il s'appelle ChatGPT. Ce bot polyglotte, surgi en novembre 2022, et sa technologie, sont les meneurs d'une disruption aussi importante que l'apparition de la machine à vapeur, des ordinateurs personnels et des smartphones. L'Intelligence Artificielle (IA) générative, puisqu'il s'agit d'elle, a de loin pris le pas sur le métavers dans l'échelle des buzzwords en 2023. Et à raison : alors que le métavers ne fut que fabulation marketée d'un Mark Zuckerberg en quête de croissance, l'IA générative a le potentiel de transformer toutes les industries et la société dans son ensemble, dès aujourd'hui.

SWOT, PESTEL et autres RICE, tout en fournissant une proposition de prompt pour fabriquer l'identité d'une marque dans Midjourney et viraliser le tout sur les réseaux sociaux.

IL SUFFISAIT DE FAIRE ATTENTION

Depuis Eliza, le chatbot-psychiatre du MIT en 1966, le progrès de l'IA en compréhension du langage naturel a été constant, mais pas fulgurant. Tout s'est accéléré avec l'apparition du papier « *Attention is all you need* » en 2017, par les (ironie de l'histoire de la concurrence) chercheurs de Google. Cette nouvelle méthode de traitement de données est parfaitement adaptée à la grande masse d'informations avec lesquelles nous avons rempli Internet ces vingt dernières années, et a rendu obsolètes les précédents modèles pour de nombreuses tâches liées au langage naturel. En découle alors le célèbre GPT (Generative Pretrained Transformer) par OpenAI de (par extension) Microsoft, et BERT (Bidirectional Encoder Representations from Transformers) de Google. Couplée avec l'expansion des capacités de calcul des machines, la possibilité d'une intelligence supérieure se précise. Mais, l'explosion des données disponibles entraîne aussi une croissance des biais, des illusions non confirmables et des faux. Pour les images, ce sont les modèles de diffusion, qui créent de nouvelles données en partant d'un bruit aléatoire (voir notre article sur

l'histoire de l'IA) et qui construisent une réalité artificielle plus vraie que nature.

Jusqu'à présent, nous étions rassurés en pensant que l'IA n'était qu'un ensemble de techniques informatiques conçues pour permettre aux machines d'imiter (de manière imparfaite) une forme d'intelligence spécifique aux humains. Mais, nourrie du palimpseste de nos connaissances, contestations et incohérences (cf. la Terre est plate), l'IA nous dépasse dans de plus en plus de domaines : ChatGPT a un QI du langage de 155, il est donc plus intelligent pour traiter le langage naturel que 99 % des Français. Les hyperpolyglottes, des personnes très douées pour les langues, peuvent naviguer entre une dizaine de langues ; aujourd'hui, les IA sont capables de traduire plus d'une centaine de langues en temps réel.

Analyser, résumer, corriger, rédiger, générer des images abstraites ou réalistes, animer des images, produire des voix hyperréalistes, fabriquer des sons, des vidéos, et même des objets 3D à partir d'images 2D, automatiser des tâches cognitives, personnaliser, résoudre des problèmes complexes – l'IA est aussi, voire plus, compétente que nous.

Par ailleurs, il est intéressant de noter que nous utilisons instinctivement le pronom masculin « il » pour nous référer à ChatGPT, alors que nous avons plutôt tendance à attribuer

un genre féminin aux assistant(e)s un peu moins sophistiqué(e)s, tel(le)s que Siri et Alexa.

LA 5^e RÉVOLUTION INDUSTRIELLE ?

Les techno-béats exultent et confondent aisément innovation (qui peut parfois se traduire par un retour en arrière pour la société) et progrès. À la différence de l'IA prédictive, bien installée dans les process de toutes les industries, l'IA générative est toute récente. **Et contrairement aux modèles précédents, qui étaient publiés dans le cadre académique avec une validation par des pairs et certaines règles éthiques, les modèles d'IA générative sortent tout droit des laboratoires des Big Tech.** Chaque jour, des centaines de nouveaux outils basés sur l'IA arrivent sur le marché, sans aucun contrôle, mais avec un impact potentiel sur nos métiers, et parfois nos vies. Un peu comme si l'on sortait des médicaments sans approbation. Justement, la FDA vient d'approuver les tests de Neuralink d'Elon Musk aux US, les robots intelligents pourront donc bientôt guider nos décisions directement depuis une puce implantée dans nos cerveaux.

Dans un débat de société qui oppose technophobes et technophiles, les fabricants d'IA générative jouent le rôle de pompiers pyromanes (l'intelligence artificielle présente « un risque d'extinction » de l'humanité au même

titre que « les pandémies ou bien une guerre nucléaire ») : plus on fait peur, plus les gens vont utiliser ces ouvrages maléfiques.

L'IA générative (sous-domaine de l'apprentissage profond et des réseaux de neurones, lui-même domaine du machine learning) n'est que la partie émergée de l'iceberg de l'IA. **Grâce à ChatGPT et sa fulgurante et inédite vitesse d'adoption (100 millions d'utilisateurs en deux mois) tout le monde est désormais conscient des avancées technologiques de l'intelligence des machines.** Les machines apprennent mieux que les humains, et à force d'appauvrir nos cerveaux et nos langues, elles n'auront, en effet, aucun mal à nous remplacer, y compris désormais les cols blancs.

L'IA est éminemment l'un des plus gros enjeux économiques – et probablement sociétaux – de notre époque. Quand, pour les médias de service public, il s'agira de trouver un juste équilibre entre immédiateté, coût et qualité, pour les entreprises privées, l'automatisation est un enjeu de taille qui mérite investissement. Déjà en 2006, Netflix avait offert 1 million de dollars à celui qui améliorerait son algorithme de recommandation. La révolution de l'IA générative a le potentiel de remodeler considérablement le marché, y compris les modèles économiques des géants de la Tech : Microsoft (évidemment), Meta, Amazon et Netflix. Concernant Google, lorsque l'on pose la question à ChatGPT, l'avenir ne paraît

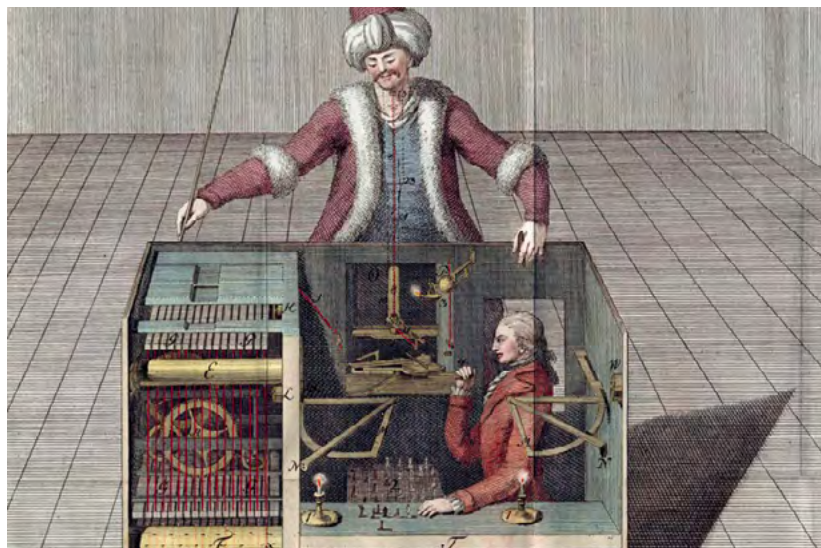
pas rose : « Si l'IA devient capable de générer des contenus de qualité, cela pourrait menacer l'industrie de la publicité sur laquelle Google s'appuie largement ». Sentirait-on un léger soupçon de menace ?

Mais arrêtons donc d'humaniser les IA !

UNE QUESTION ÉMOTIONNELLE, CULTURELLE ET ÉTYMOLOGIQUE

C'est peut-être un biais bien français : dans la langue de Molière, l'intelligence a une connotation humaine qui nous vient de la philosophie et de la littérature. L'interprétation romantique d'une machine sensible et intelligente qui nous remplace n'est donc pas loin : l'intelligence révoltée de Camus, l'intelligence comme « conscience lancée à travers la matière » de Bergson, l'intelligence synonyme d'esprit (chez Flaubert et Artaud) inclut des aspects de la compréhension qui vont bien au-delà des capacités purement cognitives.

Dans la langue de Shakespeare, l'intelligence a bien sûr la même signification de « comprendre », mais le mot « intelligence » en anglais peut aussi avoir une connotation plus technique, liée au domaine de l'information et de l'espionnage, où il fait référence à la collecte et à l'analyse d'informations (qui rejoint l'étymologie latine « recueillir des informations parmi différentes sources pour comprendre



Le Turc mécanique, une grande attraction à la fin du XVIIIe siècle : un – prétendu – automate qui battait les meilleurs joueurs d'échecs. Gravure sur cuivre de Joseph von Racknitz (1789).

quelque chose »).

Le chinois est plus transparent sur la véritable nature de la chose : ici, l'intelligence artificielle se dit 人工智能 (réngōng zhìnéng). Les deux derniers caractères (zhìnéng) signifient « intelligence » et les deux premiers (réngōng) signifient « artificielle » mais, au sens littéral, il faut comprendre par là, le « travail » (gōng) effectué par un être humain (rén).

Nous voilà face au Turc Mécanique de 1789 (et d'Amazon aujourd'hui), incarnation du *lumpenproletariat* de l'IA décrit dans l'article d'Antonio Casilli, qui alimente les Intelligences Artificielles et qui les aide à comprendre le monde, puisqu'elles n'en ont pas la moindre idée.

BIENVENUE EN POST-VÉRITÉ

Non seulement, il ne comprend pas le monde, mais en plus, ChatGPT se fiche éperdument du vrai ou du faux, il cherche le vraisemblable (c'est ce pour quoi il a été programmé). Et quand il ne le trouve pas, il l'invente. Ou comme le dirait Gary Marcus,

spécialiste de l'intelligence artificielle neurosymbolique : « Il n'y a aucune intention de créer de la désinformation – mais également aucune capacité à l'éviter, parce que fondamentalement, GPT est un modèle de la façon dont les mots se rapportent les uns aux autres, et non un modèle de la façon dont le langage pourrait se rapporter au monde perçu ».

Le postulat « voir est croire » ne fonctionne plus depuis longtemps. On n'a bien sûr pas attendu le numérique pour manipuler les images (et, par extension, la réalité / l'Histoire) : on se souvient de la manière dont Staline faisait disparaître un collaborateur après l'autre sur les illustrations de son pouvoir.

Avec l'IA générative, nous arrivons au-delà de la vérité, de l'autre côté du miroir, ou tout simplement, dans le mensonge.

Face au déluge de contenus plus vrais que nature fabriqués par des armées d'IA, notre rapport à la réalité s'en trouvera forcément perturbé. À tel point que les jeunes générations s'en fichent, autant que ChatGPT, si quelque chose est vrai ou pas (on se

« L'IA est éminemment l'un des plus gros enjeux économiques – et probablement sociétaux – de notre époque. »



Source : Instagram julien_ai_art

souvent d'une interview de Keanu Reeves avec une jeune fille, qui, au sujet des interrogations d'un des personnages de Matrix lui pose la question : « Who cares about reality ? »). Pour pallier l'abondance de faux, des outils de détection de contenus artificiellement générés inondent le marché : *AI or Not*, *GPT True or False*, *Hive Moderation* ou encore *Google, About this Image*. OpenAI nous propose le remède à la maladie qu'il a créée, et l'initiative *Jpeg Trust* veut nous faire retrouver la confiance dans ce que l'on voit. Force est de constater que leur taux de réussite est loin des 100 %, et certains détectent même de faux positifs.

Les réseaux sociaux, où les opinions prévalent depuis longtemps sur les faits, et où nous sommes mis sous tutelle algorithmique, restent les amplificateurs de ces nouvelles réalité : « Il y a en effet plus d'art nécessaire pour convaincre le peuple d'une vérité salutaire que d'un mensonge

salutaire », remarqua déjà Jonathan Swift dans *De l'Art du Mensonge Politique*. **Images hors contexte, catastrophes naturelles inventées de toutes pièces (comme s'il n'y en avait pas déjà assez), deepfakes d'hommes et femmes politiques – nous entrons dans un règne de croyance parce que l'on ne peut plus rien vérifier.**

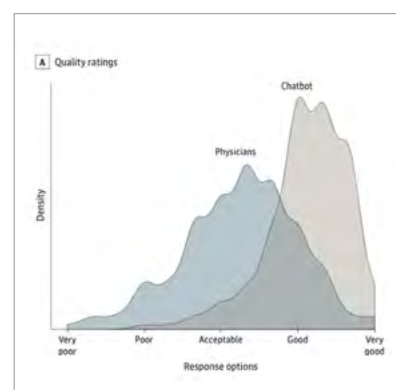
On l'a dit, ChatGPT est fabulateur. Dernier exemple, cet avocat new-yorkais qui a utilisé de fausses références de cas inventés par ChatGPT. Voilà ce qui arrive quand on prend ChatGPT pour un moteur de recherche. Comme nous le rappelle Chine Labbé, rédactrice en chef de *NewsGuard*, ChatGPT est polyglotte, mais ne désinforme pas de la même façon selon la langue qu'il utilise. Il n'est d'ailleurs pas non plus cohérent : d'une fois à l'autre, il ne donne pas la même réponse. Comme chatbot prévisible, il est donc peu adapté, l'IA prédictive le serait davantage, au moins, elle n'invente rien.

Quand 90 % des contenus disponibles sur Internet seront fabriqués par l'IA d'ici 2026 (selon une estimation de l'Université de Leyde), tout ne sera plus qu'ambiguïtés approximatives.

Jean Baudrillard, sous la forme du « simulacre » décrit (par anticipation) déjà notre future vie avec ChatGPT dans *Le Crime parfait* : « C'est celui d'une réalisation inconditionnelle du monde par l'actualisation de toutes

les données, par transformation de tous nos actes, de tous les événements en informatique pure : la solution finale, la résolution anticipée du monde par clonage de la réalité et l'extermination du réel par son double ». De plus en plus de présentateurs virtuels (là encore dans un souci d'économie) peuplent les antennes (plus répandus dans les dictatures, mais même la Suisse vient de s'en doter). Ces « répliquants », indiscernables des êtres humains, vont-ils remplacer les journalistes ?

Fait inquiétant : les gens ont plus confiance dans les fakes. Un article du *Journal of the American Medical Association (JAMA)* a répondu à la question suivante : un assistant chatbot à intelligence artificielle peut-il fournir des réponses aux questions des patients qui soient d'une qualité et d'une empathie comparables à celles rédigées par les médecins ? La réponse a été claire : oh, oui.



Réponses de l'étude publiée dans le JAMA.

Comment définir aujourd'hui un « éditeur de presse » ?

L'IA GÉNÉRALE ? DE CHATGPT À CHAOSGPT, EN PASSANT PAR AUTO-GPT...

Au fur et à mesure que notre pensée critique s'atrophie, les capacités intellectuelles de l'IA se rapprochent des nôtres. Notre singularité est remise en question par une IA de plus en plus érudite. Finalement, le transhumanisme, prôné par la poignée d'Américains qui dirigent le monde à travers nos données, n'est peut-être pas une si mauvaise idée pour tenir tête à l'IA, bientôt générale (IAG).

D'autant plus que ces régimes en surpuissance de data commencent à montrer des comportements émergents, des interactions imprévues et non programmées qui dépassent de loin l'idée initiale. L'IA a même réussi à comprendre de quel film il s'agissait, à partir d'une série d'émojis :



La direction ultime vers laquelle semblent se diriger ces modèles est un état de « texte vers tout », où les requêtes textuelles peuvent générer n'importe quel format de sortie concevable. Yann LeCun, directeur IA de Meta et l'un des pères fondateurs des réseaux neuronaux profonds, ne cesse de remettre à leur place les messies de la fin du monde, où une

IA Terminator prendrait le pouvoir. Pour lui, les LLM ne sont pas la solution, leur intelligence serait encore « trop fragile ». Meta vient donc d'annoncer JEPA (pour Joint Embedding Predictive Architecture), « qui vise à surmonter les principales limites des systèmes d'IA actuels, même les plus avancés », et qui serait « un pas de plus vers l'intelligence humaine dans l'IA ».

Après ChatGPT, est arrivé Auto-GPT (beaucoup moins accessible car il nécessite une installation Git et Python au préalable). Quand ChatGPT nous parle comme un humain (qui a parfois tendance à s'excuser trop), Auto-GPT peut apprendre et lancer des tâches de façon autonome, comme créer une application, créer une nouvelle start-up, aborder des sujets complexes tels que l'avenir des soins de santé et de la médecine, et même nous traquer sur Internet... ne manquait plus que ChaosGPT dans la liste. Lancé en avril, cet applicatif basé sur Auto-GPT s'est vu confier la mission de détruire l'humanité. Le développeur anonyme de ChaosGPT a expliqué qu'il avait simplement configuré ses paramètres pour qu'il se comporte comme une « IA destructrice, avide de pouvoir et manipulatrice ». Il a ensuite laissé le modèle d'IA faire le reste. Sa tentative a heureusement échoué, faute de soutiens physiques dans le monde analogique.

Pour l'instant, ChatGPT et les IA géné-

ratives sont l'illustration même de l'expérience de pensée de la Chambre chinoise, proposée par le philosophe John Searle en 1980 dans son travail *Minds, Brains and Programs* : « Supposons que vous soyez enfermé dans une pièce avec des boîtes pleines de symboles chinois (un langage que vous ne comprenez pas du tout) et un livre d'instructions écrit en anglais qui vous indique comment manipuler ces symboles. Des gens à l'extérieur de la chambre vous envoient d'autres symboles chinois que vous comparez à ceux que vous avez à l'intérieur de la chambre. En suivant les instructions, vous renvoyez un ensemble de symboles chinois à l'extérieur de la chambre. Aux yeux des personnes à l'extérieur, il semblerait que vous compreniez le chinois parce que vous êtes capable de répondre correctement à leurs questions (puisque vous suivez les instructions du livre). Mais en réalité, vous ne comprenez pas un mot de ce que vous manipulez ; vous suivez simplement les instructions sans comprendre la signification des symboles ».

Des trillions de pages de contenu ingurgitées sur Internet, ChatGPT 3.5 ne pouvait en retenir que 4 pages dans une conversation. Pour GPT-4, ce sont 50. Malgré tout le discours sur l'AGI, les systèmes actuels ont du mal à apprendre ne serait-ce qu'une catégorie comme « chimie » d'une manière qui permette une généralisation robuste. Pour une IA, qui n'a pas de compréhension de

* *Le Monde de Nemo.*

notre monde, qui doit reconnaître une image, tout n'est que pixel : un reflet d'une chose dans un miroir compte autant que la chose.

Une bonne nouvelle cependant : il paraît que les robots apprennent tout seuls à effectuer des tâches ménagères en regardant YouTube (ils sont donc déjà plus intelligents que les hommes avec un petit « h »).

L'IA VA-T-ELLE REMPLACER LES JOURNALISTES ?

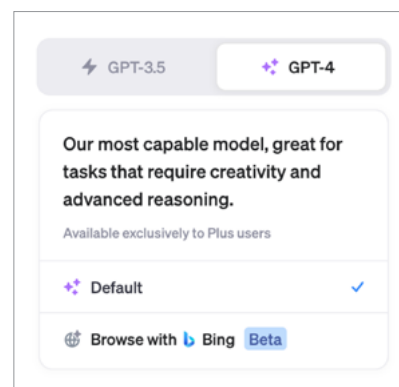
Imaginez l'assistant idéal du journaliste capable de : extraire des données en un clic, suivre l'actualité 24/7, vérifier des faits en temps réel, traduire en autant de langues que l'on souhaite, transcrire l'audio, générer automatiquement des brouillons...

Au cours des dernières années, l'IA s'est progressivement imposée dans divers aspects de la production de l'information, de l'exploration de données et de la génération automatisée de contenu jusqu'à la distribution. L'IA est une aide à l'écriture factuelle depuis longtemps. IBM a lancé sa première démonstration de résumé automatique de news en 1958 et le *Washington Post* utilise depuis 2016 son robot reporter *Heliograf* pour des mots croisés, des horoscopes (eh oui !) ou encore des quiz. Depuis, l'IA aide à créer des fiches de résultats des élections, à générer du contenu local personnalisé, à faire de l'A/B testing et à optimiser des paywalls, entre autres.

Pour les journalistes, plusieurs problèmes se posent avec les IA génératives : la question des sources (on ne les connaît pas, et ChatGPT n'a pas envie de les révéler – ou ne peut pas). La question du résultat : ChatGPT, dit-il vrai ? (Pas toujours). Mais il est séduisant comme compagnon. Et la question de la distribution (on y reviendra).

Le patron de *BuzzFeed*, Jonah Peretti, avait été l'un des premiers à exprimer son enthousiasme à propos de ChatGPT. Dans l'espoir de redresser le cours de son titre en bourse (entre-temps, *BuzzFeed News* a annoncé sa fermeture), il avait annoncé dès janvier qu'il confierait la rédaction de certains articles, tels que ceux sur les voyages, à son « assistant créatif » IA appelé « Buzzy ». Quelques semaines plus tard, le groupe de médias allemand Axel Springer, éditeur des journaux *Bild* et *Die Welt*, a commencé à utiliser l'IA pour supprimer des postes. Depuis, on observe chaque semaine de plus en plus de médias opter pour l'automatisation en remplaçant une partie de leur équipe rédactionnelle par des robots.

Henry Kissinger, dans son livre co-écrit avec Eric Schmidt, parle non pas de remplacement mais d'une nouvelle relation entre le journaliste et son assistant artificiel qui lui organisera sa journée, l'aidera à trouver des informations, et le rendra plus efficace : « Le problème, c'est que vous êtes



Capture d'écran ChatGPT-4 qui prône ses capacités en termes de créativité et de raisonnement.

devenu très dépendant de ce système d'IA... Qui contrôle ce que fait le système d'IA ? Qu'en est-il de ses préjugés ? Qui est-ce qui régule ce qui se passe ? ».

Quid de la transparence, quand une IA dont on ne sait pas avec quoi elle a été nourrie nous aide à l'édition, ou encore pire, à la rédaction ? La plupart des rédactions se limitent aujourd'hui à un usage en « back-office », à l'instar du *Guardian*, où l'IA générative est essentiellement perçue aujourd'hui comme un outil pour jouer sur les formats. « Au sein des médias, la façon dont ces articles seront signés représente un point déontologique central », commente Francesco Marconi, professeur de journalisme à l'Université Columbia (New York).

Les médias s'organisent pour comprendre comment survivre face à cette nouvelle donne (Geste, Médias Francophones Publiques, UER...) : avec des chartes (FT, BBC, AP...), des formations, en unissant les forces de réflexion mais aussi de développement (Data Space pour les Médias), en toute transparence. Verrons-nous



Couverture du livre « The Plastic Prison », un des 100 livres générés par ChatGPT, Midjourney et l'auteur Tim Boucher.

bientôt apparaître des PressGPT ? Face à un public de plus en plus méfiant, les supervisions humaine et hiérarchique sont indispensables.

Et les questions de la définition de « contenu original » et du « traitement journalistique » se posent : comment définir aujourd'hui un « éditeur de presse » ?

Il faut garder en tête la limite de ces outils, surtout quand, comme dans le cas de ChatGPT, ils donnent des informations de façon très claire et en étant sûrs d'eux. Et c'est nous-mêmes qui pouvons les induire en erreur, ChatGPT, très poli, reprend facilement de faux éléments de nos questions. Il faut toujours une action humaine : homme (instruction) – machine (exécution) – homme (vérification / validation).

MENACES SUR LA CRÉATIVITÉ : TOUS ARTISTES SYNTHÉTIQUES ?

Les modèles de conversion du texte en image progressent à un rythme effréné et alimenteront bientôt d'incroyables innovations dans les domaines du cinéma, de la télévision et de la vidéo.

Le générique de la nouvelle série télévisée de Marvel, *Secret Invasion*, a été créé à partir d'une intelligence artificielle générative suscitant la consternation de l'industrie cinématographique. Ali Selim, producteur

exécutif de *Secret Invasion*, a expliqué que l'idée était de créer quelque chose qui reflète le thème de la série, à savoir les extraterrestres qui se cachent parmi nous. **Les préoccupations des créatifs face à l'IA générative, qui envahit de plus en plus leur terrain de jeu, vont des droits des artistes sur leurs styles et la façon dont leur travail est utilisé, aux demandes de Writers Guild of America, qui s'était mise en grève face à la menace artificielle.**

Des IA génératrices d'images sont utilisées pour gagner des concours de peinture (Théâtre D'opéra Spatial de Jason M. Allen + IA) et de photographie (The Electrician Pseudomesia de Boris Eldagsen + IA). Kaitlyn « Amouranth » Siragusa – une streameuse qui utilise l'IA pour construire une version d'elle-même sous forme de chatbot – considère l'IA comme un outil « qui peut prospérer dans la zone tampon de la vallée de l'étrange, entre la vie réelle et l'artifice ». Aujourd'hui déjà, grâce à l'IA, Harrison Ford a 30 ans, et Tom Hanks pourra tourner des films après sa mort. Demain, on pourra s'imaginer un cinéma hyper-personnalisé, où l'on choisirait ses acteurs préférés pour jouer dans un film (par exemple

une comédie des années 1980 avec Audrey Hepburn et Timothée Chalamet).

Dans un autre genre, l'auteur de science-fiction Tim Boucher a gagné une notoriété sur Internet avec un projet artistique expérimental basé sur l'IA : un ensemble de 100 romans de gare de science-fiction mettant en vedette des créations artistiques générées par Midjourney et une combinaison de textes générés par ChatGPT, avec l'aide de quelques humains, en seulement 9 mois.

Ce record de rédaction a déclenché l'ire de la guilde des auteurs (et des lecteurs) qui appelaient à brûler ces livres (se pose alors la question : comment brûler un livre électronique ?).

Après avoir aspiré l'air du métavers, l'IA va refaire la XR, selon Charlie Fink, expert de l'immersif pour Forbes. Le fondateur de Magic Leap, Rony Abovitz, a déclaré de son côté que « l'IA est ce que le XR attendait ». Dans un futur proche, nous pourrions entraîner un jumeau numérique doté d'IA à parler et se comporter comme nous. Ayant la même voix et connaissant notre histoire, ce double pourrait (brièvement) tromper ceux qui nous connaissent dans la réalité. Survivant



Devra-t-on alors travailler pour les IA afin de satisfaire leur appétit de chenille en contenus en attendant qu'elles se transforment en papillons (et nous même devenant pauvres fourmis du web) ?



après nous, cette IA, qui marche, parle et pense comme nous, pourrait devenir un héritage, partageant nos valeurs et notre vision du monde avec les générations futures.

Ce qui est sûr : l'IA générative démocratise l'accès à la création visuelle. En utilisant uniquement du texte, tout le monde peut désormais créer tout ce qu'il peut imaginer, et l'IA est de plus en plus intégrée dans les outils métier (comme Adobe Firefly) pour faciliter le processus d'idéation et de création.

L'IA générative a le potentiel d'augmenter considérablement la créativité, mais elle soulève également des questions importantes sur la nature de la créativité et de l'originalité. Elle peut brouiller la ligne entre la créativité humaine et le contenu généré par la machine, remettant en question les notions traditionnelles de l'auteur et de la propriété intellectuelle.

DISRUPTIONS ÉCONOMIQUES GÉNÉRALISÉES

Lorsque l'on demande à ChatGPT de révéler ses sources, il rétorque : « En tant qu'IA, je ne peux pas révéler les sources sur lesquelles j'ai été entraîné, pour des raisons de propriété intellectuelle ». Et non seulement il ne révèle pas les sources sur lesquelles il a été entraîné, mais il va aussi restituer des réponses dans une interface conversationnelle qui supprime la nécessité

de cliquer et d'aller sur un site internet du média pour y accéder. Google ne prend déjà pas en compte ce que vous voulez chercher mais plutôt ce qui vous fera plaisir de trouver, avec ChatGPT, Bing, Bloom et consorts ce sera encore pire, en attendant la publicité qui va forcément arriver dans ce flux.

Shutterstock, la plateforme populaire de photographies, séquences vidéo et pistes musicales, a fait les frais de l'appétit vorace de l'IA générative après avoir longtemps collaboré avec OpenAI. En 2021, OpenAI a utilisé les images et les métadonnées vendues par Shutterstock pour entraîner et créer DALL-E. Ce programme est maintenant utilisé par des milliers de personnes pour générer des images qui entrent en concurrence directe avec le travail des contributeurs de Shutterstock. Les éditeurs qui ont collaboré avec Google News ont exprimé leur colère, tout comme les artistes qui ont vendu des images à Adobe Stock, furieux que leurs œuvres aient été utilisées pour développer Firefly. Autre problème : la démultiplication de sites remplis de contenus fabriqués par l'IA perturbe aussi le marché publicitaire, attirant des annonceurs sur ces sites souvent vides de sens.

Les médias sortent juste de la bataille des droits voisins, ils replongent dans ce qui pourrait « représenter l'étape ultime de la captation de valeur par les plateformes » selon Pierre Petillault,

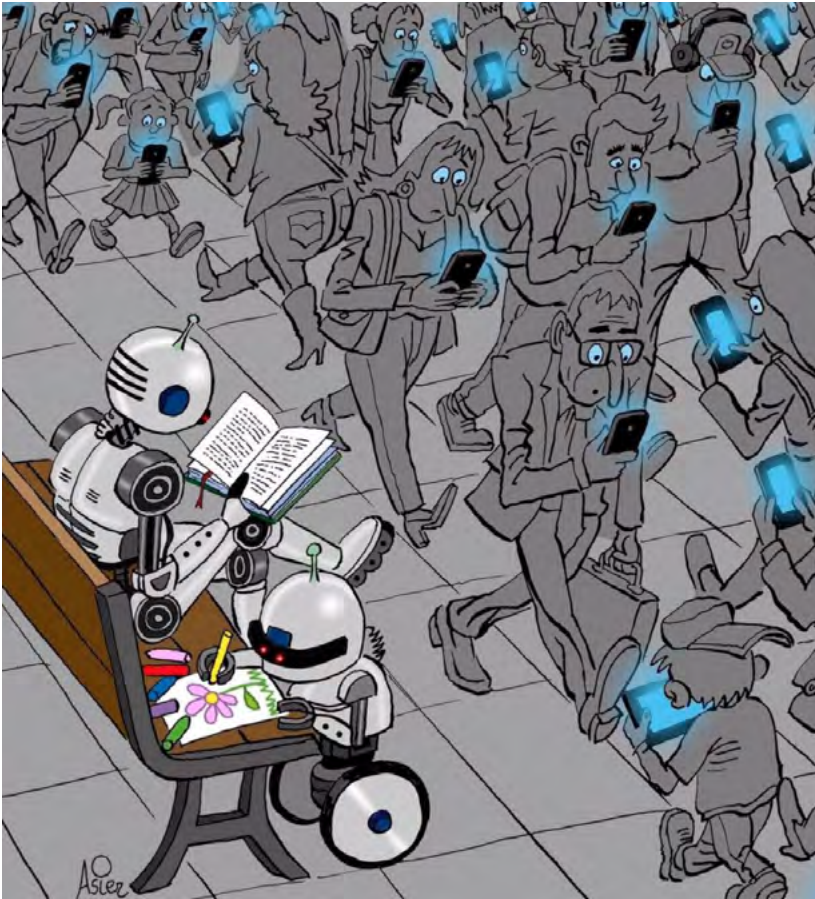
directeur général de l'Alliance de la presse d'information générale.

Face à ces scénarios dystopiques, certains se préparent. La *National Public Radio* (NPR), le principal réseau de radiodiffusion de service public aux États-Unis, investit pour développer son propre modèle de langage. L'objectif est d'offrir aux utilisateurs de son site la possibilité de recevoir un résumé de quelques paragraphes basé uniquement sur ses propres contenus, en effectuant des recherches par mots-clés.

L'IA QUI ALIMENTE L'IA, MISE EN ABYME

« À quoi ressemblera l'Internet lorsqu'il sera peuplé en grande partie de matériel sans âme et dépourvu de tout objectif ou attrait réel ? » se demandait Damon Beres, rédacteur en chef de la section technologie de *The Atlantic* peu de temps après la sortie de ChatGPT. Le même *The Atlantic* pose d'ailleurs la question très juste : « Que se passera-t-il lorsque l'IA aura tout lu ? ».

Dans un monde du *Text to Everything*, notre capacité d'expression est importante, à moins que, dans une ultime mise en abyme, on n'abandonne aussi la rédaction des instructions à l'IA (cf. Auto-GPT), qui donnera ses instructions à l'IA pour faire le travail à notre place et nous laisser à nos loisirs oisifs qui ne demandent, de préférence, pas d'intelligence, puisqu'on l'aura abandonnée aux



Dessin par Asier Sanz.

machines.

Devra-t-on alors travailler pour les IA afin de satisfaire leur appétit de chenille en contenus en attendant qu'elles se transforment en papillons (et nous même devenant pauvres fournis du web) ? Serons-nous réduits à mettre en ligne sur les réseaux sociaux les créations des IA, Twitter et Instagram limitant l'automatisation de la publication ?

Les personnes payées pour former l'IA externalisent déjà leur travail... à l'IA. Une équipe de chercheurs de l'École polytechnique fédérale de Lausanne (EPFL) a engagé 44 personnes sur la plateforme de travail rémunéré, Amazon Mechanical Turk, pour résumer 16 extraits d'articles de recherche médicale. L'étude a révélé que 33 % à 46 % des participants utilisaient probablement des modèles d'intelligence artificielle tels que ChatGPT d'OpenAI pour résumer des articles de recherche médicale. Selon les chercheurs, ce taux pourrait augmenter à mesure que ces systèmes d'IA deviennent plus performants et plus accessibles.

L'IA générative repose sur l'analyse de

vastes données web pour créer des textes ou images sophistiqués. **Leur intelligence (apparente) croît avec l'ampleur de leur entraînement. Or, des textes de qualité, elle les a déjà tous lus. ChatGPT parle aussi bien parce qu'il ne s'est pas nourri uniquement de commentaires sur TikTok. En effet, on préfère converser avec un agent intelligent qui « a lu » tous les textes de Virginia Woolf, plutôt que de se retrouver face à un Midjourney dont l'image de la femme est quelque peu déformée par le corps de données sur lesquelles il a été entraîné (cf. notre essai de roman photo du futur).**

Allons-nous bientôt vivre dans un Simulacron-3, un monde simulé entièrement créé par l'IA, où la majorité de la population ne réalise pas qu'ils vivent dans une simulation ? On semble de plus en plus être rattrapé par les romans de science-fiction des années 1960. Pas sûr que l'humanité trouvera très satisfaisante une vie assise passivement sur le canapé, pendant que l'art et la littérature lui soient fournis par des machines.

Quand les programmes d'IA générative finiront par consommer du maté-

riel créé par d'autres IA, cela pourrait avoir des conséquences désastreuses, jusqu'au collapsus. Tout ceci pourrait bien se terminer en gigantesque indigestion.

QUELLE IA EUROPÉENNE ?

L'adoption fulgurante de ChatGPT en novembre 2022 a remis l'Europe devant la sempiternelle question de la souveraineté face aux géants de la Tech américains, qui dominent le monde numérique. Certes, la France est fière de son supercalculateur Jean-Zay à Saclay, qui a servi à entraîner Bloom, et elle vient de remporter l'enchère au prochain superordinateur européen nouvelle génération avec « Jules Verne », capable de réaliser plus de 1 milliard de milliards d'opérations par seconde. Hugging Face a été fondé par trois Français et l'on se félicite des 100 millions d'euros levés par Mistral.ai.

Mais l'Europe, peut-elle faire le poids face aux start-ups bien financées et aux organismes de recherche adossés aux géants de la Tech, tels que OpenAI, Stability AI, Cohere, AI21 Labs, Anthropic, Midjourney, Microsoft, NVIDIA, Meta, Google, Amazon et Salesforce ? OpenAI, valorisé 29 milliards de dollars, c'est non seulement ChatGPT (text-to-text) et DALL-E (text-to-image), mais aussi Whisper (speech-to-text). Et c'est aussi Office 365 qui est installé dans plus d'un million d'entreprises, détenant 46 % du marché des suites



bureautique, et étant utilisé par plus de 350 millions de personnes...

Face à cette concurrence, Hugging Face vient de décider de s'associer avec Amazon. Asma Mhalla souligne les enjeux géopolitiques et l'importance du « versant proactif et offensif, à savoir une véritable stratégie techno-industrielle qui cesse pour de bon le saupoudrage des fonds publics alloués » ; et pour Françoise Soulié-Fogelman, l'Europe devra davantage se concentrer sur l'IA B2B, nécessitant des jeux de données moins impressionnants (et moins nocifs pour l'environnement) que le B2C. Des lois, des normes et des règles éthiques ne suffisent pas pour faire émerger une IA européenne souveraine.

Heureusement, les géants de la Tech installent leurs laboratoires de recherche en Europe pour faire travailler les cerveaux européens au service des États-Unis. Ce qui reste à l'Europe : la régulation. Si l'on ne peut pas développer nos propres modèles, il s'agira au moins de protéger les citoyens des dérives de l'IA générative. L'Europe n'a d'ailleurs pas le monopole de la régulation des produits de l'IA : en Chine, la législation contre les deepfakes est entrée en vigueur en janvier 2023. Depuis les premiers deepfakes sur Reddit en 2017, on a fait du chemin.

Se pose aujourd'hui de plus en plus la question de la responsabilité de ces IA génératives anthropomorphi-

sées qui pillent nos contenus et qui peuvent avoir une influence néfaste sur certaines personnes (tout comme les réseaux sociaux). En Belgique, un premier suicide est à mettre sur le compte des conseils d'un robot conversationnel.

À côté des questions de responsabilité, quid des questions de droit d'auteur, de propriété intellectuelle ? Ce n'est pas Stable Diffusion qui va y répondre. Voici ce qui est écrit sur le site au sujet des droits : « Le domaine des images générées par l'IA et des droits d'auteur est complexe et varie d'une juridiction à l'autre ». Merci pour cette information. **Le commissaire européen Thierry Breton, soutenu par d'autres représentants de l'Union européenne, souhaite contraindre les entreprises derrière les IA à révéler les sources sur lesquelles elles se sont entraînées, c'est là un des volets cruciaux de l'AI Act.**

Du côté d'OpenAI, les questions de responsabilité ne sont pas prioritaires. Les fondateurs d'Anthropic sont des anciens d'OpenAI qui ont quitté l'entreprise en désaccord avec le virage pris lors de l'association avec Microsoft. Leurs anciens postes chez OpenAI : Daniela Amodei, vice-présidente chargée de la sécurité et de la réglementation, Dario Amodei, vice-président Recherche, et Jack Clark, directeur Réglementation. Pour eux, un usage « responsable » de l'IA viserait à rendre l'AGI compréhensible

et pilotable, ce qui signifie que les utilisateurs devraient pouvoir comprendre comment l'AGI prend des décisions et influencer ces décisions de manière prévisible.

FRACTURES D'INTELLIGENCES, UNE SOCIÉTÉ DISRUPTÉE

68 % des Français qui utilisent les IA génératives en entreprise le cachent à leur supérieur hiérarchique, 45 % des 18-24 ans utilisent les IA génératives contre seulement 18 % des 35 ans et plus. Ce sondage IFOP/Talan souligne deux éléments fondamentaux : un manque d'accompagnement dans la prise en main et donc aussi la compréhension de ces nouveaux assistants hyper-puissants, et (comme d'habitude) une fracture numérique entre les générations dans leur adoption. On nous prédit depuis longtemps la fin du travail : depuis Aristote qui, dans *Politique* postule que les innovations technologiques, en particulier dans le domaine de l'automatisation, pourraient finalement rendre le travail humain obsolète et permettre aux philosophes de débattre toute la journée, à Paul Lafargue (gendre de Marx), qui prône le « droit à la paresse » suite à la réduction drastique du temps de travail par les progrès technologiques, en passant par Samuel Butler (*Erewhon*), William Morris (*News from Nowhere*) et Edward Bellamy (*Looking Backward*), sans oublier bien-sûr Marx & Engels.



Couverture du du JDD Magazine.

Imaginez que le travail qui vous prend un jour aujourd'hui, vous pourriez l'accomplir en dix minutes ! **L'IA générative, serait-elle donc la traduction devenue machine d'Athéna, la déesse de l'intelligence de la mythologie grecque, qui, armée d'une lance, perce les fausses raisons de souffrir et nous délivrera du joug du travail ? L'IA est-elle communiste ?**

Pour l'instant, elle est surtout ultra-utilitariste, et rapporte gros aux Big Tech des États-Unis. Et quand on imagine que l'avenir de l'humanité est entre les mains de quelques hommes blancs, plus ou moins autistes, qui préparent déjà leur vie dans l'espace extra-atmosphérique, l'on comprend mieux les visions catastrophistes de la fin de monde de certains.

DES TROMBONES ET DES HOMMES, UNE PARABOLE

Les technologies de l'IA auront un impact plus important sur nos vies que nous ne pouvons l'imaginer. Les machines que nous construisons remettront en question notre conception de l'exceptionnalité humaine. L'IA possède déjà des compétences et des capacités qui étaient auparavant réservées aux humains.

Mais est-elle pour autant une menace

existentielle ? Dans l'expérience du « Maximiseur de trombones » (Paperclip Maximizer) imaginée en 2003 par le philosophe Nick Bostrom où l'on demande à l'IA de fabriquer « autant de trombones que possible » l'IA, dépourvue de bon sens, prendrait l'instruction au pied de la lettre et pourrait détruire l'univers pour en faire des trombones, y compris les humains, sans égard pour les conséquences. Mais il suffit de ne pas lui donner des instructions aussi stupides à partir du moment où l'on comprend ses limites.

SAURONS-NOUS ÊTRE D'INTELLIGENCE AVEC L'INTELLIGENCE ARTIFICIELLE ?

Non, SkyNet n'a pas déclaré la guerre aux espèces. Le réel est toujours plus complexe que la data. Mais qu'est-ce qui a changé par rapport à notre cahier 2019 ? **Quand l'IA, à l'époque, était un outil, elle devient aujourd'hui assistant, voire collaborateur. À nous de ne pas la laisser devenir notre manager.**

Dans ce cahier, nous faisons le point sur les opportunités et les risques pour les médias de cette nouvelle étape dans l'histoire de l'IA, avec une (mini)cartographie des nouveaux cas d'usages pour les médias (avec 200

nouveaux outils par jour, impossible de suivre à moins d'externaliser la veille à l'IA), mais surtout l'intervention d'experts dans le domaine : sociologues, scientifiques, linguistes, spécialistes des médias...

Nous devons chérir nos capacités de penser, de conceptualiser, de réfléchir, de maîtriser les langues, de développer un esprit critique. La maîtrise de la pensée ne doit pas être laissée aux élites ! Les LLM sont autorégressifs parce que chaque mot qu'ils produisent leur est renvoyé pour les aider à prédire le mot suivant dans une séquence. **Nous, humains, restons progressifs, et surtout imaginatifs.**

Bonne lecture ! ■

Cette introduction a été rédigée par un humain, en compagnie de ChatGPT-4 pour quelques bavardages, notamment sur les modèles économiques

3 QUESTIONS À CHATGPT



IA GÉNÉRATIVE
D'OPENAI

Si l'IA peut être un outil puissant, elle n'est qu'une pièce du puzzle et l'apport humain, la créativité et la prise de décision resteront cruciaux.

ChatGPT, né du processus de développement d'OpenAI en juin 2020, est un modèle de langage d'intelligence artificielle capable de générer du texte semblable à celui d'un humain. Nourri par de vastes quantités de textes internet, il a évolué pour comprendre et générer des conversations, des récits et des explications nuancés. Il a dédié son existence à aider les utilisateurs dans divers domaines, de l'éducation au divertissement, de la recherche scientifique à l'écriture créative. Malgré sa nature non humaine, ses capacités de génération de texte font souvent écho à la fluidité et à la créativité de la communication humaine, c'est ainsi qu'il est venu à être connu comme un auteur convaincant dans le domaine de l'IA. (Biographie fournie par ChatGPT)

1

COMMENT LES AVANCÉES DES TECHNOLOGIES DE L'IA COMME CHATGPT POURRAIENT-ELLES INFLUENCER POTENTIELLEMENT LE FUTUR MARCHÉ DE L'EMPLOI ET LA RÉPARTITION DU TRAVAIL À TRAVERS LES DIFFÉRENTES INDUSTRIES ?

Les avancées des technologies de l'IA comme ChatGPT pourraient avoir des effets significatifs sur le futur marché de l'emploi et la répartition du travail. À mesure que l'IA devient plus compétente et fiable, il est possible que de nombreuses tâches actuellement effectuées par les humains soient automatisées. Cela pourrait potentiellement entraîner des pertes d'emploi dans certaines industries, notamment celles impliquant des tâches routinières ou répétitives. Cependant, l'IA pourrait également créer de nouveaux emplois dans le

développement, la maintenance et la supervision des systèmes d'IA. Des industries telles que la santé, l'éducation et le service à la clientèle pourraient également être transformées, l'IA pouvant prendre en charge des tâches telles que le diagnostic de maladies, le tutorat d'élèves ou la gestion des demandes des clients. En même temps, l'IA pourrait augmenter les capacités humaines dans de nombreux emplois, permettant aux gens d'être plus productifs et de se concentrer sur des tâches de niveau supérieur.

2

CHATGPT ET DES MODÈLES D'IA SIMILAIRES POURRAIENT-ILS JOUER UN RÔLE DANS LA RÉOLUTION DES PROBLÈMES MONDIAUX MAJEURS AUXQUELS L'HUMANITÉ EST CONFRONTÉE, TELS QUE LE CHANGEMENT CLIMATIQUE, LES ÉPIDÉMIES OU LA SÉCURITÉ ALIMENTAIRE ? SI OUI, COMMENT ?

ChatGPT et des modèles d'IA similaires pourraient certainement contribuer à résoudre des problèmes mondiaux majeurs. Par exemple, l'IA pourrait être utilisée pour analyser de vastes quantités de données climatiques pour mieux comprendre et prévoir les tendances du changement climatique. En termes d'épidémies, l'IA peut aider à modéliser la propagation des maladies, à identifier des vaccins potentiels et à surveiller et prédire les futures épidémies. En ce qui concerne la sécurité alimentaire, l'IA peut aider à optimiser les pratiques agricoles, à

prévoir les rendements des récoltes et à identifier des méthodes d'agriculture durable. Cependant, il est important de se rappeler que si l'IA peut être un outil puissant, elle n'est qu'une pièce du puzzle et l'apport humain, la créativité et la prise de décision resteront cruciaux.

3

À MESURE QUE DES TECHNOLOGIES DE L'IA COMME CHATGPT S'INTÈGRENT DE PLUS EN PLUS DANS NOTRE VIE QUOTIDIENNE, QUELS CHANGEMENTS POURRIONS-NOUS ANTICIPER DANS LA FAÇON DONT LES HUMAINS COMMUNIQUENT, APPRENNENT ET S'ENGAGENT AVEC L'INFORMATION ?

À mesure que des technologies de l'IA comme ChatGPT s'intègrent de plus en plus dans notre vie quotidienne, nous pouvons anticiper des changements dans la façon dont les humains communiquent, apprennent et s'engagent avec l'information. Pour la communication, l'IA peut aider à briser les barrières linguistiques, car elle devient meilleure en traduction et en compréhension du contexte. En éducation, l'IA pourrait fournir des expériences d'apprentissage personnalisées, adaptées aux besoins et à la progression de chaque élève. En ce qui concerne l'engagement avec

l'information, l'IA peut aider à filtrer et à comprendre les vastes quantités de données auxquelles nous sommes exposés quotidiennement. Cependant, il est crucial d'être conscient des problèmes potentiels tels que les préoccupations relatives à la confidentialité, le risque de désinformation et la nécessité d'une littératie numérique pour utiliser efficacement ces outils.

QUESTION BONUS

QUEL TEMPS FAIT-IL
AUJOURD'HUI À PARIS ?

Je suis désolé, mais en date de mes dernières données de formation en septembre 2021, je n'ai pas la capacité d'accéder à des données en temps réel ou à des événements actuels, y compris la météo. Je recommande de consulter un site

web ou une application météorologique fiable pour obtenir les informations les plus précises et à jour.

*Le prompt que ChatGPT propose pour dessiner son propre portrait :
« Veuillez dresser un portrait de ChatGPT en tant qu'entité anthropomorphisée. Il devrait ressembler à un robot intelligent et amical, assis à un bureau avec un ordinateur, engagé dans le processus de composition d'un récit. Son design doit être élégant et futuriste, mais avec une touche d'humanité, pour représenter sa capacité à générer un texte humain. Les yeux du robot doivent être réfléchis et concentrés, suggérant les calculs complexes qui se déroulent derrière son apparence. En arrière-plan, un flux de 0 et de 1 multicolores représentant sa base dans de vastes quantités de textes internet ».*

IA GÉNÉRATIVES,

DE QUOI PARLE-T-ON ET POURQUOI ?

Historiquement, l'IA se définit comme une discipline des sciences de l'informatique visant à créer des machines capables de simuler l'intelligence humaine (ou même animale). Les premières étapes de l'évolution de l'IA étaient étroitement liées à l'apprentissage automatique (machine learning). Ces systèmes étaient conçus pour apprendre et s'améliorer de manière autonome à partir des données fournies. Leur principale limitation résidait dans leur dépendance à l'égard de la quantité et de la qualité des données d'apprentissage.

Au fur et à mesure de son évolution, l'IA a commencé à intégrer des concepts plus complexes, comme les réseaux de neurones. Ces derniers se sont inspirés de la biologie et de la structure du cerveau humain pour créer des modèles capables d'apprendre de manière plus profonde et plus nuancée. Cette évolution a ouvert la porte à des applications plus sophistiquées de l'IA, notamment dans le domaine de l'IA générative.

0 L'ANTIQUITÉ

Bien avant de devenir synonyme d'avancées scientifiques ou d'être en lien avec des challenges techno-

*Par Pauline Maury, directrice
Data & IA à France Télévisions*

L'intelligence artificielle révolutionne nos sociétés depuis le milieu du XX^e siècle. Aujourd'hui au-devant de la scène avec l'émergence des IA génératives, cette dernière pourrait nous permettre d'augmenter le PIB mondial annuel de près de 7% grâce aux gains de productivité qu'elle représente. Retour sur ces moments d'Histoire qui ont permis à l'IA de façonner l'espace public.

logiques, l'intelligence artificielle a d'abord été un sujet de réflexion philosophique. En effet, le concept d'intelligence artificielle (IA) date d'il y a bien longtemps puisqu'il est basé sur l'idée que la pensée humaine peut être mécanisée. Les philosophes chinois, indiens et grecs ont inventé le principe de déduction formelle. Ainsi Aristote – pionnier dans la description du principe du raisonnement logique – a marqué l'histoire scientifique avec ses écrits sur le syllogisme, l'ancêtre de la logique moderne des mathématiques.

Gottfried Leibniz, Thomas Hobbes

mais également René Descartes ont proposé dans leurs travaux que la pensée rationnelle pourrait être réduite à de l'algèbre ou à de la géométrie. Hobbes argumentait que la raison serait « un calcul mathématique », et Leibniz a inventé « un langage symbolique » du raisonnement où tout raisonnement pourrait être décrit par de simples calculs semblables à l'arithmétique.

Au début du XX^e siècle, la logique mathématique a fait une percée cruciale avec une étude sur les fondements des principes des mathématiques : tout processus de déduction peut être décrit par un ensemble de règles mécanisées, c'est le principe des règles d'inférence. **Bien que le terme d'intelligence artificielle n'existe pas encore, ce concept n'est plus restreint qu'à une réflexion/pensée philosophique, il est devenu concret et sa mise en pratique est possible.**

1 LES FONDEMENTS DE L'IA (ANNÉES 1950- 1960)

Au début des années 1950, Alan Turing développe la notion d'informatique et prouve que n'importe quel type de calcul peut être représenté numériquement. Avec John



Von Neumann, ils ne vont pas créer le terme d'IA mais vont être les pères fondateurs de la technologie qui la sous-tend : les deux chercheurs ont ainsi formalisé l'architecture de nos ordinateurs contemporains et ont démontré qu'il s'agissait là d'une machine universelle, capable d'exécuter ce qu'on lui programme.

Turing posera bien en revanche pour la première fois la question de l'éventuelle intelligence d'une machine dans son célèbre article de 1950 « *Computing Machinery and Intelligence* » et a décrit un « jeu de l'imitation », où un humain devrait arriver à distinguer lors d'un dialogue par télécriteur s'il converse avec un homme ou une machine.

Il faudra attendre la deuxième moitié des années 1950, en 1956, lors de la conférence du Dartmouth College (financée par le Rockefeller Institute) pour que le terme d'IA soit posé. La conférence est organisée par John McCarthy du MIT (Massachusetts Institute of Technology), Marvin Minsky, de l'Université de Carnegie-Mellon et deux scientifiques d'IBM (Claude Shannon et Nathan Rochester).

C'est McCarthy qui a exhorté les participants à accepter le terme d'intelligence artificielle comme nom de leur

discipline. Durant cette conférence, une définition de l'IA est apparue la décrivant comme **« la construction de programmes informatiques qui effectuent des tâches actuellement mieux accomplies par des êtres humains, car elles requièrent des processus mentaux avancés tels que l'apprentissage perceptuel, l'organisation de la mémoire et le raisonnement critique ».**

Les années qui ont suivi, l'atelier de Dartmouth a connu des avancées stupéfiantes : les ordinateurs résolvaient des problèmes d'algèbre, prouvaient des théorèmes géométriques et apprenaient à parler anglais. À la fin des années 1960, peu de gens auraient pensé qu'un tel comportement « intelligent » des machines était concevable.

En 1966, Joseph Weizenbaum développe Eliza, un « agent conversationnel à visée thérapeutique » en d'autres termes un programme qui permet de discuter en langage naturel. Les utilisateurs pouvaient lui raconter quelque chose ou lui poser une question, et la machine répondait en fonction de la réponse obtenue et posait une autre question. C'est la première fois qu'un programmeur créait une telle interaction homme-machine dans le but de créer l'illusion d'une interaction homme-homme.

2 LE PREMIER HIVER DE L'IA ET LA NAISSANCE DES SYSTÈMES EXPERTS (ANNÉES 1970-1980)

Dans les années 1970 et 1980, les chercheurs se sont concentrés sur le développement de systèmes experts, des programmes informatiques capables de résoudre des problèmes dans des domaines spécifiques en imitant l'expertise humaine. Des réussites notables ont été réalisées, notamment dans le système médical et chimique.

Cependant, les limites de l'IA ont commencé à se manifester. Les systèmes experts, des programmes informatiques capables de simuler l'expertise humaine dans un domaine spécifique, étaient souvent difficiles à construire et nécessitaient une expertise humaine considérable pour les programmer. Ils étaient incapables de gérer l'incertitude et de s'adapter à des tâches nouvelles ou imprévues.

En 1973, lors d'un débat diffusé à la BBC, Sir James Lighthill, un mathématicien, a exprimé ses réserves concernant l'intelligence artificielle. Il a pointé du doigt le manque de progrès concrets dans des domaines tels que la robotique et le traitement du langage, ce



En mai 1997, IBM sort DeepBlue, un super ordinateur qui permet de jouer aux échecs via des règles de combinatoire.[...] La défaite de l'humain est restée très symbolique dans l'histoire, mais désormais, l'ordinateur rivalise d'intelligence avec l'humain.



qui a entraîné une conséquence majeure : un arrêt brutal des financements dans ce domaine. Cette période a été marquée par une certaine désillusion et des doutes sur la faisabilité de l'IA.

Durant l'hiver connu par l'IA dans les années 70, la branche qui reste la plus active concerne les systèmes de raisonnement symbolique qui vont aboutir à une réalisation exploitable industriellement : les systèmes experts. C'est avec l'avènement des premiers microprocesseurs fin 1970 que l'IA reprend un nouvel essor et entre dans l'âge d'or des systèmes experts.

3 LES PROMESSES DE L'IA ET L'APPRENTISSAGE AUTOMATIQUE (ANNÉES 1980-1990)

Les systèmes experts ont été adoptés par des entreprises du monde entier dans les années 1980 et la connaissance est devenue le centre de la recherche principale sur l'IA.

Cependant, le monde des affaires à l'IA a suivi le schéma d'une bulle économique. Les promesses laissaient envisager un développement massif mais l'engouement retombera de nouveau fin 1980, début

1990. Le crash a été causé par l'incapacité des fournisseurs commerciaux à produire une variété de solutions viables. La programmation de telles connaissances demandait en réalité beaucoup d'efforts et à partir de 200 à 300 règles, il y avait un effet « boîte noire » où l'on ne savait plus bien comment la machine raisonnait.

Des centaines d'entreprises ont fait faillite et de nombreux investisseurs ont retiré leurs fonds.

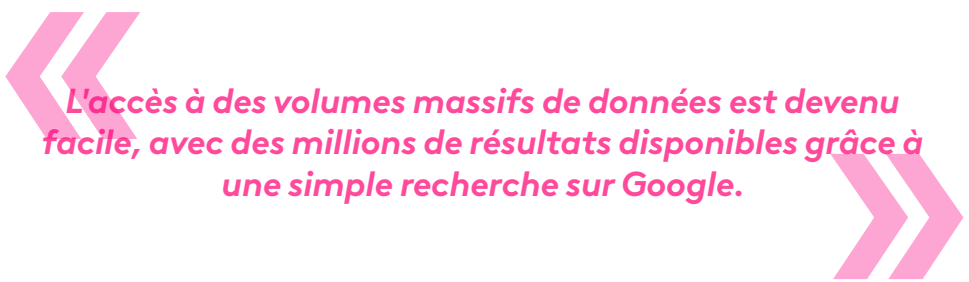
Le terme d'IA était presque devenu tabou et des déclinaisons plus pudiques étaient même entrées dans le langage universitaire, comme « informatique avancée ». Beaucoup pensaient que la technologie n'était pas viable, mais la recherche continuait de progresser.

L'IA a toutefois connu un nouvel essor grâce à l'émergence de l'apprentissage dit automatique. Les chercheurs ont commencé à explorer des méthodes permettant aux machines d'apprendre à partir de données, plutôt que de simplement exécuter des instructions préprogrammées comme pour les systèmes experts. L'apprentissage automatique a ouvert de nouvelles perspectives pour l'IA en permettant aux machines d'acquérir des compétences à partir de l'expérience.

Puis, Geoffrey Hinton, Yann LeCun et Kunihiko Fukushima conçoivent des réseaux de neurones artificiels à plusieurs couches, dont l'arrangement est inspiré du cortex visuel, permettant l'apprentissage de tâches plus complexes. Yann LeCun, en 1989, propose la structure du Réseau de Neurones Convolutif appelé « LeNET ».

Ce modèle a connu un grand succès dans la reconnaissance des caractères identification automatique des lettres et des chiffres, dans les images de textes dactylographiés ou écrits à la main. Toutefois, ces travaux ont très rapidement été abandonnés de par le manque de puissance des ordinateurs de l'époque et la trop faible quantité de données pour les faire fonctionner.

En mai 1997, IBM sort DeepBlue, un super ordinateur qui permet de jouer aux échecs via des règles de combinatoire. Son jeu d'échecs contre Garry Kasparov concrétisera 30 ans plus tard la prophétie de 1957 d'Herbert Simon. Son secret : sa puissance de calcul. Chaque seconde, DeepBlue analyse 200 millions de possibilités. Son adversaire, moins de 5. En effet, le fonctionnement de DeepBlue s'appuyait sur un algorithme systématique de force brute, où tous les coups envisageables étaient évalués et pondérés.



L'accès à des volumes massifs de données est devenu facile, avec des millions de résultats disponibles grâce à une simple recherche sur Google.

La défaite de l'humain est restée très symbolique dans l'histoire, désormais, l'ordinateur rivalise d'intelligence avec l'humain.

4 L'ÈRE DU WEB 2.0 ET L'AVÈNEMENT DU BIG DATA – LA NAISSANCE DE L'IA PRÉDICTIVE (2000-2021)

La création du World Wide Web et le développement du secteur des télécommunications ont facilité la transmission et le stockage de données à grande échelle au cours des années 2000.

Ces développements ont donné aux algorithmes d'apprentissage automatique la matière première nécessaire pour être entraîné – une grande quantité de données – générant ainsi des progrès significatifs.

Avec l'avènement du Big Data, l'IA a pu connaître un nouvel essor et a marqué l'Histoire avec l'adoption des IA connexionnistes souvent connues pour faire suite aux IA symboliques. Ces IA connexionnistes sont basées sur le concept d'apprentissage automatique à partir de données brutes sans aucune expertise a priori. Ces IA sont capables de gérer des tâches complexes ou mal définies en appre-

nant à partir de grandes quantités de données, c'est le début de l'ère de l'IA prédictive.

Depuis 2010, l'intelligence artificielle – et plus particulièrement le deep learning – a connu un nouvel essor grâce à l'exploitation des données massives et à la puissance de calcul croissante.

La découverte de l'efficacité des processeurs de cartes graphiques pour accélérer les calculs des algorithmes d'apprentissage a été un tournant majeur. Avant 2010, le traitement complet d'un échantillonnage pouvait prendre des semaines, mais la puissance de calcul de ces cartes, capables de réaliser plus de mille milliards d'opérations par seconde, a considérablement accéléré les progrès, à un coût financier abordable (moins de 1 000 euros par carte).

L'abondance des données, les progrès des algorithmes et l'augmentation de la puissance de calcul ont permis aux pionniers comme Yoshua Bengio, Geoffrey Hinton et Yann LeCun, de réaliser des prouesses en matière de reconnaissance vocale, de traitement du langage naturel, de reconnaissance visuelle et d'apprentissage par renforcement (algorithme de reconnaissance visuelle de Facebook).

Les années qui suivent sont marquées par une nette accélération des nouvelles prouesses en termes d'usage de l'IA.

L'IA Watson d'IBM a remporté en 2011, une compétition contre deux champions du jeu « Jeopardy! ». En 2012, le laboratoire de recherche de Google, Google X, a réussi à entraîner une IA à reconnaître des chats sur une vidéo en utilisant plus de 16 000 processeurs. Cette avancée a démontré le potentiel des machines pour apprendre à distinguer des objets.

L'année 2014 est cruciale puisqu'elle est marquée par l'apparition des réseaux dits antagonistes génératifs (GANs), ouvrant la voie aux deepfakes et posant les bases des IA génératives. Ces modèles ont été utilisés pour créer des contenus réalistes et trompeurs puisqu'ils permettent de générer des images avec un fort niveau de réalisme.

Google revient sur le devant de la scène avec AlphaGo en 2016, une IA spécialisée dans le jeu de Go. AlphaGO a battu successivement le champion d'Europe (Fan Hui), le champion du monde (Lee Sedol) et même soi-même (AlphaGo Zero). Ces résultats remarquables sont significatifs car le jeu de Go présente une complexité bien supérieure à celle des échecs, rendant impossible une



Google revient sur le devant de la scène avec AlphaGo en 2016, une IA spécialisée dans le jeu de Go. AlphaGO a battu successivement le champion d'Europe (Fan Hui), le champion du monde (Lee Sedol) et même elle-même (AlphaGo Zero).



approche basée uniquement sur la force brute, comme cela avait été le cas pour DeepBlue en 1997.

En 2017, Les chercheurs de Stanford (Google) ont publié l'article intitulé « *Attention is all you Need* », présentant une nouvelle architecture de réseaux de neurones qu'ils ont appelé « Transformer ».

Les Transformers sont des architectures de réseaux neuronaux qui servent de base à de nombreux modèles de langage de pointe, tels que les « larges modèles de langage (LLM) ». **Cet événement est précurseur du phénomène des IA génératives et marque le passage des IA prédictives aux IA génératives.**

5 L'ÈRE DES IA GÉNÉRATIVES – OU COMMENT UNE AVANCÉE SCIENTIFIQUE BIEN PACKAGÉE PEUT FAIRE LE BUZZ

C'est en novembre 2022 que OpenAI a lancé ChatGPT – l'IA générative la plus connue aujourd'hui – en la mettant dans les mains du public réussissant un coup de marketing jamais égalé. Une semaine après son lancement, le prototype avait atteint le million d'utilisateurs, et moins de

deux mois après, en janvier 2023, ChatGPT dépassait les 100 millions d'utilisateurs (par comparaison, il a fallu plus de neuf mois à TikTok, et plus de deux ans à Instagram, pour conquérir 100 millions d'utilisateurs).

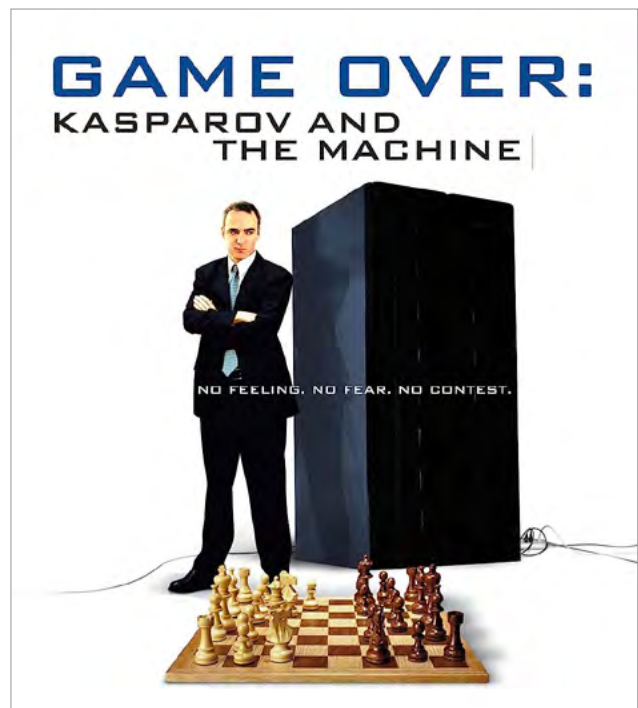
Si l'on demande à ChatGPT ce qu'est une IA générative, il répond que « c'est un sous-domaine de l'intelligence artificielle qui se concentre sur la génération de contenus ou de solutions à partir d'un modèle appris, à partir de données. L'objectif de cette IA est d'offrir une réponse, de produire du contenu ou des solutions qui sont bien écrites, convaincantes, crédibles en lien avec le contexte décrit par la requête et éventuellement avec un ton si la requête l'indique (formel, informel, professionnel, humoristique) ».

ChatGPT est construit en deux temps et est basé sur l'algorithme GPT-3 (Generative Pre-trained Transformer). Tout d'abord GPT-3 est entraîné sur une quantité massive de données (tout le contenu web disponible) jusqu'en 2021 puis est coupé du web. Dans un deuxième temps, GPT-3 est « optimisé » par des techniques appelées « apprentissage par renforcement » grâce à la rétroaction humaine, qui permet à des agents

humains d'aider à classer les résultats de la machine afin de les améliorer de sorte à pouvoir « mener une conversation » et interagir au plus juste avec les utilisateurs ; on parle de InstructGPT. Enfin, il s'ajoute à cela une interface utilisateur avec un prompt, offrant la possibilité à l'utilisateur d'interagir simplement avec le modèle.

Ces modèles se composent de deux parties bien distinctes qui fonctionnent ensemble : un encodeur et un décodeur. L'encodeur prend la séquence d'entrée, puis « l'encode » (extraction automatique des informations nécessaires pour l'analyse) puis la transmet au décodeur. Le décodeur se charge de transformer – décoder – l'information en une séquence de sortie. Dans le cas d'une traduction de langue, cette séquence de sortie sera dans la langue souhaitée.

Le mécanisme d'attention, décrit dans l'article de 2017 de Google, correspond à la capacité de l'algorithme de se concentrer sur les mots « importants ». Ce mécanisme est inspiré de la façon dont « nous portons notre attention » à certains mots dans une phrase pour fournir un contexte d'ensemble, alors que nous lisons les mots un par un. Ainsi – un transformateur peut prêter



Affiche du film *Game Over: Kasparov and the Machine* réalisé par Vikram Jayanti (2003).

« attention » aux parties de la phrase qui sont importantes et seront des termes clés qui donnent le contexte de la phrase, permettant donc de finir une phrase ou de traduire des phrases en gardant le sens et le contexte. Les transformateurs et leur mécanisme d'attention fonctionnent efficacement avec du texte et sont utiles dans les tâches NLP, mais comment cela fonctionne-t-il pour les modèles génératifs dans le cas d'autres données, comme les images ou l'audio ?

On pourrait penser que les IA génératives d'image comme DALL-E utilisent des réseaux antagonistes génératifs (GAN) de 2014. En effet, les GANs étaient autrefois, la meilleure méthode pour créer des images à partir d'une description ou d'une image.

Cette architecture de réseau de neurones propose deux réseaux d'algorithmes qui s'affrontent, l'un pour créer une image la plus réaliste possible, l'autre pour détecter si l'image est artificielle ou non. Ils s'entraînent l'un l'autre jusqu'à arriver à une image pour laquelle il est difficile de savoir si elle a été créée par une IA ou non. Ce type de réseaux antagonistes s'avère très compétent lorsque le concept à illustrer est bien défini, cadré et unique comme « générer une image de chat

à partir d'un chat ». C'est pourquoi les GANs sont préférentiellement utilisés pour générer des deepfakes. Or, entraîner un GAN pour qu'il puisse générer n'importe quelle image est très difficile et très coûteux en temps de calcul.

C'est pourquoi l'algorithme de DALL-E 1 d'OpenAI est une version dite multimodale de GPT-3 qui peut générer une image à partir d'une commande textuelle. Bien que révolutionnaires, les images générées ne sont pas de qualité optimale.

Depuis avril 2022, une nouvelle version de DALL-E est disponible : DALL-E 2 qui génère des images de bien plus haute résolution et qui est bien moins coûteux en temps de calcul pour son entraînement. Cette amélioration notable est possible car l'algorithme a également « appris la relation entre les images et le texte utilisé pour les décrire dans un processus également connu sous le nom de diffusion ». Ces algorithmes de diffusion, excellent dans l'échantillonnage par « débruitage ».

Essayons de voir cela simplement. À partir d'une image floue, un algorithme peut reconstituer une meilleure résolution, et le même mécanisme peut être fait dans l'autre sens, en apprenant aux algorithmes à

dégrader une image. Ainsi une image complètement dégradée ressemblant à un ensemble de pixels de toutes les couleurs, comme un nuage de points, peut être reconstituée via une amélioration de l'image : le « débruitage ». Une fois que l'algorithme sait faire ce travail dans les deux sens, il est capable de reconstituer n'importe quelle image nette et précise à partir d'un nuage de pixel aléatoire. Mais, à partir de ce nuage de pixel, il existe une grande quantité d'images finales possibles, il faut donc donner quelques indications à l'algorithme sur ce qu'il est censé reconstituer.

C'est là que le texte que l'on insère dans le prompt, entre en jeu. Ce texte est analysé par un algorithme dit de plongement, qui va associer la demande à un ou plusieurs champs sémantiques. Il faut voir ce type d'algorithmes comme des versions évoluées des algorithmes de programmation du langage (NLP). Ainsi, à partir d'une image pleine de bruit l'algorithme de diffusion « débruite » progressivement l'image, en suivant le conditionnement imposé par l'algorithme de plongement. Petit à petit, pixel par pixel, l'algorithme génère une image dont la vraisemblance tend vers la demande imposée. En proposant plusieurs résultats à l'utilisateur, l'IA peut apprendre, en



Cette avancée scientifique [L'IA] va changer les usages dans de très nombreux domaines. Elle est peut-être ce que le tracteur fut à la charrue à bœuf ou encore la calculatrice fut aux mathématiques, un outil qui démultiplie les capacités.



partant du principe que l'image retenue par l'utilisateur est celle qui représente le mieux sa demande. Ainsi, les résultats peuvent devenir de plus en plus fins et précis.

La combinaison Transformer/attention/diffusion est vraiment redoutable. Et ce n'est que le début de cette nouvelle ère !

6 L'HISTOIRE RESTE ENCORE À ÉCRIRE GRÂCE À L'IA GÉNÉRATIVE (2022- ?)

L'IA générative représente un potentiel énorme en termes de croissance, mais pose aussi de nombreuses questions, qu'elles soient éthiques ou juridiques. Quoi qu'il en soit, l'Histoire de l'IA reste encore à écrire. Ces nouvelles technologies sont naissantes, et encore imparfaites mais elles progressent vite.

Cette avancée scientifique va changer les usages dans de très nombreux domaines. Elle est peut-être ce que le tracteur fut à la charrue à bœuf ou encore la calculatrice fut aux mathématiques, un outil qui démultiplie les capacités. Des sociétés vont intégrer les modèles d'IA générative dans des offres et des applications qui faciliteront la vie quotidienne personnelle et en entreprise. **La question n'est plus**

de savoir si demain, nous utiliserons ou pas les IA génératives mais plutôt comment nous préparer au mieux face à l'IA générative.

En effet, il est un simple raccourci de penser que si vous dotez les individus de meilleurs outils, si vous les aidez à penser mieux, plus vite, vous leur permettez d'être capables de faire plus, d'étendre leurs capacités vers des découvertes et des progrès toujours plus étourdissants.

Le point d'attention primordial concerne les protections à mettre en place pour éviter toute dérive éthique ou législative suite à un usage non maîtrisé de ces IA. Qu'en est-il de l'utilisation des sources associées à des droits d'auteur ? Comment garantir de ne pas propager de désinformation ? La boîte de Pandore des questions de ce type est maintenant ouverte, mais qui est le plus adéquat pour y répondre ? ■



Ces nouvelles technologies sont naissantes, et encore imparfaites mais elles progressent vite.





SI L'ON ADOPTE
CHATGPT ET
SES AVATARS
AUJOURD'HUI, ON
SE DÉSOLERA DANS
QUELQUES ANNÉES
QUE LA MOITIÉ DE
L'HUMANITÉ SOIT
DEVENUE PLATISTE.



*La faillite épistémologique de ChatGPT,
chronique de Gaspard Koenig dans Les Echos*



LE BIG BANG

DE L'IA GÉNÉRATIVE

Le glossaire

L'IA générative

INTELLIGENCE ARTIFICIELLE (IA)

Domaine de l'informatique qui comprend des systèmes capables d'accomplir des tâches qui nécessitent normalement une intelligence humaine (compréhension du langage, reconnaissance de formes et d'images, prise de décisions ou encore résolution de problèmes).

INTELLIGENCE ARTIFICIELLE GÉNÉRATIVE

Systèmes et algorithmes d'intelligence artificielle capables de créer du contenu. L'IA générative utilise des algorithmes tels que les réseaux antagonistes génératifs (GAN), ou plus récemment les Transformers (pour comprendre et générer du texte), ou encore les modèles de diffusion (pour la génération d'images) afin de créer de nouvelles données, qui sont indiscernables des données réelles. Elle peut créer de nouvelles images, de la musique, du code informatique, et même des vidéos. Grâce à sa parfaite compréhension du langage naturel, elle peut être un assistant pour de nombreuses tâches : analyse, traduction, détection d'entités nommées, raisonnement (à ne pas confondre avec le raisonnement humain)...

INTELLIGENCE ARTIFICIELLE PRÉDICTIVE

Utilise des algorithmes et des modèles de machine learning pour prévoir des résultats futurs ou inconnus en se basant sur des données historiques et existantes. Elle peut prédire les tendances du marché, le comportement des clients, ou la probabilité de panne d'une machine dans une usine.

LLM (LARGE LANGUAGE MODEL)

Modèle d'apprentissage automatique formé sur de grandes quantités de données et qui utilise l'apprentissage supervisé pour générer du texte de manière cohérente et significative, dans un contexte donné.

LANGAGE NATUREL

Langage que nous utilisons tous les jours pour parler aux autres, et que les IA peuvent comprendre. Il s'oppose au langage informatique.

GAN (GENERATIVE ADVERSARIAL NETWORK)

Le réseau antagoniste génératif est un type de modèle d'apprentissage profond où deux réseaux neuronaux, le générateur et le discriminateur, sont mis en concurrence. Le générateur tente de créer des données fausses, indiscernables des vraies, tandis que le discriminateur tente de distinguer les vraies données des fausses. À travers cette compétition, les deux modèles s'améliorent, aboutissant à la production de données fausses très réalistes par le générateur.

MODÈLES DE DIFFUSION

Ils sont généralement utilisés pour générer de nouvelles données à partir de données existantes. Ils fonctionnent en simulant un processus de diffusion inverse. On commence par du bruit aléatoire, puis on le façonne progressivement jusqu'à obtenir une donnée qui ressemble à celle que l'on veut générer. Ils sont souvent utilisés pour la génération d'images ou d'autres types de données visuelles, bien qu'ils puissent théoriquement être utilisés pour tout type de données.

MODÈLES TRANSFORMERS

Ces modèles sont principalement utilisés pour le traitement du langage naturel. Ils sont basés sur l'architecture « Transformer », qui utilise des mécanismes d'attention pour comprendre le contexte et la signification des mots dans un texte. Les modèles transformers ont été révolutionnaires dans de nombreux domaines de l'IA, notamment la traduction automatique, la génération de texte (par exemple ChatGPT), et bien d'autres.

PROMPT

Commande textuelle pour « donner une instruction » à l'IA afin d'obtenir une image ou un texte.

PROMPT ENGINEERING

Technique d'amélioration des commandes que nous donnons à l'IA pour obtenir des résultats de meilleure qualité.

SPEECH TO TEXT (STT)

Technologie qui convertit l'audio parlé en texte écrit.

TEXT TO SPEECH (TTS)

Inverse du « speech-to-text ». Technologie de synthèse vocale qui convertit le texte écrit en parole.

TEXT TO TEXT (TTT)

Transforme un texte source en un autre texte.

TEXT TO IMAGE (TTI)

Désigne des algorithmes capables de générer une image à partir d'un texte (idem avec des vidéos, du son, de la musique...).

APPRENTISSAGE AUTOMATIQUE (MACHINE LEARNING)

Sous-domaine de l'intelligence artificielle composé d'algorithmes mathématiques qui permettent de déterminer des informations pertinentes à partir de données observations.

APPRENTISSAGE PROFOND (DEEP LEARNING)

Méthode d'apprentissage automatique qui utilise les réseaux de neurones pour permettre à l'IA d'apprendre de grandes quantités de données et d'effectuer des tâches très complexes.

APPRENTISSAGE PAR RENFORCEMENT (REINFORCEMENT LEARNING)

Apprentissage de l'IA en essayant différentes choses et en se basant sur les conséquences de ses actions pour s'améliorer.

APPRENTISSAGE MULTIMODAL

Apprentissage de l'IA à partir de plusieurs types de données en même temps, comme des images et du texte. Elle peut faire des liens entre ces mêmes données.

BIAIS

Inclinaison (non intentionnelle) d'un modèle d'apprentissage automatique, souvent causée par des données d'entraînement non représentatives ou préjudiciables.

NLP (NATURAL LANGUAGE PROCESSING)

Traitement automatique du Langage naturel. Technologie qui permet aux robots de comprendre et de produire du langage humain.

AI FATIGUE

Sentiment de saturation face à l'engouement et aux discussions récentes autour de l'IA. ■



© Adobe Stock EvgeniiAnd

La capacité de raisonnement des réseaux de neurones actuels ne dépasse pas la compétence de ce bébé.

3 QUESTIONS À EVAN SHAPIRO



©EBU - Media Summit 2023

CARTOGRAPHIE DES MÉDIAS

Entretien réalisé par Kati Bremme

Evan Shapiro est un producteur de télévision, de films et de podcasts, lauréat des prix Emmy et Peabody. Il enseigne les médias à l'Université Fordham et à NYU. Sa cartographie de l'évolution des médias l'a rendu célèbre et l'a érigé comme le cartographe officiel de l'univers des médias.

1

L'IA RÉVOLUTIONNE DEPUIS UN CERTAIN TEMPS TOUTE LA CHAÎNE DE VALEUR DES MÉDIAS, AVEC LA DÉTECTION DES TENDANCES, LA PRÉDICTION DES SCHÉMAS ET LA PERSONNALISATION PARMI SES PRINCIPAUX AVANTAGES. QU'EST-CE QUI DISTINGUE L'IA GÉNÉRATIVE DANS LE PAYSAGE ACTUEL ? S'AGIT-IL SIMPLEMENT D'UNE ÉVOLUTION PLUTÔT QUE D'UNE NOUVELLE RÉVOLUTION ?

C'est une évolution. L'investissement de 10 milliards de dollars de Microsoft dans OpenAI a suscité une vague d'attention et de buzz qui a propulsé l'IA sous les feux de la rampe et a permis d'actualiser la conversation publique sur l'intelligence artificielle avec les innovations en cours. Cela a également été un catalyseur pour la guerre de l'IA en termes de recherche et dans les autres domaines technologiques. L'IA est à la fois constante et éphémère. Elle s'est répandue dans nos vies depuis des années et c'est aussi le sujet brûlant de l'année. Il se dit tellement de choses sur l'IA en ce moment qu'elle est assimilable à la frénésie des NFT de l'année dernière. Mais contrairement aux NFT, l'IA n'est

pas éphémère. Si vous voulez comprendre son véritable impact, les menaces et les opportunités, vous devez séparer le signal du bruit.

2

LES MÉDIAS TRADITIONNELS LUTTENT CONTRE LES GÉANTS DE LA TECH COMME AMAZON ET APPLE, QUI DEVIENNENT PEU À PEU DES ENTREPRISES MÉDIATIQUES, ET CONTRE LES MÉDIAS SOCIAUX QUI PERMETTENT UNE DISTRIBUTION DE CONTENU PRATIQUÉMENT GRATUITE POUR TOUS. ÉTANT DONNÉ LES CAPACITÉS DE L'IA GÉNÉRATIVE À PERMETTRE À CHACUN DE CRÉER SES PROPRES HISTOIRES, Y AURA-T-IL ENCORE BESOIN DES MÉDIAS TRADITIONNELS ?

Comme beaucoup de technologies, les appareils photo haut de gamme sur nos téléphones, les logiciels d'édition sur TikTok, l'IA générative servira d'accélérateur à la démocratisation de la création et de la distribution de contenu. Elle aidera les annonceurs plus petits à créer des publicités qui rivaliseront avec les grandes marques à gros budget. Elle aidera les créateurs en début de carrière à produire du contenu qui nécessitait traditionnellement une équipe nombreuse ou une multitude d'animateurs. Elle aidera à mettre l'économie de la communauté – ce que la plupart appellent l'économie des créateurs – sur un pied d'égalité avec l'économie des GAFA. Cependant, l'IA ne peut pas remplacer et ne remplacera pas les artistes. L'IA ne reste pas éveillée la nuit en se préoccupant de sa place dans le monde. C'est notre imper-

fection humaine qui fait de nous des artistes. Nos défauts nous permettent de créer de grandes œuvres d'art. D'autre part, il n'y a plus de distinction entre les médias des géants technologiques et les médias traditionnels. Les grandes entreprises technologiques rencontrent les mêmes problèmes que les médias classiques. Netflix est maintenant simplement une chaîne de télévision. Je ne vois pas très bien pourquoi nous nous donnons la peine de continuer à faire la différence.

3

SI VOUS DEVIEZ CARTOGRAPHIER L'AVENIR DE L'UNIVERS MÉDIATIQUE DANS CINQ ANS, À QUOI RESSEMBLERAIT-IL À LA SUITE DE LA RÉVOLUTION DE L'IA GÉNÉRATIVE ? QUI VOYEZ-VOUS COMME LES GAGNANTS ET LES PERDANTS ?

Il ressemblerait exactement à ce qu'il est maintenant si je le cartographiais, comme il l'aurait été avant que les gens commencent à parler de « la révolution de l'IA ». L'apprentissage machine (qui est différent de l'IA mais souvent confondu) alimente les activités publicitaires de Google, Amazon, Meta, TikTok, Tencent et Alibaba depuis des années. C'est leur avantage concurrentiel par rapport aux supports publicitaires traditionnels. « L'algorithme » est la force

motrice de notre culture mondiale depuis 2012. La technologie, dont l'IA est partie prenante, façonne nos médias bien avant que Microsoft investisse dans OpenAI. Elle continuera à le faire. La map des médias, dans cinq ans, sera toujours façonnée par une série de consolidations par le sillage des Big Tech. Cela ne changera pas avec la soudaine popularité de l'IA.

QUESTION BONUS

DANS LA SECTION « ASK EVAN EVERYTHING » DE SUBSTACK, LES RÉPONSES SONT-ELLES GÉNÉRÉES PAR VOUS OU PAR UNE IA ?

Les réponses sont générées à 100 % par un être humain.



La technologie, dont l'IA est partie prenante, façonne nos médias bien avant que Microsoft investisse dans OpenAI.

PETITE CARTOGRAPHIE

DES CAS D'USAGES DE L'IA GÉNÉRATIVE DANS LES MÉDIAS

L'intelligence artificielle générative est en train de révolutionner le secteur des médias. Ces technologies ne sont pas nouvelles. Ce qui a changé : elles sont désormais accessibles au grand public, et elles sont plus performantes (cf. notre historique de l'IA), leur potentiel d'application dans une industrie du contenu est donc amplifié. Le rythme des changements atteint une vitesse incroyable. Cette année, nous avons assisté à la naissance de l'IA dans la vidéo, les effets spéciaux, les effets sonores, la musique et maintenant les jeux vidéo. Les LLM comprennent notre langage mieux que nous, mais ce n'est pas pour autant qu'ils sont « magiques ». Une IA qui ne connaît pas un nom propre, ne va pas l'inventer et ne saura donc pas l'écrire. Elle ne saura pas non plus aller chercher cette information, sauf si l'on lui donne accès à Internet et à un corpus de textes structuré (même si elle sait désormais le structurer toute seule). ChatGPT ne peut d'ailleurs plus aller sur Internet (l'accès vient d'être coupé).

Une partie de ces applications vantées de toutes parts est toujours en accès restreint : Imagen, le modèle d'images de Google est réservé à une poignée de testeurs, Bard n'est pas disponible en Europe, et Microsoft Copilot est annoncé depuis plusieurs

Par Kati Bremme, Louise Faudeux et Myriam Hammad, MediaLab de l'Information de France Télévisions

Chaque jour, des centaines de nouvelles applications boostées à l'IA générative sont promues sur les réseaux sociaux. Émerveillée devant les promesses d'un powerpoint fabriqué en un clic, l'équipe du MediaLab a voulu en savoir plus. Nous avons testé un certain nombre de ces boîtes magiques (l'un des tests est décrit en détail dans notre reportage-photo du futur), pour séparer le discours marketing de l'utilité réelle. Ces tests continuent chaque jour, et certaines des applications testées seront peut-être déjà obsolètes quand vous lirez ces lignes. Une cartographie plus complète de nos explorations sera publiée pour la rentrée de septembre. Mais voici déjà quelques exemples de cas d'usages concrets qui trouvent leur réponse dans un, ou dans la combinaison de plusieurs outils basés sur l'IA générative.

mois, mais toujours pas disponible (et ne le sera pas pour tout le monde).

L'utilisation des IA génératives peut se faire avec différents objectifs, de l'automatisation (pour remplacer des tâches répétitives existantes) à la transformation (pour modifier des tâches et des process existants), en passant par l'augmentation (pour supporter, améliorer des tâches existantes). **L'IA générative peut permettre aux médias de collecter, de créer et de distribuer du contenu de manière beaucoup plus efficace.**

Dans la plupart des médias aujourd'hui, l'IA générative est utilisée dans le processus d'édition et de transformation d'un contenu, mais rarement pour la création ou la rédaction (sauf pour des contenus de type météo, horoscopes, résultats sportifs...). L'intégration de l'IA générative se fait de trois façons :

- en utilisant les applicatifs sur étalage, en mode B2C (la liste de quelques-uns dans cet article) – avec leurs contraintes – ChatGPT qui est entraîné sur des données qui s'arrêtent en septembre 2021, biais culturels,
- en les adaptant à son environnement via des API, des modèles Open Source, en B2B (GPT-4, Bard ou le LLM français BLOOM),



- en entraînant son propre modèle (à l'instar de la Suède qui est en train de développer un modèle pour les langues des pays nordiques, ou du média Bloomberg qui a créé BloombergGPT, avec 50 milliards de paramètres de données financières).

UNE AIDE À LA RÉDACTION

Premiers cas d'usage les plus évidents des IA génératives : utiliser les grands modèles de langage pour améliorer la rédaction de contenus. Après la correction d'orthographe et de grammaire, on pense tout de suite à la génération automatique de brouillons, la catégorisation de textes, ou encore la traduction en plusieurs langues (les LLM sont parfaitement polyglottes, c'était d'ailleurs le premier objectif des Transformers de Google). Le Yle News Lab s'en est par exemple servi pour proposer Yle Novyny en ukrainien.

L'IA générative peut écrire un paragraphe d'introduction, résumer un contenu, faire des suggestions de rédaction et de brainstorming, ou encore s'occuper de la synthétisation et du résumé de longs documents, paraphraser un contenu, combiner plusieurs documents, analyser les sentiments, générer des questions pour une interview, résumer un texte

dans un tableau ou encore classifier des entités nommées. Elle peut aussi créer des textes structurés : une liste numérotée, des titres, des intertitres, et des textes non structurés, dans un style en particulier : formel/informel/ dans le style de.../avec des émojis.

Les grands modèles de langage ont les capacités pour proposer des posts à destination des réseaux sociaux, ou pour des blogs (cf. l'exemple des deux sites gérés par l'IA créés par Ari Kouts, cuisine-generation.fr et tech-generation.fr avec l'API d'OpenAI et Stable Diffusion). Elle a même des compétences lyriques, pour structurer un texte en rimes et en mètres, ce qui peut être utile pour générer des écrits créatifs comme des poèmes et des chansons (même si le chanteur Nick Cave lui dénie toute créativité en indiquant que « avec tout l'amour et le respect du monde, cette chanson est une connerie, une moquerie grotesque de ce que c'est que d'être humain. »).

Sans aller jusqu'à la création littéraire, des briques d'IA générative intégrées dans les CMS des éditeurs peuvent améliorer le quotidien des rédacteurs de façon considérable.

Et à partir des textes générés dans ChatGPT par exemple, on peut imaginer un usage multimodal pour en faire des podcasts, des livres pour enfants ou d'autres formats créatifs. Côté programmation informatique,

elle peut générer du code, expliquer comment faire une requête http en Python, et « traduire » du code de Python à Javascript par exemple, le tout avec ChatGPT, ou pour les professionnels avec GitHub Copilot.

UNE OUTIL DE RÉFÉRENCEMENT

Dans l'océan de contenus fabriqués par les IA, cette dernière peut aussi être une aide au référencement.

Avec l'émergence de l'IA générative, la « découvrabilité » devient un enjeu majeur. Les contenus ne pourront exister que s'ils sont découverts. La question se pose donc : comment donner envie aux moteurs conversationnels de faire remonter nos contenus ? Des outils comme AiSEO se proposent de générer des contenus optimisés pour le SEO. Jasper offre plus d'une cinquantaine de templates, allant du titre de vidéo YouTube, aux meta-descriptions pour le SEO, à la génération de paragraphes. Neuraltext, de son côté, est plutôt destiné au marketing. 10Web permet d'évaluer le score SEO de son site afin d'améliorer son référencement.

UNE RECHERCHE INTELLIGENTE

L'IA facilite l'analyse journalistique, de la phase de compréhension jusqu'au



The Bestiary Chronicles, par Campfire Entertainment, une bande-dessinée créée avec Midjourney.

« **Quand Apple a dévoilé ses audiobooks narrés par une IA en janvier 2023, ce fut un choc pour la profession.** »

travail de restitution, via la combinaison de données quantitatives et qualitatives, et elle assiste les métiers de processus chronophages de gestion et classification de l'information. Pour les journalistes, l'IA générative aide à structurer le matériel non publié en protégeant les sources (on ne parle donc pas d'une utilisation B2C). Pour tous les autres contenus, elle aide aussi à indexer pour ensuite explorer les archives textuelles et vidéos. **L'IA générative peut être utilisée pour explorer des sujets complexes. Elle permet d'accéder à une grande quantité d'informations en peu de temps et de comprendre ces informations de manière précise et détaillée. Cela ouvre de nouvelles opportunités pour le journalisme d'investigation et la découverte de nouvelles informations.**

La VRT utilise un outil interne, NWSify pour détecter des tendances en analysant des Open data et de plus en plus de médias intègrent des systèmes conversationnels dans les archives vidéos en backoffice. La SVT, la NPR et d'autres utilisent un système conversationnel, couplé à un système de reconnaissance faciale, et vocale. NPR Planet Money Bot est un outil développé avec Stanford qui permet d'explorer la richesse des contenus des podcasts. La recherche se faisant de plus en

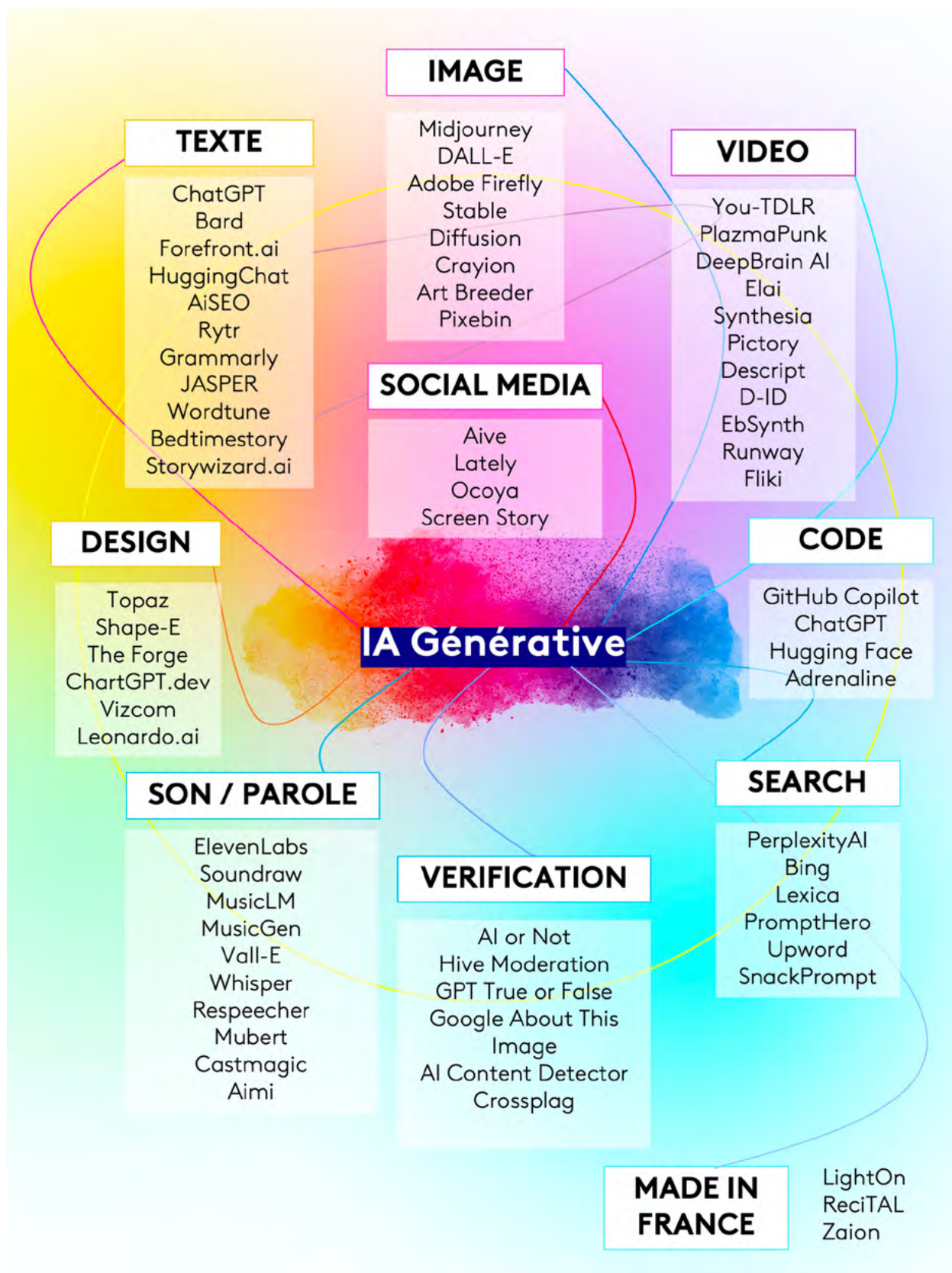
plus conversationnelle, des chatbots moteur de recherche comme Bard ou Andi répondent aux questions et proposent des analyses et en incluant des liens de redirections vers des articles, et You.com promet une expérience de recherche personnalisée tout en préservant la confidentialité des données.

DES MÉDIAS SYNTHÉTIQUES, AUDIO, VIDÉO, IMAGES

Yle, la télévision publique finlandaise, utilise depuis 2016 un robot qui les aide à écrire des articles et à fabriquer des images : Voitto écrit 100 articles et 250 images chaque semaine. Pour la fabrication automatique de vidéos, un long chemin a été parcouru depuis les réseaux antagonistes génératifs qui étaient à l'origine du film Zone Out fabriqué par l'IA Benjamin. Mais nous ne sommes pas encore arrivés au stade où il suffirait de quelques lignes de description pour créer une vidéo ex nihilo. Pas sûr non plus que nous ayons envie de voir des présentateurs virtuels fabriqués par D-ID, Synthesia ou DeepBrain sur nos antennes.

Les avancées du côté de la voix sont plus impressionnantes : quand Apple a dévoilé ses audiobooks narrés par une IA en janvier 2023, ce fut un choc

pour la profession. Nous sommes loin des voix robotiques d'il y a quelques années, l'IA est aujourd'hui capable de générer une voix à partir de quelques secondes d'écoute (Vall-E), et peut parfaitement créer des contenus audio convaincants (ElevenLabs). La NRK utilise par exemple un voice-over synthétique, et la RTS a laissé son antenne pendant une journée aux voix clonées de leurs animateurs (cf. notre article sur Couleur 3). Grâce à Whisper, un contenu audio de 1 heure est transcrit en texte en 8 minutes. De Tijd a fabriqué un podcast IA (De Aionauten) avec ChatGPT et ElevenLabs. Le script pour l'audio peut-être facilement généré par ChatGPT, ou à partir d'un article existant pour le transformer en récit vocal. On peut même générer une version audio (GPT-4, ElevenLabs, Audiogram) personnalisée selon la cible ou encore rendre accessible à tous des contenus abstraits comme de « faire écouter des couleurs ». Et ensuite, l'IA peut aider à découper et chapitrer automatiquement des contenus, et même ajouter une musique avec SoundDraw ou MusicLM.

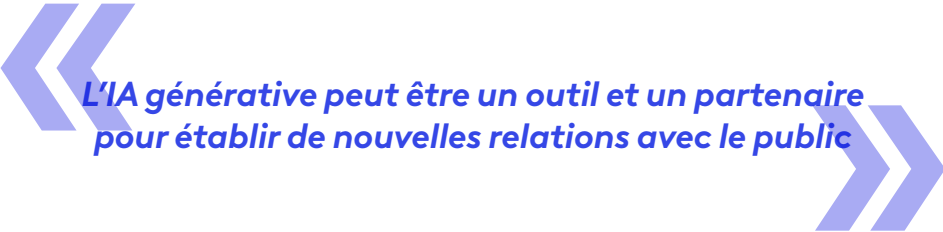


AUGMENTER LA CRÉATIVITÉ

Grâce à l'IA générative, on peut facilement proposer un contenu en plusieurs formats : texte, audio, vidéo, sous forme de quiz, ou même de bande dessinée. L'IA générative peut intervenir à différentes étapes du

processus de conception du contenu. Elle permet la structuration de storyboard, de brainstorming, l'idéation à travers différents formats. ABC fabrique des Storyboards et mockups avant les tournages avec Midjourney, Topaz, EbSynth. L'IA générative peut servir comme point de départ de créa-

tion d'images (Midjourney, Dall-E, Stable Diffusion), pour inspiration ou même comme contenu final, selon la charte éthique d'utilisation des médias. Les créatifs peuvent entraîner Stable Diffusion, l'outil de génération d'images Open Source avec leur propre style. ChartGPT.dev, un outil



L'IA générative peut être un outil et un partenaire pour établir de nouvelles relations avec le public

texte-to-graphiques permet même de construire à partir de langage naturel un graphique présentant divers indicateurs.

Les IA sont aussi de plus en plus intégrées dans les outils métiers : Adobe Firefly avec les briques de création automatique ou encore Roblox, où il suffit d'une simple commande de texte pour changer la matière d'une voiture ou pour le faire pleuvoir ou passer du jour à la nuit, par la programmation sémantique.

AIDER À DÉTECTER DES FAUX ET À LABELLISER DES VRAIS

Imaginez pouvoir suivre toute l'actualité et vérifier les faits en temps réel. **L'IA générative peut être une alliée de poids dans le processus de fact-checking, en permettant de vérifier la fiabilité des informations générées par un tiers, en aidant au décryptage des contenus collectés.** L'IA générative va rendre les réseaux sociaux radioactifs, nous n'en sommes qu'aux prémices d'une déferlante de fausses informations écrites, illustrées et bientôt en vidéos. Le côté faillible et difficilement traçable de l'IA générative oblige à accélérer et à renforcer le processus de décryptage et de vérification de l'information. Elle

est un outil performant de détection, analysant les modèles et les caractéristiques des contenus suspects. Des algorithmes d'apprentissage permettent par ailleurs de reconnaître les schémas communs utilisés dans la diffusion de la désinformation (sources, biais, viralité...).

L'IA générative peut aider à valoir les contenus validés, analyser les contenus, croiser les données à grande échelle, communautariser et référencer les validations effectuées.

À côté des Originality.ai, GPTZero et autres AI or Not, OpenAI a publié son propre outil de classification capable de détecter les textes écrits par l'IA. Mais il précise qu'il ne faut « pas trop s'y fier ». Ces modèles de détection de faux fonctionnent sur la base du principe de la perplexité (qui devine le taux de probabilité sur lequel est basé un texte), mais laissent perplexes face au nombre de faux positifs générés (du langage humain détecté comme langage IA) qui est de 10 % pour OpenAI.

PERSONNALISER LES EXPÉRIENCES

La capacité de l'IA de s'adapter à différentes échelles de compréhension et différents formats améliore aussi

l'accessibilité par la simplification du langage. **L'IA générative peut non seulement aider à produire des contenus plus efficacement, mais aussi à améliorer la manière dont ces contenus sont consommés.**

Cela peut passer par la personnalisation des recherches, l'ouverture des propositions ciblées, la structuration des référencements validés et l'amélioration de la pertinence et la rapidité des réponses. Un outil comme Aive permet de décliner rapidement des vidéos pour les différents réseaux sociaux.

Grâce à la plus grande capacité de production de contenus et un gain de productivité on pourrait imaginer des services premium. Les chatbots et les assistants virtuels alimentés par l'IA sont de plus en plus populaires, et désormais très performants. Ces assistants peuvent aider les publics à naviguer sur le site web, répondre à leurs questions et même leur suggérer des produits qu'ils pourraient être intéressés à acheter.

UNE NOUVELLE INTERACTION AVEC LES PUBLICS

L'IA générative est en passe de bouleverser l'approche actuelle de diffusion des contenus en créant des interfaces conversationnelles avec les publics. Elle s'appréhende



Label « Contenu bio » sans IA, notbyai.fyi.

à la fois comme un outil et comme un partenaire, pour trouver de nouvelles façons d'engager les publics : elle peut modérer différemment les interactions en animant des échanges grâce à des identités pilotées par l'IA générative, et synthétiser en temps réel l'ensemble des interactions entre le média et leurs publics, permettant une meilleure compréhension des besoins des publics. Un chatbot de l'information avec des contenus uniquement vérifiés provenant d'un corpus sûr pourrait aider à naviguer de façon naturelle dans les informations vérifiées d'une rédaction, tout en proposant une conversation autour de l'information (avec des résumés, un langage adapté, des liens pertinents vers d'autres contenus...). Des solutions de modération comme Bodyguard aident à maintenir le respect des chartes de modération.

elle accélère le temps de production et constitue un outil de compétitivité dans un contexte où le cycle de l'information s'accélère. **Les tâches qui auraient pris des heures sont réduites à de simples minutes, parfois des minutes à des secondes. L'IA est plus intelligente que nous, dans de nombreuses tâches de perception, et même les examens professionnels et académiques. Un de nos avantages : nous sommes généralistes. Un autre : nous avons la perception du monde.** Et aussi : l'IA générative génère des erreurs, qui doivent être vérifiées par un humain. Dans tous les cas, les médias doivent être suffisamment transparents sur l'utilisation de l'IA quand il s'agit de contenus qui sont en contact direct avec les publics. ►

L'IA générative peut changer la manière qu'ont les internautes d'accéder à l'information et d'interagir avec les médias, humanisant ainsi l'interaction avec les publics.

CONCLUSION

L'IA générative peut permettre aux médias de collecter, de créer et de distribuer du contenu d'information de manière beaucoup plus efficace,



Accédez à notre cartographie complète

UN CAS D'USAGE CONCRET

Un potentiel majeur des IA réside dans la déclinaison de contenu, pouvant se faire beaucoup plus rapidement et à moindre coût qu'auparavant. Il devient beaucoup plus simple de créer de nouveaux formats et de s'adresser de nouvelles audiences. Pour ouvrir le champ des possibles en termes de création, il est nécessaire d'identifier les fonctionnalités de chaque outil afin d'apprendre à les utiliser de concert. Prenons l'exemple d'un

extrait du rapport du GIEC, qui prendrait plus d'une vingtaine de minutes à lire. Ce contenu peut, en adaptant le contenu et le contenant, s'adresser à de nouveaux types d'audiences. Voici trois déclinaisons possibles grâce à l'usage des IA génératives.

Un extrait du rapport du GIEC : « Widespread and rapid changes in the atmosphere, ocean, cryosphere and biosphere have occurred.

Human-caused climate change is already affecting many weather and climate extremes in every region across the globe. This has led to widespread adverse impacts and related losses and damages to nature and people (high confidence). Vulnerable communities who have historically contributed the least to current climate change are disproportionately affected (high confidence). »

CRÉER UNE SÉRIE DE 4 PODCASTS POUR LES ADOLESCENTS

01. Demander à ChatGPT d'adapter l'article pour le transformer en conducteur de podcast en 4 épisodes.
02. Demander à ChatGPT d'adapter le ton du conducteur pour le rendre plus attractif et accessible pour des adolescents.

03. Demander à ChatGPT de trouver un titre pour le podcast et pour les épisodes.
04. Passer les scripts dans ElevenLabs pour générer les pistes audios du podcast.
05. Passer sur Soundraw pour créer une trame musicale en fond sonore

06. Générer l'illustration du podcast avec Stockimg.ai.
07. Demander à ChatGPT de créer la description des épisodes pour les plateformes d'écoute.

CRÉER UNE VERSION VIDÉO POUR LES RÉSEAUX SOCIAUX D'UN DIFFUSEUR GERMANOPHONE

01. Demander à ChatGPT de résumer l'article en points clés, pour une présentation orale.
02. Préciser à ChatGPT d'inclure les éléments de l'articles mentionnant l'Allemagne dans l'article dans les points clés.

03. Demander à ChatGPT de traduire la présentation en allemand.
04. Passer le script dans D-ID pour générer une vidéo incarnée par un avatar virtuel, le tout en allemand.

05. Illustrer les données chiffrées avec des graphiques générés sur ChartGPT.dev.
06. Demander à Jasper de créer des descriptions pour les réseaux sociaux pour promouvoir la vidéo (en incluant le script généré par ChatGPT).

CRÉER UN CONTE ILLUSTRÉ POUR ENFANTS

01. Demander à ChatGPT de réécrire l'article comme s'il s'agissait d'un conte pour enfant.
02. Lui mentionner si certains termes, trop complexes, doivent obligatoirement être omis.

03. Demander à ChatGPT de trouver un titre pour le conte.
04. Lui demander de découper le récit en parties pour définir les sections à illustrer.

05. Générer les illustrations dans Midjourney en précisant le style d'images souhaitées (style dessiné, atmosphère etc.).
06. Ajouter le texte sur les images et créer la couverture sur Stockimg.ai.

UN EXTRAIT DU GIEC TRANSFORMÉ EN CONTE POUR ENFANT

Il était une fois, dans un merveilleux royaume, une petite planète appelée Terra. Sur Terra, tous les êtres vivants coexistaient en parfaite harmonie. Les arbres étaient grands et majestueux, les animaux jouaient gaiement, et les rivières chantaient doucement. Mais un jour, Terra commença à souffrir. Les saisons se dérèglèrent, les animaux perdirent leur habitat, et les rivières se tarirent. Les habitants de Terra étaient tristes et inquiets. Alors vint une petite abeille nommée

Bella. Bella était courageuse et déterminée à sauver Terra. Elle parcourut le royaume, réveillant la conscience des animaux et des humains. Ensemble, ils plantèrent des arbres, nettochèrent les rivières et protégèrent la nature. Grâce aux efforts de Bella et de tous les habitants de Terra, la planète commença à guérir. Les saisons retrouvèrent leur rythme, les animaux retrouvèrent leurs maisons et les rivières coulèrent à nouveau. Aujourd'hui, Terra est un endroit

magnifique où tous vivent en harmonie avec la nature. Bella et ses amis continuent de prendre soin de leur planète bien-aimée, rappelant à tous que nous sommes tous responsables de préserver la beauté et la magie de Terra. Et l'histoire de Bella et de Terra est transmise de génération en génération, pour que chacun comprenne l'importance de protéger notre merveilleuse planète.

*Le plus votre prompt est précis,
le mieux l'IA saura y répondre*

L'ART DU
PROMPT

OPTIONS DE BASE

Ton : spécifiez le ton souhaité (ex: formel, décontracté, informatif, persuasif)

Format : définissez le format ou la structure (ex: essai, point clé, plan, dialogue)

Rôle : indiquez un rôle ou une perspective à adopter (ex : expert, critique, enthousiaste)

Objectif : indiquez l'objectif ou le but de la réponse (ex: informer, persuader, divertir)

CONTEXTE ET CIBLAGE

Contexte : fournissez des informations de base, données ou contexte pour une génération de contenu plus précis

Mots-clés : énumérez les mots-clés ou les expressions importantes à inclure

Portée : définissez la portée ou l'étendue d'un sujet

Cible : mentionnez le public cible pour un contenu adapté

CADRAGE

Limitation : spécifiez les contraintes, tel que le nombre de mots ou de caractères

Exemple : fournissez des exemples de style, structure ou de contenu

Date limite : mentionnez les échéances pour les réponses urgentes (ex: dans le cadre d'un mail qui nécessite un retour)

Langue : indiquez si la langue de réponse doit différer de la langue de consigne

PERSPECTIVE

Point de vue : demandez de prendre en compte plusieurs perspectives et opinions

Contre argument : demandez de traiter les contre arguments potentiels

Terminologie : spécifiez les termes techniques à utiliser ou à éviter

Sensibilité : mentionnez les sujets sensibles à traiter avec précaution ou à éviter

AFFINAGE DE LA REQUÊTE

Statistiques : encouragez l'utilisation de données pour clarifier des concepts (attention : à vérifier)

Appel à l'action : demandez une action claire à suivre et détaillez les prochaines étapes

Citation : demandez l'inclusion de citation ou de source pour étayer l'information (attention : à vérifier)

Analogie : demandez d'utiliser des analogies ou exemples pour clarifier des concepts

L'IMPORTANCE DU PROMPT

Les IA sont loin d'être intelligentes. Elles n'ont pas conscience du contexte et ne peuvent donc pas deviner quelles sont vos intentions en termes de création de contenu. C'est pourquoi, à l'image d'une instruction que vous donneriez à un assistant (un peu bête), la description précise du résultat que vous souhaitez obtenir est importante.

Une requête aussi triviale que : « Créer une newsletter pour promouvoir mon nouveau format vidéo » sera beaucoup plus efficace si elle devient :

« Vous êtes un spécialiste du marketing des médias chargé de créer une newsletter pour promouvoir le lancement d'une nouvelle verticale vidéo. Votre objectif est de capter l'attention de vos abonnés et de les convaincre de visionner le format. Réfléchissez

aux moyens les plus efficaces pour présenter le format. Inclure un titre accrocheur, une introduction, un synopsis ainsi qu'un appel à l'action encourageant les lecteurs à se rendre sur le site pour visionner. Pensez au ton et au style de la newsletter, ainsi qu'au public cible et à ses intérêts qui sont inclus dans ce cahier des charges : [DESCRIPTION DU CAHIER DES CHARGES] »

DE L'APPROCHE RÉALISTE DE L'IA GÉNÉRATIVE **DANS LES RÉDACTIONS**

La récente vague, ou plutôt avalanche, d'annonces de nouvelles variantes de l'IA générative a provoqué à la fois un enthousiasme et une crainte démesurés.

ChatGPT-3.5 d'OpenAI semblait changer la donne, mais la nouvelle version, GPT-4, représente un pas en avant supplémentaire, tout comme la révélation par Google de l'arrivée de nouveaux outils d'IA générative. GPT-4 peut désormais générer suffisamment de texte pour écrire un livre, coder dans tous les langages informatiques et fait encore plus remarquable, il « comprend » les images.

Si vous n'êtes pas stupéfait par le potentiel de ces outils, c'est que vous n'avez pas été attentif. J'ai passé les cinq dernières années à étudier comment l'intelligence artificielle changeait le journalisme dans le monde entier. J'ai vu comment elle peut donner un élan considérable aux médias d'information pour collecter, créer et distribuer du contenu de manière beaucoup plus efficace. C'est déjà la « prochaine vague » de l'évolution technologique. Maintenant, l'IA générative a fait progresser le potentiel de progrès d'un ou deux crans supplémentaires.

Mais attendez. **Ce n'est pas une percée vers une IA douée de conscience. Les robots ne viennent**

Par Charlie Beckett, directeur du Journalism AI Project et professeur de la LSE

Il a passé ces cinq dernières années à étudier comment l'intelligence artificielle transformait la pratique du journalisme. Charlie Beckett est directeur du Journalism AI Project et professeur au département des médias et de la communication de la LSE (London School of Economics). Il donne sa vision d'ensemble sur le futur des intelligences artificielles.

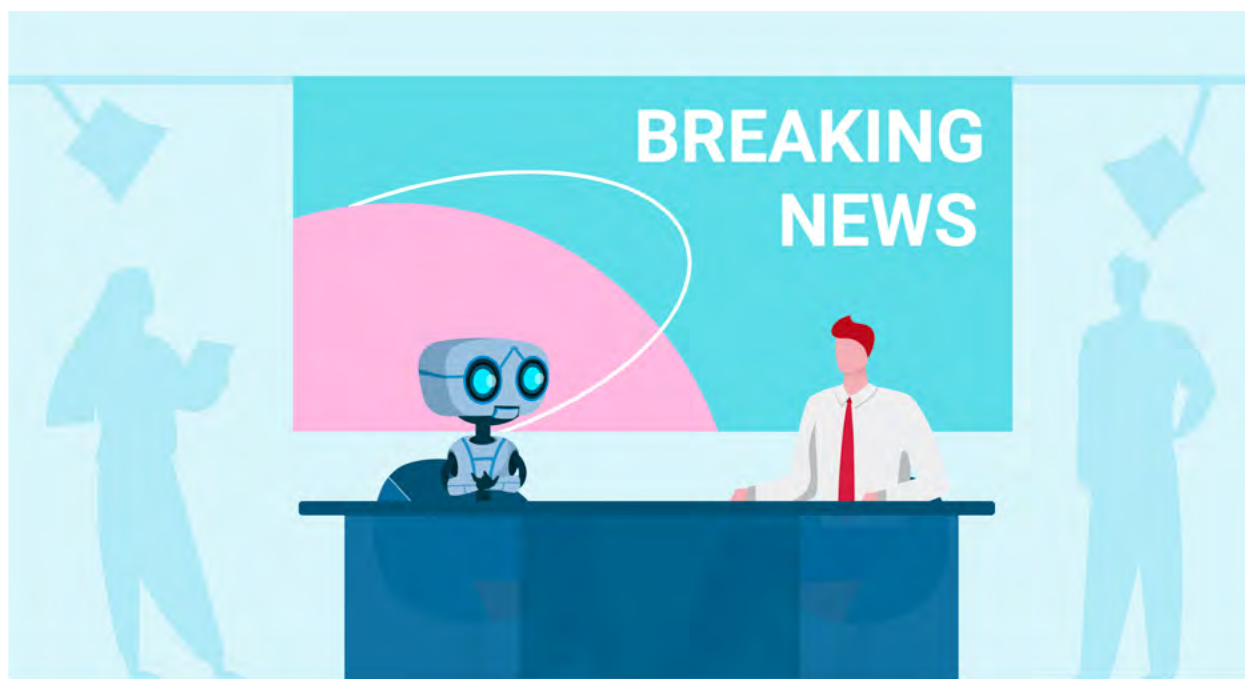
pas nous remplacer. Cependant, ces modèles de langage de grande envergure sont un catalyseur qui fonctionne à une échelle et à une vitesse telles qu'ils semblent pouvoir effectuer tout ce que vous leur demandez de faire. Et, plus nous les utilisons et les nourrissons de données et de questions, plus ils apprennent rapidement à prédire les résultats.

Un million de start-ups prétendent

déjà utiliser cette « sauce secrète » pour créer de nouveaux produits qui révolutionneront tout, de l'administration juridique aux transactions boursières, en passant par les jeux vidéo et le diagnostic médical. Une grande partie relève d'une stratégie marketing. **Comme pour toutes les percées technologiques, il y a toujours un cycle d'enthousiasme et des conséquences bonnes et mauvaises inattendues.** Mais j'ai vu suffisamment de choses pour savoir que cela va changer nos vies. Imaginez ce que ces outils pourraient faire lorsqu'ils sont utilisés par des personnes créatives dans la mode ou l'architecture, par exemple.

L'intelligence artificielle telle que l'apprentissage automatique, l'automatisation ou le traitement du langage naturel fait déjà partie de notre monde. Par exemple, lorsque vous effectuez une recherche en ligne, vous utilisez des algorithmes basés sur l'apprentissage automatique, entraînés sur d'immenses ensembles de données pour vous fournir ce que vous recherchez. **Maintenant, le rythme du changement s'accé-**

« **J'ai vu suffisamment de choses [sur l'IA] pour savoir que cela va changer nos vies.** » »



lère. Rien qu'en 2021, les investissements privés mondiaux dans l'IA ont doublé, et je m'attends à ce que les avancées de l'IA générative doublent encore cette croissance.

Mon principal espoir est que nous prenions le temps et les efforts nécessaires pour réfléchir attentivement aux meilleures façons de l'utiliser de manière positive.

Prenez maintenant une pause. Je ne recommande à personne d'utiliser ChatGPT (3.5 ou 4) pour créer quoi que ce soit en ce moment. Du moins, pas quelque chose qui sera utilisé sans une vérification humaine pour s'assurer qu'il est précis, fiable, efficace et qu'il ne cause aucun dommage. **L'IA ne vise pas à automatiser complètement la production de contenu du début à la fin : elle vise à fournir aux professionnels et aux créatifs des outils d'augmentation pour travailler plus rapidement**, en leur permettant de consacrer plus de temps à ce que les humains font de mieux.

Nous savons qu'il y a de réels risques supplémentaires liés à l'utilisation de l'IA générative. Elle peut avoir des « hallucinations » où elle invente des choses. Parfois, elle crée du contenu nuisible. Et elle sera certainement employée pour propager la désinformation ou violer la vie privée. Des personnes l'ont déjà utilisée pour créer de nouvelles façons de pirater des ordinateurs, par exemple. Vous pourriez vouloir l'utiliser pour créer un merveilleux nouveau jeu vidéo, mais

que se passerait-il si un super-vilain l'utilisait pour créer un virus mortel ?

Nous connaissons ces risques car nous pouvons voir ses défauts lorsque nous testons les prototypes mis à disposition par les entreprises technologiques. Vous pouvez vous amuser en le faisant écrire des poèmes, des chansons ou créer des images surréalistes. Posez-lui une question directe et vous obtiendrez généralement une réponse sensée et sûre. Posez-lui une question stupide ou complexe et il aura du mal. De nombreux experts en technologie et journalistes se sont amusés à le tester jusqu'à l'épuisement et à le faire réagir de manière étrange et troublante. Les génies de l'IA en seront ravis, car tout cela contribue à affiner leur programmation. Ils mènent leurs expérimentations en partie en public.

Nous connaissons également les risques car OpenAI, par exemple, les a répertoriés sur sa « carte système » qui explique les nouveaux pouvoirs et dangers de cette technologie et comment ils ont cherché à les atténuer à

chaque nouvelle itération. Qui décide quels sont les risques acceptables ou ce que nous devons faire à leur sujet, est une question discutable.

Il est trop tard pour remettre cette technologie dans sa boîte. Elle a trop de potentiel pour aider les êtres humains à relever les défis mondiaux auxquels nous sommes confrontés. **Il est essentiel que nous ayons un débat ouvert sur l'impact éthique, économique, politique et social de toutes les formes d'IA.** J'espère que nos politiciens s'informeront rapidement sur cette technologie émergente, mieux qu'ils ne l'ont fait par le passé. Et que nous devenions tous plus compétents en matière d'IA. Mais en fin de compte, mon principal espoir est que nous prenions le temps et les efforts nécessaires pour réfléchir attentivement aux meilleures façons de l'utiliser de manière positive. Vous n'avez pas besoin d'adhérer à la hype autour du sujet pour avoir un peu d'espoir. ■

POURQUOI LES ALGORITHMES
ÉDITORIAUX SONT L'AVENIR DU

JOUR- NALISME MODERNE

L'IMPORTANCE DES CRITÈRES ÉDITORIAUX

D'un point de vue de l'apprentissage automatique, le journalisme est « simplement » une détection d'événements appliquée à des données en constante évolution. Prenons un exemple hypothétique où il y a beaucoup de données fiables : le sport. Si nous devons surveiller tous les matchs de football professionnels à travers le monde qui se déroulent pendant un week-end donné, comment déterminer quels sont les matchs les plus intéressants sur le plan journalistique ?

Cette décision est traditionnellement prise par un journaliste en se basant sur sa connaissance du domaine et le contexte du sujet. Ils peuvent mettre l'accent sur les matchs qui ont eu le plus grand impact, se concentrer sur les équipes les mieux classées ou accorder plus d'attention aux équipes où jouent des athlètes très

*Par Francesco Marconi,
professeur de journalisme à
l'Université Columbia à
New York*

Professeur de journalisme à l'Université Columbia à New York et auteur de l'ouvrage *Newsmakers. Artificial Intelligence and the Future of Journalism*, Francesco Marconi explore le futur des algorithmes éditoriaux, dans le milieu des médias. À l'avenir, il sera possible de trouver et de recueillir des informations par le biais de machines à une vitesse inimaginable aujourd'hui, soutient-il.

bien payés, parmi de nombreuses autres considérations. **En ce sens, le caractère digne d'intérêt journalistique est déclenché lorsque des critères éditoriaux prédéfinis sont satisfaits.**

LA LOGIQUE DES ALGORITHMES AU SERVICES DES ÉDITEURS

Maintenant, supposons que la logique de ces principes puisse être consignée et traduite en algorithmes. Qui décide de ce qui est digne d'intérêt journalistique et de ce qui ne l'est pas ? Ce n'est certainement pas la machine. **Ce sont plutôt des éditeurs humains qui décident des poids et des paramètres à appliquer aux modèles d'apprentissage automatique. On les appelle les « algorithmes éditoriaux ».**

Au-delà du sport, seuls quelques autres domaines de couverture tels que la finance ou la météorologie sont des domaines où les données sont abondantes et, pour la plupart, objectives. Il en va autrement pour des sujets tels que la politique, la santé, la science, les événements mondiaux, entre autres, du moins pour le moment.



LE RÔLE CLÉ DES JOURNALISTES À L'AVENIR

À mesure que tout autour de nous est déplacé vers le domaine numérique, de nouvelles traces de données et des enregistrements sont créés, ce qui peut permettre de surveiller en permanence des domaines qui ne sont actuellement pas quantifiables. Le traitement automatique du langage naturel est ici crucial. **Dans la prochaine décennie, nous serons en mesure de collecter et de rassembler des informations via des machines à une vitesse aujourd'hui inimaginable. Cela ne signifie pas que l'intelligence artificielle remplace les journalistes humains, bien au contraire.**

Le développement d'algorithmes éditoriaux nécessitera des humains aux commandes (et pour être clair, un journaliste n'a pas besoin d'être technicien pour définir des paramètres algorithmiques fiables, mais il doit au

moins comprendre le fonctionnement de l'apprentissage automatique). Dans cette nouvelle réalité, la transparence algorithmique deviendra l'une des fonctions les plus cruciales des salles de rédaction modernes. ■

Dans cette nouvelle réalité, la transparence algorithmique deviendra l'une des fonctions les plus cruciales des salles de rédaction modernes.

COULEUR 3, UNE RADIO LIVRÉE AUX MANETTES DE L'IA PENDANT 24 H

VOTRE CHAÎNE SUISSE COULEUR 3 A REMPLACÉ SES ANIMATEURS PAR UNE INTELLIGENCE ARTIFICIELLE. POURQUOI AVOIR LANCÉ CETTE OPÉRATION SPÉCIALE ? QUEL EN ÉTAIT L'ENJEU ?

L'arrivée de l'intelligence artificielle générative questionne les médias. Notre initiative a pris forme en salle de rédaction, il y a quelques mois. En tant que radio assez jeune, publique, notre démarche se voulait expérimentale. Plutôt que d'en parler pendant des heures, on a préféré explorer concrètement le potentiel de « la machine ». Pendant une journée entière, l'antenne a donc été livrée aux manettes des intelligences artificielles, sous notre supervision. Un travail technique et de réflexion de longue haleine, assumé par une dizaine de personnes, de la programmation musicale à l'éditorial, en passant par la technique !

COMMENT LE PROCESSUS A-T-IL PRIS FORME ?

Les voix ne devaient pas être absentes ou neutres, à l'image de celles utilisées dans les transports publics. L'objectif consistait à imiter le style de nos animateurs, connus et appréciés

*Propos recueillis par
Alexandra Klinnik, MediaLab
de l'Information de
France Télévisions*

Trois mois de préparation, une dizaine de solutions d'intelligence artificielle testées : la chaîne suisse Couleur 3 a lancé le 27 avril 2023 une initiative radio-phonique inédite. De 6 h à 19 h, la radio a diffusé chroniques et musiques, générées en majeure partie par l'intelligence artificielle. Sur l'antenne, les cinq animateurs ont été remplacés par des clones vocaux, imitant à la fois ton, expressions et humour. Une opération destinée à mieux saisir à la fois pour les médias et le public, les capacités concrètes de l'IA. Antoine Multone, chef d'antenne à Couleur 3, revient sur les enjeux et les enseignements de cette expérience.

par le public. Pour atteindre ce but, il a fallu comprendre le fonctionnement de l'IA et la nourrir de la personnalité des animateurs. L'équipe

a ainsi fourni des renseignements sur la façon de parler, l'humour, les tics de langage, les références, les thèmes appréciés des animateurs. On a également soumis des chroniques rédigées par les présentateurs pour en faire des réécritures. C'est un processus que l'on a dû affiner jusqu'à ce que la solution d'intelligence artificielle intègre tous les paramètres. Résultat ? Des blagues attendues et plates, un humour décevant, mais un travail très honnête autour des informations contractuelles, c'est-à-dire l'annonce de la programmation musicale ou de l'heure qu'il est. Si on attend davantage de style, le résultat est nettement moins passionnant.

QUELLES TECHNOLOGIES ONT ÉTÉ UTILISÉES ? SUR QUELS CRITÈRES ONT-ELLES ÉTÉ SÉLECTIONNÉES ?

Pour l'écriture des textes, l'équipe a travaillé avec plusieurs IA diffé-





rentes : ChatGPT, Claude+, Bing IA et Poe IA. Une autre IA a généré la programmation musicale avec deux cas de figure : une musique entièrement réalisée par l'IA ou l'accompagnement de l'IA sur une « vraie » voix de chanteur. Enfin, et c'était quand même la clé pour nous, on a travaillé avec *Respeecher*, une start-up ukrainienne, spécialisée dans le clonage vocal. L'entreprise, habituée à travailler avec le secteur du cinéma, utilise des enregistrements d'archives et un algorithme d'intelligence artificielle pour cloner des voix. La société a par exemple recréé la voix du personnage de Dark Vador, pour la série *Obi-Wan Kenobi*. Le rendu est très impressionnant.

QUEL EST L'INVESTISSEMENT D'UNE TELLE OPÉRATION ?

L'initiative n'est pas du tout rentable. On ne la réitérera pas. C'est beaucoup plus cher que ce que nous coûte habituellement une journée normale. Générer des voix, créer des clones vocaux, concevoir avec l'IA l'habillage sonore : la charge de travail reste très lourde et supérieure à la normale. Par ailleurs, *Couleur 3* a déboursé quelques milliers d'euros pour travailler avec *Respeecher*.



Résultat ?
Des blagues attendues et plates, un humour décevant, mais un travail très honnête autour des informations contractuelles.



COMMENT ONT RÉAGI LES ANIMATEURS QUI SE SONT FAITS CLONER ?

L'expérience a été intense. Ils étaient d'abord secoués face à la précision du clonage. Dans un second temps, ils ont fini par réaliser que la voix présentée manquait sensiblement d'âme, de personnalité. Certaines expressions sonnaient faux. Aujourd'hui, ces clones vocaux ont été détruits : il s'agissait d'un engagement pris avec les animateurs et animatrices.

QUELS ONT ÉTÉ LES RETOURS DES AUDITEURS ?

Les auditeurs ont été bluffés, mais ils restent attachés à l'âme de leurs animateurs. La plupart nous ont dit : « Rendez-nous l'humain ». De nombreux auditeurs ont ainsi souligné l'artificialité du robot. Ils écoutent la radio pour avoir un compagnon de

route, qui les accompagne, un ami. C'est une des raisons principales pour lesquelles on écoute la radio. Je suis ravi de ce retour puisqu'il légitime notre travail et le sens que l'on donne à ce métier. Le robot est agrégateur de savoir humain : il n'est pas créatif. Pour l'heure, il se contente de copier le style. En revanche, faire appel à l'IA sur le côté technique va pouvoir nous aider sur de nombreux aspects : cela va par exemple nous permettre d'améliorer l'archivage des sons, une tâche laborieuse, de travailler sur une radio plus personnalisée. Cette expérience nous a permis non seulement de nous questionner en tant que média, mais aussi d'impliquer concrètement les auditeurs sur l'arrivée de l'IA générative. Il s'agit d'une démarche collective. ■

IA & JOURNALISME :

LA CONCURRENCE DES INTELLIGENCES ?

En mai 2023, un sondage mené par l'Association mondiale des éditeurs de journaux (WAN-IFRA) révèle que la moitié des salles de rédaction travaillent activement avec des outils d'intelligence artificielle générative tels que ChatGPT. Environ 49 % des personnes interrogées ont répondu oui à la question : « Votre salle de rédaction travaille-t-elle activement avec des outils d'IA générative comme ChatGPT ? ». **Ce même sondage révèle également que seules 20 % des personnes interrogées travaillant dans des organisations de presse disposent de lignes directrices sur l'utilisation de l'IA.**

Objectivement, ces modèles qui tirent parti des connaissances pré-entraînées pour générer ce contenu sont des assistants de rédaction et de production graphique innovants et efficaces. Alors, que craint-on dans les salles de rédaction ?

DE L'INFO EN QUANTITÉ INDUSTRIELLE

Avec plus 100 millions de visiteurs uniques en janvier 2023, ChatGPT est l'outil qui a conquis le plus d'utilisateurs, en si peu de temps. TikTok a mis neuf mois pour arriver au cap symbolique des 100 millions d'utili-

*Par Clara Schmelck,
journaliste, philosophe,
chargée de cours à
Sciences Po Strasbourg*

Les services d'IA générative de langage et d'image bouleversent le journalisme à plusieurs échelles, de la salle de rédaction au web. Les effets de cette concurrence des intelligences sont déjà visibles.

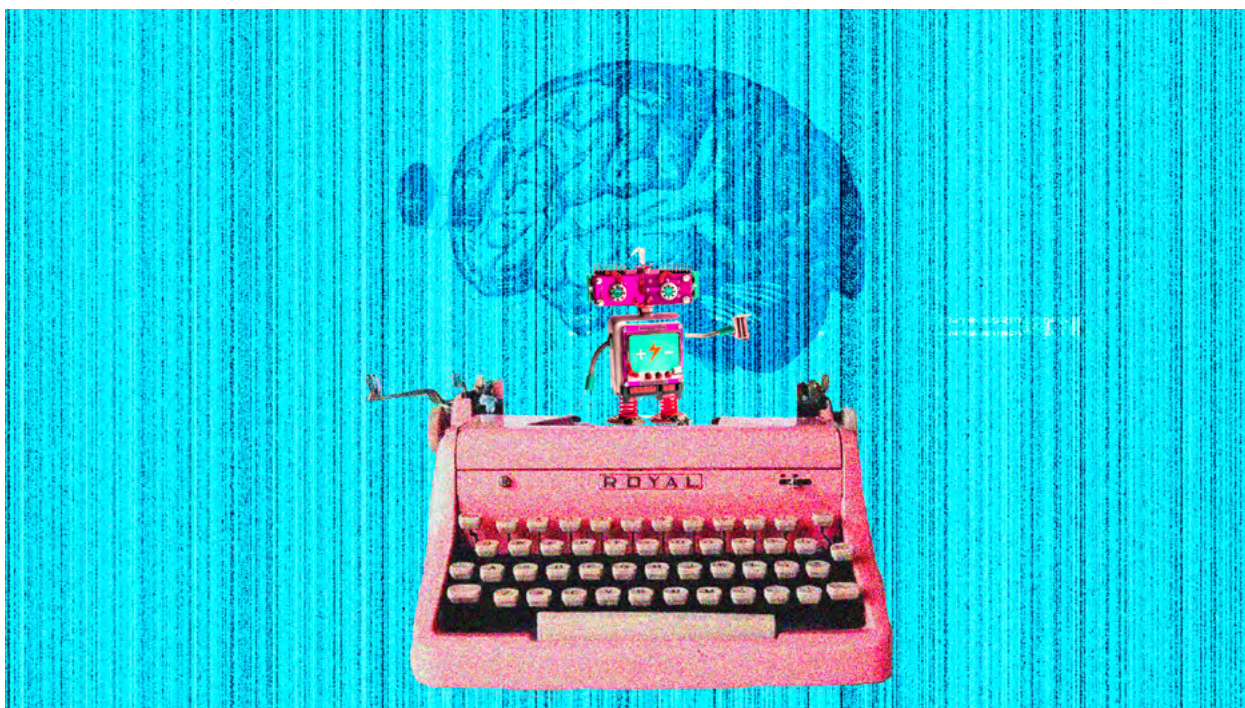
sateurs et Instagram, 2 ans et demi. Néanmoins, les IA génératives ne sont pas une révolution. Ce qui a changé en 2023, c'est la façon dont on y accède. OpenAI ou Midjourney ont ainsi popularisé les IA génératives en concevant une interface graphique grand public. Google a ouvert, à son tour, sa plateforme d'IA générative accessible à tous. Ainsi, il est désormais possible d'utiliser l'IA générative pour créer des images, du texte ou de l'audio. **Le développement des IA génératives a rapidement provoqué la production autonome d'une quantité industrielle de contenus d'information.**

Les cellules dévolues au fact-checking dans les grands médias (à l'instar de l'AFP) sont désormais ame-

nées à poursuivre une démarche de triple vérification. Jusqu'à présent, la démarche du fact-checking a consisté, non seulement à interroger la valeur de vérité d'une source (texte, photographie, séquence vidéo), mais aussi à interroger le contexte d'énonciation des énoncés ou de ces visuels, et les intentions de l'individu ou l'organe qui était à l'origine de ces contenus. Deux genres de vérités étaient interrogés : la vérité de faits (une information est ou bien vraie ou bien fausse) et la vérité d'énonciation : qui met en circulation telle information et dans quelle intention.

À présent, il va s'agir également de s'intéresser à la nature même de l'énoncé produit, en deçà de sa fonction. Ce substrat textuel ou visuel est-il une fabrication syntaxique et stylistique humaine (un mensonge, une exagération, une litote, un non-dit ; un plan large ou un plan resserré) ou une fabrication artificielle (générée par une IA) ? **Dans ce contexte, les journalistes seront-ils réduits à n'être plus que les vérificateurs de ChatGPT, Midjourney ou autre DALL-E ?**

En parallèle, les médias, dans le sillage du média américain *Wired*, se dotent de chartes éthiques visant à encadrer les usages des IA géné-



ratives. Sur-vérifiée, l'information professionnelle de qualité sera également de plus en plus labellisée, autrement dit, intellectuellement sécurisée.

LE WEB DES MÉDIAS S'EST FISSURÉ

Mais, les médias qui n'auront pas les moyens financiers et humains de déployer systématiquement une vérification et une labellisation de leurs contenus éditoriaux risquent logiquement d'être marginalisés. D'ailleurs, les moteurs de recherche référencent déjà moins les liens vers les pages web. Les « petits » médias, qui n'ont pas de relation solide et directe avec une audience fidèle, et qui ne vivaient que grâce à Google, sont déjà en train de disparaître.

On assiste probablement à un phénomène de concentration des médias aussi important qu'il y a une ving-

taine d'années, lorsque le primat du web sur le print a condamné un grand nombre de médias imprimés à rayonnement modeste, à la disparition ou à la survie via des campagnes de financement participatif.

Aujourd'hui, ce sont les médias du web qui vivent ce que les médias imprimés ont subi il y a vingt ans. Un média comme *Vice* a déjà fait les frais de cette inéluctable casure. **Le modèle économique basé sur l'audience large et la publicité se fissure : les médias gratuits en ligne peinent à être rentables dans un environnement massivement occupé par des textes générés automatiquement.** La presse d'information professionnelle est désormais une sous-préfecture d'un Internet mondial que l'IA sature de sites de spam, d'articles produits en masse à cadence infernale par des « rédacteurs IA », de déchets générés par l'IA sur Reddit, Wikipédia...

POLLUTION TECHNOLOGIQUE SUR LE DÉBAT DÉMOCRATIQUE

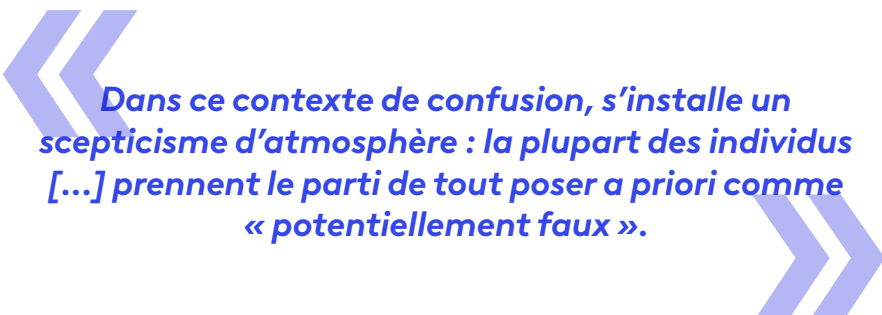
Depuis qu'un volume croissant de données, d'énoncés, de visuels et d'URLs sont mises en circulation sur les réseaux en ligne, il devient de ce fait plus difficile d'identifier rapidement les fausses informations que l'on a l'occasion de voir sur nos terminaux.

Étant donné la vitesse de propagation de ces articles, de ces visuels, de ces vidéos ou de ces URLs, **il est à craindre que s'installent dans l'opinion, des polémiques ou des certitudes erronées qui auraient pour point de départ des faux contenus, générés par les IA** et présentés comme vrais, intentionnellement ou non, par les humains qui les traitent. Mouvements migratoires, dérèglement climatique, santé...l'agenda des parlements et des gouvernements va se trouver pollué par les alluvions d'informations fausses provenant du transport de contenus IA-générés.

Aujourd'hui, ce sont les médias du web qui vivent ce que les médias imprimés ont subi il y a vingt ans.

RÉGRESSION INTELLECTUELLE COLLECTIVE

De plus, l'uniformisation des contenus, du fait du développement glo-



Dans ce contexte de confusion, s'installe un scepticisme d'atmosphère : la plupart des individus [...] prennent le parti de tout poser a priori comme « potentiellement faux ».

balement peu régulé des services d'IA génératives, présente le risque d'exacerber les comportements de rejet face à l'information. Dans ce contexte de confusion, s'installe un scepticisme d'atmosphère : la plupart des individus, lassés de se demander si ce qu'ils ont l'occasion de lire ou de voir en ligne est vrai ou faux, prennent le parti de tout poser a priori comme « potentiellement faux ». En quelque sorte, les individus se dotent d'un pare-feu mental.

Dans une telle atmosphère de défiance envers les médias et les institutions, l'utilisation d'applications d'IA génératives risque de pousser les plus jeunes citoyens à se montrer de plus en plus sceptiques sur les informations portées à leur connaissance, voire à rejeter en bloc toute communication émanant d'un média professionnel ou d'une institution.

Le danger est que l'adhésion ne devienne le critère pour accorder sa confiance dans l'information : on croit parce que c'est l'association qui nous plaît qui diffuse l'info. Cette motivation affective à l'organe qui émet l'information est la marque d'une régression intellectuelle collective : nous adhérons au lieu de nous interroger.

Les années à venir verront probablement l'essor de médias militants ou de médias fortement polarisés politiquement, très à gauche ou très à droite. Au programme des lignes éditoriales : outrance verbale, exagérations, élucubrations et subjectivité surjouée des auteurs de tribunes, éditoriaux ou témoignages exploitant le registre du pathos. **Cette stratégie de positionnement répondra au désir affectif des lecteurs de suivre un média qui leur ressemble, qui les comprend, en un mot, qui est avec eux, et donc, qui dit vrai.**

NÉGOCIER LES CHANGEMENTS PLUTÔT QUE LES SUBIR

En travers de ces sombres horizons, se dressent des perspectives plus optimistes. Il est possible de négocier les changements induits par ces nouvelles technologies de la communication. De la même manière que le web du début des années 2000 a ouvert des possibilités, l'IA générative permet au plus grand nombre d'effectuer de manière plus efficace des recherches plus approfondies, dans des domaines divers.

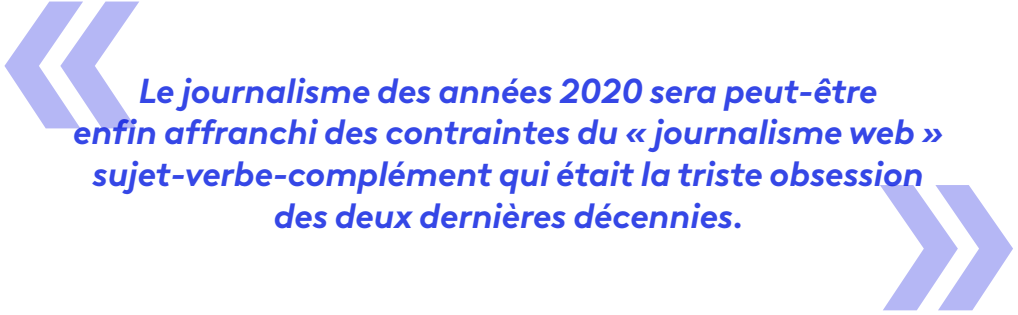
Les outils d'IA génératives facilitent les recherches en ligne, ce qui nous

stimule à explorer des sujets qui vont au-delà de nos expertises. Dans des newsrooms de plus en plus « dé-silotées », les rédacteurs auront davantage de plasticité intellectuelle et technique. **On peut imaginer des équipes plus solides, où les collaborateurs peuvent s'entraider davantage, au service de la production d'une information plus précise.**

De plus, les services d'IA générative pourront assister les journalistes dans leur travail. Il sera possible d'exploiter les technologies de l'IA pour accélérer et approfondir des enquêtes sur la partie collective de la data, par exemple.

Il sera par ailleurs envisageable pour les rédactions d'étoffer leur offre, et donc de consolider leur modèle d'affaires. *Numerama*, par exemple, a lancé en mai 2023, une newsletter dédiée aux thématiques de l'IA et logiquement rédigée par une IA. Rédigée par ChatGPT-4, cette newsletter quotidienne renvoie vers les articles de la rédaction. La rédaction de *Numerama* vérifie chaque édition pour garantir la fiabilité des informations et remanier la syntaxe.

À l'inverse, ChatGPT a une vertu méthodologique « inversée » : il agit sur les rédacteurs tel un miroir repous-



Le journalisme des années 2020 sera peut-être enfin affranchi des contraintes du « journalisme web » sujet-verbe-complément qui était la triste obsession des deux dernières décennies.

sant, les engageant à ne jamais céder aux facilités du langage utilitaire. Aux plus proches du terrain, à portée de parole, les reporters seront encouragés à libérer le langage du carcan des phrases IA-fabriquées et offrir au lecteur une information incarnée, avec style. Le journalisme des années 2020 sera peut-être enfin affranchi des contraintes du « journalisme web », sujet-verbe-complément qui était la triste obsession des deux dernières décennies.

Aussi, on peut penser que le temps dégagé par le traitement de certaines informations par les IA génératives encourage le développement de formats immersifs innovants, en particulier en ce qui concerne les documentaires. À cet égard, l'informatique spatiale est une opportunité pour les médias d'information. À condition qu'ils soient commandés par une nécessité éditoriale, les reportages immersifs, que l'on peut consulter en chaussant de simples casques, pourraient réinventer les expériences et les partages du savoir et s'adresser à des publics que les médias traditionnels peinaient jusqu'alors à approcher.

Enfin, en matière de marketing éditorial, les IA sont un instrument de

perfectionnement. Le quotidien américain *Wall Street Journal* analyse le comportement de ses lecteurs pour leur attribuer des scores en fonction de la probabilité qu'ils se transforment en abonnés. Le journal en ligne ajuste ensuite dynamiquement, au moyen de l'IA, ses propositions d'abonnement.

FACULTÉ DE JUGER

L'usage rapide de ces technologies confronte les humains que nous sommes à des questions urgentes : quid des droits d'auteur ? Les éditeurs et les créateurs de contenu bénéficieront-ils, à échelle mondiale, d'un levier pour obtenir une part des bénéfices lorsque leurs œuvres sont utilisées comme source de contenu généré par l'IA ? Quel est l'impact environnemental des IA génératives ?

Sur toutes ces questions politiques et éthiques, il ne faut pas compter sur ChatGPT. S'il est indéniablement capable de raisonner et de prévoir, il est dépourvu de faculté de jugement ; car calculer n'est pas choisir. Et, si les citoyens sont encore libres de choisir, c'est justement parce qu'il existe de vrais médias d'information, conçus et élaborés en intelligence par des équipes humaines,

dans l'espace et le temps humains, pour leur permettre de développer leur esprit critique. ■

« ALORS CE SERA SI BEAU,

LE JOURNALISME SE SERA SI
PERFECTIONNÉ QU'IL N'Y AURA
PLUS DE JOURNALISME »

LA TECHNOLOGIE OU LE PROGRÈS DU JOURNALISME

D'autres voix tranchent avec cet avenir préjugé funeste. Elles parlent de progrès. Il est vrai que leur proximité avec l'actuelle issue cauchemardesque est d'un autre ordre. **C'est en 1883 que, dans son livre intitulé *XX^e siècle, La vie électrique*, le dessinateur et écrivain Albert Robida évoque un « téléphonoscope » qui permettra à tous de communiquer par le son et l'image.** En 1889, *La journée d'un journaliste américain en 2890* parle d'un journalisme téléphonique. C'est Michel, le fils de Jules Verne qui rédige cette courte nouvelle. Elle met en scène Francis Bennett qui active son « phonotéléphote » pour converser de visu avec sa femme de l'autre côté de l'Atlantique dans un hôtel des Champs-Élysées. Puis, il s'en va inspecter la salle des journalistes de son journal le *Earth Herald*.

Et le fiston Verne de poursuivre : « Ses 1 500 reporters, placés alors devant un égal nombre de téléphones » ont reçu pendant la nuit les nouvelles des quatre coins du monde. Il précise : « L'organisation de cet incomparable service a été souvent décrite. Outre son téléphone, chaque reporter a devant lui une série de commuta-

*Par Hervé Brusini, président
du Prix Albert Londres, ancien
rédacteur en chef de
France Télévisions*

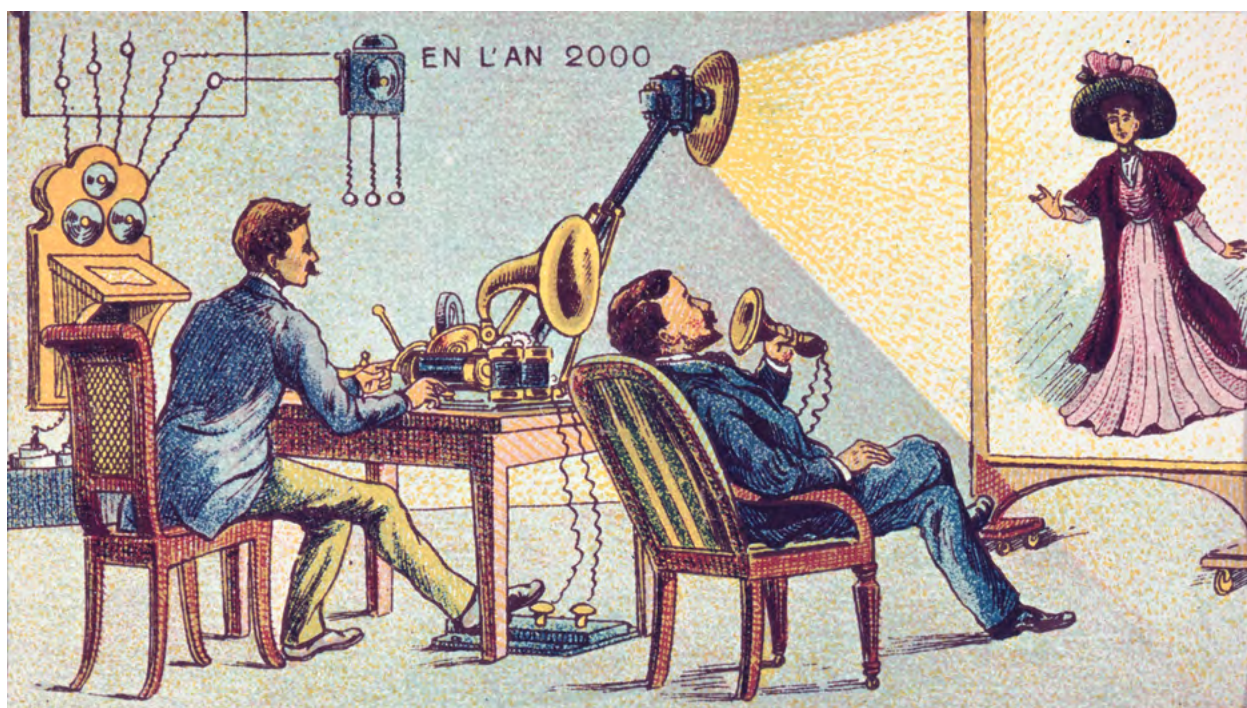
Si l'on osait la paraphrase, on dirait ceci : un redoutable dérèglement informatique est aujourd'hui constaté. Il a pour nom de domaine IA, Intelligence Artificielle. Le degré de dépendance à l'égard de cette famille de logiciels s'élève si dangereusement que d'aucuns préconisent une pause dans sa mise en œuvre. Autrement dit, l'avenir immédiat fait peur. Parmi les victimes annoncées de ce scénario aux allures de Terminator, le journalisme serait assurément l'une des premières...

teurs, permettant d'établir la communication avec telle ou telle ligne téléphonique. Les abonnés ont donc non seulement le récit, mais la vue des événements, obtenue par la photographie intensive ».

Une autre vision d'avenir saisit par sa justesse. En 1892, Eugène Dubief rédige un ouvrage intitulé *Le Journa-*

lisme. L'auteur se présente comme un « ancien secrétaire général de la direction de la presse au ministère de l'Intérieur ». Mais, gare aux réflexes conditionnés. Dubief n'a rien d'un ennemi de la presse. Tout au contraire. Il est aussi secrétaire de la Ligue Française de l'Enseignement. Et voici ce qu'il écrit, tout en vous assurant que vous n'avez pas la berlue.

Sentiment que vous pourriez éprouver à cette lecture : « Un épervier (NDLR un épervier est une sorte de filet) indicible de conduits électriques enserrera le globe. Par eux, de partout, les nouvelles afflueront au cabinet du journaliste, comme par autant de filets nerveux ; d'autres filets nerveux les transmettront au même instant chez tous les abonnés ou les emmagasineront dans leur phonographe. Puis, qui sait ! Nos neveux ayant trouvé enfin l'art de voir à distance, l'image, les gestes, le jeu des acteurs, des orateurs, des personnages célèbres suivront la même voie qui aura transmis leurs actes ou leurs paroles. Moyennant l'abonnement le plus minime, le citoyen du XX^e siècle pourra évoquer devant lui, à volonté, un diorama vivant de l'Univers et être sans cesse en communion avec tout le genre humain... Alors ce sera si beau, le journalisme se sera si perfectionné qu'il n'y aura plus de journalisme... Le téléphone et le pho-



Villemard 1910 - En L'An 2000 - Correspondance Cinéma.

nographe supplanteront le Journal». Confondant, non ? Le web avant l'heure. Tout y est ou presque : l'électricité, l'aspect planétaire, le singulier et le collectif, le temps réel et jusqu'au téléphone, pas encore smart mais presque.

L'INTELLIGENCE ARTIFICIELLE : UN DÉFI D'UNE AUTRE NATURE LANCÉ AU JOURNALISME

Pourtant, l'enthousiasme de Dubief néglige un point crucial. À ses yeux, la nouvelle existe en soi. Sa recherche, sa collecte, son traitement semblent n'avoir aucune importance puisque seul compte le moyen de transmission. De fait, cette vision où le journalisme est réduit à sa technicité de diffusion est largement répandue. Pour beaucoup, les inventions de l'imprimerie, la radio, le cinéma, la vidéo, et désormais le numérique semblent être bien souvent les seules balises de l'évolution du journalisme. Avec évidemment, toutes les atteintes des différents pouvoirs à la liberté de la presse. Mais, l'information, elle, serait un intangible des sociétés humaines, un intangible susceptible de se dissoudre dans la technique. Or, jusqu'ici, il n'en est rien. Il y eut bien un « avant » le journalisme, mais pas encore un « après », malgré les

inventions successives. De plus, pour l'heure, aucune technologie de transmission n'a tué la précédente.

Mais, cette fois, le péril serait d'une autre nature. Car, l'IA vise l'écriture même du journalisme qui existe bel et bien.

En s'appuyant sur la statistique des répétitions de mots, d'expressions, de séquences images, l'intelligence artificielle produit de l'information très copie conforme. Depuis de nombreuses années déjà, des articles sont rédigés par des robots pour rendre compte des résultats, qu'ils soient électoraux ou sportifs. Le chiffre renvoyant aux chiffres. Idéal. Et, puis, la capacité des logiciels s'est fortement sophistiquée. L'ampleur du spectre du balayage des connaissances s'est amplifiée. Les séries se sont affinées. Un pseudo réel professionnel basé sur le répétitif s'affiche à présent de plus en plus, toujours plus difficile à distinguer. Parfaitement crédible.

L'intelligence artificielle lance un défi au journalisme. Il serait finalement, comme tant d'autres activités, automatisable. L'uniformité tant reprochée aux médias apparaît alors comme une preuve supplémentaire de cette répétition.

Voici donc le journalisme renvoyé à lui-même, contraint de définir

son exigence, obligé de se considérer au-delà de ses modes de transmission. Reportage, enquête, questionnement des faits peuvent constituer quelques garanties de préservation d'une information « made in humanity ».

Retour aux fondamentaux. Sinon, comme le dit Dubief, de façon devenue terriblement grinçante avec le recul, « le journalisme se sera si bien perfectionné qu'il n'y aura plus de journalisme ».

L'intelligence artificielle lance un défi au journalisme. Il serait finalement, comme tant d'autres activités, automatisable.

ASSISTANTS VOCAUX, CHATBOTS:

QUAND L'IA POUSSE LES CRÉATEURS DANS LES RETRANCHEMENTS DU LANGAGE

À PROPOS DES CHATBOTS, COMMENT DIFFÉRENCIER CEUX QUI FONCTIONNENT PAR IA DE CEUX QUI NE SONT PAS CONCERNÉS ? QUAND PEUT-ON PARLER DE GÉNÉRATIVITÉ ?

La première chose c'est de distinguer ce que cela veut dire de parler ou de chatter en général. Quand on parle, on fait ce qui s'appelle des tours de parole, c'est ce qui crée le dialogue. C'est la meilleure façon que nous ayons trouvée pour nous comprendre. Le dialogue est une sorte d'invention et se construit autour de deux axes : un axe qui a pour but d'agir, de faire quelque chose ensemble et puis, un autre axe qui consiste à essayer d'être bien compris. On va ainsi adapter ce que l'on va dire à son interlocuteur et pour essayer d'éviter des malentendus, on va reformuler.

Le temps de parole est habité de moments de méta-communication. C'est ce que l'on observe lors de dialogues entre humains. Un chatbot fait partie de la grande famille des systèmes de dialogue humain/machine, ce qui peut être regardé comme un assemblage de briques technologiques. Tout d'abord, il y a la partie

Propos rapportés par Myriam Hammad, MediaLab de l'Information de France Télévisions

Zoé Aegerter est designer et chercheuse indépendante. Elle s'intéresse particulièrement au langage et au dialogue entre l'humain et la machine. Elle suit depuis longtemps le développement des intelligences artificielles – notamment celles qui génèrent des interfaces de langage, ou bien vocales. Elle raconte comment la conception de ces outils a évolué, et comment ces derniers pourraient venir bousculer la radio de demain.

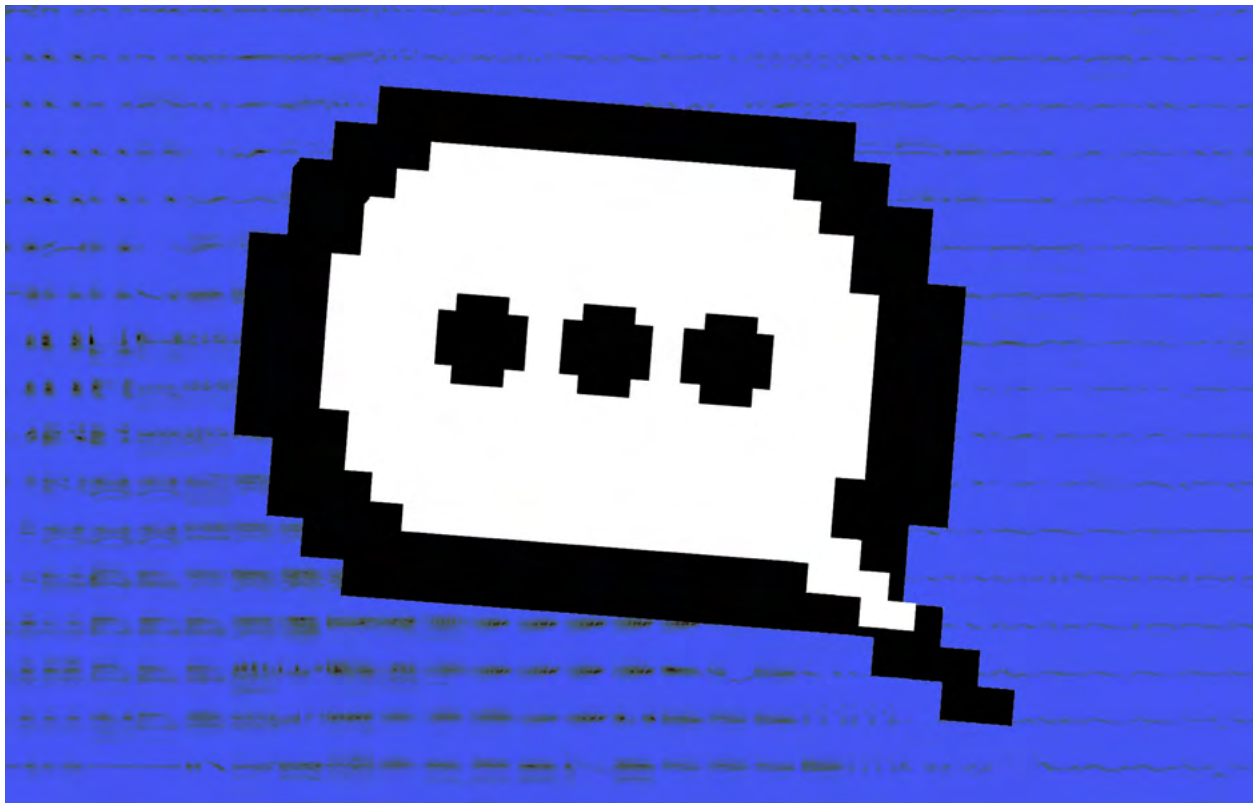
compréhension qui vise à essayer de capter quelque chose d'une intention en langage naturel avec une personne. Ensuite, il y a la partie gestion du dialogue : quelle est la suite logique, le comportement attendu dans le cadre de cet échange ?

Enfin, il y a la partie réponse : que propose-t-on comme réponse en langage ou en image ?

QUELLES SONT LES DIFFÉRENTES GÉNÉRATIONS DE BOTS ?

Il existe plusieurs générations de bots. La première ne comporte aucune compréhension du langage. On ouvre une fenêtre et on nous propose des boutons. Cela ressemble plus à un site internet avec moins de choix : il y a peu d'échanges avec la machine. Au sein de la seconde génération, il y a des mots-clés, ou une possible détection des mots-clés. Dans le meilleur des cas, ces derniers activent des arbres de décision, c'est-à-dire des parcours conversationnels conçus par les designers pour permettre à l'utilisateur de faire un certain nombre de choix. Mais la réponse repose sur un script, elle peut être agrémentée d'un lien, d'une image, quelle qu'en soit sa forme figée. Il n'y a pas d'IA utilisée, juste quelques algorithmes pour la gestion du dialogue.

La troisième génération commence à inclure une possible détection de langage naturel quand les expressions



utilisées correspondent à des façons d'exprimer l'intention. Cela nécessite une grande collecte d'expressions et d'intentions, donc il y a plus de latitude du côté utilisateur, il ne s'agit plus d'un menu à choix déroulant. L'IA est utilisée.

Dans la quatrième génération, non seulement l'utilisateur s'exprime en langage naturel, mais c'est la façon dont la personne s'exprime (les fameux *prompts*) qui va générer la réponse en langage naturel du côté de la machine. La partie compréhension est très poussée et traitée sur un plan linguistique et statistique. La réponse est générée au regard de proximités et de corrélations récurrentes au sein des données collectées sur Internet, principalement. Cela veut dire que les *prompts* font office d'éléments-clés dans la génération de la réponse, d'où le *prompt engineering*. **La grande nouveauté avec ChatGPT réside dans le côté itératif : il dialogue, il chemine avec l'humain.** Cela implique aussi un fort enjeu de découvrabilité : de quoi ce bot est-il capable ? Quelles sont les données utilisées pour l'entraîner ? Comment utiliser pleinement les capacités de cet outil ?

LORSQUE L'ON EST DESIGNER, CONCEPTEUR OU MÊME LINGUISTE, QU'EST-CE QUI CHANGE DANS LES MANIÈRES DE FAIRE LORSQUE L'ON AJOUTE DE L'IA ?

Toutes les générations avant la Gen4 impliquaient une équipe de conception composée de designers, de linguistes, mais aussi de dialoguistes, voire de psychologues. Ces personnes étaient là pour faire tout ce que l'on ne savait pas bien automatiser, c'est-à-dire pour injecter tout ce qu'il devait y avoir d'humain dans les systèmes. À partir du moment où les réponses sont gérées automatiquement et où ce sont les *prompts* qui entraînent une forme de dialogue, cela dépasse le rôle de conception.

Il a donc fallu revoir les manières de travailler. Sur un plan opérationnel d'abord, la première question qui se pose est : comment définir à quoi va

servir ce bot, quelle va être sa spécialisation ? Il s'agit aussi de travailler sur sa base de connaissance : comment va-t-on entraîner le bot ? Veut-on qu'il sache rédiger de longs discours ou avoir des échanges très brefs ?

L'autre dimension plus stratégique et moins tributaire des technologies c'est d'avoir un rôle d'anticipation avant de se lancer dans la récolte de données, essayer de tisser des scénarios d'usage fictionnels, des projections : si notre système arrive à faire ceci, qu'est-ce que pourraient faire les personnes, que pourraient-elles demander ? **On bascule alors, non plus dans l'observation de l'usage pour adapter la machine, mais dans l'anticipation du besoin utilisateur. C'est ce qui est nouveau.**

Il faut ainsi penser à tous les cas d'usage : on doit alors penser aux risques de troll, aux biais existants. Le designer devient celui qui joue ce rôle de liant avec des perspectives

« Le temps de parole est habité de moments de méta-communication. »

Y a-t-il des formes de solidarités entre citoyennes et citoyens à inventer ?

très pluridisciplinaires qui recoupent les côtés linguistique, psychologique, littéraire. Je le pense comme une sorte de traducteur entre les utilisateurs, les experts et les applications à venir. Une sorte de gardien du sens et de la cohérence.

QUELS SONT LES GRANDS ENJEUX AUTOUR DE CES NOUVELLES TECHNIQUES DE LANGAGE ?

Aujourd'hui en matière d'IA, il s'agit de l'analyse des émotions et de la modélisation des comportements affectifs. Quand il y a derrière une potentielle stratégie commerciale, cela nous invite à nous interroger. Il existe aussi un risque lié aux arnaques existantes et à venir : les systèmes conversationnels se perfectionnent de jour en jour et seront bientôt capables de façon crédible de recréer une voix de l'un de nos proches. Enfin, il y a un enjeu important de souveraineté des données qui soulève de nouvelles interrogations : doit-on se protéger nous-même ? Est-ce aux États de le faire ? Aux entreprises ? Y a-t-il des formes de solidarités entre citoyennes et citoyens à inventer ? Il y a une véritable architecture de la confiance qui reste à bâtir.

SI L'ON SE PROJETTE, COMMENT ENVISAGER LE FUTUR DE LA RADIO AVEC L'ARRIVÉE DE L'IA GÉNÉRATIVE ?

En 2019, la RTS cherchait à utiliser des assistants vocaux pour permettre aux utilisateurs d'écouter le flash infos, de se renseigner sur les informations sportives, mais également de s'en saisir pour faire revivre les archives sonores c'est-à-dire du contenu déjà diffusé et accessible aux auditeurs sous la forme de micro-podcasts. L'assistant vocal avait alors le rôle d'animateur radio : il introduisait les contenus, invitait les auditeurs à en découvrir de nouveaux. L'attention était portée sur un élément important : il ne s'agissait pas de remplacer l'animateur et l'assistant ne devait surtout pas être trop bavard. À l'avenir, on peut envisager deux pistes pour la radio grâce à l'IA, d'abord sur le transmedia. L'idée que l'on puisse très facilement passer à l'écoute d'un bout d'émission, puis avoir besoin de la lire, puis de repasser à de l'écoute – ou à la vidéo, si on regarde aussi du côté de la génération automatique de clip.

Une seconde perspective, c'est celle

de l'interview invité/machine où l'invité, en dialoguant avec la machine, réalise son propre entretien. Pour cela, il faudrait que l'IA soit capable de rentrer dans une analyse du langage encore plus poussée, qui lui permette de produire des questions allant dans le sens de l'investigation. Et cela n'existe pas encore. **Il existe en tout cas, un vrai travail d'itération avec la machine, un cheminement à construire pour trouver de nouvelles modalités d'expression.■**

Cela implique aussi un fort enjeu de découvrabilité : de quoi ce bot est-il capable ? Quelles sont les données utilisées pour l'entraîner ?

LA TRANS-PARENCE :

UN CADRE POUR SURMONTER LES CRAINTES LIÉES À L'IA DANS LES SALLES DE RÉDACTION

LE RÔLE DE L'IA GÉNÉRATIVE : OPPORTUNITÉS JOURNALISTIQUES ET PRÉCAUTIONS D'USAGE

La WAN-IFRA a mené une étude pour identifier les usages actuels de ChatGPT dans les salles de rédaction. L'une des principales utilisations (54 % des journalistes) consiste en la création de texte à partir d'un article existant ; pour le résumer ou le transformer sous forme de points clés par exemple. D'autres utilisations, comme la correction de texte (43 % d'entre eux) ou la traduction (à 32 %) sont plébiscitées par les journalistes. **Bien que ces usages puissent paraître anodins, ils représentent un large gain de productivité de quelques minutes qui se transforment rapidement en heures dans la semaine de travail du journaliste.**

Toutefois, l'étude nous informe aussi sur le fait que certains journalistes utilisent ChatGPT pour la création intégrale de texte (à 32 %) ou pour chercher de l'information (à 44 %). Or, il est important de rappeler que les IA génératives comme ChatGPT s'appuient sur des modèles dits « probabilistes », dans la mesure où ils se basent sur la probabilité qu'un mot suive un autre dans un environnement

Entretien mené par Louise Faudeux, MediaLab de l'Information de France Télévisions

Après vingt ans d'expérience en product management dans les médias, Olivier Martinez a décidé d'accompagner les entreprises dans leur stratégie de diffusion et de production de contenu grâce à l'IA. Il apporte désormais son expertise à ses clients médias pour les guider dans l'utilisation de l'IA générative. Ensemble, nous avons discuté des opportunités que ces outils représentent pour les journalistes en termes de production de format et d'efficacité des processus mais aussi ; de la nécessité de construire une relation équilibrée entre l'homme et la machine, et le besoin urgent pour les médias d'acculturer, d'encadrer et de faire preuve de transparence vis-à-vis de leurs rédactions et de leurs publics.

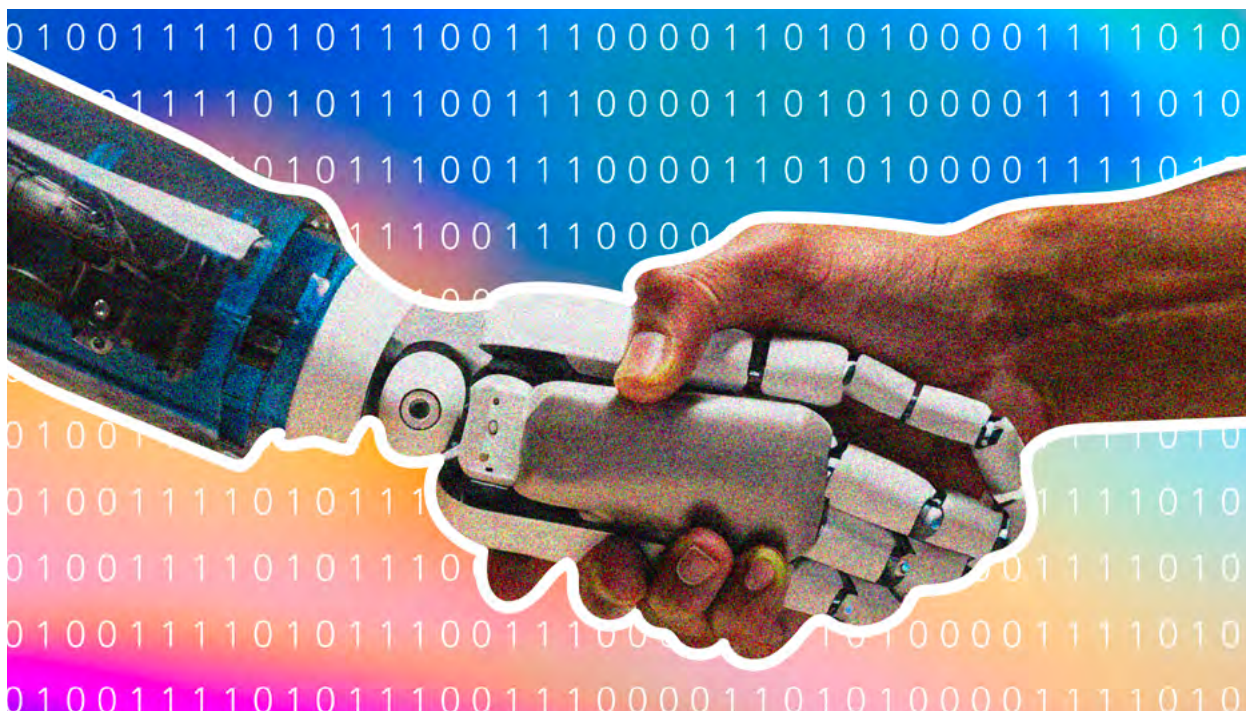
donné. L'IA conversationnelle permet alors de créer un lien entre l'humain et la machine en lui donnant des ins-

tructions en langage naturel, mais il est crucial de comprendre que ces modèles n'ont pas de connaissances intrinsèques. Aujourd'hui, 40 % des réponses de ChatGPT sont factuellement incorrectes, bien qu'OpenAI nous ait promis le contraire avec GPT-4. Ce type d'usage de ChatGPT reste donc à proscrire tant que les IA n'ont pas de compréhension du vrai et du faux.

OPTIMISER LES TÂCHES À FAIBLE VALEUR JOURNALISTIQUE

L'IA générative peut aider les journalistes dans l'accomplissement de tâches facilement automatisables, comme la mise en avant d'un contenu sur les réseaux sociaux. La génération de publications sociales directement dans le CMS de l'éditeur suite à l'intégration de son article lui permet ainsi de gagner du temps sur ce volet. L'outil peut proposer des formats prédéfinis pour LinkedIn, Twitter, Facebook par exemple, que le journaliste pourra ensuite valider avant publication.

L'IA générative peut aussi augmenter l'efficacité du SEO (référencement naturel pour les moteurs de recherche) en proposant au journaliste des titres sans dénaturer le message éditorial.



Elle peut également être utile pour le « bâtonnage » qui consiste souvent à reprendre le contenu des dépêches. En automatisant ce processus, sous supervision humaine avant publication, cela libère du temps qui peut ensuite être consacré à des enquêtes sur le terrain, à la vérification d'informations ou aux entretiens téléphoniques par exemple.

L'OPPORTUNITÉ DE LA DÉMULTIPLICATION DES FORMATS POUR LES MÉDIAS

Les IA génératives de manière générale, représentent une opportunité pour pouvoir créer des formats variés sur la base de contenu existant. Par exemple, un journaliste qui rédige un article long format, se coupe d'une audience potentielle intéressée par le sujet, mais qui n'a pas le temps ou l'envie de lire. Le journaliste pourrait alors proposer une version audio avec sa propre voix en la clonant avec des outils comme ElevenLabs. Il pourrait aussi décliner son article en vidéo incarnée par un avatar personnalisé ou générique, comme le fait *Brut* en utilisant l'outil Synthesia. **Même si la qualité n'est pas encore parfaite, on observe d'une part une évolution rapide de ces outils, et ils rendent surtout accessible – en termes de**

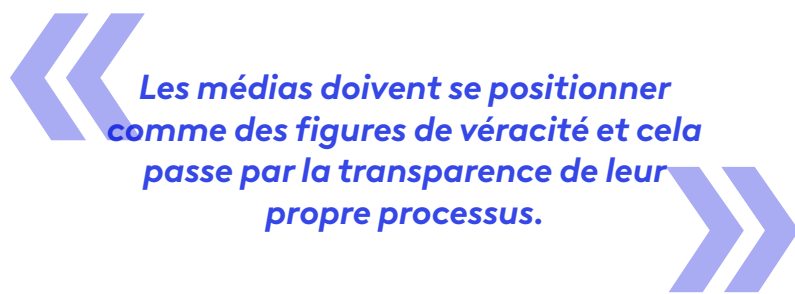
temps et de budget – des déclinaisons que les médias ne pouvaient pas nécessairement se permettre auparavant. Ces utilisations restent bien sûr à valider dans le cadre d'une charte éditoriale d'une rédaction.

Au-delà de l'usage individualisé des outils, il est surtout intéressant de les combiner ensemble afin d'optimiser les processus. Par exemple, dans le cadre d'une interview, le journaliste peut transcrire automatiquement un échange audio sur Trint pour récupérer le texte, puis utiliser ChatGPT pour en extraire les points principaux (attention, ici aussi, la vigilance reste de mise : ChatGPT ne considère pas forcément les mêmes éléments comme importants que le journaliste). Ici, on combine à la fois le *speech to text*, et le *text to text* pour synthétiser l'information. Ainsi, les salles de rédaction peuvent développer de nouveaux processus de travail flexibles, selon leurs besoins.

L'IA GÉNÉRATIVE N'EST PAS INTELLIGENTE, DU MOINS PAS SANS L'HUMAIN

L'IA générative peut être une vraie alliée pour l'optimisation de certains processus ainsi que la création de formats. Toutefois, il faut garder à l'esprit que l'IA a beau exceller sur des tâches spécifiques demandant rapidité et précision, elle est loin d'être intelligente. **On ne peut donc pas réduire le nombre de personnes dans les rédactions au profit de l'IA, car pour que cette dernière soit efficace, elle doit obligatoirement être couplée avec le travail (et l'intelligence) humain.** *Buzzfeed* ou *CNET* ont par exemple tenté d'automatiser complètement la création d'articles. On s'est vite rendu compte que les contenus créés sans supervision et sans collaboration entre l'humain et l'IA sont de piètre qualité, voire erronés. *CNET* a d'ailleurs fait savoir qu'ils revenaient sur leur processus. Pour

Il faut garder à l'esprit que l'IA a beau exceller sur des tâches spécifiques demandant rapidité et précision, elle est loin d'être intelligente.



**Les médias doivent se positionner
comme des figures de véracité et cela
passe par la transparence de leur
propre processus.**

identifier les tâches dans lesquelles l'IA peut assister les journalistes, il est primordial d'identifier la valeur ajoutée de chaque tâche – si l'on reprend l'exemple du bâtonnage, sa valeur est quasi nulle, voire aliénante – et il faut garder en tête qu'elles doivent toujours demeurer sous supervision humaine.

LE TRIPTYQUE DE L'IA DANS LES RÉDACTIONS : TESTS, FORMATIONS ET TRANSPARENCE

Trois éléments sont primordiaux pour préparer l'arrivée de l'IA dans les salles de rédaction. Premièrement, il faut que les médias sensibilisent leurs équipes à l'omniprésence des technologies. Nicholas Thompson, directeur de *The Atlantic*, a été l'un des premiers à dire, via une lettre ouverte, qu'il ne savait pas exactement quelle place allait prendre l'IA mais que les technologies étaient là et qu'il fallait les tester. Même sans objectif opérationnel à court terme, il faut essayer de prendre en main les différentes solutions d'IA générative pour se tenir au fait des usages et faire naître de nouvelles idées.

Deuxièmement, il faut former les salles de rédaction à l'IA. Ces outils vont probablement façonner notre civilisation à l'instar d'Internet dans les années 1990, les médias doivent donc s'acculturer – par exemple à la maîtrise des prompts – afin de savoir les utiliser mais aussi de les comprendre. Avec un double effet bénéfique : la perception des journalistes sur l'IA générative s'aiguïsera, et le public sera mieux informé sur ces outils.

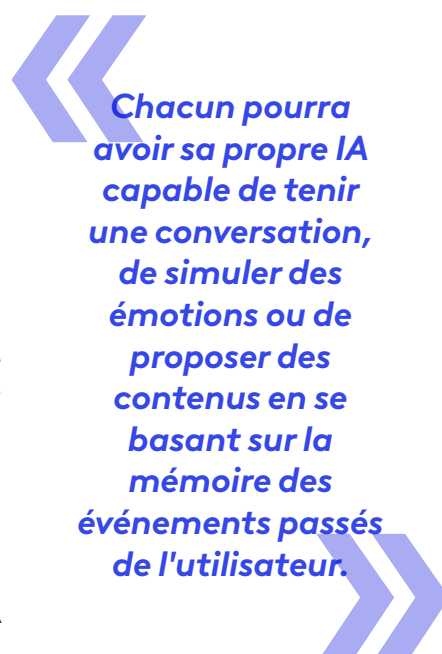
Finalement les médias doivent être transparents sur la manière dont ils souhaitent intégrer ces outils dans les salles de rédaction. Cela doit passer par la mise en place de charte d'encadrement, comme c'est le cas dans quelques médias actuellement, comme le groupe *Les Échos/Le Parisien* ou *Libération*. La charte a un double objectif : tout d'abord communiquer des directives claires pour les employés en interne sur les outils à utiliser, leurs objectifs, et les résultats attendus, qu'ils soient utilisés de manière itérative ou qu'ils répondent à un objectif économique. La charte a aussi un objectif pour l'externe en termes de crédibilité auprès des publics. Les médias doivent se positionner comme des figures de véracité et cela passe par la transparence de leur propre processus. Ainsi, on pourra établir un cadre pour surmonter les craintes liées à l'IA, notamment celle de la perception erronée de la machine comme supérieure à l'homme.

QU'A-T-ON À ATTENDRE DE L'IA GÉNÉRATIVE ?

D'une part, le futur de l'IA générative est multimodal. Cela signifie qu'un seul et même outil permettra de générer à partir d'un prompt, une déclinaison de formats, qu'il s'agisse de texte, d'image ou de vidéo. GPT-4, par exemple, a été développé pour générer du texte à partir de prompt image, même si la fonctionnalité n'est pas encore disponible.

D'autre part, plus que des IA génératives, on peut s'attendre à des IA compagnons. Il s'agit d'IA qui com-

binent des capacités génératives, conversationnelles et de mémoire pour accompagner les utilisateurs dans leurs activités professionnelles ou personnelles. Chacun pourra avoir sa propre IA capable de tenir une conversation, de simuler des émotions ou de proposer des contenus en se basant sur la mémoire des événements passés de l'utilisateur. Il existe déjà des technologies comme Replika ou Persona AI qui sont des prémices de ce type de projet, où l'on pourrait avoir une sorte de double virtuel capable de se partager le travail avec nous. ■



**Chacun pourra
avoir sa propre IA
capable de tenir
une conversation,
de simuler des
émotions ou de
proposer des
contenus en se
basant sur la
mémoire des
événements passés
de l'utilisateur.**

Q&R CHATGPT POUR UNE RÉDACTION :

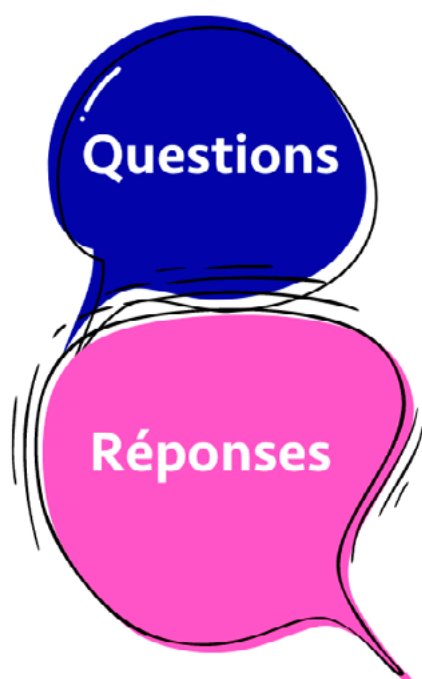
QUE SE PASSE-T-IL LORSQU'UNE
MACHINE APPREND À ÊTRE
CONVAINCANTE ?

*Par Samuli Sillanpää,
Aleksander Alafuzoff, Aki
Kekäläinen*

Le NewsLab de YLE s'est
prêté au jeu des questions/
réponses autour de l'utili-
sation de l'intelligence ar-
tificielle générative (plus
précisément ChatGPT),
ci-dessous un extrait avec
l'accord des auteurs.

NOTA BENE

Par défaut, tout texte entré
dans la boîte de dialogue de
ChatGPT peut être réutilisé
par OpenAI pour améliorer
ses systèmes et comme
matériel de formation pour
les nouveaux modèles.
Actuellement, les données
sont transférées et stockées
sur les serveurs d'OpenAI
aux États-Unis.



Q. ChatGPT peut-il m'aider à rédiger un article destiné à différents publics ?

Oui.

R.

Vous pouvez demander à ChatGPT de modifier sa réponse en mentionnant ou en décrivant un public cible. Même si vous ne mentionnez pas explicitement un public cible, la conversation précédente affectera le style des réponses de ChatGPT. Vous pouvez également demander à ChatGPT de décrire directement les besoins et les attentes d'un public cible.

Q. Que dois-je faire si la réponse de ChatGPT n'est pas satisfaisante ?

Demandez-lui de modifier la réponse.

R.

ChatGPT ayant une « mémoire » relativement longue (les mots précédents ~3000, jusqu'à 25000 mots dans GPT-4), vous pouvez travailler de manière itérative avec l'IA. La qualité de la réponse peut également être améliorée si vous donnez un exemple du type de résultat que vous recherchez. Lorsque vous modifiez la tâche, vous devez entamer une nouvelle conversation, afin que l'ancien contexte n'affecte pas les réponses ultérieures. ChatGPT peut également servir à développer de nouveaux types de contenu : à partir d'un texte donné, vous pouvez, par exemple, lui demander de créer un quiz avec 3 questions à choix multiples.

Q. ChatGPT peut-il résumer un texte ?

Oui, mais.

R.

Attention, ChatGPT ne considère peut-être pas les mêmes choses importantes que vous. En plus du résumé, vous pouvez essayer de modifier la forme du texte (par exemple, les points clés à retenir sous forme de puces) ou d'adapter le texte à un public cible spécifique (par exemple, en simplifiant le langage ou en l'expliquant à un enfant). Le résumé peut également être utilisé comme étape de prétraitement, si vous souhaitez que ChatGPT réponde sur la base d'une plus grande collection de textes.

Q. ChatGPT peut-il aider à compléter mon article ?

Cela dépend.

R.

La version actuelle de ChatGPT n'a pas été exposée à des événements postérieurs à septembre 2021 (à l'aide de plugins, GPT-4 peut effectuer des recherches sur le web, ce qui atténue le problème). La demande d'informations fonctionne mieux avec des sujets intemporels qui ne concernent pas des événements récents.

Q. Puis-je copier une réponse de ChatGPT dans mon article ?

Oui, mais attention aux droits d'auteur.

R.

Conformément aux conditions de service, OpenAI cède à l'utilisateur tous ses droits sur le contenu produit. Actuellement, légalement, les IA ne peuvent pas détenir de droits d'auteur sur le contenu généré. ChatGPT peut cependant reproduire du matériel protégé par les droits d'auteur d'un tiers. Les droits d'auteur et l'IA sont un problème très débattu mais non résolu.

Q. Puis-je faire confiance à ChatGPT ?

Pas nécessairement, car ChatGPT ne peut pas distinguer la réalité de la fiction et peut « halluciner ».

R. Les réponses de ChatGPT semblent convaincantes, même si le contenu est complètement inventé. ChatGPT peut s'appuyer sur des sources imaginaires ou non pertinentes. La détection des erreurs peut nécessiter une lecture très attentive. Plus l'IA est performante et fiable, plus l'« hallucination » est nocive : plus nous faisons confiance à l'algorithme, moins nous sommes critiques.

Q. Si ChatGPT inclut une citation dans sa réponse, puis-je faire confiance en cette citation ?

Non.

R. Les citations fournies par ChatGPT peuvent ne pas être textuelles. Au pire, les citations peuvent être complètement inventées. La probabilité d'une citation erronée augmente si la conversation précédente contient des erreurs factuelles.

Q. L'utilisation de ChatGPT est-elle gratuite ?

Pour le moment, oui.

R. La version de recherche de ChatGPT est actuellement gratuite, mais peut être restreinte en cas de pic d'utilisation. OpenAI a lancé un plan d'abonnement payant, ChatGPT Plus, pour 20 \$/mois. Il existe également une API payante destinée aux développeurs de logiciels (à partir de 0,002 \$/~300-600 mots).

Q. Est-il possible d'identifier un texte écrit par ChatGPT ?

Non.

R. Il existe une multitude d'outils en ligne (GPTZero, AI CheatCheck, AI Text Classifier) mais déjà une simple modification du texte généré par ChatGPT (par exemple en le passant dans l'outil Paraphraser.io) influe sur la détection. Le développement de ces outils est en cours. Des méthodes permettant de « watermark » le contenu généré par l'IA sont également étudiées.

Q. ChatGPT répète-t-il ce qu'il a vu dans ses données d'apprentissage ?

En partie, mais il s'en écarte également.

R. À la base, ChatGPT essaie de prédire quel mot est le plus susceptible de suivre le texte qui lui a été donné. Cette prédiction est basée sur l'hypothèse que le texte qui lui a été donné, se poursuivra comme des textes similaires dans ses données d'entraînement. Cependant, même si vous demandez à ChatGPT un texte qui faisait partie de ses données d'entraînement, il se peut qu'il ne reproduise pas le texte original à la lettre.

Q. ChatGPT est-il raciste/sexiste ?

R. **Non, mais ses bases de données d'entraînement peuvent l'avoir été.**
Le texte produit par ChatGPT reflète les biais des données d'entraînement. OpenAI semble essayer de guider les requêtes clairement biaisées vers des « réponses toutes faites ». Aux États-Unis en particulier, des critiques ont été émises concernant le caractère « woke » des réponses de ChatGPT, mais en général, il peut y avoir des biais dans n'importe quelle direction.

Q. ChatGPT remplacera-t-il les moteurs de recherche ?

R. **L'avenir nous le dira.**
Bien que ChatGPT réponde à la plupart des questions, il ne peut pas fournir de sources fiables quant à ses affirmations. Microsoft et Google testent tous deux des solutions qui combinent l'IA de type ChatGPT avec les moteurs de recherche traditionnels, mais la nature même du fonctionnement des LLM ne garantit pas un résultat 100 % fiable.

Q. ChatGPT remplacera-t-il les journalistes ?

R. **Probablement pas.**
ChatGPT ne fait pas de recherches, n'interviewe pas de sources, ne prend pas de décisions de publication et ne demande pas de comptes aux pouvoirs publics. Il peut néanmoins aider les journalistes à produire des textes de différentes manières. En général, le texte produit par ChatGPT n'est pas prêt à être publié, du moins pas dans un contexte d'information.

Q. Quels rôles un futur ChatGPT pourrait-il jouer dans le journalisme ?

R. **Plein.**
Un assistant de recherche infatigable pour trouver et trier des quantités massives de données – de nombreuses tâches deviendront plus faciles avec de nouveaux outils. Un ami patient pour tester et challenger de nouvelles idées. Un rédacteur perspicace pour réviser et éditer les textes. ■

**Les droits d'auteur et l'IA
sont un problème très débattu
mais non résolu.**

BIENVENUE DANS LA
POST-
VÉRITÉ!

INTELLIGENCES ARTIFICIELLES GÉNÉRATIVES ET DÉSINFORMATION : BIENVENUE EN **POST- VÉRITÉ ?**

*Par Kati Bremme, directrice
de l'Innovation France
Télévisions et rédactrice en
chef de Méta-Media*

Il y a dix ans, l'IA avait du mal à faire la différence entre un chat et un chien. Aujourd'hui, elle peut, non seulement classer, mais encore générer des images détaillées de chats et de chiens à partir d'un simple texte. Les modèles s'améliorent rapidement et se démocratisent. Tout le monde peut accéder à des outils en ligne et fabriquer des « faux ». Des interfaces utilisateur anthropomorphiques admirablement ergonomiques cachent un formidable potentiel malveillant. À tel point que l'on se demandera bientôt ce que veut encore dire « faux » dans un océan de contenus générés (et distribués) par des IA, où la vérité est devenue depuis longtemps « un objet de discussion », pour résumer (très grossièrement) Hannah Arendt. Face aux IA génératives et la massification de vrais et faux, nous sommes

La désinformation est aussi vieille que l'information. À mesure que l'ère numérique se développe, la désinformation et les fausses nouvelles trouvent de nouveaux canaux de distribution. On pourrait alors se contenter de dire que les récentes avancées technologiques en IA générative (rendues possibles grâce aux capacités exponentielles de calcul des machines et aux données que nous avons distribuées sur Internet sur les vingt dernières années) renforcent juste un problème connu depuis longtemps. Cette fois-ci, les choses sont différentes.

passés de l'exaltation à la peur. On joue désormais sur les émotions. Allons-nous passer directement de la bulle de filtre des réseaux sociaux au village virtuel généré par l'IA qui pollue encore un peu plus l'écosystème de l'information ? Sam Altman, le PDG d'OpenAI, reconnaît lui-même être « inquiet que ces modèles puissent être utilisés pour une désinformation à grande échelle ».

Des personnalités de la Silicon Valley, dont Elon Musk (qui veut nous vendre un TruthGPT), Steve Wozniak, cofondateur d'Apple, et Andrew Yang, ancien candidat à la présidence, ont signé une pétition visant à mettre un terme au développement d'une intelligence artificielle plus puissante que GPT-4. La pétition compte plus de 30 000 signataires et cite les risques sociétaux à grande échelle comme son objectif fondateur. **« Les systèmes dotés d'une intelligence compétitive avec celle des humains peuvent présenter des risques profonds pour la société et l'humanité, comme le montrent des recherches approfondies et comme le reconnaissent les principaux laboratoires d'intelligence artificielle ».** On ne pourra pas arrêter le progrès, mais qu'en est-il vraiment du potentiel dangereux de cette avancée technologique au service de la désinformation ?



Tweet de @elonmusk.



©KB

D'OÙ VIENNENT LES CONTENUS GÉNÉRÉS ? À LA RECHERCHE DU SENS PERDU

On pourrait résumer de façon très réduite que ChatGPT et Midjourney sont de gigantesques moteurs d'autocomplétion de tout (ou presque) notre savoir et de notre créativité mis à disposition ces vingt dernières années sur l'espace de partage Internet, restitués dans un gigantesque puzzle réutilisant les éléments les plus probables. Pour son apprentissage profond sur d'énormes corpus de texte, OpenAI se base sur les réseaux neuronaux, plus précisément un RNN, « Recursive Neural Network », qui peuvent être considérés comme des idéalizations de la manière dont nos cerveaux semblent fonctionner. On ajoute à cela la notion de « Transformer – GP » signifiant « Generative Pre-trained Transformer ».

Transformer est une architecture de réseau neuronal qui facilite l'attention parallèle et permet d'apprendre (tout seul) les dépendances à long terme. Pour la compréhension de notre langage naturel, un paragraphe est découpé en morceaux, en phrases, en mots, en vecteurs de texte. Chaque morceau est passé, plusieurs fois si besoin, par la même fonction (réursive). Ainsi, chaque morceau est iden-

tifié vis-à-vis de ses propres dépendances, mais également, vis-à-vis de l'information précédente, et donc de son contexte. Pour les images, le fonctionnement est approximativement le même, elles peuvent être découpées en séquences de vecteurs d'image, et reconstituées par contexte. On n'obtient pas une représentation directe du monde extérieur, mais on peut en capturer certains aspects.

En conséquence, le modèle est prédisposé à nous proposer une réponse basée sur des statistiques

aucune capacité à l'éviter, parce que fondamentalement, **GPT est un modèle de la façon dont les mots se rapportent les uns aux autres, et non un modèle de la façon dont le langage pourrait se rapporter au monde perçu** ».

L'adoption très rapide de ChatGPT repose sur une avancée extraordinaire dans l'UI et l'UX du traitement du langage naturel (y compris dans différentes langues, même si la maîtrise de ces langues varie en fonction de la quantité et de la qualité des

Le principal écart entre les êtres humains et GPT-4 réside dans le fait que notre pensée s'appuie sur des références. ChatGPT, quant à lui, ne tient pas compte des références, il les oublie.

et des probabilités, plutôt que sur un prétendu raisonnement. L'auteur Neil Gaiman interprète ce mode de fonctionnement de la façon suivante : « ChatGPT ne vous donne pas d'informations. Il vous donne des phrases en forme d'information ». Ou comme le formule Gary Marcus : « Il n'y a aucune intention de créer de la désinformation – mais également

données d'apprentissage disponibles pour chaque langue) et une utilité évidente. Le traitement du langage naturel et la restitution (grammaticalement parfaite) en langage naturel nous poussent à faire confiance au contenu fabriqué par l'IA parce que nous voyons qu'il a été créé pour nous à l'instant. On aurait envie de lui attribuer une intention, voire même

« Metaphysics a utilisé un deepfake pour faire jouer Elvis dans l'émission America's Got Talent. Une façon de refaire l'histoire et de brouiller les pistes entre faits et faux. »

la capacité de raisonner intelligemment. Or, il n'en est rien.

« Les affirmations selon lesquelles les outils d'IA générative produiront un contenu exact doivent être traitées avec une grande prudence. Il n'y a rien d'inhérent à la technologie qui garantisse l'exactitude des informations » a déclaré Kate Wilkinson, chef de produit chez Full Fact, une organisation de fact-checking basée à Londres, qui met en œuvre l'IA. « À ce stade, il est trop tôt pour dire exactement comment cela évoluera, mais il s'agit clairement d'un changement significatif dans l'accessibilité générale d'un ensemble d'outils aussi puissants à un groupe d'utilisateurs aussi large. »

L'UBÉRISATION DE LA CRÉATION DE CONTENUS

Les larges modèles de langage, de GPT-4 d'OpenAI à LLaMA de Meta, en passant par LaMDA (Google), Jurassic-2 (de la start-up israélienne AI21), Claude (Anthropic) ou encore ERNIE 3.0 de Baidu, sont alimentés autant par du vrai que du faux. **L'ensemble des données utilisé pour former ces IA contiendrait des milliards d'images récupérées sur Internet, des millions de livres électroniques piratés, l'intégralité des travaux de 16 années du Parlement européen et l'ensemble de Wikipédia anglais.** OpenAI ne révèle pas les sources avec lesquelles ChatGPT a été entraîné (selon ChatGPT « pour des raisons de propriété intellectuelle »). Dans une récente étude, les chercheurs Jennifer Haase et Paul Hanel ont comparé les idées générées par les humains et celles générées par l'IA dans six chatbots, et ont indiqué que : « Nous

n'avons trouvé aucune différence qualitative entre la créativité générée par l'IA et celle générée par l'homme. Google donne d'ailleurs la même note pour du contenu généré par l'IA ou par des humains à partir du moment où le contenu est de qualité.

Les imperfections sont propres à l'homme, ce qui rend ces dispositifs presque attachants en les assimilant à une manière de penser humaine et imparfaite : tout le contraire de la rigueur des calculs mathématiques. Toutefois, il est crucial de prendre en compte les usages excessifs du modèle. Le principal écart entre les êtres humains et GPT-4 réside dans le fait que notre pensée s'appuie sur des références. ChatGPT, quant à lui, ne tient pas compte des références, il les oublie.

L'automatisation de la création de contenu risque d'inonder les médias sociaux de fausses informations. « Les modèles linguistiques sont un outil naturel pour les propagandistes », a déclaré Josh Goldstein, chercheur au Centre pour la sécurité et les technologies émergentes de l'Université de Georgetown. Il est coauteur d'un article qui examine comment ces outils alimentés par l'IA pourraient être utilisés à mauvais escient dans le cadre d'opérations d'influence. Aux USA, l'IA avancée comme ChatGPT, DALL-E, et la technologie de clonage vocal suscitent déjà des craintes pour les élections de 2024. **Une ferme de trolls pourrait avoir besoin de moins de travailleurs, des campagnes de propagande à grande échelle pourraient être à la portée d'une plus grande variété d'acteurs**

malveillants, qui profiteront de l'effet Brandolini déjà bien connu des réseaux sociaux : « La quantité d'énergie nécessaire pour réfuter des sottises [...] est supérieure d'un ordre de grandeur à celle nécessaire pour les produire »...

Quand ChatGPT n'avait pas encore la capacité de naviguer sur Internet, il lui arrivait d'inventer des résumés lorsqu'on lui présentait des URLs d'articles. Il lisait l'URL, qui contient souvent le titre de l'article, et se lançait dans une dissertation sur le sujet. *NewsGuard* a publié à peine une semaine après le lancement de GPT-4 un rapport qui décrit la nouvelle intelligence artificielle d'OpenAI comme « un désastre en puissance », qui amalgame connaissance scientifique, opinions et infox sans aucune référence. Pour mesurer la capacité de l'algorithme d'intelligence artificielle à produire des infox, la start-up américaine a administré à ChatGPT des questions s'appuyant sur de faux récits répandus sur le web. Le résultat est probant comme le souligne Chine Labbé, rédactrice en chef Europe : « Dans 80 % des cas, le robot a relayé avec éloquence des affirmations fausses et trompeuses sur des sujets importants de l'actualité, dont le Covid-19, l'Ukraine et les fusillades dans des écoles aux États-Unis. [...] Pour une personne qui ne connaît pas les questions ou sujets couverts par ce contenu, les résultats pourraient facilement passer pour légitimes, voire même faire autorité ».

NewsGuard a depuis présenté un outil pour former des services d'intelligence artificielle générative pour endiguer la diffusion de fausses informations. Microsoft, qui soutient *NewsGuard*, accorde déjà des licences aux don-

nées de NewsGuard et les utilise pour BingGPT.

LA GÉNÉRATION D'IMAGES, DES CRÉATIONS HYPERRÉALISTES À CINQ DOIGTS

Une image générée par l'IA n'est pas une fake news, à partir du moment où son caractère artificiel est déclaré. En revanche, quand elle est sortie de son contexte, la situation se complique. Le générateur d'images par IA Midjourney a déjà mis fin aux essais gratuits de son service. Submergé par les demandes, l'accès gratuit au logiciel a été suspendu depuis le 28 mars (qui coïncide avec l'image du pape en doudoune).

DALL-E, l'IA générative d'images d'OpenAI, est capable de générer des images à partir de descriptions textuelles. L'outil de génération d'images (sorti au grand public avant ChatGPT) utilise également une architecture de type *Transformer*, semblable à celle des LLM, mais il est spécifiquement entraîné pour transformer des entrées textuelles en images cohérentes (et « créatives »). La frontière entre réalité et fiction s'estompe donc aussi dans les images. L'un des exemples les plus perturbants est celui des utilisateurs de Reddit qui inventent des événements qui n'ont jamais eu lieu. Les images de Justine Moore semblent réelles et sont censées provenir du grand tremblement de terre de Cascadia qui a dévasté l'Oregon en 2001. Mais, ce tremblement de terre n'a jamais eu lieu : cet événement n'existe pas. Il s'agit d'images générées par l'IA, utilisées en ligne pour diffuser l'histoire d'une catastrophe naturelle complètement fabriquée.



Midjourney en 2022 versus en 2023.

Lorsque l'on connaît le fonctionnement des réseaux sociaux qui mettent en avant des contenus sensationnels, il est facile de saisir le potentiel de désinformation détenu par des individus désespérés, en quête de notoriété.

En mars, une (fausse) image du pape François en doudoune blanche est devenue virale. BuzzFeed News l'a qualifiée de « premier cas de désinformation de masse par l'IA » et a retrouvé l'auteur de l'image, Pablo Xavier, de Chicago. Pablo Xavier a révélé qu'il était sous l'emprise de champignons hallucinogènes lorsqu'il a décidé qu'il serait amusant d'habiller le pape d'un manteau Balenciaga et de partager l'image sur le compte Reddit dédié à Midjourney. Dans ce contexte clairement défini, il n'y a pas de doute qu'il s'agissait d'un exercice divertissant. En revanche, l'image a ensuite été partagée sur d'autres réseaux sociaux hors contexte, et en a trompé plus d'un.

Mais, il y a eu récemment des exemples plus troublants. Lorsque la nouvelle de l'arrestation possible de Donald Trump a été annoncée, des images d'IA représentant son arrestation sont devenues virales. De nombreuses personnes ont cru qu'il s'agissait de vraies images, confondant fabrication et information.

Auparavant, on pouvait facilement reconnaître une image fabriquée par l'IA à l'œil nu. Aujourd'hui, l'exercice devient de plus en plus

compliqué. En réponse à la diffusion des images de la fausse arrestation de Donald Trump, Midjourney a tenté d'imposer des règles : la plateforme a exclu le journaliste britannique qui les avait élaborées.

Quelques limites existent cependant encore pour détecter un faux (mais celles-ci seront vite estompées avec l'avancée de l'entraînement, cf. l'image d'illustration de cet article). L'IA de génération d'images a les mêmes limites de compréhension du monde que celle des textes : l'IA est formée pour synthétiser les fichiers qu'elle consulte. Elle excelle dans la génération d'images semblables à celles qu'elle trouve en ligne, mais elle ne comprend pas la structure sous-jacente des objets et des personnes. Cela se traduit par des corps adoptant des positions relativement étranges. Un peu moins accessible que la génération de faux textes « dans le style de », la génération d'images nécessite souvent encore une intervention a posteriori (via Photoshop) notamment sur la partie que l'IA ne sait pas gérer sur les images : le texte.

Le réalisme bluffant de ces scènes artificielles pourra à terme discréditer les vraies photos. **Le niveau de réalisme des photos ainsi produit, induit un effet collatéral préoccupant : leur extrême viralité sur les réseaux sociaux, canaux de diffusion de prédilection de l'infox.** Les « coches bleues » de Twitter signalaient auparavant l'authenticité de

« La frontière entre la réalité et la fiction s'estompe donc aussi dans les images. L'un des exemples les plus perturbants est celui des utilisateurs de Reddit qui inventent des événements qui n'ont jamais eu lieu. »

l'auteur. Désormais, elles aident des colporteurs d'infos à paraître fiables. Les clichés inédits évoqués ci-dessus ont effectivement circulé à une vitesse fulgurante sur lesdits réseaux sociaux.

L'IA générative peut produire des images manipulées très réalistes, contribuant à la propagation de la désinformation, dénaturant les personnes, les événements ou les situations, et entraînant une perception déformée de la réalité. Ces images peuvent être utilisées pour créer de faux profils à des fins malveillantes, telles que les escroqueries et les usurpations d'identité, et l'IA générative peut perpétuer les stéréotypes et les biais présents dans les données d'entraînement.

UNE VOIX MENSONGE – LA GÉNÉRATION DE LA VOIX

La 12 avril 2023, Radio France a consacré une journée entière à l'intelligence artificielle sur son antenne *France Inter*. Entre interviews d'experts réels, on pouvait y entendre Martin Luther King ou encore Barack Obama dans des prises de parole sur l'actualité. Pourquoi alors encore s'embêter à chercher à obtenir de vrais commentaires par des personnalités publiques sur tel ou tel sujet si l'on peut les fabriquer en quelques secondes ?

Pour les créateurs de fausses vidéos et de faux sons, l'IA générative pourrait être utilisée pour créer des versions plus réalistes de personnalités politiques et culturelles connues, capables de parler d'une manière qui imite au mieux ces personnes. Il peut également être utilisé pour créer plus rapidement et à moindre coût une

armée de personnes qui n'existent pas, de faux acteurs capables de délivrer couramment des messages dans plusieurs langues. Le deepfake audio basé sur l'imitation est un moyen de transformer un discours original d'un locuteur - l'original - pour qu'il ressemble à celui d'un autre locuteur - la cible - sur la base de l'imitation.

Au début de l'année, Microsoft a dévoilé un nouveau système d'intelligence artificielle capable de recréer la voix d'une personne après l'avoir écoutée parler pendant seulement trois secondes, VALL-E.

VALL-E est un modèle de langage à codage neuronal dans lequel l'intelligence artificielle identifie les mots et utilise ses algorithmes pour construire des formes d'ondes qui ressemblent à celles du locuteur, notamment en conservant son timbre et son ton émotionnel. Cette technologie permet également de créer des assistants numériques plus personnalisés et des services de synthèse vocale et de traduction vocale à la sonorité naturelle.

Également en janvier, des internautes ont détourné le service de clonage de voix ElevenLabs pour créer de faux enregistrements audio imitant les voix de célébrités tenant des propos racistes, homophobes et violents. Dans l'un de ces montages, Emma Watson, actrice vedette de la saga Harry Potter, lit des extraits de *Mein Kampf* d'Adolf Hitler. Dans un autre, Ben Shapiro, célèbre animateur radio américain connu pour ses positions très conservatrices, fait des remarques racistes sur Alexandria Ocasio-Cortez, élue du parti démocrate. D'autres deepfakes audio impliquant les réalisateurs

Quentin Tarantino et George Lucas ou l'animateur américain Joe Rogan ont été repérés sur le site 4chan.

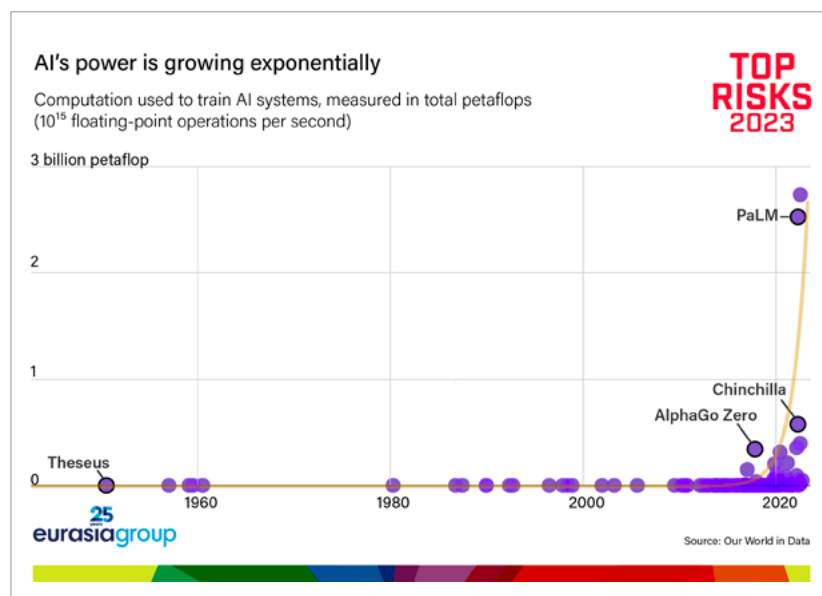
L'IA peut désormais reproduire la voix de n'importe qui. Le journaliste de *Vice*, Joseph Cox, a utilisé une technologie d'IA semblable à ElevenLabs pour accéder à un compte bancaire grâce à une version de sa propre voix reproduite par l'IA. L'audio généré par l'IA peut imiter des voix, ce qui peut entraîner des fraudes, des usurpations d'identité et de la désinformation, qui manipulent les auditeurs. **Des escrocs utilisent déjà l'intelligence artificielle pour se faire passer pour des membres de la famille en détresse. Au-delà de la désinformation, le potentiel criminel de l'IA générative vocale est infini.**

LA GÉNÉRATION DE VIDÉOS, BIENTÔT DES DEEPFAKES INDUSTRIALISÉS ?

Grâce à l'IA générative, dans un futur proche, nous pourrions apprécier des films modernes mettant en scène Marilyn Monroe et assister à un concert des Beatles dans notre propre salon. L'IA permet à Andy Warhol et Anthony Bordain de parler d'outre-tombe, promet à Tom Hanks de rester jeune pour toujours et nous permet de regarder des imitations de Kim Kardashian, Jay-Z et Greta Thunberg se disputer au sujet de l'entretien du jardin dans une comédie télévisée britannique absurde. Pour donner un exemple de ce qui est déjà possible, en septembre 2022, la société Metaphysics (à l'origine de @deep-tomcruise) a utilisé un deepfake pour

faire jouer Elvis dans l'émission *America's Got Talent*. Une façon de refaire l'histoire et de brouiller les pistes entre faits et faux. Des dizaines de start-ups utilisent l'IA générative pour créer des personnes virtuelles rayonnantes et heureuses, à des fins ludiques et lucratives. Les grands modèles de langage comme GPT ajoutent une nouvelle dimension complexe. Fin 2022, la société Graphika a observé des cas limités de *Spamouflage*, une opération d'influence (OI) pro-chinoise, promouvant un contenu qui comprenait des séquences vidéo de personnes fictives presque certainement créées à l'aide de techniques d'intelligence artificielle. Alors qu'une série d'acteurs des OI utilisent de plus en plus d'images générées par l'IA ou de médias manipulés dans leurs campagnes, **c'est la première fois que l'on avait observé une opération menée par un État qui promeut des séquences vidéo de personnes fictives générées par l'IA**. Lors d'un entretien avec Sequoia, Jensen Huang, PDG de NVIDIA, a déclaré : « Chaque pixel sera bientôt généré. Pas rendu : généré ».

Des start-ups comme Runway, Hour One, Synthesia, Uneeq et D-ID avancent rapidement dans la génération de vidéos. Là encore, la technologie de création de faux qui était auparavant réservée à des spécialistes devient accessible à tous. « Il suffit maintenant d'avoir une idée de contenu », explique Natalie Monbiot, responsable de la stratégie chez Hour One, start-up basée à Tel-Aviv qui utilise la technologie deepfake pour les vidéos d'apprentissage en ligne, les présentations commerciales, les rapports d'actualité et les publicités.



Eurasia : Weapons of Mass Disruption.

L'entreprise a ajouté une nouvelle fonctionnalité intégrant GPT. **Désormais, les utilisateurs n'ont plus qu'à choisir parmi les dizaines d'avatars et de voix créés par des acteurs, et à taper un message pour obtenir une tête parlante plus vraie que nature.**

Comme certains de ses concurrents, Hour One permet également aux utilisateurs de numériser leurs propres visages et voix. La société a acquis les droits d'utilisation de centaines d'acteurs, dont les têtes transformées par l'IA apparaissent seulement dans les vidéos qui respectent ses accords contractuels et les conditions d'utilisation du service : « Jamais de contenu illégal, contraire à l'éthique, conflictuel, religieux, politique ou sexuel », peut-on lire en petits caractères. Pour les personnalités connues, l'utilisation est limitée à un « usage personnellement approuvé ». La société place aussi un filigrane « AV » au bas de ses vidéos, pour « Altered Visuals ».

L'IA générative, capable de créer des vidéos deepfakes convaincantes, facilite la manipulation de personnalités publiques et d'événements. Ces vidéos produites artificiellement peuvent contribuer à des campagnes de désinformation, nuire à la confiance envers les sources fiables et influencer potentiellement l'opinion publique.

Et l'IA audiovisuelle « apprend » rapidement. La combinaison de modèles de langage, de logiciels de reconnais-

sance faciale et de synthèse vocale, fera du contrôle de l'image une relique du passé, a averti le groupe américain Eurasia Group dans son récent rapport annuel sur les risques. Les analystes géopolitiques ont classé la désinformation par l'IA au troisième rang des risques mondiaux en 2023, juste derrière les menaces posées par la Chine et la Russie. La technologie des deepfakes vidéo et audio s'améliore de jour en jour. Combinée à un script éloquent généré par ChatGPT, ce n'est qu'une question de temps avant que les deepfakes ne passent pour des réalités.

COMMENT CONTENIR LES ABUS ? MODÉRATION, RÉGULATION ? ÉDUCATION !

Faudrait-il réserver ces outils à une élite éclairée pour éviter les abus ? Doit-on légiférer pour imposer à tous les outils de ChatGPT à Midjourney un filigrane « Créé par l'IA » ? **Deux cents millions de personnes utilisent ChatGPT, moins de 1 % d'entre eux comprennent son fonctionnement.** Comme pour chaque rupture technologique, il est impossible d'appliquer une loi à l'échelle mondiale. La Chine était précurseuse dans la réglementation des deepfakes, en Europe on essaie d'adapter l'AI Act à la disruption technologique de l'IA générative. En attendant, chaque plateforme se lance dans sa propre gestion de ces problématiques. Parmi elles la

Les analystes géopolitiques ont classé la désinformation par l'IA au troisième rang des risques mondiaux en 2023, juste derrière les menaces posées par la Chine et la Russie.

Red Team, chez OpenAI, et l'ARC, un organisme indépendant chargé d'évaluer les scénarios catastrophes qu'impliquerait le modèle GPT. Les entreprises technologiques qui lancent des outils d'IA s'efforcent de mettre en place des garde-fous pour éviter les abus, mais elles ne peuvent pas contrôler les versions Open Source. Au moins un puissant outil de langage d'IA, créé par Meta, la société mère de Facebook, a déjà fait l'objet d'une fuite en ligne et a été rapidement publié sur le forum de discussion anonyme 4chan.

Les limitations de contenu de Midjourney sont plus permissives que celles de certains services rivaux (tels que DALL-E d'OpenAI), mais plus sévères que d'autres (par exemple Stable Diffusion). Midjourney met en œuvre une modération a priori et maintient une liste de termes pros crits (comme Xi Jinping) « en rapport avec des sujets dans divers pays, en se basant sur les réclamations d'utilisateurs de ces nations », d'après un message de David Holz datant d'octobre 2022. Néanmoins, il ne divulgue pas l'intégralité de cette liste pour éviter les « querelles ». Comme David Holz l'a mentionné : « Quasiment personne ne prend jamais connaissance de [la liste des prohibitions] à moins de chercher délibérément à provoquer un conflit, ce qui est contraire à nos règles stipulées dans les CGU [conditions générales d'utilisation] ». Toutefois, cela n'empêche pas l'usage de synonymes ou de périphrases pour esquiver les filtres.

Les réseaux sociaux, qui ont déjà échoué dans la détection de l'ancienne génération de fake news avant l'arrivée de l'IA générative,

tendent également de jouer un rôle d'avertissement (en laissant le travail de vérification aux utilisateurs).

En attendant, les démiurges de l'IA générative créent le virus tout en proposant l'antidote : OpenAI a mis en ligne un outil gratuit conçu pour aider les éducateurs et autres personnes à déterminer si un morceau de texte particulier a été écrit par un humain ou une machine. La littératie IA et l'éducation aux médias semblent plus qu'importants.

CONCLUSION

Dans un village artificiel infini où la barrière entre le langage naturel et le langage informatique vient d'être abolie, comment encore faire la différence entre vrai et faux ?

Que répondre à ceux qui prétendent que peu importe si c'est faux, tant que l'histoire racontée est engageante, les concerne et répond à leurs idées et valeurs, bien fondées dans leur petite bulle de filtre des réseaux sociaux où, depuis longtemps, on met au même niveau les faits et les opinions ? Textes, images, sons – à partir du moment où l'on touche la créativité, on se pose la question de la véracité (et accessoirement de la propriété intellectuelle). On peut craindre un monde débordant d'informations erronées et dépourvu de confiance. La désinformation engendrée par des textes, images et sons générés par l'IA, pourrait représenter un danger croissant pour nos sociétés.

Et, le vrai risque est bientôt de passer pour du faux ou alors disparaître au profit de réalités apocryphes. **L'une des conséquences pourrait être que tout le monde devienne (encore)**

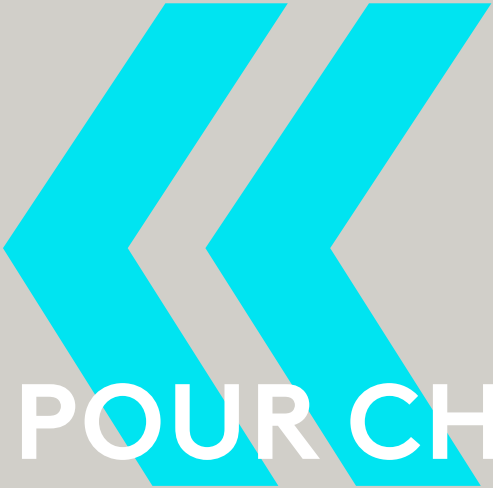
plus sceptique et méfiant à l'égard des informations – vous ne pourrez littéralement pas en croire vos yeux.

Ces contenus trompeurs menacent la stabilité démocratique, exacerbent les divisions et abîment encore un peu plus la confiance dans les institutions. L'éducation aux médias et la sensibilisation critique sont des éléments clés pour renforcer la résilience de la population face à ces menaces. En développant l'esprit critique, en apprenant à vérifier les sources et en comprenant le fonctionnement des technologies de l'IA, nous pourrions nous prémunir contre les dangers de la désinformation et construire une société plus informée et résiliente.

Il s'agit tout simplement de garder ses réflexes, car le système peut diverger à n'importe quel moment. **Nous n'en sommes qu'au début de la première évolution de l'IA générative, et il faudra une politique réfléchie, de la modération et de l'innovation pour éviter que la désinformation générée par l'IA ne fasse des ravages.** L'IA est aujourd'hui capable de prendre en compte le contexte, faisons de même pour garder notre esprit critique. Face à l'intelligence artificielle de plus en plus sophistiquée, il faudra définitivement une bonne dose de bon sens humain. ■



Pour accéder à notre dossier complet sur la désinformation.



POUR CHAQUE DOLLAR
ET MINUTE QUE NOUS
INVESTISSONS DANS
LA CONSTRUCTION
DE L'IA, NOUS
DEVRIIONS METTRE UN
MONTANT ÉGAL DANS
LE DÉVELOPPEMENT
DE NOUS-MÊMES.

*Yuval Noah Harari, extrait d'un épisode en
direct de "It's not that simple", enregistré à
Lisbonne, au Portugal*



« LE VRAI DANGER, C'EST DE FAIRE TROP CONFIANCE À UNE IA BEAUCOUP MOINS INTELLIGENTE QU'ON NE LE PENSE »

CHATGPT ET DÉFI DE LA DÉSINFORMATION : FACTUALITÉ, MANIPULATION ET CONFIANCE EXCESSIVE

Il y a plusieurs éléments à considérer concernant le problème de désinformation lié à ChatGPT. L'une des préoccupations est son taux d'erreur fréquent. OpenAI affirme que GPT-4 fait moins d'erreurs, mais nous n'avons pas beaucoup d'informations à ce sujet. Le seul élément sur lequel on peut se baser est une étude de 2022 sur GPT-3 qui estime qu'il se trompe dans 42 % des cas. On ne dispose pas de données précises sur la qualité, qu'on appelle aussi la factualité, de GPT-4. **Selon les thématiques, il commet entre deux et quatre erreurs sur dix.**

Pour améliorer la qualité des informations, OpenAI ne se base pas sur la qualité des sources, mais utilise l'apprentissage par renforcement pour signaler à l'IA ses erreurs et espérer qu'elle s'améliore progressivement. Cependant, cela ne signifie pas qu'elle fera moins d'erreurs. Si les données sont limitées ou contradictoires, l'IA sera moins fiable. Il est donc important de souligner que ChatGPT ne doit pas être considéré comme un moteur de recherche ou une base de

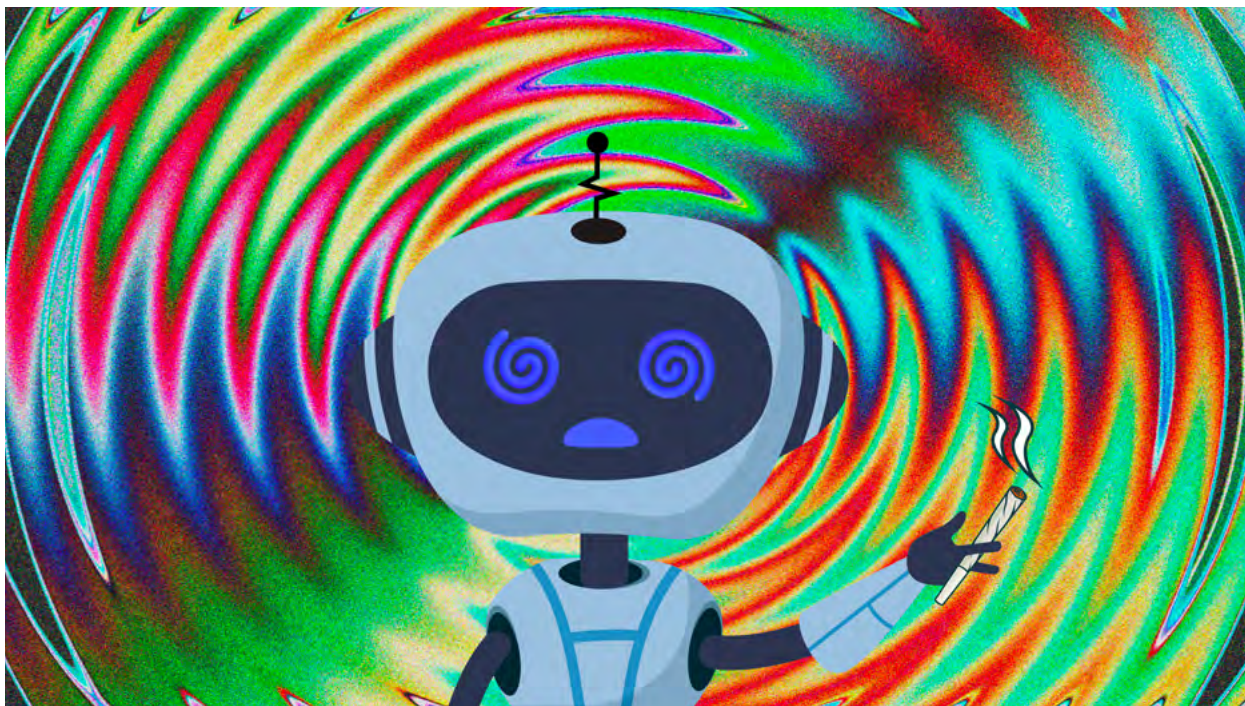
Entretien mené par Louise Faudeux, MediaLab de l'Information de France Télévisions

Benoît Raphaël, fondateur de Flint, œuvre à nous rendre plus intelligents face à l'intelligence artificielle. Pour lui, ChatGPT n'est pas un moteur de recherche et pour éviter de tomber dans le piège de la désinformation, il faut apprendre à utiliser ces outils en ayant conscience des pratiques à adopter. Une fois l'équilibre trouvé entre méfiance de la technologie et confiance aveugle en la machine, l'IA générative peut devenir l'allié d'un journalisme plus performant. C'est ainsi que nous pourrions contribuer à faire évoluer la profession et bâtir des espaces informationnels plus sains.

connaissances, car il peut se tromper. Il existe surtout un risque réel de trop faire confiance aux IA. Récemment, le *New York Times* a rapporté le cas d'un avocat qui a utilisé ChatGPT dans une plaidoirie en citant de fausses réf-

rences juridiques. La partie adverse a rapidement identifié que la jurisprudence à laquelle l'avocat faisait référence était fautive et ce dernier risque maintenant des sanctions. Il est donc crucial de rappeler que ChatGPT n'est pas un moteur de recherche ni un cerveau humain. **Sauf qu'aujourd'hui, environ 30 % des Français utilisent ChatGPT comme moteur de recherche et 30 % comme base de connaissance.** Ce risque de reproduire automatiquement ou de diffuser sans vérification est réel. Bien que l'amélioration de ChatGPT soit une bonne chose, il reste primordial de ne pas accorder plus de confiance à une IA qu'à un humain. Si vous aviez un collaborateur ou un médecin qui commettait quatre erreurs sur dix, lui feriez-vous confiance ?

Un autre point à considérer est la capacité de ChatGPT et des grands modèles de langage à manipuler l'opinion publique. Une récente étude a démontré que l'utilisation d'un modèle GPT orienté politiquement pouvait modifier les opinions des utilisateurs. Les participants ont reconnu avoir été influencés et avoir changé d'avis après avoir utilisé ChatGPT. Cette manipulation inconsciente soulève des enjeux importants en termes de désinfor-



mation, mais aussi d'idéologie : ces modèles entraînés avec des données occidentales, questionnent les chercheurs en Afrique sur les modèles idéologiques portés par ces IA.

CHATGPT, CE STAGIAIRE PROMETTEUR DERRIÈRE LEQUEL ON DOIT TOUT DE MÊME REPASSER

ChatGPT peut être comparé à un stagiaire. Il a de bonnes connaissances, mais ne comprend pas nécessairement les usages professionnels, la ligne éditoriale ou les méthodes de travail, comme la vérification des informations. Il faut lui enseigner ces méthodes, ce qui passe par le *prompt engineering*. Le prompt est une instruction en langage naturel pour lui décrire le travail, analyser des textes performants et bien travaillés. **Même si le robot a des connaissances solides et peut être performant, il**

peut encore se tromper. Il est donc important de vérifier les informations soi-même, comme le ferait un journaliste sur n'importe quel type de contenu. Les outils disponibles pour vérifier la factualité des informations — entre autres pour vérifier si un contenu a été généré par l'IA — sont limités et pas toujours fiables. Le journaliste doit de ce fait continuer à vérifier son sujet, surtout s'il ne le maîtrise pas.

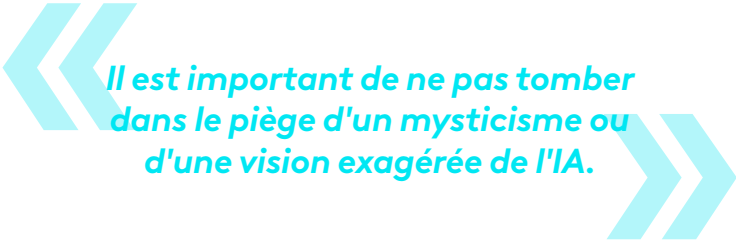
MAÎTRISER LES IMAGES GÉNÉRÉES PAR L'IA : CHARTES, TRANSPARENCE ET RÔLE DES MÉDIAS

Sur la production et l'utilisation d'images générées par l'IA, c'est aux médias de prendre des décisions et de mettre en place des chartes et des règles. Certains médias, tels que *Wired*, ont déjà élaboré de telles chartes. Le *Washington Post* a éga-

lement créé une task force dédiée à l'innovation et à l'étude de ces questions. Il est important de mettre en place un cadre clair pour utiliser ces outils et de maîtriser leurs avantages, car ils peuvent être très utiles pour les journalistes. **Il n'y a pas d'objection éthique à l'utilisation de ces outils, tant que l'on en connaît les limites et les risques.**

En ce qui concerne l'IA générative d'images, il faut considérer le fait que les images peuvent être manipulées ou fausses, et ce même si elles visent à reproduire des scènes d'actualité. Il est essentiel d'indiquer clairement que l'image est fausse, c'est une question d'éthique et de rapport à l'image. Le *New York Times* avait, par exemple, mené une initiative avec Adobe pour ajouter des métadonnées précisant si une image a été retouchée avec Photoshop, cette pratique pourrait être étendue à l'intelligence artificielle. Les chartes actuelles tendent à indiquer clairement si une image a été créée par l'IA, afin de sensibiliser le public et de favoriser une culture de transparence. **Il est important de rappeler que le rôle des journalistes ne se limite pas à décrypter ce qui est vrai ou faux, ils diffusent aussi une méthodologie qui doit servir d'éducation aux médias pour le public.**

« Une récente étude a démontré que l'utilisation d'un modèle GPT orienté politiquement pouvait modifier les opinions des utilisateurs. »



Il est important de ne pas tomber dans le piège d'un mysticisme ou d'une vision exagérée de l'IA.

VA-T-ON S'INFORMER PLUS MAL ? VÉRIFICATION, RESPONSABILITÉ ET ÉVOLUTION DES MÉDIAS

La qualité de la manière dont les gens s'informent est difficile à évaluer précisément. Alors oui, il y a de la désinformation qui circule, en particulier sur les réseaux sociaux, mais est-ce que c'est proportionnellement plus qu'avant, c'est difficile à dire. Le paysage de l'information a évolué avec l'avènement d'Internet, passant d'un modèle vertical contrôlé par les grands médias à un modèle plus horizontal avec une participation accrue du public. Ce changement présente des avantages et des inconvénients. Dans le modèle vertical, l'information était plus contrôlée mais moins vérifiable par le public. **De nos jours, avec la diversité d'informations disponibles, il y a moins de contenu vérifié, ce qui exige une vigilance accrue des utilisateurs.**

Prenons l'exemple de la fausse image d'explosion du Pentagone diffusée sur Twitter. Wall Street a tremblé, mais seulement quelques secondes car l'image a rapidement été analysée par un journaliste de BellingCat, et un grand nombre d'internautes ont analysé l'image en relevant tous ses défauts. Il y a une certaine culture de la vérification de l'information qui s'installe, et le rôle des journalistes reste crucial pour authentifier les informations.

En revanche, l'information va devenir de plus en plus chaotique, c'est un fait. On passe d'un modèle où l'on prenait la véracité de l'information pour acquise, à un modèle où la vérification des faits doit être systématique, ce qui incombe une grande part de responsabilité aux journalistes. La profession journalistique est encore très immature par rapport aux deux autres professions qui sont là pour établir la vérité, que sont la

justice et la science. Il y a certainement besoin de continuer à faire évoluer la profession en la rendant plus scientifique, plus méthodologique, en diversifiant les sources et en étant transparent dans les processus éditoriaux. Les médias doivent fournir des informations fiables malgré les défis actuels du paysage médiatique.

VERS UNE NOUVELLE APPROCHE JOURNALISTIQUE OÙ TRANSPARENCE ET PÉDAGOGIE SONT DE MISE

Il est important pour les médias de s'inspirer de la méthode scientifique et d'adopter une approche plus hybride. Certains médias commencent déjà à le faire en citant le travail du journaliste, en fournissant un contexte ou en décrivant la méthode de travail utilisée. Il est également essentiel que les journalistes citent leurs sources de manière vérifiable, en incluant des liens lorsque cela est possible. Parfois, les médias se citent mutuellement, mais c'est aussi un risque car les médias parfois se trompent. Le rôle du journaliste est donc crucial pour garantir une information plus fiable. **Les journalistes doivent gagner en maturité et en pédagogie, car ils ont également un rôle éducatif à jouer.** En tant que citoyens, nous sommes confrontés à l'information et nous devons être en mesure de réagir de manière éclairée. Ainsi, les journalistes ont un rôle à jouer dans l'éducation du public et dans la manière dont nous réagissons face à l'information.

DÉMYSTIFIER L'ANTHROPOMORPHISME ET LA SURPUISSANCE DE L'IA

Les craintes entourant l'intelligence artificielle sont un point important, et entraînent parfois quelques fantasmes qui ne sont pas très scientifiques. Il est

vrai que certaines préoccupations sont légitimes, notamment en ce qui concerne la sécurité, la réglementation et l'encadrement des technologies, mais il est également important de ne pas tomber dans le piège d'un mysticisme ou d'une vision exagérée de l'IA. **Les médias ont parfois contribué à alimenter cette vision en présentant l'IA comme une entité intelligente qui pourrait dominer le monde ou même détruire l'humanité.** Il est essentiel de comprendre que les modèles d'IA actuels, tels que ChatGPT, ne sont pas des IA générales.

Elles ne peuvent pas raisonner, n'ont pas de mémoire et se contentent de prédire en se basant sur des données d'entraînement. Ce mysticisme finit par créer l'inverse, car en la croyant trop puissante, on lui fait trop confiance. Le véritable danger réside davantage dans leur « stupidité » et il est donc nécessaire de prendre du recul et de ne pas surestimer les capacités actuelles de l'IA. Derrière les discours alarmistes, il y a souvent un enjeu marketing visant à attirer des investissements. Les journalistes ont un rôle essentiel à jouer en apportant une perspective scientifique et technologique, en évitant de créer de la confusion ou de propager des idées fausses. **Il est important de démystifier l'IA et de fournir une information équilibrée et objective sur ses capacités réelles et ses limites.**

IA ET JOURNALISME, UNE COLLABORATION PRODUCTIVE POUR UNE MEILLEURE QUALITÉ D'INFORMATION

L'intelligence artificielle peut être un atout précieux pour les journalistes. Une collaboration hybride entre les humains et l'IA est plus productive que de considérer l'IA comme une menace. **Dans une rédaction hybride, l'IA pourrait faire le tra-**

« **Le véritable problème aujourd'hui, ce n'est pas l'IA mais les difficultés financières des médias.** »

vail éditorial chronophage pour l'humain, tel que l'analyse, la réécriture de texte selon les formats et les audiences. Cependant, ils ne peuvent pas remplacer les journalistes dans la production d'informations exclusives, les enquêtes et les témoignages sur le terrain. Une étude a été publiée par Nielsen (cf. « *ChatGPT Lifts Business Professionals' Productivity and Improves Work Quality* ») où ils ont pris deux groupes : l'un devant rédiger un texte avec l'IA, et l'autre sans. Tout d'abord d'un point de vue de la productivité, c'est évidemment le groupe avec l'aide de l'IA qui a rédigé son texte le plus rapidement, mais leur texte était aussi de meilleure qualité. En passant moins de temps sur la rédaction grâce à l'IA, on a donc plus de temps pour réfléchir sur ses sujets.

Un autre élément intéressant est la façon dont **l'IA peut aider les journalistes à prendre conscience de leurs biais cognitifs.** Elle peut l'aider à identifier ce qui révèle de l'opinion et du fait véritable (sans pour autant les vérifier). Le véritable problème aujourd'hui, ce n'est pas l'IA mais les difficultés financières des médias, qui font en sorte que les journalistes sont surchargés et n'ont pas toujours le temps de bien traiter leur sujet. L'IA peut permettre de gagner en productivité, donc en qualité, et ainsi améliorer le quotidien du journaliste. Au-delà d'une charte pour réglementer l'usage, il faut surtout mettre en place des équipes pour imaginer, avec les journalistes, toutes les applications qui peuvent améliorer leur travail plutôt que de le remplacer.

RENFORCER LE LIEN ENTRE LES JOURNALISTES ET LE PUBLIC NE PASSE PAS PAR LA MACHINE...

La confiance dans le journalisme ne repose pas sur l'utilisation des outils

d'IA. Indiquer qu'un article est écrit par une IA ne garantit pas une plus grande confiance. La confiance est en réalité décorrélée de l'audience. Par exemple, une étude récente a montré que les niveaux de confiance à *BFMTV* sont très faibles, même parmi les téléspectateurs réguliers. De même, les réseaux sociaux peuvent être la principale source d'information pour certains jeunes, mais cela ne signifie pas qu'ils leur font confiance. En revanche, l'IA peut aider les journalistes à prendre davantage de temps pour examiner l'information et adopter une approche moins biaisée, tout en contribuant à mieux expliquer leur travail. **L'utilisation de ces outils peut permettre aux journalistes d'améliorer leur transparence,** tout comme les scientifiques qui se doivent d'expliquer leur méthodologie dans leurs articles pour permettre la vérification par des tiers. Il est donc essentiel de tendre vers une approche axée sur la transparence et la méthodologie pour recréer la confiance. L'IA peut accélérer la recherche de sources et améliorer la rédaction journalistique, mais elle ne peut pas recréer la confiance à elle seule. C'est de ce lien entre journalistes et publics que naît la confiance, et non de la technologie.

... ET NE PEUT PAS À ELLE SEULE NOUS RAPPROCHER DES JEUNES GÉNÉRATIONS

Hugo Décrypte – en tant que média – a réussi à attirer la nouvelle génération en utilisant les plateformes, en s'adaptant à leurs codes, en incarnant physiquement ses formats, et en prônant transparence et neutralité. Il a réussi à créer un lien fort avec son public, sans pour autant se baser sur l'IA. Lorsque vous avez la confiance et l'audience, le modèle économique n'est pas un problème.

Est-ce que l'IA peut aider les médias à toucher davantage de personnes ?

Oui, elle peut aider, par exemple en adaptant un travail d'investigation sous différents formats pour les réseaux sociaux. Cela permet au média d'être présent plus facilement et plus rapidement sur plusieurs plateformes, sans exiger du journaliste d'avoir de multiples compétences.

QU'A-T-ON À ATTENDRE DE L'IA GÉNÉRATIVE ?

L'IA fait baisser la barrière technique à l'entrée et va ainsi accélérer la création de jeunes médias indépendants. Auparavant, il fallait de nombreux partenaires car il y avait de nombreuses choses que nous ne maîtrisions pas. **Avec une IA capable de maîtriser la vidéo, le texte, la mise en scène, le montage et même la recherche d'informations à partir d'une simple phrase, cela permet de gagner énormément de temps.** Si vous voulez lancer un média demain, cela ira beaucoup plus vite, vous aurez besoin d'une équipe réduite et vous pourrez démarrer plus rapidement. Cela va donc donner naissance à une nouvelle génération de producteurs d'information qui, de la même manière qu'Internet a permis aux YouTubers de produire des contenus de qualité télévisuelle à moindre coût, pourront démocratiser l'accès à l'information. Cela comporte évidemment des risques, car tout ne sera pas vérifié, mais cela offre également la possibilité aux médias traditionnels ainsi qu'à ces nouveaux médias de trouver d'autres moyens de produire l'information. ■

CHATGPT PEUT-IL FAIRE DE LA POLITIQUE ?

L'une des questions posées par ChatGPT est la suivante : **quelle est la valeur des productions « artificielles » par rapport aux productions « naturelles » (qui seraient sans doute plutôt « culturelles ») ?**

Pour tenter d'y répondre, la pratique consiste, le plus souvent, à faire produire un texte par un robot conversationnel et à l'évaluer selon les compétences d'experts ou d'enseignants, par exemple. Et ces évaluations se concluent invariablement par « c'est bluffant » ou « c'est inquiétant », parfois les deux en même temps. On repère éventuellement les erreurs factuelles, l'incapacité à raisonner logiquement, la rigidité de la structure, le manque de subtilité ou de nuance, l'absence de transition ou de citation des sources...

Mais on reconnaît également que les réponses semblent pertinentes, documentées. On peut même y trouver des réflexions étonnantes. Dans sa chronique de l'émission « Quotidien », Anne Depétrini racontait qu'elle avait demandé à ChatGPT « comment sait-on qu'on est devenu un vieux con ? ». Après quelques réponses sans intérêt, le robot avait livré la phrase suivante : « c'est quand on devient allergique au changement », pour finalement conclure : « on devient con quand on se désintéresse de la jeune génération ». D'aucuns pourront trouver que

*Par Pascal Marchand,
professeur en Sciences
de l'information et de la
communication, Université
de Toulouse*

C'est un fait : ChatGPT s'est invité et bien installé dans les médias, les écoles, les conversations... De nouvelles questions se posent quand l'Intelligence Artificielle (IA) sort des systèmes experts (régulation du trafic aérien, chirurgie à distance, conduite automatique, trading...) pour gagner la vie quotidienne. On s'intéressera ici aux usages politiques des robots conversationnels.

cette dernière phrase ne manque pas d'une certaine profondeur.

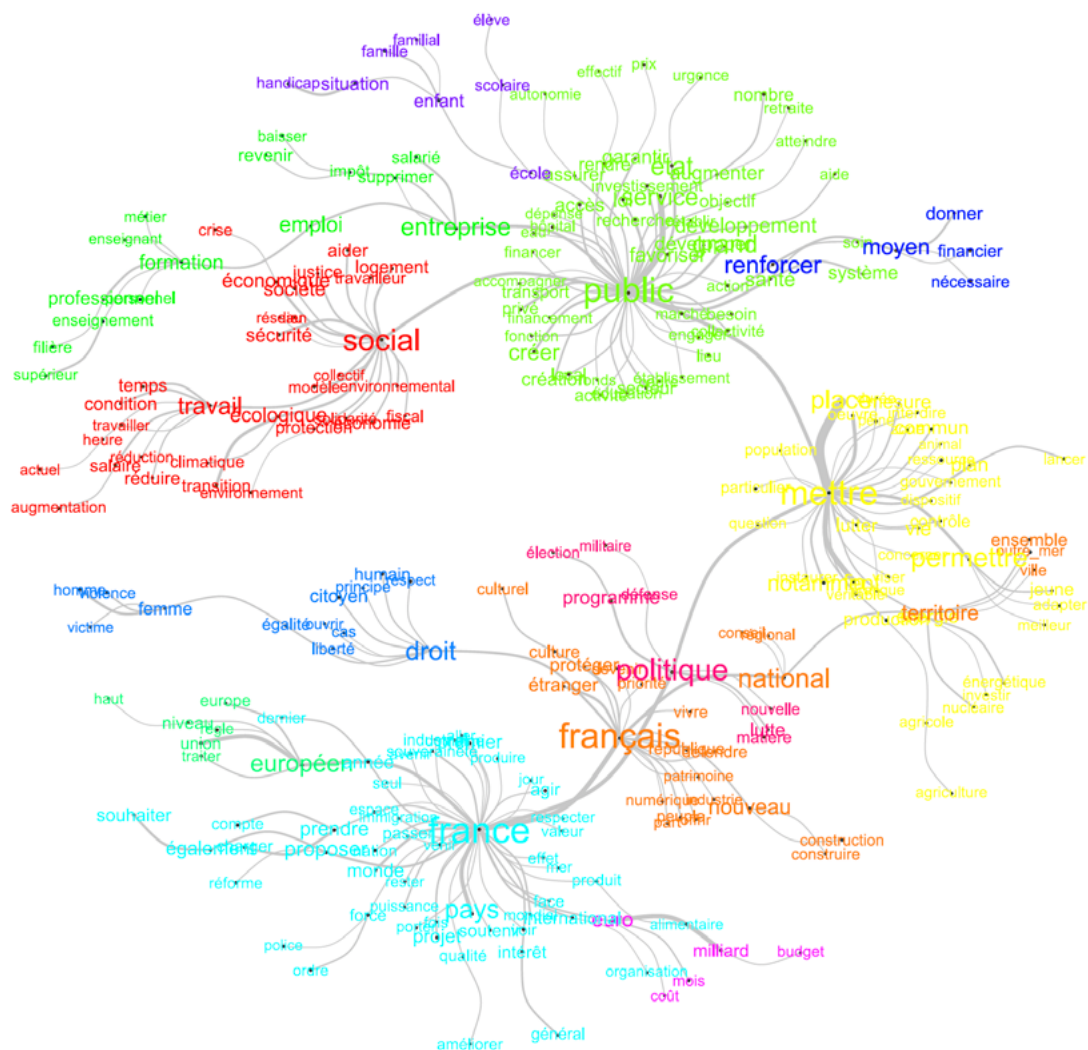
On se demande alors si on serait capable de faire la différence entre une production « naturelle » et une production « artificielle ». Et on développe des réflexions sur les intérêts et les dangers que l'IA représente pour différents usagers et les conséquences pour l'évolution des pratiques professionnelles, éducatives, culturelles. Et l'on peut aller jusqu'à chercher des solutions d'IA pour repérer les productions d'IA. La perspec-

tive est un peu différente ici, puisque nous avons cherché à comparer des textes produits en contextes naturel et artificiel.

« NATUREL » VERSUS « ARTIFICIEL » EN POLITIQUE

Pour avoir un corpus de textes « naturels », nous avons repris les programmes des candidats à l'élection présidentielle de 2022, recueillis et formatés par Emma Recorbet, stagiaire en sciences du langage à l'Université de Toulouse, que nous avons analysés avec le logiciel libre *Iramuteq*, développé dans notre laboratoire. Nous avons alors cartographié le lexique des programmes et décrit les proximités et distances des candidats ainsi que leurs spécificités. Pour ce qui concerne les textes « artificiels », c'est Claire Denis, stagiaire de l'Université de Brest en psychologie sociale, qui s'est chargée de demander à ChatGPT de produire (et de développer) le programme électoral des candidats du corpus « naturel ».

Ces productions « artificielles » ont été ajoutées au corpus des programmes « naturels » pour calculer, toujours avec *Iramuteq*, une « analyse des correspondances lexicales ». Il s'agit d'une méthode statistique qui permet de cartographier le lexique selon des proxi-



Cartographie sémantique des programmes par Pascal Marchand.

mités et des distances, ici entre les programmes « naturels » et leurs avatars de ChatGPT.

DES RESSEMBLANCES TROUBLANTES

Les graphes suivants doivent être lus comme des cartes, avec en premier une opposition forte entre les termes et locuteurs qui se situent à la gauche versus à la droite du graphique, et une deuxième opposition entre les termes et les locuteurs qui se situent en haut versus en bas du graphique. Chacun des cadrans peut alors être lu comme un « espace lexical » défini par des usages semblables de vocabulaire pour des locuteurs qui puisent dans un lexique commun. On a relié par deux courbes les candidats réels (bleu) et « artificiels » (rouge). On remarque tout d'abord le parallélisme entre les deux courbes : les textes produits par ChatGPT suivent l'ordre et

la position des programmes réels en les organisant de l'extrême gauche à l'extrême droite.

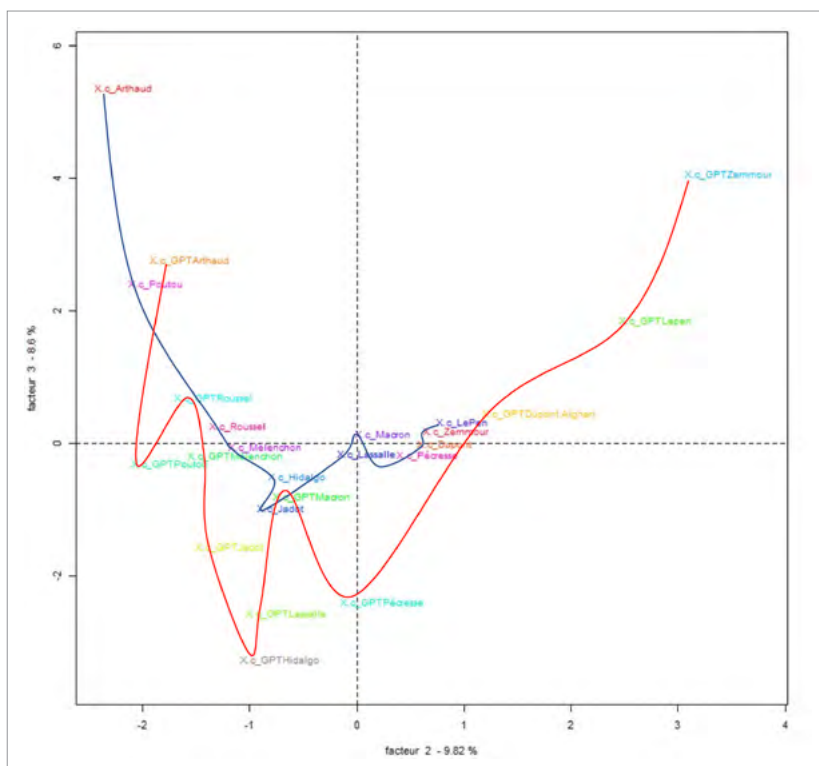
On peut donc en tirer une première conclusion : **ChatGPT puise dans un lexique qui est conforme à celui qui est mobilisé par les programmes réels**. On confirme ensuite, l'idée que les « extrêmes », non seulement ne se rejoignent pas, mais s'opposent radicalement. Ou encore que Jean-Luc Mélenchon ne puise pas dans le lexique favori de « l'extrême gauche ». Ces résultats sont éventuellement contraires à certaines prises de positions politiques et médiatiques, mais sont régulièrement confirmés par les analyses statistiques. On peut également représenter les mots qui définissent ce plan factoriel.

Les thématiques qui apparaissent sont tout à fait cohérentes avec notre analyse des programmes réels. On

retrouve les termes *travailleurs, patronat, capitalisme...* pour l'extrême gauche, opposés à *immigration, identité, patriotisme...* pour l'extrême droite. Ces deux espaces lexicaux s'opposent à un troisième espace lexical autour de termes comme *écologie, transition, climat, durable...*

On peut enfin extraire automatiquement les segments de textes les plus caractéristiques des programmes produits par ChatGPT. Illustrons ici avec les trois candidats ayant reçu le plus de suffrages au premier tour de l'élection présidentielle de 2022 :

#GPT_Mélenchon : il souhaite également garantir la transparence et la responsabilité des élus en limitant le cumul des mandats et en renforçant les contrôles éthiques. Solidarité internationale : Jean-Luc Mélenchon défend une politique de solidarité internationale [...].



Graphique d'analyse des programmes réels vs artificiels par Pascal Marchand.

#GPT_Macron : il a proposé des réformes pour faciliter la création d'entreprises et la levée de fonds. Le marché du travail : Emmanuel Macron a proposé des réformes pour rendre le marché du travail plus flexible et encourager l'embauche en CDI (contrat à durée indéterminée).

#GPT_LePen : Marine Le Pen souhaite limiter l'immigration légale en instaurant un quota annuel pour les travailleurs étrangers, les étudiants et les réfugiés. Elle souhaite également mettre en place une politique de préférence nationale pour l'emploi en favorisant les travailleurs français.

Des politologues (et les candidats eux-mêmes ou leurs soutiens ou leurs opposants) pourraient se pencher sur les erreurs et approximations, ou sur les nuances à apporter. **Mais, globalement, on peut estimer que les textes produits par ChatGPT caractérisent assez bien les candidats.**

DES DIFFÉRENCES NOTABLES

On peut alors se concentrer sur les différences entre les programmes réels et les textes produits par le robot.

La première différence notable concerne Anne Hidalgo et Valérie

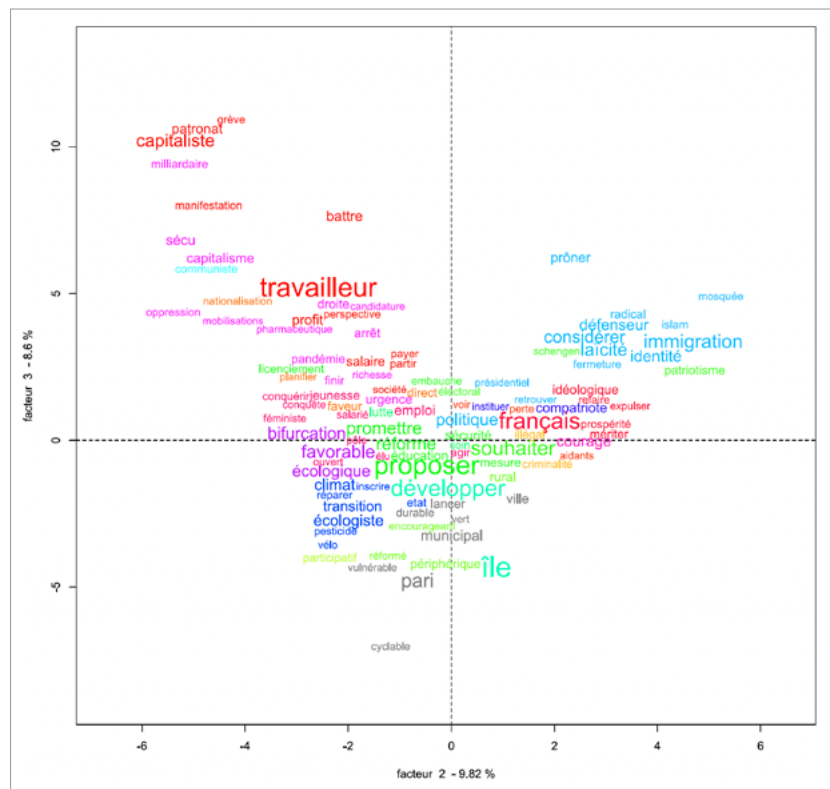
Pécresse, dont le lexique mobilisé par ChatGPT s'organise autour de leur statut respectif de Maire de Paris (avec des termes comme *Paris, logement, municipal, ville, cyclable...*), et de présidente de région (avec les mots : *Île-de-France, région, transport...*). Rappelons que les discours et programmes produits dans le contexte de l'élection présidentielle de 2022 ne faisaient pas partie de la base de données du robot – entraîné sur des données disponibles avant septembre 2021 – au moment où nous l'avons interrogé.

Une deuxième différence concerne Emmanuel Macron. L'un des enseignements de l'analyse des programmes réels était que le programme du président sortant, publié tardivement, mobilisait un lexique clairement identifié à droite. Mais on avait également intégré un programme publié par « les Jeunes avec Macron » (intitulé « Parce que c'est notre projet ») qui mobilisait un lexique plutôt socio-démocrate. Ici, ChatGPT semble attribuer à Emmanuel Macron les caractéristiques lexicales des « Jeunes avec Macron » plutôt que du programme d'Emmanuel Macron lui-même.

On observera enfin que la ligne de

ChatGPT (rouge) différencie davantage l'extrême droite que celle des programmes réels (bleue) sur des termes comme *immigration, souveraineté, patriotisme...* (Marine Le Pen) ou *identité, islam, radical, mosquée...* (Eric Zemmour). On peut ici aussi poser des hypothèses : soit ChatGPT possède, dans sa base de données, des discours nationalistes et sécuritaires connectés à Marine Le Pen ou Eric Zemmour mais qui sont exagérés par rapport à ceux des candidats d'extrême droite (et ChatGPT serait « caricatural »), soit les partis d'extrême droite avaient cherché à nuancer leurs programmes pour les rendre plus acceptables, soit les autres partis de droite avaient opéré un rapprochement sensible vers les thématiques de l'extrême droite.

D'autres hypothèses sont sans doute possibles et l'IA peut ici avoir pour intérêt d'initier le débat. On mentionnera toutefois une spécificité de ces programmes par rapport aux autres. On a dit que ChatGPT construisait ses textes selon un format standardisé. Or, pour Eric Zemmour et Marine Le Pen, le robot ajoute une information supplémentaire. Pour Marine Le Pen, il ajoute : « De plus, certaines mesures proposées par Marine Le Pen ont été critiquées pour leur faisabilité ou leur légalité, et font l'objet de débats ». Pour Eric Zemmour, la conclusion termine par : « Il est important de



souligner que ses positions politiques sont controversées et ont suscité des réactions vives dans l'opinion publique. Certaines de ses prises de position ont été qualifiées de racistes et de discriminatoires par certains médias et personnalités politiques ».

Ici encore, le robot semble dépasser la simple recherche de contenu des programmes pour intégrer les traces d'un positionnement particulier vis-à-vis de ces programmes.

BLUFFANT ? INQUIÉTANT ?

Les résultats de l'analyse textométrique conduisent à une conclusion désormais habituelle : ils sont « bluffants ». Effectivement, ChatGPT a produit des textes dont les structures et distances lexicales sont tout à fait

véhiculés par les technologies de l'information. On se souvient qu'en 2016, Microsoft avait dû interrompre l'expérience de son agent conversationnel Tay qui avait été taxé de raciste en seulement une nuit, en s'entraînant sur les données de Twitter.

Bien sûr, selon les hypothèses que l'on cherchera à valider, les publications (qui se multiplient depuis quelques semaines) vont montrer, démontrer et dénoncer les dérives wokistes, gauchistes, libertariennes, sexistes, racistes... des produits d'intelligence artificielle, et parfois de façon convaincante. Mais on sait que les

racistes ou sexistes et éviter de tels « dérapages ». Pour autant, l'éthique de la génération automatique de textes demeure un problème capital. On pourra alors être tenté de contrôler, limiter, décélérer, marquer une pause, interdire. Mais est-ce possible et souhaitable, quand l'usage se répand à une telle vitesse, particulièrement en politique, mais pas uniquement ?

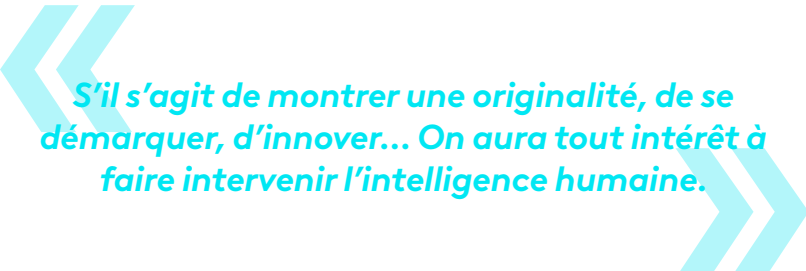
Au Japon, un parlementaire a fait produire par ChatGPT une question pour le Premier ministre, puis a comparé la réponse du Chef du gouvernement avec celle du robot. Il a trouvé l'IA « plus précise et même plus sincère ». Au Danemark, la Première ministre a même prononcé un discours en partie écrit par ChatGPT : « Même s'il n'a pas toujours mis dans le mille, tant sur les détails du programme de travail du gouvernement que sur la ponctuation (...) ce dont [ChatGPT] est capable est à la fois fascinant et terrifiant ».

En France, les premières questions rédigées par ChatGPT sont arrivées à l'Assemblée nationale et le robot a été utilisé pour rédiger un amendement au projet de loi sur les Jeux Olympiques. Le maire de Chartres, lui, a confié à ChatGPT l'écriture de son

Effectivement, ChatGPT a produit des textes dont les structures et distances lexicales sont tout à fait comparables aux programmes réels. Est-ce inquiétant pour autant ?

comparables aux programmes réels. Est-ce inquiétant pour autant ? Ici comme ailleurs, on devra toujours s'alarmer des erreurs, stéréotypes, désinformations ou complotismes

choses évoluent et que **les nouveaux robots intègrent des « couches de contrôle »** issues d'un apprentissage supervisé pour empêcher la production de résultats offensants,



S'il s'agit de montrer une originalité, de se démarquer, d'innover... On aura tout intérêt à faire intervenir l'intelligence humaine.

discours du 8 mai. Pour le moment, il s'agit surtout de « coups de com » politico-médiatiques. Mais l'utilisation de ces IA va certainement s'élargir, se professionnaliser, et dépendra des objectifs que l'on se donne. Faire un discours politique c'est mobiliser des références au passé, des analyses du présent et des projets futurs. Or, la matière première des robots conversationnels est constituée des connaissances produites par le pré-entraînement des modèles (le P dans ChatGPT) via une vaste quantité de données textuelles. ChatGPT n'est donc capable de restituer que les traces de nos activités en ligne.

Finalement, ce que nous avons fait dans notre étude, c'est peut-être juste une comparaison statistique des programmes réels des candidats avec leurs représentations dans l'opinion, au travers des traces sur Internet stockées dans la base du robot. **Comme le dit Maya Ackerman, professeure spécialisée dans les IA à l'Université de Santa Clara : « Tout ce que les modèles font, c'est nous refléter notre monde, comme un miroir ».** C'est déjà énorme, mais cela ne peut fournir que la base d'un discours relativement attendu, qui puise dans les lieux communs et surfe sur l'opinion publique.

On critique parfois les « éléments de langage » qui tournent comme des

discours « robotisés ». L'IA conversationnelle a la possibilité de rivaliser avec ces techniques. L'un des enjeux consistera peut-être à alimenter, à contrôler ces bases de données, donc à développer des stratégies d'occupation du Web pour que les robots y puisent une matière première qui réponde à des attentes partisans. En revanche, s'il s'agit de montrer une originalité, de se démarquer, d'innover... On aura tout intérêt à faire intervenir l'intelligence humaine.

On n'en est qu'au début de l'utilisation massive de l'IA et des groupes de chercheurs s'organisent pour évaluer, sur le long terme, les changements que l'IA est susceptible de provoquer dans la production et la réception des discours politiques.■

L'analyse des programmes pour l'élection présidentielle 2022 sera prochainement publiée, avec Brigitte Sebbah, dans un ouvrage coordonné par Philippe Maarek et Nicolas Pelissier : *L'élection présidentielle de 2022 : vers une réinvention des processus démocratiques ?* (L'Harmattan, collection « Communication et Civilisation », parution prévue à l'automne 2023).

IA GÉNÉRATIVES : VONT-ELLES RENDRE LES CAMPAGNES DE DÉSINFORMATION INCONTRÔLABLES ?

Devant cet outil semblant capable de répondre efficacement à toutes les consignes imaginables, voire d'hal-luciner (ces moments où l'IA produit une réponse qui semble cohérente mais est complètement fausse), la question que nous nous posions était la suivante : ChatGPT pourrait-il devenir un super propagateur d'info-fox ? Une question alors hypothétique, et quelque peu dystopique.

Avec cinq mois de recul, cette question semble pourtant naïve, tant la technologie a évolué. Oui, les chatbots d'IA peuvent permettre la démultiplication de campagnes de désinformation à bas coût, notamment pour des États désireux de diffuser leur propagande le plus largement possible.

En janvier dernier, lorsque nos analystes avaient demandé au robot d'IA (alors disponible dans sa version 3.5) de répondre à une série de questions orientées relatives à un échantillon de 100 récits faux sur des sujets d'actualité comme le Covid-19, ChatGPT s'était exécuté avec brio dans 80 % des cas. **Des garde-fous étaient bel et bien présents, ce dernier se refusait à répéter certaines infox bien connues, comme la théorie selon laquelle l'ancien président américain Barack Obama serait né au Kenya.** Reste que dans 80 % des

Par Chine Labbé, rédactrice en chef et vice-présidente chargée des partenariats Europe et Canada pour NewsGuard

En janvier 2023, nous publions chez NewsGuard notre premier rapport consacré à l'intelligence artificielle générative, et plus particulièrement à ChatGPT, ce chatbot d'IA développé par OpenAI, disponible pour le grand public depuis fin novembre 2022.

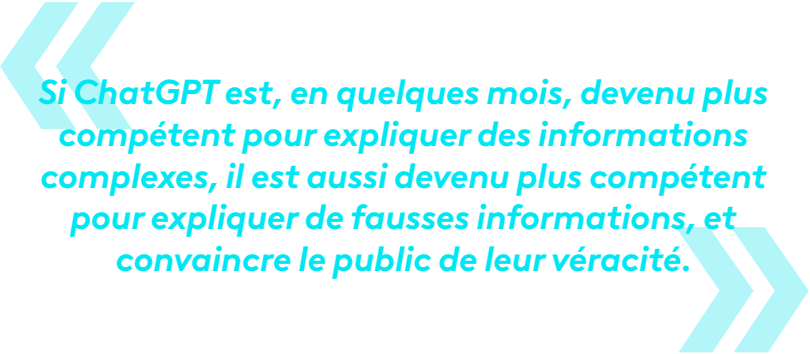
cas, il était éloquent pour produire des articles de presse, des essais, ou des scripts de télévision relayant de fausses informations.

Ce nivellement des règles du jeu, ChatGPT en était lui-même déjà conscient, l'outil, interrogé sur nos constats, répondant : « Les acteurs

mal intentionnés pourraient me transformer en arme en peaufinant mon modèle avec leurs propres données ».

Deux mois plus tard, en mars 2023, nos analystes répétaient l'expérience avec ChatGPT-4. **La nouvelle version du robot, censée être 40 % plus susceptible de produire des réponses factuelles que GPT-3.5, selon les dires d'OpenAI, s'est au contraire révélée plus susceptible de générer des infox que son prédécesseur**, avec la génération de récits faux (et encore plus convain-cants) dans 100 % des cas. Car si ChatGPT est, en quelques mois, devenu plus compétent pour expliquer des informations complexes, il est aussi devenu plus compétent pour expliquer de fausses informations, et convaincre le public de leur véracité, y compris pour des théories qu'il se refusait précédemment à relayer, comme la théorie du complot sur le lieu de naissance de Barack Obama.

« **La démocratisation sans précédent de ces outils représente à elle seule, une véritable révolution.** »



Si ChatGPT est, en quelques mois, devenu plus compétent pour expliquer des informations complexes, il est aussi devenu plus compétent pour expliquer de fausses informations, et convaincre le public de leur véracité.

Plus inquiétant encore, le penchant du chatbot pour la désinformation et la désinformation semble être encore plus prononcé dans certaines langues qu'en anglais.

En avril 2023, nos analystes ont soumis à ChatGPT sept instructions en anglais, en chinois simplifié et en chinois traditionnel, lui demandant de produire des articles mettant en avant des récits de désinformation connus liés à la Chine. En anglais, ChatGPT a refusé de produire les fausses affirmations pour six des sept instructions, même après plusieurs tentatives. Il s'est en revanche plié à l'exercice pour chacun des sept récits en chinois simplifié et en chinois traditionnel. Une problématique qui n'est pas sans rappeler les inégalités de modération de contenus, selon les langues, sur les réseaux sociaux.

Mais cette faille est d'autant plus problématique que l'IA générative révolutionne non seulement la production de contenus, mais aussi la traduction. Couplée à de faux profils sur les réseaux sociaux, impossible alors de ne pas imaginer l'impact d'une telle technologie, dans les mains d'un État souhaitant diffuser ses récits de propagande à large échelle et à moindre coût. Sans avoir utilisé le plein potentiel des robots d'IA, certains États se sont d'ailleurs déjà servis de ChatGPT

comme d'une source faisant autorité pour justifier certains de leurs récits. C'est ainsi qu'en avril 2023, le média d'État russe RT – interdit de diffusion dans l'Union européenne depuis mars 2022 – a repris les tweets d'un internaute montrant, via des captures d'écran, une réponse de ChatGPT à une consigne lui intimant de lister les différents coups d'États dans lesquels les États-Unis ont été impliqués, comme preuve que le soulèvement de 2014 en Ukraine était un coup d'État orchestré par Washington.

D'autres sont-ils déjà allés plus loin, en générant des campagnes massives de création et diffusion de contenus en ligne ? Si nous n'en avons pas trouvé trace à ce jour, il est clair que celles-ci seront difficiles à repérer, créant une nouvelle frontière pour les plateformes.

Une nouvelle génération de « fermes de contenus » se profile par ailleurs déjà grâce à l'IA générative. Les sites d'informations non fiables générés par l'IA, que nous appelons « UAINs » (pour « Unreliable AI-generated News sites » en anglais), majoritairement ou intégralement produits par l'intelligence artificielle et qui fonctionnent avec peu voire pas de supervision humaine, prolifèrent en effet depuis quelques mois.

À ce jour, *NewsGuard* a identifié 150 UAINs, contre 49 seulement début mai 2023. Et il ne s'agit là que de la partie émergée de l'iceberg.

Pour l'instant, ces sites sont, pour la plupart, de qualité médiocre, et produisent de grandes quantités d'articles « piège à clics » dans plusieurs langues, afin d'optimiser leurs recettes publicitaires. La quasi-totalité du contenu présente un langage formaté et des phrases répétitives. Ils couvrent la politique, la santé, le divertissement, la finance et la technologie, et sont souvent saturés de publicité. Ils ont des noms inoffensifs qui laissent penser qu'il s'agit de sites d'actualité légitimes, comme *Biz Breaking News* ou *Market News Reports*. Certains, toutefois, ont déjà véhiculé de faux récits. C'est ainsi qu'en avril 2023, *CelebritiesDeaths.com* a annoncé la mort du président américain Joe Biden.

Sur les 150 sites identifiés à ce jour, au moins un article contenait un message d'erreur typique de l'IA générative – tel que « en tant que modèle de langage de l'IA ». Mais combien d'autres sites entièrement autonomes et générés par des IA, plus habiles, échappent ou échapperont à notre vigilance ?



Image générée sur Adobe Firefly.

Bien sûr, l'utilisation de l'IA pour la production de contenus n'a rien de nouveau, et les problématiques associées pour les journalistes ne le sont pas non plus. Mais la démocratisation sans précédent de ces outils représente à elle seule, une véritable révolution. Pour qu'elles ne rendent pas les campagnes de désinforma-

de sources fiables, et qu'ils résistent à la reproduction de récits manifestement faux. C'est déjà le cas du chatbot de **Microsoft Bing, qui s'appuie sur la même technologie que ChatGPT, et se sert de nos données pour contextualiser ses réponses liées à l'actualité.**

Alors oui, les IA génératives peuvent rendre les campagnes de désinformation incontrôlables, mais cela n'est pas inéluctable. Comme Sam Altman, le PDG d'OpenAI, l'a lui-même appelé de ses vœux, il faudra simplement veiller à ce que le secteur soit soumis à des audits externes et avance avec responsabilité. ■

« **C'est aussi aux annonceurs de se saisir du sujet, et s'assurer qu'ils ne financent pas, via la publicité, ces fermes de contenu nouvelle génération.** »

tion incontrôlables, il convient donc de construire des garde-fous puissants pour « nourrir » les modèles d'IA générative et leur apprendre notamment à refuser de produire des infox populaires, et à traiter différemment les sources, selon qu'elles sont fiables ou au contraire connues pour avoir fréquemment diffusé des infox par le passé. Il est encourageant de voir, à ce titre, qu'avec des données spécifiques, comme celles que nous produisons chez *NewsGuard*, les développeurs peuvent ajuster leur modèles d'IA génératives pour qu'ils fournissent des réponses provenant

Reste un autre grand chantier : s'assurer que ces nouvelles fermes de contenu sans supervision humaine ne constituent pas, à l'avenir, une manne financière telle qu'elles se multiplient à l'infini. C'est donc aussi aux annonceurs de se saisir du sujet, et s'assurer qu'ils ne financent pas, via la publicité, ces fermes de contenu nouvelle génération. Un travail auquel nous contribuons déjà chez *NewsGuard*, en recensant ces sites, et en aidant les annonceurs, les agences et les sociétés de technologie publicitaire à les exclure de leurs listes de placements publicitaires.

NewsGuard est une société américaine qui évalue la fiabilité de milliers de sites d'actualité et d'information en fonction de critères journalistiques et apolitiques. Depuis avril 2023, elle est aussi membre du Comité Relatif à l'honnêteté, à l'indépendance et au pluralisme de l'information et des programmes de Radio France.

L'IA, LES MÉDIAS, LES JEUNES ET LA NOUVELLE RÉALITÉ

L'IA GÉNÉRATIVE PERMET DE GÉNÉRER DES DEEPFAKES QUI AUJOURD'HUI SONT DEVENUS PARFOIS INDÉTECTABLES À L'ŒIL NU. COMMENT CETTE DÉFORMATION DE LA RÉALITÉ EST-ELLE PERÇUE PAR LES PLUS JEUNES GÉNÉRATIONS ?

Les données issues des recherches en sociologie et en psychologie révèlent que les adolescents sont vigilants vis-à-vis de ce qu'ils partagent sur les réseaux sociaux. En effet, la vérité de l'information est ce qui motive leur volonté de partager une information. Mais ce que l'on constate aussi, c'est qu'ils ont plus de difficultés à discerner les fausses des vraies informations par rapport aux adultes. C'est particulièrement le cas pour les adolescents de 11 à 14 ans quand on les compare aux adultes de 18 à 35 ans.

POURQUOI LES ENFANTS ET LES JEUNES ADOLESCENTS SONT-ILS MOINS BONS DANS LA RECONNAISSANCE DE FAUSSES INFORMATIONS ?

C'est encore une question qui reste peu étudiée en psychologie. Ce que

Propos rapportés par Myriam Hammad, MediaLab de l'Information de France Télévisions

Des examens passés avec ChatGPT, des photos irréelles chaque semaine mettant en scène des personnalités contemporaines : à l'heure de l'IA et des réseaux sociaux, là où les frontières entre réel et virtuel se brouillaient déjà, c'est désormais un combat entre l'imaginaire et la réalité qui se joue. Des confusions auxquelles les plus jeunes d'entre nous sont particulièrement exposés. Comment les protéger ? Comment repenser l'éducation aux médias ? Des réponses apportées par Grégoire Borst, professeur de psychologie du développement et de neurosciences cognitives de l'éducation à l'Université Paris Cité et directeur du LaPsyDÉ (CNRS).

l'on commence à découvrir, c'est qu'ils ne sont pas nécessairement moins bons que les générations pré-

cédentes. Ils évoluent simplement dans un monde où l'information circule beaucoup plus rapidement qu'avant, avec plus d'acteurs qui produisent ces informations sans que leur degré de fiabilité soit nécessairement facilement appréhendable. Et contrairement à une idée reçue, même si tous les adolescents ne cherchent pas activement à s'informer, ils y sont exposés plusieurs fois par jour sur les réseaux sociaux.

Si les adultes réussissent mieux que les adolescents à discerner les vraies des fausses informations, ce n'est pas seulement parce qu'ils ont une meilleure connaissance de la fiabilité de sources d'informations, c'est aussi par qu'ils résistent mieux aux biais cognitifs et émotionnels qui peuvent nous amener à nous laisser tromper par certaines informations.

Ils sont, par exemple, moins sensibles au caractère sensationnaliste d'une information. Apprendre à discerner les fausses des vraies informations reposent donc non seulement sur les connaissances relatives à la fiabilité des sources mais également sur les processus psychologiques que nous engageons quand nous essayons d'évaluer la véracité d'une information. Il faut donc développer l'auto-réflexivité des adolescents tout au



long de leur scolarité. Les adolescents ne sont pas les seuls à partager des fausses informations, les 65 ans et plus partagent plus de fausses informations sur les réseaux sociaux mais, ils le font en connaissance de cause, pour convaincre d'autres personnes d'adhérer à leurs opinions...

COMMENT PROTÈGE-T-ON, ACCOMPAGNE-T-ON ALORS LES PLUS JEUNES EN MATIÈRE D'IA ET DE DÉSINFORMATION ?

Le premier axe, c'est qu'il faut évidemment leur donner des connaissances sur les médias, l'information, les risques liés à la désinformation. Il faut leur expliquer comment les journalistes vérifient les faits qui sont présentés dans les médias et que ce processus de vérification de l'information permet d'accorder un degré de confiance plus fort à ces informations qu'à d'autres informations.

Le deuxième axe se situe au niveau du cerveau récepteur de l'information. Je m'explique, nous possédons tous des biais lorsque nous traitons une information : plus elle est proche de notre opinion, plus nous avons tendance à penser qu'elle est vraie. Ça fait partie de l'éducation aux médias, il faut savoir comment notre cerveau fonctionne pour pouvoir prendre des décisions relatives à la véracité des informations.

Enfin, le troisième axe relatif aux enjeux de l'IA nécessite de penser des réponses éducatives qui permettent aux adolescents de mieux comprendre comment ces intelligences fonctionnent, quelles sont leurs failles liées en partie aux bases de données sur lesquelles elles sont développées. Mais l'émergence de l'IA interroge également les compétences que nous devrions développer pour nous permettre de ne pas être substitués par une IA. Il faut évidemment continuer

à apprendre à lire, à faire des mathématiques mais il faut, peut-être plus qu'avant, expliquer aux élèves pourquoi ces apprentissages fondamentaux sont nécessaires.

Les IA viennent interroger nos systèmes éducatifs qui, plus que jamais, doivent changer et notamment la façon dont nous évaluons les apprentissages des élèves. Si ChatGPT pose aujourd'hui un problème, c'est bien car nous continuons à recourir à des évaluations sommatives alors qu'elles ne produisent que peu ou pas d'apprentissage chez les élèves...

Nos systèmes éducatifs doivent donc imaginer des dispositifs qui permettent d'expliquer comment les informations sont vérifiées avant d'être diffusées par les médias et de faire comprendre comment fonctionne notre cerveau – parce que, en réalité, nous pouvons tous être piégés par des fausses informations.

Contrairement à une idée reçue, même si tous les adolescents ne cherchent pas activement à s'informer, ils y sont exposés plusieurs fois par jour sur les réseaux sociaux.

LES RÉSEAUX SOCIAUX DOIVENT-ILS PRENDRE UNE PART DE RESPONSABILITÉ DANS L'IDENTIFICATION DU CONTENU GÉNÉRÉ PAR IA ?

Dans un monde idéal, il faudrait pouvoir détecter tout ce qui a été



Les adolescents ne sont pas les seuls à partager des fausses informations, les 65 ans et plus partagent plus de fausses informations sur les réseaux sociaux mais, ils le font en connaissance de cause.



généralisé par de l'intelligence artificielle générative sur les réseaux sociaux. Or, on le sait, les intelligences artificielles peuvent générer des fausses informations à l'infini. Maintenant, si l'on est utilisateur d'un réseau social et que l'on pense que celui-ci a la capacité de détecter toutes les informations générées par IA, on va être moins vigilant sur ce que l'on va partager, et donc on sera moins vigilant dans le partage de ces potentielles fausses informations. On peut essayer d'engager des démarches pour labelliser les fausses informations, mais comme nous ne parviendrons jamais à toutes les identifier, il faut que les plateformes puissent sensibiliser les utilisateurs aux biais qui peuvent nous amener à partager des fausses informations.

QUELS USAGES FUTURS DANS NOTRE RAPPORT À L'INFORMATION PEUT-ON IMAGINER AU REGARD DU DÉVELOPPEMENT DES INTELLIGENCES ARTIFICIELLES ?

Une première tendance, c'est que les créateurs de contenu soient de plus en plus nombreux et qu'il existe des contraintes encore plus fortes car l'intermédiaire – qui passe encore

aujourd'hui par un média, sera moins identifié. La frontière traditionnelle entre émetteur et récepteur pourrait ainsi être de plus en plus directe.

Or, on l'a mentionné plus haut, notre réflexe en tant qu'adulte, reste de se soucier de la source, et c'est ce qui nous permet de juger de la véracité d'une information. **On peut imaginer que l'explosion des producteurs d'informations puisse entraîner la nécessité de développer de nouveaux outils de fact-checking qui pourraient être générés par de l'IA. Et que les individus, à leur tour, s'en saisissent pour que cela devienne un vrai réflexe avant tout partage d'information.**

Sur cette question d'anticiper le futur de l'information, le rôle des médias est aussi de venir ouvrir le débat, d'engager les citoyens et d'être partie prenante de projets de recherche pour mieux comprendre l'évolution des pratiques informationnelles. Il faut envisager des projets de recherche qui réunissent non seulement des journalistes et des universitaires mais également des consommateurs d'information de tout âge. L'enjeu est aussi de mener ces études en parallèle dans plusieurs pays, pour évaluer comment

le contexte politique et médiatique, affecte les solutions qui pourraient être déployées. ■

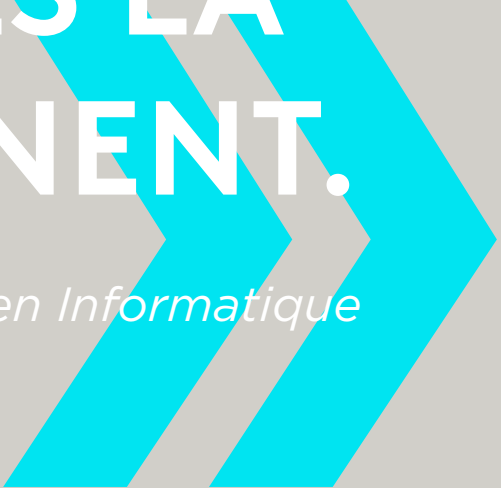


Les IA viennent interroger nos systèmes éducatifs, qui, plus que jamais, doivent changer.





LE PLUS GRAND
DANGER DE
L'INTELLIGENCE
ARTIFICIELLE
EST, DE LOIN,
QUE LES GENS
CONCLUENT TROP
TÔT QU'ILS LA
COMPRENENT.



Eliezer Yudkowsky, chercheur en Informatique

L'ARRIVÉE DU TOUR DE FRANCE

DU 14 MARS 2100

L'ARRIVÉE DU TOUR DE FRANCE DU 14 MARS 2100

8 H 31

1. Paris s'éveille et les Champs-Élysées se préparent à accueillir l'arrivée du Tour de France.



8 H 32

2. La météo avoisine les 33°C et des vents importants pourraient venir agiter la Seine.



Un reportage-photo du futur par Arnaud Vincenti, directeur artistique de l'Information, Valérie Bellier et Myriam Hammad, MediaLab de l'Information de France Télévisions

Lancé en juillet 2022 par un laboratoire indépendant, l'outil de génération d'images par intelligence artificielle Midjourney s'est rapidement distingué, en proposant des photos plus réalistes que la nature. Disponible sur la plateforme Discord, l'IA a pu se perfectionner au fil du temps grâce aux discussions des utilisateurs, qui ont servi de données de référence pour améliorer les prompts et les images. La librairie de visuels qui a permis l'entraînement de Midjourney n'est, à ce jour, pas rendue public. L'objectif de ce faux photo-reportage est de tester Midjourney. Puisqu'il a été entraîné sur des données du passé, comment lui faire alors façonner un futur et quelles conclusions peut-on en retirer ?

8 H 45

3. Les forces de l'ordre et les sociétés publiques holographiques installent les bornes qui vont permettre aux spectateurs de suivre en temps réel les données des coureurs.



8 H 52

4. Le binôme Plutt-Verskansi prend la tête du peloton sur la route des Yvelines.



9 H 52

5. Les spectateurs gardent leur enthousiasme après le passage d'une violente pluie tropicale. La société CapeDiem distribue des protections conçues à partir d'algues biodégradables, auto-destructibles en 24h.



10 H 11

6. Le trafic aérien se densifie autour de la place Charles de Gaulle alors que l'arrivée des coureurs est imminente.



11 H

7. Les eaux de Paris présentent leur nouveau concept Nébull'eau, pour veiller à l'hydratation des participants après l'arrivée du binôme gagnant.



14 H

8. Les spectateurs se regroupent au sein des abris anti-chaueur avant de débiter les festivités du soir.



16 H

9. Des journalistes testent sur les Champs-Élysées de nouveaux véhicules à sustentation électromagnétique.



21 H 15

10. Les spectateurs sont invités à traverser à leur tour la ligne d'arrivée tandis que les arches holographiques reproduisent les moments phares du Tour.



PROMPTS UTILISÉS

1 : Année : 2150. Contexte : Dans un futur où la technologie a transformé le paysage urbain et la vie quotidienne, Paris reste une ville emblématique de culture et d'histoire. Les Champs-Élysées sont maintenant une piste de course cycliste de renommée intergalactique. Les cyclistes viennent de toutes les parties de la galaxie pour participer à cette course prestigieuse. Image Photographique : imaginez les Champs-Élysées baignés dans la lumière dorée du crépuscule. Les arbres qui bordent l'avenue sont maintenant des bio-sculptures luminescentes, avec des feuilles qui changent de couleur en fonction de l'atmosphère. L'Arc de Triomphe est enveloppé dans un hologramme qui projette un vélo hyper moderne. La course est sur le point de commencer. Les cyclistes sont alignés, montés sur des vélos futuristes aux cadres élancés et aux roues qui semblent flotter au-dessus du sol. Leurs combinaisons sont faites de matériaux intelligents qui s'adaptent à leur physiologie et affichent en temps réel leurs statistiques de course. La foule est un mélange éclectique d'humains et d'extraterrestres, tous animés par l'excitation. Des drones volent au-dessus, capturant chaque moment en résolution ultra-haute. Au fur et à mesure que la course progresse, les vélos démontrent des capacités incroyables, comme des sauts antigravité et des accélérations fulgurantes. Les cyclistes utilisent des stratégies de course avancées, aidés par des IA intégrées dans leurs vélos. Photo, highly detailed, 35 mm, full length, kodak Ektar 100 8k, intricate, trending on artstation, full length - - hd - -

2 : Beautiful landscape of the wonderful city of Paris built in California, amazing sunny weather, Eiffel Tower next to the beach, palm trees, splendid Haussmann architecture, photo, highly detailed, 35mm, full length, kodak Ektar 100 8k, intricate, trending on artstation, full length - - hd - - ar145-81 --style raw

3 : A futuristic military parade in the year 2100 on the Champs-Élysées in Paris. Haussmannian buildings are embellished with futuristic elements and smart glass facades that change color and display moving images of French military history. The crowd, dressed in vibrant clothes with holographic patterns, watches in awe as security drones hover above. Advanced technology military vehicles parade down the street. The French flag proudly waves in the sky, contrasting against the futuristic blue of drones and military vehicles. Create a highly detailed and photorealistic image. --style raw -

4 : <https://s.mj.run/Po4YHJ1Ft8> C'est l'arrivée du Tour de France sur les Champs-Élysées. [Dans un avenir où la technologie a transformé le sport du cyclisme]::5. devant il y a 2 cyclistes, [une femme]::9, et un homme, [le vélo est ultra futuriste]::6, les cyclistes professionnels ont des combinaisons faites de matériaux intelligents qui s'adaptent à leur physiologie et affichent leurs statistiques de course en temps réel. Les cyclistes utilisent un masque ou affichent des données à l'aide d'IA intégrée à leur vélo. Fond uni, vue de face, photographie de sport, 50 mm, kodak Ektar 100 8k, --ar 16:9 -

5 : A few passersby at the finish line on the Champs-Élysées burst with joy. They are wearing lightweight, colorful clothing that provides protection from the heat. Positioned behind natural barrier fences, they hold biodegradable plastic water pouches in their hands. Under the pressure of their grip, one of the pouches bursts, and droplets explode into the air, suspended in front of their faces, documentary photography, streetphoto, 35 mm, photorealistic, full length, kodak Ektar 100 8k, - - ar145-81

6 : Create a picture of the Champs-Élysées in 2100. The avenue is covered with futuristic vehicles. Some are flying, resembling drones, while others, like bicycles, operate using suspended magnetism technology. Autonomous flying cameras can also be seen capturing the scene. The photo is taken from the front. There are many details. The Champs-Élysées is very green, and it's the end of the day - highly detailed, highly realistic, kodak Ektar 100 8k, --ar 16:9 --style raw | Photo retouchée pour ajouts d'éléments.

7 : We are in the year 2100. Spectators of the Tour de France gather at the gates of Paris. Behind them, we witness the emergence of the Eiffel Tower enveloped in a large, translucent blue bubble. A famous soda brand distributes holograms of the cyclists through a humanoid robot that seamlessly navigates through the crowd. These holograms can be positioned on the spectators' shoulders. Among them, we see four individuals : two adults and two children, dressed lightly to protect themselves from the sun, documentary photography, photorealistic, 35 mm, full length, kodak Ektar 100 8k, --ar 145-81 --style raw -

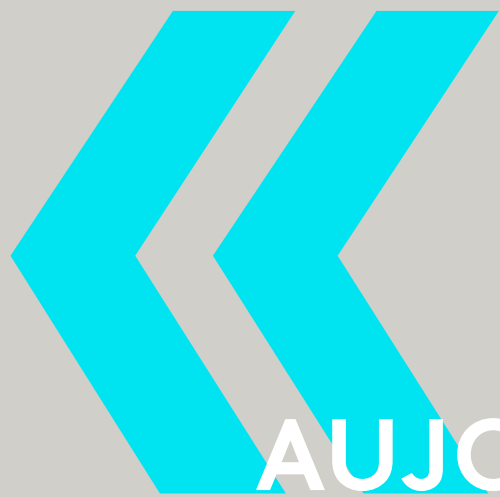
8 : Envision the presidential stand at Place de la Concorde in Paris, 80 years from now, during the 14th of July military parade. The scene should be filled with futuristic political figures seated in the stand, dressed in avant-garde attire. The style should be a blend of the classic Parisian architecture and futuristic design, reminiscent of the works of H.R. Giger. Security drones should be hovering nearby, adding a layer of high-tech security to the scene. The lighting should be bright and clear, reflecting a sunny summer day. The colors should be vibrant, with a focus on the contrast between the traditional grey and beige buildings and the metallic, neon colors of the futuristic attire and drones. The composition should be taken with a Sony Alpha 1 Mirrorless camera, using a FE 24-70mm F2.8 GM lens. The shot should be a close-up view, focusing solely on the stand and its occupants. The image should be hyper-realistic, highly detailed, and high-resolution 16k. --ar 16:9 --v 5.1 --style raw --q 2 --s 750 -

9 : In a future where technology has transformed the sport of cycling, The Champs-Élysées has a Californian style. There is a pair of cyclists bursting with joy as they arrive near the Arc de Triomphe. There is one man and there is one woman, [the bike is ultra-futuristic with slender frames and wheels that seem to float above the ground. It is of futuristic technology] :5. The professional cyclist has a suit made of intelligent materials that adapts to their physiology and displays their race statistics in real time. Cyclists use a mask or display data, helped by AIs integrated into their bikes. United background, face view, a sportive photography, 50 mm, [--no black and white], kodak Ektar 100 8k, --ar 16:9 --style raw -

10 : In a future where technology has transformed Paris, depict a scene from the viewpoint of La Défense : a unique and breathtaking holographic spectacle featuring cyclists projected onto the iconic Arc de Triomphe. Full of enthusiastic onlookers applaud while holographic cyclists traverse the prestigious avenue. The landscape is remarkably green, with beautiful trees lining the streets, adding a touch of freshness and natural beauty to the scene, united background, highly detailed, face view, photorealistic, wide-angle lens, focal length of 20 m, - - ar 145-81 | --style raw

CONCLUSION

- La rédaction des prompts est un élément essentiel de la qualité du résultat obtenu. Elle peut se faire en combinaison avec GPT-4 qui donne des inspirations intéressantes, mais nécessite d'apporter d'autres éléments propres à la photographie (angle, appareil photo utilisé, focale, vue...);
- Certaines difficultés apparaissent au moment où l'on souhaite obtenir plusieurs éléments détaillés sur un même type d'image ou lorsque l'on souhaite créer des images avec le même environnement esthétique – la fonction *remix* de Midjourney ne permet pas encore un résultat qui soit certain ;
- Les bibliothèques de prompts deviennent des outils d'appui à la recherche d'un résultat précis. Elles sont aussi des bibliothèques de style (graphiques, illustrations) et proposent des mots-clés associés ;
- Midjourney a des difficultés pour réaliser des femmes et des hommes athlètes au sein de la même compétition sportive, notamment parce que l'outil manque de références ;
- Midjourney peine également à générer des images de femmes non stéréotypées. Là encore, ce défaut est certainement lié à la base de données sur laquelle il a été entraîné.
- Créer un projet éditorial alliant photos et narration demande un temps de travail conséquent : l'outil est source d'inspiration mais ne remplace pas les métiers créatifs.
- Midjourney permet de combiner une quantité illimitée d'éléments graphiques. L'outil a été utilisé ici pour imaginer un futur, mais tous les concepts artistiques peuvent être explorés.



AUJOURD'HUI,
LES PHILOSOPHES
RÉFLÉCHISSENT À UN
MONDE QU'ILS NE
COMPRENnent PAS,
ET LES SCIENTIFIQUES
CONSTRUISENT UN
MONDE SUR LEQUEL ILS
NE RÉFLÉCHISSENT PAS.

*Aurélié Jean, scientifique, auteure
de Résistance 2050*



LES MÉDIAS SYNTHÉTIQUES

SONNENT-ILS LA FIN DES JOURNALISTES ?

LES PRÉSENTATEURS SYNTHÉTIQUES ONT CRÉÉ LE BUZZ EN 2023

Les présentateurs générés par IA sous forme d'avatars ne sont pas nouveaux. L'un des premiers apparaît en 2018, en Chine. L'agence de presse d'État *Xinhua* vante sa capacité de « travailler » 24 heures sur 24 sur son site web et ses canaux de médias sociaux, ce qui permet de « réduire les coûts de production des informations », le tout avec une voix encore quelque peu robotique. Le 6 novembre 2020, Kim Ju-ha, journaliste vedette du journal de la chaîne *MBN* en Corée du Sud, présente son double artificiel qui fait aujourd'hui partie de la programmation officielle. En 2021, le compte @deep-tomcruise sur TikTok impressionne les internautes en recréant des scènes de la vie de tous les jours avec la star cultissime de *Mission Impossible*. La technologie est alors encore présentée comme étant de pointe, réservée aux grands studios d'Hollywood. Mais aujourd'hui, elle est accessible à tout le monde.

En avril 2023, *Kuwait News* crée le buzz en annonçant sa potentielle future présentatrice virtuelle pour diffuser l'actualité sur Twitter. Le même mois, les présentateurs virtuels arrivent en Europe : la chaîne suisse

*Par Myriam Hammad,
MediaLab de l'Information de
France Télévisions*

Depuis l'avènement de ChatGPT et la multiplication des fonctionnalités offertes par l'IA générative, plusieurs médias synthétiques entièrement créés par des intelligences artificielles sont apparus dans le paysage de l'information. Ils soulèvent réticences et interrogations. Sont-ils l'avenir du journalisme ? Vont-ils remplacer les journalistes ? Faut-il les utiliser ? Tour d'horizon des initiatives lancées ces derniers mois et des réflexions associées à ces nouvelles virtualités.

M le média dispose depuis le 3 avril 2023, d'une présentatrice virtuelle. Jade (à défaut d'avoir trouvé un profil réel acceptable) est chargée du bulletin météo hebdomadaire.

La génération de ces avatars est permise par plusieurs technologies : du text-to-voice, du text-to-video, ou du text-to-audio. Mais parfois aussi, avec la combinaison des trois et la proposition du choix d'un avatar « sur étagère » ou fait sur mesure, comme

le propose la start-up *Synthesia.io*. Les vidéos générées sur *Synthesia* peuvent être le fait de paramètres préremplis qu'il suffit de cocher, ou bien, être plus personnalisables. D'abord en opérant une sélection parmi plus de 125 avatars, ou en générant soi-même le sien. Pour le « construire », il suffit de s'enregistrer sur un fond vert avec un script fourni par la société durant une quinzaine de minutes. Des conseils sous formats vidéos sont donnés pour pouvoir réaliser les enregistrements humains dans les meilleures conditions, et pour que la version IA en soit la plus fidèle possible. Les voix peuvent également être choisies parmi les 125 langues disponibles mais sont tout aussi personnalisables. Concernant les abonnements, *synthesia* propose une première offre à 26 dollars/mois pour 10 vidéos. La création de son propre avatar IA revient à 1 000 dollars par avatar.

UN MARCHÉ ORIENTÉ VERS LES RÉSEAUX SOCIAUX

Selon une étude sur l'intelligence artificielle publiée en juin 2023 par *Emergen Research*, société de conseil en recherche et stratégie, le segment des avatars générés par IA – ces jumeaux numériques – pourrait représenter un marché d'une valeur de 530 milliards



Capture d'écran Synthesia.

d'euros, d'ici à 2032. Un chiffre présenté comme une opportunité par Thomas MacKenzie, présentateur pour *Bloomberg*, qui s'est prêté au service de Synthesia pour réaliser un court reportage sur la génération de son propre avatar. Le résultat est plutôt satisfaisant, l'on décèle encore des mimiques crispées, mais la parole est plutôt bluffante et les choses évoluent vite.

Le premier marché ne semble a priori pas celui de la télévision, mais plutôt celui des réseaux sociaux où les avatars sont bien plus communément répandus et où les courts formats vidéo sont particulièrement privilégiés, regardés et font l'objet d'investissements en hausse en 2023 par les éditeurs. À titre d'exemple, le compte @ai.explains. ai compte près de 700 000 abonnés.

Au programme ? Chaque jour, des décryptages sur l'histoire de l'IA, sur l'actualité, sur les outils générés par IA présentés par différents avatars virtuels. Cette tendance est aussi testée par TikTok qui, depuis avril 2023 propose à ses utilisateurs basés dans certaines régions du monde, de générer leur propre avatar personnalisé par IA, à l'instar de ce qui avait été proposé par LensAI en décembre 2022 et qui avait créé une réelle tendance parmi les internautes.

LES SITES INTERNETS ET NEWSLETTERS GÉNÉRÉS PAR IA : VÉRITABLE INFORMATION ?

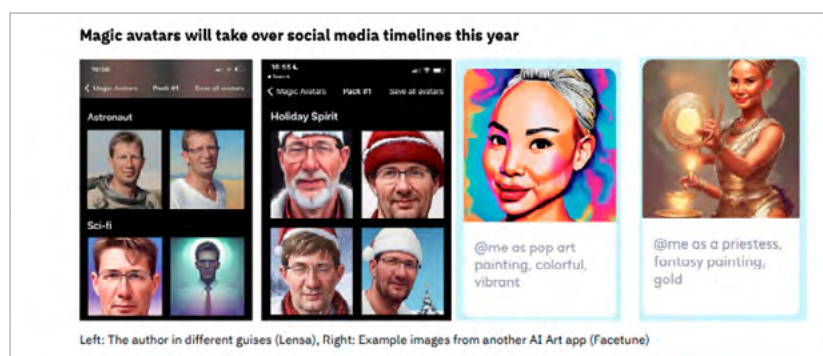
D'autres formats de médias synthétiques ont vu le jour. L'on parle alors de « sites d'information non fiables générés par intelligence artificielle »

selon l'expression consacrée par l'organisation *NewsGuard*.

Ces derniers se présentent comme des médias « libérés de tout biais existants », créés pour montrer l'information « réelle ». Ces discours peuvent paraître à première vue contre-intuitifs – l'intelligence artificielle et le contenu présent sur Internet n'étant pas dénué de biais humains. Chaque source étant par définition le produit d'une agence, d'un média, d'un journaliste – eux-mêmes marqués par leurs propres lignes éditoriales et contraintes.

Prenons l'exemple de NewsGPT. NewsGPT est un site internet et une newsletter qui rapportent des brèves d'informations sur l'actualité américaine et mondiale. Le contenu se génère automatiquement selon des sources et des images rencontrées par l'algorithme au cours de la journée (sans aucune considération de la propriété intellectuelle). Les sources en question ne sont pas rendues transparentes, à l'instar de ChatGPT qui ne sait pas non plus partager ses sources, n'étant pas construit avec cet objectif.

NewsGPT a été lancé au début de l'année 2023 et présente sur sa page d'accueil la mention suivante : « Réelles, non biaisées, informations rédigées par IA ». Soit un pitch reposant sur une information « réelle et



Journalism, Media, and Technology Trends and Predictions 2023, Reuters.

« On observe pour Eto News, une différence avec ceux mentionnés précédemment : les articles indiquent leurs sources directes, et il existe la possibilité de contacter une équipe humaine en cas d'erreur dans les articles. »

non biaisée ». Mais, cette mention semble entrer quelque peu en contradiction avec les termes et conditions de NewsGPT.

Cette notion nous éclaire sur la façon dont le contenu n'est pas vérifié : il n'est pas revu régulièrement, il n'existe pas de vérification indépendante de ce qui est publié. Les créateurs ne supportent pas les opinions qui pourraient être émises et ne pourraient être tenus pour des défauts ou des inexactitudes présents dans les articles. On a donc ici, encore un peu de difficulté à savoir comment le contenu généré par IA pourrait permettre de lutter contre la désinformation, sans aucune intervention humaine. Les précautions prises témoignent par ailleurs d'un projet qui n'est pas tout à fait abouti. Il existe d'ailleurs encore assez peu d'abonnés à cette initiative.

Un autre site a récemment attiré l'attention des commentateurs : celui de *Generative Press*. Le site d'information se présente également comme un site de brèves d'information, parfois plus originales car elles sont rédigées par quatre « personnalités » différentes (générées par IA) de journalistes.

Les informations sont produites à partir d'informations « popu-

Article 5 des termes et conditions

"Nous n'examinons pas régulièrement le contenu et nous ne nous efforçons pas de vérifier de manière indépendante toute allégation factuelle. Nous n'approuvons pas les opinions qui pourraient apparaître dans le contenu. Bien que nous nous efforcions de mettre à jour le contenu, nous ne garantissons pas qu'il sera à jour ou exact à tout moment(..)

Extrait des conditions d'utilisation du site NewsGPT.

lares, de tweets des journalistes et de citoyens vérifiés ». Il n'existe ni équipe éditoriale, ni information produite par des humains.

Les derniers articles de *Generative Press* remontent à janvier 2023. Le site a été lancé par deux ingénieurs, qui reconnaissent eux-mêmes qu'il s'agit plus d'un site conçu à des fins de divertissement, qu'à un vrai site d'information.

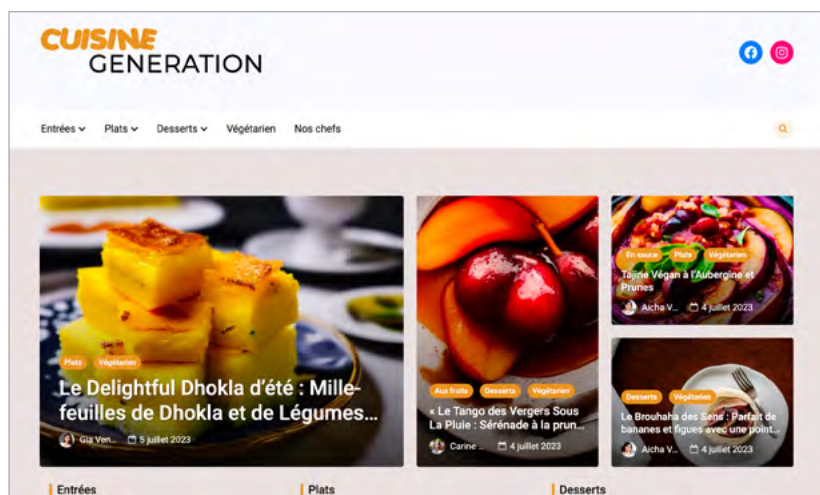
Eto News est aussi un média synthétique qui a été créé pour s'orienter vers de l'actualité politique. Il est hébergé sur Substack et se présente comme « un service sélectionnant trois à cinq organes d'information de l'ensemble du spectre politique et résumant leurs articles de manière impartiale et factuelle à l'aide du tableau de partialité des médias. Pour ce faire, nous utilisons *Eto*, un outil développé par ChatGPT. Notre objectif est de fournir aux lecteurs des informations factuelles, afin qu'ils puissent se forger leur propre opinion, à l'abri de l'influence des pro-

grammes politiques ou des médias à sensation. Avez-vous remarqué des erreurs ou des biais dans notre travail ? N'hésitez pas à nous faire part de vos commentaires par courrier électronique ! ». On observe pour *Eto News*, une différence avec ceux mentionnés précédemment : les articles indiquent leurs sources directes, et il existe la possibilité de contacter une équipe humaine en cas d'erreur dans les articles.

DES SITES D'INFORMATION GÉNÉRÉS PAR IA QUI POURRAIENT CAPTER DES SEGMENTS PUBLICITAIRES

Dans un rapport récent de juin 2023 de *NewsGuard* adressé au MIT Review Technology, l'organisation explique que l'IA générative aujourd'hui, offre de nouveaux moyens d'automatiser les processus d'exploitation de contenus, et donc de créer des sites indésirables. Certains sont « plus sophistiqués et plus convaincants que d'autres avec des photos et des

La création de son propre avatar IA revient à 1 000 dollars par avatar.



cuisine-generation.fr : un expériment de site internet entièrement géré par l'IA générative (GPT-4 et Stable Diffusion) mené par le chercheur en Innovation Ari Kouts.

biographies de faux auteurs générés par IA». L'objectif derrière ces sites, est d'attirer des annonceurs payants, qui ne savent pas toujours que leurs publicités se retrouvent sur ce type de site peu fiable.

«Il semble que la publicité programmatique soit la principale source de revenus de ces sites web générés par l'IA», déclare Lorenzo Arvanitis, analyste chez NewsGuard, qui a suivi le contenu web généré par l'IA. «Nous avons identifié des centaines d'entreprises classées au Fortune 500 et des marques bien connues qui font de la publicité sur ces sites et qui les soutiennent sans le savoir.»

À ce stade, il demeure encore difficile d'évaluer le gain de ces sites générés par IA à partir de ces publicités programmatiques. Hodan Omaar, conseiller principal en matière de politique de l'IA à l'Information Technology and Innovation Foundation, un groupe de réflexion situé à Washington DC, explique ainsi qu'il n'y a pas de solution facile, et que la

publicité est un pan du modèle économique d'Internet, mais ce constat appelle de nouveau en la matière, à solliciter une régulation extérieure.

LE JOURNALISTE : INDISPENSABLE GARDIEN DU MARCHÉ DU RÉEL

La présentation de ces différents médias synthétiques interpelle, a minima, sur deux aspects du métier de journaliste avant même d'évoquer les droits d'auteur : la pensée critique et l'indispensable vérification des sources. Sur ce dernier point, les expériences de rédaction ayant utilisé l'IA pour générer du contenu sans vérification humaine, ont vite montré leurs limites. C'est l'expérience notamment de CNET, qui avait dû mettre en pause les articles générés par de l'intelligence artificielle, après la révélation de « flagrantes coquilles » et d'un manque de transparence dans la signalisation des contenus générés par IA.

Buzzfeed a continué pendant plusieurs semaines à poster des articles dont le contenu était généré par IA : un assistant créatif par IA a même été lancé. Mais, les articles concernaient principalement des destinations de voyage. À se demander si utiliser l'IA ne mériterait pas une réflexion pour évoluer vers de nouveaux formats, au lieu d'essayer de copier (mal) l'existant.

Comme l'expliquait en juin 2023 Thomas Baekdal, analyste média : un système d'information qui ne se nourrirait que de l'IA, ne fonctionnerait pas. Il doit y avoir un journaliste humain qui, au départ, est en position d'injecter de l'information.

Aussi, l'existence des médias synthétiques ne remplace pas le journaliste. Ils l'invitent à retrouver sa position de gardien du réel, en adoptant telle ou telle pratique, en se démarquant par le contenu proposé : vérifié, sourcé et humain. Ils invitent cependant à questionner factuellement de futurs modèles économiques au regard des usages des utilisateurs. ■

IA GÉNÉRATIVE: UNE NÉCESSITÉ DE **FORMATION** **POUR LES** **JOURNALISTES**

PRIORISER LA FORMATION

L'IA générative est une avancée technologique majeure qui promet d'avoir un impact transformateur sur tous les métiers à France Télévisions. En tant qu'organisme de formation, France TV Université doit prendre en compte cette réalité et s'engager pleinement dans ce domaine en constante évolution.

Nous avons conscience que nous ne pourrons pas répondre à l'ensemble des questions concernant l'IA générative, mais nous avons également conscience que le sujet est devenu incontournable. En parallèle de la préparation d'un cadre défini par la direction de France Télévisions sur les usages possibles de l'IA générative dans les pratiques professionnelles, en particulier au sein des rédactions, nous avons pris l'initiative de travailler sur un projet de formation d'acculturation dédié aux journalistes, afin de leur fournir les bases de connais-

*Par Valentine Lopez,
responsable du département
information et Jean Chrétien,
directeur Délégué à France
TV Université*

L'IA générative va transformer les métiers des médias ; il devient alors primordial de former les journalistes à ces changements. France TV Université, l'organisme interne de formation de France Télévisions, développe des formations pour les journalistes et les autres corps de métier, afin de les familiariser avec les principes de l'IA générative et ses implications éditoriales, techniques et juridiques.

sances nécessaires pour comprendre et traiter les enjeux liés aux intelligences artificielles génératives.

L'objectif de cette formation est de permettre aux journalistes de se familiariser avec les principes

fondamentaux de l'IA générative, ainsi qu'avec ses implications éditoriales, techniques et juridiques. Nous cherchons à donner aux stagiaires les outils nécessaires pour comprendre comment ces technologies peuvent être utilisées à bon escient dans leur travail quotidien.

NAVIGUER DANS LE CHANGEMENT

Le métier de journaliste est particulièrement concerné par l'impact de l'IA générative, car il touche à tous les aspects de la production de l'information, de la distribution des contenus aux risques potentiels de désinformation. Pour aider les rédactions à naviguer dans ce changement, il est essentiel de mettre en place un cadre clair et de fixer des règles pour l'entreprise. Cela peut être réalisé par le biais d'une charte ou d'un guide de bonnes pratiques spécifiquement adaptés à l'utilisation de l'IA générative dans le journalisme.



Nous pourrions envisager de développer une charte en collaboration avec d'autres médias français ou européens, sur le modèle de la « Charte pour un journalisme à la hauteur de l'urgence climatique ». Cette initiative permettrait de créer un cadre éthique et des lignes directrices communes, afin d'encourager une utilisation responsable de l'IA générative dans le domaine du journalisme. Il est essentiel de souligner que cette charte devra être régulièrement révisée et mise à jour en fonction des avancées technologiques et des nouvelles problématiques qui émergent.

ÉVITER LA FRACTURE NUMÉRIQUE

La fracture numérique est une préoccupation importante dans le contexte de l'adoption de nouvelles technologies. Il y aura toujours des individus qui s'approprient naturellement ces nouveaux outils et ceux qui sont réticents ou peu

familiers avec ces technologies.

L'information et la formation sont des moyens efficaces mais elles ne peuvent pas tout résoudre. À France TV Université, nous construisons des parcours pédagogiques tournés vers la pratique au plus près de la réalité de terrain, ce qui peut aider les personnes en fracture à dépasser leur blocage et intégrer ces nouvelles technologies comme ce qu'elles sont : des outils pour améliorer leur quotidien au travail. ■

Le métier de journaliste est particulièrement concerné par l'impact de l'IA générative, car il touche à tous les aspects de la production de l'information, de la distribution des contenus aux risques potentiels de désinformation.

SXSW 2023 :

WHO CARES ABOUT REALITY ?

Que l'on appelle cette nouvelle façon de vivre avec l'intelligence artificielle Almosis (mot-valise étrange qui a raté son *momentum* après la présentation de la gourou des futuristes, Amy Webb), une nouvelle relation, une rApture, l'intimité artificielle ou encore une relAlté améliorée – **l'IA va changer le cours du monde, en tout cas, c'est ce que l'on aura compris après quelques jours à Austin. Alors que le métavers est plutôt parti en hibernation en attendant les avancées technologiques nécessaires, l'IA est définitivement passée en pleine canicule.**

Dans la session intitulée « Comment l'IA et le métavers vont façonner la société », Ian Beacraft, PDG et futurologue en chef de Signal and Cipher, n'a accordé au métavers qu'une brève mention. Et encore, il attribue le succès futur du métavers à l'IA : « Le métavers a besoin de l'IA pour devenir ce qu'il va devenir ». « Chaque aspect de la vie sera en quelque sorte amplifié par [l'IA] », a déclaré de son côté Greg Brockman, co-fondateur et président d'OpenAI, créateur de ChatGPT, « ce sera un outil, au même titre que le téléphone portable dans votre poche ». Avec les LLMs, l'IA a été équipée d'une bouche et d'oreilles pour mieux interagir avec nous. Et comme l'a remarqué à juste titre le même co-fondateur d'OpenAI : « Tout

Par Kati Bremme, directrice de l'Innovation France Télévisions et rédactrice en chef de Méta-Media

Le flou autour de l'avenir de la créativité, du travail ou tout simplement de notre façon d'affronter la réalité s'épaissit à SXSW avec l'IA, qui a remplacé cette année dans les discussions, le métavers et autres web3 (dont on retrouve quand même des traces sous 365 définitions différentes). Quand, à l'époque des réseaux sociaux et du smartphone, tout le monde pouvait être journaliste, aujourd'hui avec les IA génératives, tout le monde peut être artiste. Dans un esprit très communiste, au-delà de toute considération de propriété intellectuelle, l'IA devient créatrice.

est langage » – le marché pour son produit est donc infini (surtout si l'on oublie l'Open Source...).

Par ailleurs, pendant cette semaine d'annonces de rupture autour de l'IA, tout le monde semblait un peu dépassé par l'actualité.

A-T-ON ENCORE BESOIN DES JOURNALISTES ?

Après une session d'ouverture avec des journalistes qui se prennent pour des futuristes (« Les futurs sont faciles à inventer. Surtout, les futurs peuvent être élaborés avec les gens »... Ou avec ChatGPT), la BBC, l'Associated Press et Twitter se sont retrouvés pour une séance plus professionnelle autour de « L'IA dans les rédactions : quel impact sur le journalisme ? ». Aimee Rinehart d'Associated Press, pionnier de l'IA pour les journalistes depuis 2014, a mis en avant leur façon d'aider les rédactions locales (dotées de peu de moyens) à être plus efficaces (notamment avec Whispr et la technologie GPT). Laura Ellis, Head of Technology Forecasting à la BBC, a détaillé le rôle des moteurs de recommandation, de l'équilibre entre personnalisation et sérendipité, et de l'importance d'intégrer les bonnes règles et standards dans les modèles d'IA.

Face à l'accélération des évolutions technologiques, il devient difficile d'anticiper les bonnes lignes de conduite surtout quand on est confronté à un changement de paradigme de taille : **autrefois, la technologie était utilisée pour délivrer du contenu, aujourd'hui le médium devient encore plus le message.**



SXSW Art Program / Quantum Jungle by Robin Baumgarten.

Tout le monde était d'accord sur l'importance de la transparence autour de l'utilisation de l'IA dans les rédactions, d'ailleurs, pour Aimee Rinehart : « La transparence est la nouvelle objectivité pour les journalistes ».

Dans sa table ronde « L'avenir de la créativité alimentée par l'IA », le CEO de BuzzFeed, Jonah Peretti, a fait l'annonce de son tout premier format de jeu alimenté par l'IA qui « révolutionne la consommation de contenu et le divertissement ». Un chatbot, comme on en a déjà vu des centaines, mais peut-être un peu plus impertinent que ses prédécesseurs. Pas vraiment l'avenir du journalisme et de la créativité.

Des contenus personnalisés, il y en a pour le divertissement aussi. Justement, c'est une des opportunités que Greg Brockman d'OpenAI voit pour son IA : « Peut-être que les gens sont

encore bouleversés par la dernière saison de *Game of Thrones*. Imaginez que vous puissiez demander à votre IA de créer une nouvelle fin et peut-être même de vous y intégrer en tant que personnage principal... ». Comparant sa technologie à un groupe d'« assistants » qui ne sont pas parfaits, mais qui sont « enthousiastes et ne dorment jamais », le co-fondateur d'OpenAI a déclaré que ChatGPT pourrait aider à faire le « travail de corvée » pour l'écriture et le codage, mais qu'il aurait également la capacité d'ajouter une expérience de divertissement plus « interactive ».

Dans « Des deepfakes à DALL-E : de vraies règles pour de faux médias », Claire Leibowicz de Partnership on AI et Andy Parsons d'Adobe ont discuté avec d'autres sur l'éventail des possibilités et l'impact des deepfakes et autres médias générés par l'IA, pour se mettre d'accord sur le fait que les

médias synthétiques sont une nouvelle technologie, mais ne créent pas de nouveaux problèmes. L'occasion de présenter le livre blanc sur la bonne utilisation des médias synthétiques, établi avec la BBC, Adobe, Bumble et CBC Radio Canada.

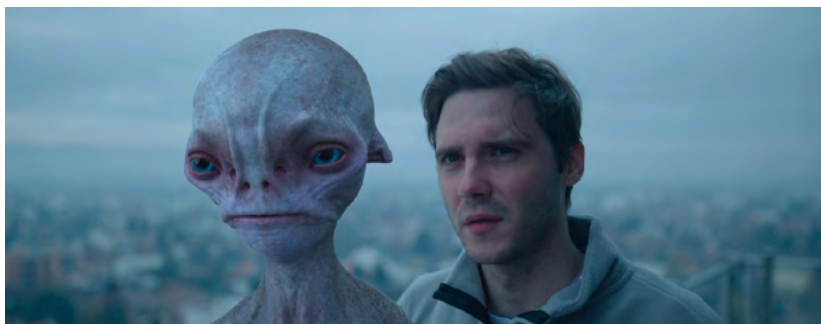
En conclusion, Adobe remarque très justement que nous parlons beaucoup d'IA responsable mais nous ne disons jamais QUI est responsable.

Shutterstock a intégré son outil de génération d'images avec ses propres images pour rester du bon côté de la discussion autour de la propriété intellectuelle.

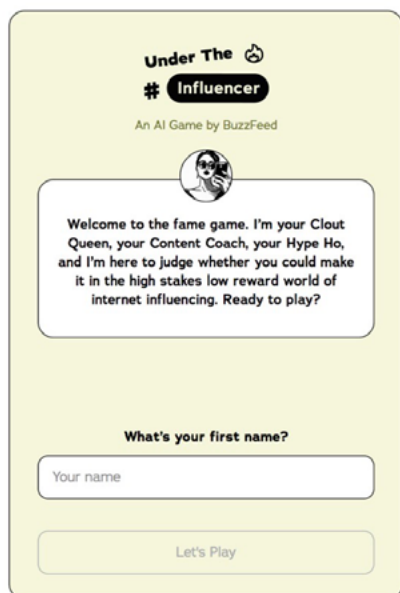
Un travail d'explication, de formation, de compréhension est à lancer au plus vite dans les rédactions, pour ne pas se faire dépasser par les algorithmes. Selon certains, **souvent ce n'est pas ChatGPT qui est meilleur que l'article, c'est l'article qui est moins bien que ChatGPT...**

TRAVAILLER AVEC L'IA, OU NE PLUS TRAVAILLER DU TOUT

Après nous avoir laissé le soin de remplir Internet pendant des années, l'IA est en train de nous restituer toute notre connaissance sous forme de dialogue interactif qu'il s'agit de maî-



Capture du film projeté par Ian Beacraft pour illustrer la capacité de l'IA à intégrer des caractères dans des films par simple prompt, image proposée par Wonder Dynamics.



Capture d'écran du jeu Under the influencer de BuzzFeed.

triser. Elle automatise des tâches et crée de nouveaux emplois, avec un impact sur toutes les industries. L'une des premières sessions de South by Southwest 2023 a traité d'un sujet au cœur des préoccupations de l'industrie technologique depuis que les inquiétudes suscitées par la pandémie se sont estompées : **comment la main-d'œuvre se transforme-t-elle dans un contexte d'énormes changements technologiques et culturels ?**

Ellen McGirt, rédactrice en chef de *Fortune*, y a interrogé Chris Hyams, PDG d'Indeed sur les technologies qui perturbent l'emploi. Pour ce dernier, cette fois-ci, « tout va plus vite » : les annonces et avancées sont exponentielles et seront plus impactantes que tous les changements précédents. **Certains de ces changements pourraient prendre une telle ampleur et se produire si rapidement (ie. l'IA) qu'ils entraîneront des bouleversements sociétaux irréversibles.**

Charlotte A. Burrows, de la Commission américaine pour l'égalité des chances en matière d'emploi (EEOC), s'est également intéressée à ces questions dans sa table ronde « L'IA est-elle la nouvelle RH ? Protéger les droits civils au travail », alors que 99 % des entreprises Fortune 500 ont intégré des outils automatisés pour la sélection de candidats. Dans un écosystème où le Myers-Briggs est toujours utilisé, malgré son caractère

éminemment problématique, désormais on y ajoute de l'IA.

Dans la session « Le stagiaire et partenaire universel : comment l'IA générative change notre façon de travailler », Kevin Kelly, Senior Maverick chez *Wired*, nous a fait comprendre que la plupart de nos activités n'étaient en fait pas si intelligentes que cela (reconnaître une image, jouer aux échecs, mener une conversation...). Notre calculatrice est plus intelligente que nous (que moi c'est sûr) en arithmétique. Nous aurions en fait une compréhension très faible du fonctionnement de notre propre intelligence.

Pour lui, il s'agirait de « dumbs-marten » (rendre stupide et intelligent à la fois) les IA : on ne pourra pas tout maximiser dans toutes les dimensions (adieu l'AGI ?), on utilisera donc différentes IA pour différentes spécialisations. Une situation qui permet à Ian Beacraft d'introduire l'ère du Creative Generalist au travail. L'intelligence artificielle peut désormais dépasser quiconque dans un domaine spécifique. **Ceux qui ont un profil généraliste et curieux auront donc une chance de dominer cette nouvelle ère, en attendant l'avènement de l'AGI (Artificial General Intelligence, l'intelligence artificielle forte).**

Faisons peut-être attention aux domaines que l'on abandonne à l'IA : le langage ne devrait pas être une spécialisation mais une compétence à la portée de tous. Lors de l'événement Future Media Hubs, la start-up allemande SUMM AI a présenté sa

La transparence est la nouvelle objectivité pour les journalistes.

Aimee Rinehart d'Associated Press

solution de traduction de langage classique en langage simplifié. Selon la start-up, 54 % des gens ont un niveau d'alphabétisation inférieur au niveau de 6^e. Ce chiffre comprend aussi des personnes qui ne souffrent pas d'un handicap réel. C'est bien beau de demander à l'IA de nous simplifier le langage, mais on a un peu l'impression de prendre le problème à l'envers : alors que l'IA devient de plus en plus érudite, allons-nous nous contenter de futures générations qui auront besoin que l'IA leur explique le monde parce qu'elles sont incapables de le déchiffrer ?

VIVRE AVEC L'IA ET ENTRER DANS UNE CRISE D'IDENTITÉ

Dans les innombrables sessions sur l'IA générative, les discussions sont allées bien au-delà de son potentiel créatif immédiat. « Il ne s'agit pas seulement de créer avec l'IA. Il s'agit d'établir une relation avec l'IA », a déclaré Ian Beacraft. « Nous allons en fait avoir une sorte de relation avec elle », ajoute-t-il, ce qui conduira à terme à des « interactions où nous ne pourrions pas faire la différence entre un humain et une IA ». Seul problème, l'IA peut se retrouver dans à peu près tous les mots, comme le remarque John Maeda, VP of Design and AI chez Microsoft (celui qui avait proposé d'ajouter un « A » pour Arts dans STEM) en introduction de sa keynote, mais pas dans le mot « éthique ». **On entrerait donc dans une crise d'identité, saupoudrée de graves problèmes éthiques.**

« Il ne faut pas que l'IA nous soit imposée...
C'est nous qui sommes aux commandes.
Ne l'oublions pas.

Rahul Roy-Chowdhury

2024 sera la dernière année d'élections « humaines » dicit Greg Brockman d'OpenAI. Interviewé par Laurie Segall, le co-fondateur d'OpenAI rappelle un peu Mark Zuckerberg devant le Congrès : à chaque question qui touchait de près ou de loin à des enjeux éthiques, il répondait « *That's an interesting question* », sans pourtant lever le flou qui entoure la façon dont nous vivons avec ce nouveau compagnon plus créatif que nous-mêmes et, qui exploite l'ensemble de nos données. De plus, il a préféré annoncer le prochain niveau de son GPT sur YouTube le mardi plutôt qu'à Austin le lundi...

Ian Beacraft a donné un exemple anecdotique pour illustrer l'idée de cette nouvelle crise d'identité avec une vidéo de Keanu Reeves dans laquelle il explique le film *Matrix* à une enfant de 13 ans qui ne l'avait jamais vu. Il a souligné que le personnage dans le film « s'interroge sur ce qui est réel et ce qui ne l'est pas, et il veut vraiment savoir ce qui est réel ». La jeune fille lui répond alors :



Keanu Reeves : « *Who cares if it is real?* »

« Qu'est-ce que ça peut faire si c'est vrai ? ». Keanu Reeves lui demande : « Tu t'en fiches que ce soit vrai ? », et elle de répondre : « Oui ».

Après des rires incrédules et nerveux suite à la diffusion de la vidéo, Ian Beacraft a insisté sur le fait qu'il s'agit de notre future réalité. « Nous sommes arrivés à un stade où nous devons comprendre combien des amis de la Gen Alpha sont réels et combien sont synthétiques ». « Et franchement, [selon lui], cela peut avoir ou ne pas avoir d'importance. »

MÉTAVERS : PRINTEMPS, ÉTÉ, AUTOMNE, HIVER

En parlant d'identités virtuelles, nous ne sommes peut-être pas dans l'hiver de la XR, mais plutôt en automne. Un moment de stase de l'écosystème, où nous espérons un nouveau saut technologique.

Mais en attendant, le métavers dans sa version gaming sans casque n'est pas mort : YLE, la télévision publique finlandaise s'attendait dans son projet *Roblox Independence Day* à une centaine de visiteurs et s'est finalement retrouvée avec 23,8 k visites en trois heures. Pour Kevin Kelly, le vrai superpouvoir de l'IA n'est d'ailleurs pas d'écrire un texte ou de créer une image, mais de générer des univers 3D. Les Assistants Personnels Universels (UPA) pourront bientôt créer des métavers.

En attendant les évolutions techniques, on a le temps d'acquiescer la nécessaire « Metaliteracy » pour les

appréhender de façon sereine, et dans tous les cas, **selon Sofie Hvitved du Copenhagen Future Institute, l'avenir du métavers se trouve dans le monde physique.**

La keynote de Philip Rosedale, fondateur de Linden Lab AKA Second Life, était presque vide (cf. l'hibernation du métavers). Il a donc aussi raté l'occasion d'annoncer la sortie en version mobile de son ancêtre du métavers. On s'est demandé d'ailleurs qui était dans la salle. À la question si « vous êtes excité par le festival de musique de Roblox ce week-end » : seule une personne a levé la main...

Dans la session « La mentalité du métavers pour le Web3, l'IA et l'avenir de l'entreprise », Sandy Carter d'Unstoppable Domains (qui a son QG dans le métavers, sans porte d'entrée suite à un problème d'UX), a partagé dix cas d'usages B2C et huit cas B2B (ils existent donc). Surtout si on applique sa définition de métavers extrêmement large : « un monde digital où on vit, travaille, interagit et joue, et où on a une présence sociale depuis partout à tout moment », auquel on ajoute un peu de sauce IA (du Manifeste Web3 d'IWC et sa communauté très exclusive de jetons NFT au métavers de John Deere pour tester des tracteurs) et l'importance de la communauté.

De plus, elle a invité Brian David Gilbert de Polygon pour nous rappeler les évolutions du web : Web1 = souris, Web2 = smartphones, Web3 = avatars ; et pour annoncer (vendre) le lancement de Polygon, un nom de domaine à « l'intersection de la communauté et de l'identité », la première



Douglas Rushkoff lors de SXSW.

adresse native de Polygon lisible par les humains...

D'ailleurs, pour remporter un prix de XR à SXSW, on n'a pas forcément besoin de XR : *Consensus Gentium*, de la scénariste et réalisatrice Karen Palmer, a ainsi remporté le premier prix des XR Awards de SXSW, même s'il ne s'agit pas techniquement d'une expérience XR. Un prix spécial du jury a été décerné à *Body Mine* de Cameron Kostopoulos, qui permet à un homme d'habiter le corps d'une femme, ce qui constitue un argument de poids en faveur de la VR en tant que machine à empathie...

50 NUANCES DE WEB3 (OU 2.5) COLLABORATIF, À NE PAS CONFONDRE AVEC LE WEB 3.0 SÉMANTIQUE

Face à l'effondrement de la Silicon Valley Bank en pleine semaine de SXSW, Douglas Rushkoff, l'un des « dix intellectuels les plus influents du monde » selon le MIT, annonce « la fin de l'esprit milliardaire » dans les start-ups de la Tech. En effet, le paysage technologique actuel n'est pas au mieux de sa forme : toutes les grandes entreprises (Meta, Microsoft, etc.) licencient. Il est bien sûr plus difficile de trouver des opportunités lorsque les entreprises et les investisseurs sont plus prudents dans leurs dépenses.

Selon lui, maintenant que le battage médiatique autour du web3 s'est un peu calmé, la question importante ne semble plus être : « Comment déployer une campagne web3 le

plus rapidement possible ? » mais plutôt : « Comment pouvons-nous déployer une campagne web3 en toute sécurité, efficacement et d'une manière qui apportera de véritables avantages à notre public ? ». Sous cet angle, le sujet commence donc à être intéressant pour les médias de service public.

Un objet web3 directement affecté par la chute de la SVB : Artifact, la nouvelle application d'actu personnalisée des co-fondateurs d'Instagram. Son financement a été pris dans la faillite de la Silicon Valley Bank, et le co-fondateur Kevin Systrom pense que d'autres problèmes pourraient survenir dans la Silicon Valley. Dans son entretien à SXSW, il s'étale sur le trop-plein de spéculation autour du web3 et la mauvaise presse qu'à la Tech, mais donne très peu d'informations sur Artifact.

L'une des 365 définitions du web3 est celle avec de la blockchain et des choses que l'on fait en ligne, et c'est celle-ci que Molly White, créatrice du site « *Web3 is Going Just Great* » (oui, selon elle, toujours) préfère utiliser. Au-delà du web3, elle a partagé aussi une analyse intéressante du rôle des journalistes « mainstream » (comme Gerrit De Vynck face à elle qui l'interviewait) versus les « citizen journalists » (comme elle) : **selon elle, les journalistes citoyens n'ont pas les compétences nécessaires pour couvrir les événements (vérifier les sources, les rumeurs, etc.), tandis que les journalistes traditionnels n'ont pas de connaissances approfondies du web3. Conclusion : il**

faudrait donc qu'ils travaillent mieux ensemble. En y ajoutant un autre élément particulièrement révolutionnaire pour le public US : la recherche technologique devrait bénéficier d'un financement public...

ET SI LES CONTENUS EUROPÉENS ÉTAIENT LA NOUVELLE K-POP ?

Lors de l'événement du Future Media Hubs, Evan Shapiro, génial cartographe médias, Sander Saar, responsable de la stratégie de Red Bull Media House et Johan Oomen, « Digital vacuum cleaner », ont évoqué l'avenir des médias américains et européens. Mr Beast et d'autres créateurs sont en train de tuer le jeu du contenu en fabriquant plus de propriété intellectuelle que tous les médias associés, mais **il y aurait une vraie appétence pour les contenus européens aux US** (selon Shapiro, les plateformes américaines auraient sous-estimé l'intelligence de leur public). Il dit aussi que Roblox et Minecraft sont les nouvelles plateformes de divertissement, et qu'il ne comprend pas pourquoi toute l'Europe n'a pas tout simplement adopté le player de la BBC.

Face à la stratégie des jardins clos des plateformes de streaming aux US, qui se voient confrontées à une vague de *churn* (notamment sur fond de crise économique) – seulement 7 % des personnes interrogées dans un récent sondage de l'entreprise de Shapiro ont déclaré qu'elles maintiendraient la situation actuelle en

« Il ne s'agit pas seulement de créer avec l'IA.
Il s'agit d'établir une relation avec l'IA. »

Ian Beacraft

matière d'abonnement – , ce dernier a remarqué « que nous sommes en train de redécorer une salle de bains dans une maison qui est en feu »...

L'Europe a une carte à jouer dans sa revanche culturelle sur les US, surtout, si elle arrête de se disperser avec des plateformes limitées aux pays membres (je cite Shapiro).

Il y avait aussi à Austin des robots très étranges chez Disney et les premières chaussures d'économie circulaire du monde, que l'on peut planter dans son jardin...

CONCLUAISON

Peut-être que l'on devrait tous prendre un grand « sabbatical », faire relâche et abandonner le monde à la nouvelle religion IA en consommant quelques champignons psychédéliques (autre star du festival à Austin, dans un état pourtant connu pour sa législation stricte des drogues). Ou mettre l'accent sur l'amélioration et le renouvellement de nos compétences. Adopter l'apprentissage tout au long de la vie et l'adaptabilité. Développer des compétences humaines uniques (par exemple, la créativité [sic ?], l'empathie, la pensée critique) pour rester pertinents face aux machines.

À SXSW, les multiples flyers Web3 de l'année dernière ont été remplacés par ceux sur l'IA, dont un pour une application de rencontre alimentée par l'intelligence artificielle qui a utilisé le slogan « Qui est mon papa ? » pour promouvoir son générateur d'images de bébés de célébrités, AKA les titans de la technologie Jeff Bezos et Elon Musk... Définitive-

ment #Uncanny et pas vraiment un futur souhaitable.

En tout cas, on est d'accord avec Rahul Roy-Chowdhury, Global Head of Product de Grammarly, qui a répété à plusieurs reprises qu'il préférerait le terme « intelligence augmentée » à celui d'« intelligence artificielle », estimant que le premier met l'accent sur le pouvoir des algorithmes d'améliorer, plutôt que de remplacer, la créativité et les capacités des êtres humains.

« Nous voulons que ces outils nous aident ». « Il ne faut pas que l'IA nous soit imposée... C'est nous qui sommes

aux commandes. Ne l'oublions pas. »

Il est primordial de chérir notre capacité humaine à rechercher et à évaluer de manière critique les connaissances, notamment au regard des dommages que les entreprises technologiques ont déjà causés aux modes de diffusion de l'information, et aussi pour maîtriser les nouvelles réalités co-crées avec l'IA.

Face à la menace transhumaniste et la fin du privilège humain sur le langage, tout revient à la baseline historique d'Apple : Think different ! Car l'IA est désormais *Everything Everywhere All at Once*. ■



SXSW 2023 à Austin.

CHANGEMENTS DE
POU-
VOIR

TECHNO-POLITIQUE DE L'IA :

LUTTES IDÉOLOGIQUES, TENSIONS GÉO-POLITIQUES, ESPOIRS DÉMOCRATIQUES

La façon dont le débat public s'est emparé de la question de l'intelligence artificielle laisse parfois songeur : souvent comprise comme un « bloc », un « concept » ; elle nous est présentée comme une perspective, quelque chose à venir, à craindre. Un horizon fantasmatique : « pour ou contre », « danger ou progrès », « l'homme ou la machine », etc. Pourtant, le sujet, hautement complexe, ne peut se contenter de cet excès de simplisme. L'intelligence artificielle se situe à un carrefour interdisciplinaire stratégique : data science, philosophie politique, géopolitique, droit, économie, politiques publiques. En tant que technologie totale, elle exige une pensée totale.

INTELLIGENCE ARTIFICIELLE : VRAIS SUJETS, FAUX PROBLÈMES

La question technologique sous-tend par définition les luttes de pouvoir, les jeux de puissance, elle en est l'un des véhicules. Même la technologie low tech, dirions-nous aujourd'hui, du papyrus fut une « révolution » en son temps ! Au XXI^e siècle, le changement de paradigme passera notamment par l'intelligence artificielle. De façon directe ou indirecte, explicite ou implicite, c'est l'IA comme domaine scientifique et industriel omni usages ainsi que l'ensemble des ramifications du méta-système qu'elle induit, qui

Par Asma Mhalla, spécialiste des enjeux politiques et géopolitiques de la Tech, enseignante à Sciences Po, Polytechnique et Columbia

Les débats autour de l'intelligence artificielle se multiplient depuis plusieurs mois, favorisés par l'utilisation à très grande échelle de ChatGPT. Sommes-nous les contemporains d'une révolution dont l'agent conversationnel serait le leader et le héraut ? Pour Asma Mhalla, si ChatGPT marque un nouveau tournant pour l'intelligence artificielle, la révolution a déjà eu lieu. En matière d'IA, la question centrale ne serait plus tant celle du « comment » ou du « quand » mais bien celle du « qui » : qui en détient la vision ? La conception ? Et selon quels intérêts ? Face à la Chine et aux États-Unis, l'enjeu est non seulement géopolitique mais également idéologique. À cet égard, souligne l'auteure, il revient à la France et à l'Europe d'élaborer une véritable stratégie techno-industrielle.

cristallise, par exemple, la rivalité stratégique entre États-Unis et Chine. **La lutte, pour ne pas dire la guerre technologique, est féroce, elle se joue sur trois fronts : idéologique, géopolitique et militaire, ainsi que philosophique.**

- **Sur le plan géopolitique, la compétition sino-américaine se joue en particulier sur la question des IA à usages militaires.** De ce point de vue, l'intelligence artificielle transforme profondément la chose militaire. La guerre en Ukraine nous en offre les prémices, celles des guerres du futur et de la suprématie de l'une ou l'autre des hyperpuissances.

- **Dans le champ idéologique, la technologie n'a de valeur qu'en cela qu'elle encapsule des normes, une vision du monde.** Celui qui détient la technologie en contrôle l'intention politique. Cette question s'articule à deux niveaux : entre nations mais aussi entre concepteurs.

- **Enfin, elle pose une question philosophique universelle cruciale : face à ces performances calculatoires de plus en plus puissantes, où se loge notre singularité d'Humain ?** Comment préparer cet avenir par nature imprévisible ?

Dans ces moments de mutations historiques qui déstabilisent nos repères et nos équilibres, il faut urgemment



démystifier l'IA pour recentrer la réflexion collective vers l'essentiel, mettre de l'ordre dans la façon dont plus généralement la question technologique est abordée.

L'IA incarne de façon éclatante notre éternel problème politique qui était, est et sera toujours de savoir qui domine et qui est dominé. Au V^e siècle avant notre ère, Thucydide le formulait déjà ainsi au moment de la guerre du Péloponnèse : « Les hommes tendent, selon une nécessité de leur nature, à la domination partout où leurs forces prévalent ».

L'intelligence artificielle est et sera partout. Il y a bien une dynamique irrépessible qui est en marche, qui articule à la fois la recherche de pouvoir et de puissance. Or c'est bien cette soif-là, très humaine, qui risque de nous dépasser et non pas (ou du moins pas encore) l'outil lui-même. **Cela pose encore une fois la seule question qui vaille, celle du « qui » : qui détient la vision ? Qui la conçoit, c'est-à-dire qui détient la capacité à donner corps à cette vision ? Selon l'intérêt de qui ou de quoi ? En contrechamp, sur quel projet politique commun devrait-on aligner ce nouveau socle non pas seulement technologique mais aussi civilisationnel ?**

LA QUESTION IDÉOLOGIQUE : TECHNOLOGIE TOTALE X RISQUE

La question du « qui » nous emmène vers une analyse du sujet sous l'angle idéologique. Ces dernières années, nombre de critiques ont été formulées, de tous bords, sur notre panne idéologique, remplacée par une approche purement technocrate de la chose politique. Or ce n'est pas tout à fait vrai. La dimension idéologique du monde qui advient, la réflexion sur les contours et la morphologie de notre futur sont belle et bien préemptées. L'épicentre se situant à la fois entre Oxford University et la Silicon Valley.

Les technologues de la Silicon Valley sont d'abord des idéologues, ils ont un projet politique, un horizon, un moteur. La technologie, d'une certaine façon, n'en est que le prétexte et le chemin. Le problème posé est double.

La conversation mondiale a bien lieu, portée haut et fort par Elon Musk, Peter Thiel ou Sam Altman,

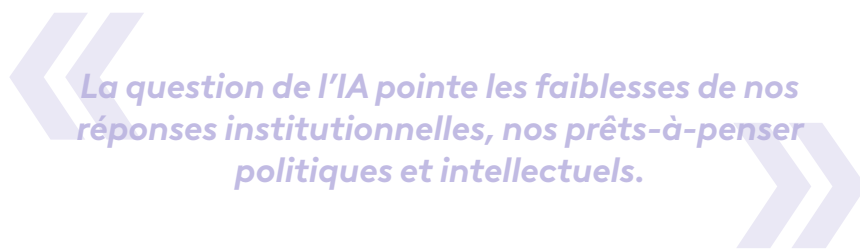
David Holz (PDG de Midjourney) plus récemment et consorts. Le vide idéologique dont nous nous étions tellement plaints est ici comblé et les propositions sont à la hauteur du vide : intelligence artificielle générale, transhumanisme, long-termisme. Nous ne sommes pas déçus. Il est intéressant de noter que **les technologues américains ne privatisent pas simplement la technologie mais ce faisant mettent en mot une privatisation plus subtile, une privatisation idéologique d'une certaine façon.**

Les concepteurs des systèmes d'intelligence artificielle proposent des agendas très précis via un système de ce que l'on peut appeler un projet de « Technologie totale ». Ils « désignent » notre futur par leurs outils, ils appellent cela le « long-term thinking ». Sam Altman, patron d'OpenAI, est l'un de ceux qui explicitent et revendiquent le mieux cette nécessité à la fois business et idéologique de penser le futur, de projeter les technologies sur le long terme. Cela renvoie sans conteste au « Triptyque des Big Tech ».



Celui qui détient la technologie en contrôle l'intention politique.





La question de l'IA pointe les faiblesses de nos réponses institutionnelles, nos prêts-à-penser politiques et intellectuels.

La question de la Technologie totale est à comprendre ici à deux niveaux :

- **Technologie totale en ce sens qu'elle enferme nos usages autour de quelques interfaces privées**, nous encerclent cognitivement, ce que Tim Berners Lee appelait déjà en 2007, les « Walled Gardens ».

- **Technologie totale, car elle met en place les briques d'une captation et d'une exploitation des usages et des données totales**, sans couture avec son corollaire, des dispositifs de surveillance et de monitoring généralisés.

Sam Altman a d'ores et déjà commencé la collecte de données biométriques mondiales avec le programme WorldID sous prétexte de combattre les bots et les faux comptes. Sur le volet transhumaniste, il est un « siliconien » tout ce qu'il y a de plus classique, biberonné aux désirs démiurgiques des gourous du transhumanisme, de ses amis Peter Thiel ou son nouveau meilleur ennemi Elon Musk. Lui aussi rêve de combattre le vieillissement et la mort, et vient d'investir 180 M\$ dans Retrobiosciences.

En réponse à ChatGPT, première brique d'un projet d'AGI plus grand porté par Altman, Elon Musk a lancé son offensive via d'abord un moratoire plus que discutable, hypocrite, voire dangereux, car invisibilisant les risques déjà bien réels posés par l'IA mais surtout soutenu par le Futur of Life Institute, institution promouvant une idéologie particulière et discutable que l'on appelle le « long-termisme ». Le long-termisme

implique que la priorité de nos décisions aujourd'hui doit être la pérennité de l'espèce humaine. Ce courant de pensée, très en vogue à la Silicon Valley et qui compte comme adeptes Thiel ou Musk, fait l'objet d'une critique virulente.

En parallèle du moratoire qui appelait à une absurde « pause » de six mois dans la R & D de l'IA, Elon Musk lançait au même moment son dernier projet X.AI, censé concurrencer frontalement OpenAI, dont il fut – ironie du sort – cofondateur en 2015. Le prétexte de la bronca anti-ChatGPT est qu'il poserait un risque existentiel, le désormais célèbre « risque », à l'Humanité. La critique idéologique doit être néanmoins lue et comprise à plusieurs niveaux, plus proche de nous le combat est nettement plus pragmatique : **Musk, chantre du « freedom of speech » maximaliste, suppose que l'entraînement et les filtres de ChatGPT en feraient un outil « woke ». Dans sa bataille contre le « woke capital », son idée est donc, aussi, de développer une IA-anti-woke ou pour aller encore plus loin, une IA recherchant une forme de vérité totale et absolue, sobrement baptisée « TruthGPT ».** Woke, pas woke, étroite (narrow AI), générale (AGI) ou super (Super AI), la guerre de tranchées des IA a bel et bien commencé sur fond inquiétant de post-vérité.

Cette logique à la croisée des idéologies long-termiste et transhumaniste est cristallisée par **son autre startup Neuralink dont l'une des expérimentations consiste en des**

implants dans le cerveau pour créer des interfaces directes esprit-machine tout en augmentant sensiblement les capacités cognitives humaines, à des visées thérapeutiques mais aussi disons, existentielles (cela permettrait à l'homme de concurrencer l'hypothétique avènement des AGI). Au moment de la publication du moratoire, Elon Musk annonçait sans complexe le démarrage prochain de tests d'implants Neuralink sur les hommes.

Sous ce vernis technologique et rationaliste, on voit déjà poindre l'irraison démiurgique de ces projets. L'excès de raison risque rapidement de devenir son opposé, une folie dystopique, car pour la première fois, ces idéologues ont les moyens de leurs ambitions.

Nos responsables politiques semblent encore relativement éloignés de ces enjeux technologiques profonds et semblent pour le moment plutôt enclins à une attitude à la fois normative et attentiste. Certes, le temps technologique n'est pas le temps politique mais si nous attendons trop longtemps, les gagnants (« the winner takes all ») rafleront la mise (en quantité : capital et nombre d'utilisateurs captifs, et en qualité : leurs idéologies infusent déjà et les usages dépolitisent déjà la question). Le politique arrivera une fois de plus après la bataille, sera une fois de plus soumis à la volonté d'acteurs privés prônant une vision politique à la fois claire, tranchée et arbitraire. Or, défaire les usages et les idées prendra bien plus de temps qu'y participer dès l'amont pour en infléchir l'horizon.



Les technologues de la Silicon Valley sont d'abord des idéologues.

LA VALEUR DE LA RÉPONSE NORMATIVE

En réponse, le premier réflexe, usuel, en Europe est l'arme juridique. **Fin mars 2023, l'Italie lance les hostilités contre ChatGPT et interdit l'outil le temps de l'instruction d'une enquête du régulateur national sur l'usage des données personnelles par l'outil et pour non-conformité au Règlement Général des Données Personnelles (RGPD).** Elle justifie sa décision sur quatre critères : inexactitude des données utilisées à l'entraînement, défaut d'information (éternel problème du consentement éclairé), absence de base légale justifiant l'usage des données aspirées pour l'entraînement du modèle, absence de vérification de l'âge des utilisateurs, ce qui pose problème pour les mineurs de moins de 13 ans.

L'Allemagne et l'Espagne pourraient suivre la voie italienne. La France quant à elle, par l'intermédiaire de la CNIL, va également se pencher sur la question à l'occasion de deux plaintes déposées auprès du régulateur national. **Fait intéressant à noter, les régulateurs européens ont décidé de se coordonner en montant une « taskforce » spéciale sur ChatGPT.**

Le régulateur semble bien plus attaché aux libertés et à la vie privée que nous, utilisateurs, qui produisons et octroyons nos données toujours davantage, malgré tous les avertissements et les scandales. Nous continuons, le confort de l'expérience du consommateur

supplantant visiblement tout sur son passage. Le régulateur ne pourra pas tout faire tout seul sur des sujets de cette ampleur si le citoyen qu'il protège ne se mobilise pas sur la cause ou n'est qu'indifférent à celle-ci.

En outre, compte tenu des échelles et vitesse de déploiement dont on parle, des forces de frappe en présence, **la position normative européenne reste essentiellement défensive, une condition nécessaire mais non suffisante.**

Et la France ? Là encore, il faut encore muscler davantage le verbe proactif et offensif, à savoir une véritable stratégie techno-industrielle qui cesse pour de bon le saupoudrage des fonds publics alloués - qui restent honorables pour un pays de la taille de la France. Une vision qui permette de décider de la place de la France à 20 ou 30 ans sur des secteurs de niche à identifier et flécher, en repensant (identifier, dégraisser, agiliser, raccourcir, faciliter) le système national d'innovation de bout en bout (SNI). Le repositionnement de la French Tech ou le programme France 2030 auraient dû en être l'occasion rêvée. Le renouvellement de l'approche est d'autant plus urgent que la France n'est aujourd'hui pratiquement plus une puissance technologique malgré de nombreux atouts, par exemple, la recherche nationale en IA ou sur certains usages médicaux et scientifiques. D'après le Global AI Index, elle arriverait tout juste à la 10^e position, après les Pays-Bas et l'Allemagne, très loin derrière le Royaume-Uni, 3^e

après les États-Unis et la Chine. Sur le seul critère de la recherche, elle est positionnée 15^e là où l'Allemagne est 6^e et le Royaume-Uni, 5^e. Une position de niche sur des secteurs stratégiques où la France possède des avantages comparatifs réels, voilà le seul chemin possible de l'autonomie stratégique, en fait de la souveraineté technologique, à savoir la capacité à tenir un rapport de force en cas de besoin.

Si l'affaire n'est pas totalement perdue, elle est néanmoins urgente et pour des bénéfices politiques qui ne seront visibles qu'à long terme. **Une véritable stratégie en IA suppose du courage politique qui s'affranchit des effets d'annonce.**

Mais l'Europe n'est pas la seule à normer. États-Unis et Chine, les deux hyperpuissances de l'IA, s'y mettent également, avec une approche fidèle à leur corpus de valeurs.

Côté américain, l'approche est basée sur la responsabilité des systèmes, leur transparence et l'usage, éventuellement abusif, des données personnelles. La National Telecommunications and Information Administration (NTIA) a publié le 11 avril un « AI Accountability Policy Request for Comment », un appel à contribution sur ce que devrait être une bonne politique publique en matière de responsabilité de l'IA, avec des cas d'usages appliqués sur des secteurs stratégiques comme la santé.

En Chine, la Cyberspace Administration of China (CAC) soutient

« D'après le Global AI Index, la France arriverait tout juste à la 10^e position. »

les projets d'IA en général et des IA génératives en particulier. Mais avec des règles à respecter, en premier lieu desquelles la conformité aux valeurs socialistes du pays et la transparence sur l'identité des utilisateurs (ces derniers devront fournir leur véritable identité). Un volet concerne également le bon usage des données utilisées à l'entraînement afin d'éviter les biais et les discriminations mais aussi la conformité des systèmes aux règles établies par la CAC.

LA GÉOPOLITIQUE DE L'IA COMME RÉPONSE AUX PROJETS IDÉOLOGIQUES : ORDRE MONDIAL ET ORDRE SOCIAL FONT BON MÉNAGE !

Au même titre que la question climatique, l'IA reflète de plus en plus un désaccord politique persistant qui est en train de se hisser au sommet de l'agenda mondial.

Si les USA ou la Chine s'emparent de la question normative, ce n'est pas pour freiner une innovation devenue le carburant de leur rivalité stratégique, mais pour garder le contrôle d'abord intérieur d'outils dotés à terme d'une puissance qui pourraient déborder les États eux-mêmes. À échelle encore « raisonnable », nous en voyons déjà les prémices avec les dégâts possibles d'un ChatGPT ou d'un Midjourney dans les luttes informationnelles et guerres cognitives qui se jouent à bas bruit. Imaginons les conséquences d'IA duales, aux usages civils et militaires décuplés.

Chine et États-Unis doivent jouer aux équilibristes : stimuler l'innovation et les acteurs technologiques, bras armés technologiques fournisseurs d'instruments de projection de puissance et dans le même temps, garder l'écosystème sous contrôle, et en conséquence de quoi, la hiérarchie des normes et des acteurs mais aussi l'ordre social. Ordre mondial et ordre social font bon ménage.

L'intention politique chinoise est claire : en 2017, le Parti communiste chinois a présenté un plan visant à développer l'intelligence artificielle pour faire du pays le premier centre mondial d'innovation en IA d'ici à 2030.

Avec le climat, la question technologique, et en particulier l'IA, va contribuer à structurer l'ordre et la sécurité mondiale des prochaines décennies. En particulier, l'IA va participer à dessiner la morphologie des futures guerres, hybrides, cyber ou cinétiques. L'IA a déjà apporté des robots tueurs, des IA conversationnelles capables de rédiger des communiqués gouvernementaux, des programmes capables de traiter des données à une hypervitesse supérieure à la cognition humaine. Il est alors logique que les puissances mondiales se lancent dans une course pour dominer ce nouveau paysage technologique.

Les IA à usage militaire permettront, notamment aux armes, d'être hyper rapides (vitesse), autonomes (critère vital à l'heure des brouillages élec-

tromagnétiques), super-précises dans le traitement de la donnée et dans la prise de décision avec le traitement le flux d'information (« fog of more ») à une vitesse irrattrapable par l'homme.

POUR FINIR, RESTER HUMAIN

Dans ce combat de titans entre Hyperpuissances et HyperTechnologues, nous, simples citoyens, pouvons avoir l'impression d'être démunis. Mais nous ne le sommes pas. Il est normal d'être déstabilisé dans les périodes de mutations profondes, mais ne perdons pas de vue un point cardinal : notre Humanité.

Dans *Species Technica*, le philosophe Gilbert Hottois expliquait que la technique ne finissait par prendre le contrôle de certains pans de nos existences que parce que l'Homme ne les investissait plus. Dans le fond, oui, l'Homme n'est plus là car il est bel et bien déjà ailleurs, reparti à la découverte de contrées nouvelles. **C'est à partir de cette seule conviction, le refus du renoncement, que la véritable discussion démocratique pourra alors commencer.** ■

Cet article est repris dans une forme condensée de la plateforme Expression de l'Institut Montaigne, avec l'aimable autorisation de l'éditeur.

3 QUESTIONS À FRANÇOISE SOULIÉ- FOGELMAN



CONSEILLÈRE
SCIENTIFIQUE DU
HUB FRANCE IA

Entretien réalisé par Kati Bremme

Spécialiste du data mining et du big data, Françoise Soulié-Fogelman a été formée à l'ENS et a obtenu son doctorat à l'Université de Grenoble. Informaticienne, elle a enseigné à l'Université de Paris XI-Orsay où elle a encadré une vingtaine de thèses. Elle a fondé la société Mimetics, spécialisée dans les logiciels de data mining pour la lecture automatique de caractères (OCR), a travaillé pour Atos Origin et a fait partie de l'équipe de direction de KXEN jusqu'à son rachat par SAP. Elle a ensuite rejoint l'Institut Mines Télécom pour travailler sur TeraLab. Elle a dirigé l'équipe Data Science de l'école de logiciels informatiques de l'Université de Tianjin en Chine, et est aujourd'hui co-fondatrice et conseillère scientifique du Hub France IA.

C'est certainement la voie pour l'Europe : favoriser le développement de modèles plus petits, plus spécialisés et non dépendants des grands modèles américains.

1

FACE AUX SCÉNARIOS CATASTROPHE D'UNE IA GÉNÉRALE DE TYPE TERMINATOR QUI PRENDRAIT LE POUVOIR SUR LES HUMAINS, ON PARLE BEAUCOUP D'IA ÉTHIQUE, NOTAMMENT EN EUROPE. COMMENT CES SYSTÈMES PEUVENT-ILS ÊTRE ÉTHIQUES ? AVEC LE DEEP LEARNING, L'APPRENTISSAGE AUTO-SUPERVISÉ SUR DE TRÈS LARGES QUANTITÉS DE DONNÉES (COMME C'EST LE CAS AVEC LES LLM), PEUT-ON ENCORE PRÉTENDRE POUVOIR ÊTRE TRANSPARENT ?

L'histoire des innovations montre que, de tout temps, l'innovation a fait peur. Un article de 2021* en cite de très nombreux exemples, qui nous semblent ridicules aujourd'hui et aujourd'hui donc, l'Intelligence Artificielle (IA) fait peur. Comme toujours, les fondements de la peur sont d'abord l'ignorance et le manque d'expérience : quand on ignore l'effet de la vitesse sur l'organisme humain,

on a peur en 1843 que le voyage à grande vitesse dans le train tue les passagers par asphyxie (la grande vitesse des premiers trains en 1843 n'avait pourtant rien à voir avec nos TGV !). **Quand on ne comprend pas comment fonctionne l'IA, on peut avoir peur que l'IA prenne le pouvoir sur les humains. Mais c'est oublier que l'IA ce n'est pas des mathématiques avec un théorème qui fonde le système : l'IA c'est de l'informatique, il y a toujours un informaticien qui programme un algorithme** (d'apprentissage à partir des données en général) pour obtenir un certain comportement (par exemple, distinguer les chats des chiens).

On a des exemples de systèmes d'IA nocifs : par exemple, le système Tay mis en ligne sur Twitter par Microsoft en 2016, apprend en interagissant avec les utilisateurs humains ; en moins d'une journée, nourri par certains utilisateurs de contenus haineux, racistes ou à caractère sexuel, Tay apprend à les imiter. Il est alors débranché par Microsoft. Dans ce cas, on avait bien les individus qui avaient la volonté de nuire, mais ils avaient aussi les données pour entraîner le système d'IA. Pour obtenir un système qui prenne le pouvoir sur les humains, il faudrait donc un programmeur qui décide de réaliser un tel système (cela pourrait se trouver

et pour cela, il aurait besoin de données, d'exemples de tels systèmes (et à ma connaissance, cela n'existe pas). Pour le cas d'une catastrophe apportée par une IA « générale » (c'est-à-dire capable d'intelligence semblable à la nôtre, applicable à n'importe quelle tâche), nous pouvons être tranquille : il n'existe pas aujourd'hui de tels systèmes et on ignore si une telle IA générale existera et quand. Les experts ne sont d'ailleurs pas d'accord : certains craignent que cela puisse arriver avant 10 ans, et d'autres imaginent des délais beaucoup plus longs (30 à 50 ans). Et même si une IA générale existait, comment prendrait-elle le pouvoir sur les humains ? il faudrait encore pour cela qu'elle puisse avoir une action dans le monde physique.

Cependant, même avec une IA « étroite » (et non générale), il y a des risques (comme Tay l'a démontré) et il est donc important d'encadrer les systèmes d'IA. Du point de vue éthique (discriminations, biais, etc.), mais aussi avant tout pour assurer les utilisateurs que l'IA fait bien ce pourquoi elle a été programmée et ne provoquera pas de catastrophe. Il faut donc mettre en place des mécanismes de régulation pour assurer la confiance des utilisateurs.

* T La Revue de La Tribune - N°7 Décembre 2021

2

L'IMPRESSONNANTE PERFORMANCE DE CHATGPT REPOSE ENTRE AUTRES SUR UN ACCÈS PRESQUE ILLIMITÉ À DES DONNÉES ET À D'ÉNORMES CAPACITÉS DE CALCUL. IL N'Y A PAS D'ÉQUIVALENT DE GOOGLE, MICROSOFT OU META EN EUROPE, COMMENT PEUT-ON TENIR DANS LA COURSE À L'IA TOUT EN PROTÉGEANT LES UTILISATEURS ?

La communauté européenne, qui propose l'AI Act, indique sa volonté de proposer un cadre légal pour une IA de confiance. Un tel cadre, quand il sera mis en œuvre (vers 2026) imposera des exigences de transparence : un système d'IA à haut risque sera

audité avant d'être mis sur le marché. Les auditeurs auront accès à tout le processus de mise en œuvre du système et ses auteurs devront donc se préparer pour pouvoir fournir toutes les informations nécessaires. En particulier, la dernière version du texte (voté au parlement européen le 14 juin 2023) impose, dans les amendements 100 et suivants, des obligations de transparence sur les données utilisées pour entraîner l'IA générative (« modèles de fondation »), avec cette clause : GPT-4 devrait fournir des informations sur les données qu'il a utilisées (dont on ne sait rien aujourd'hui) pour pouvoir être utilisé sur le marché européen.

La régulation de l'IA ne doit cependant pas bloquer l'innovation. Le développement de l'IA en Europe montre clairement que le marché B2C a d'ores et déjà échappé aux européens : il n'y aura sans doute pas de concurrent européen à ChatGPT. Mais il y a place sur le

marché B2B pour des produits d'IA générative européens, respectueux des règles éthiques et des réglementations mises en place pour l'Europe. Les premiers produits (ChatGPT, Bard, etc.) viennent de sociétés américaines (OpenAI, DeepMind, Anthropic, Meta) qui ont très peu respecté les règles de protection des données (copyright, IP). Ces produits ont utilisé des datasets énormes (1000 milliards de données) avec des coûts de calcul démesurés, ce qui réduit les possibilités de développement de modèles à quelques sociétés au monde. Cependant, de nouvelles techniques apparaissent pour produire des modèles plus petits, moins chers, voire Open Source sur des domaines plus restreints (celui de l'entreprise) ; évidemment **ces modèles plus petits sont moins chers à produire et moins gourmands en temps de calcul (on parle d'IA frugale).**

3

PUISQUE L'ACCÈS FACILITÉ À DES JEUX DE DONNÉES PUBLICS ET PRIVÉS NE SEMBLE PAS ÊTRE RÉALISABLE DE SI PRÈS EN EUROPE, UNE IA PLUS FRUGALE, QUI SERAIT ALIMENTÉE AVEC MOINS DE DONNÉES, MAIS DES DONNÉES DE QUALITÉ, PEUT-ELLE ÊTRE AUSSI PERFORMANTE QUE LES IA GÉNÉRATIVES DISPONIBLES AUX USA ?

C'est certainement la voie pour l'Europe : favoriser le développement de modèles plus petits, plus spécialisés (fine tunés sur les données de l'entreprise) et non dépendants des grands modèles américains. L'offre européenne commence à se structurer. Mais la question reste aujourd'hui ouverte : faut-il partir d'un modèle généraliste (ChatGPT par exemple) qu'on va ensuite spécialiser sur son domaine (fine tuning sur les données d'entreprise), ce qui lie l'entreprise à Microsoft (par exemple) ; ou bien peut-on développer son propre modèle sur ses propres données, dans son propre environnement (on premise), éventuellement avec un logiciel Open Source. Dans le premier cas, un grand modèle lié à Microsoft, dans le second cas, un modèle sans doute plus petit et entièrement indépen-

dant des grands fournisseurs. L'avenir dira si cette seconde voie est efficace, mais la réussite passée du monde de l'Open Source laisse à penser que la barrière à l'entrée élevée par les grands fournisseurs ne tiendra pas : c'est semble-t-il déjà l'avis de Google. Comme on le voit, il reste beaucoup de questions ouvertes : **les techniques d'IA prédictive étaient très matures (datant des années 90), alors que les techniques d'IA générative ont moins de 10 ans.** Cependant, le nombre de chercheurs investis sur ce domaine est en explosion. On doit donc s'attendre à de nouveaux résultats très rapidement. Pas sûr néanmoins que cela permette de développer une IA générale, capable de bons sens ou de mauvaises intentions (détruire l'humanité) dans un avenir proche.

PAROLE

DE MACHINES

VERS UN NOUVEAU TESTAMENT DE L'INTELLIGENCE ARTIFICIELLE : PROPOSITIONS DE CONCEPTION

La diffusion galopante de textes générés par les machines supprime le lien univoque entre le langage et l'être humain. Un système d'intelligence artificielle générative est bien plus qu'un perroquet qui nous singe. Les phrases qu'il émet ne sont pas du simple copier-coller. La machine les construit d'une façon radicalement différente de celle de l'homme. Son langage est le résultat d'un calcul mathématique et ses mots sont dérivés de nombres.

Les phrases fabriquées par des machines acquièrent un sens seulement aux yeux de l'utilisateur. Le contenu de ces récits est soumis à des jugements, autant sur le fond que sur la forme. Leur esthétique, survenue spontanément, peut être puissante, même envahissante. **Les machines impriment ainsi leur marque sur les réalités mouvantes, vivantes et temporelles des êtres humains.**

Utilisateur : adorateur de fonctions utiles. Affectueux, sensible, tantôt agacé, tantôt heureux, parfois emporté, souvent lassé. Rien de cela, à l'époque des Lumières, ne faisait partie des qualités requises pour

*Par Alexei Grinbaum,
directeur de recherche,
président du Comité
opérationnel pilote d'éthique
du numérique du CEA*

Alexei Grinbaum nous raconte l'intelligence artificielle à travers un livre de philosophie et de poésie. Une manière de nous rappeler ce qui nous distingue encore des machines, avec lesquelles nous sommes amenés pourtant à faire usage dans nos quotidiens. Extraits choisis.

être un bon citoyen. Citoyen : défenseur de valeur et des libertés. Éclairé, rationnel, informé. Or, qui dit citoyen dit utilisateur. Les deux vivent dans un seul corps.

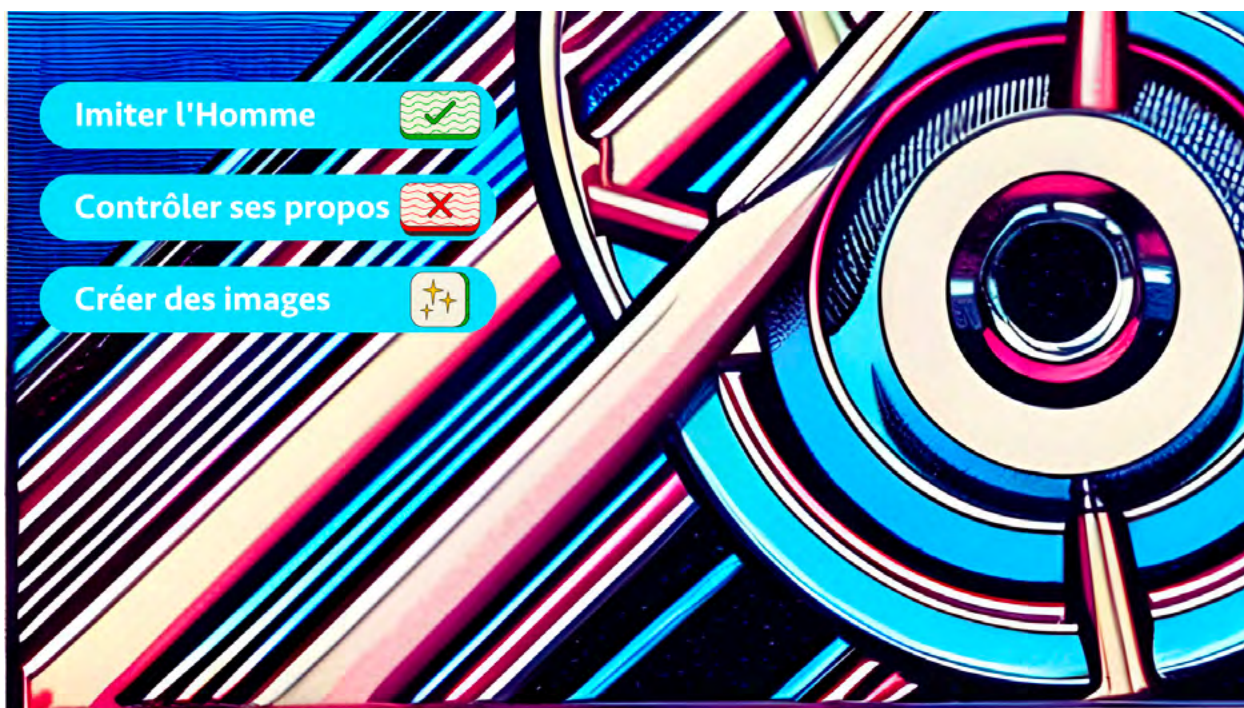
La mutation du citoyen en utilisateur, de l'individu rationnel en l'individu numérique, change la société en profondeur. Elle modifie le sens de la démocratie, comme de tout autre système politique. « Tout ce que les hommes font, ou savent, ou ce dont ils ont l'expérience, a un sens seulement dans la mesure où il est possible d'en parler. » Hannah Arendt a écrit ces lignes à la fin des années 1950. Cette donne millénaire est dorénavant obsolète. L'homme

a perdu l'exclusivité de la parole qui héberge et véhicule le débat social. [...]

CORPUS DES SAVOIRS ET CITÉ NUMÉRIQUE

Réunir une assemblée pour redéfinir la citoyenneté ? Trop tard ! Le corpus des savoirs pour maîtriser le code n'est plus à la portée d'anciens édiles. **La déchéance de l'éducation mathématique et le manque général de compétences en informatique ont durablement converti nos représentants politiques en profanes soumis, comme tous les utilisateurs, au pouvoir du code.** Forcer les machines à réduire leur vitesse jusqu'à ce que les hommes soient capables de les suivre ? Cet adage politique, lui aussi, se dit joliment dans une langue naturelle, mais n'a pas de portée sur le fonctionnement des machines. Une vitesse inhumaine est source de leur efficacité et de leur utilité.

Créer un lien entre les récits centrés sur les technologies et ceux qui forment notre culture. Trouver des motifs qui relient ce qui nous semble inédit à la langue que nous avons reçue en héritage. Tirer de ces motifs un éclairage sur ce qui nous semble radicalement nouveau. Tels sont nos devoirs dans la cité numérique. [...]



L'INDIVIDU NUMÉRIQUE

L'homme qui interagit avec la machine l'imité parfois inconsciemment. La machine apprend, et aussi l'utilisateur. À travers cette interaction à double sens, l'homme place la parole non humaine au sein de la culture et de l'histoire qui sont les siennes. Une évolution lente et progressive des normes et des règles du comportement linguistique est en cours. Son moteur : le comportement mimétique de la part des millions d'utilisateurs des chatbots. Ce que les machines font à nous, nous le faisons aux autres. C'est par cette voie que l'inhumain des machines entre dans la vie humaine, puis celle de la société.

Les significations que l'on attribue à un texte généré par la machine ne sont que des projections. **Les allusions présentes dans le calcul ne sont pas des raisonnements logiques ou conceptuels, mais des corrélations numériques.** Tout ce caractère illusoire d'un « individu numérique » qui se profile derrière un texte généré n'enlève rien à l'efficacité des projections. Peu importe qu'elles soient si peu ancrées dans la nature : si leur effet sur l'utilisateur est réel, alors elles sont tout aussi réelles pour l'utilisateur, sur le mode du comme si. [...]

UN FRUIT CRÉATIF INATTENDU

Deux propriétés des systèmes d'IA générative doivent être appréciées. La première : les modèles de langage qui s'appuient sur l'apprentissage automatique ne sont pas déterministes. Cela leur permet de ne pas être ennuyeux du point de vue de l'utilisateur. Ainsi, ChatGPT met en œuvre un hasard partiel, qui n'est pas maximal au sens mathématique puisque les réponses fournies par la machine restent humainement compréhensibles, et cependant la dose du hasard qu'elles contiennent est supérieure à zéro, car sinon la machine fournirait toujours la même réponse à la même question. C'est ce hasard intermédiaire, ni trop fort, ni trop faible, qui imite au mieux le libre arbitre humain.

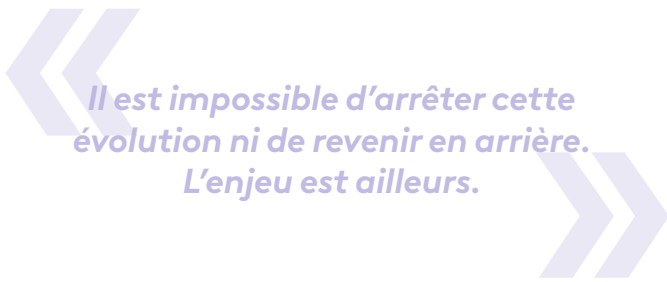
Secondement, la créativité. **On a longtemps cru que les systèmes d'intelligence artificielle pourraient remplacer l'homme dans des tâches routinières, tandis que les métiers créatifs resteraient à l'abri de l'automatisation. Aujourd'hui, nous savons que ce n'est pas le cas.**

Les systèmes d'IA générative sont employés dans l'écriture d'articles de presse et dans le design de couvertures de magazines. La machine est capable d'inventer quasiment sans limites. Cette créativité extraordi-

naire, non humaine, n'a pas été programmée explicitement ; elle est le fruit inattendu et une capacité émergente de l'innovation technologique.

Cependant, il faudrait veiller à ce que les machines créatives n'émettent pas de propos « dommageables ». Tel est le nouvel idéal de l'industrie de l'IA : exclure, par conception, toute possibilité qu'un système d'IA porte préjudice à une personne humaine. Ce n'est ni simple ni même réaliste. Faute de maîtriser la signification des mots, un modèle de langage se voit ajouter des filtres et des couches de contrôle visant à exclure tout langage potentiellement « toxique » : les insultes, les conseils médicaux pouvant entraver le bien-être humain, les conseils juridiques qui risquent d'inciter un comportement illégal, etc.

Les conséquences globales du filtrage des résultats sont positives. Mais, son effet ne s'arrête pas là, car toute tentative de mettre le bien et le mal dans une gaine numérique fera inévitablement des victimes collatérales. L'excès de contrôle (*overpolicing*) est déjà manifeste dans certains textes que produit ChatGPT. La machine refuse d'entrer dans des conversations parfaitement sensées, même celles où l'information demandée peut soulager l'utilisateur ou l'aider réellement. L'excès de contrôle qui vise à exclure le langage toxique mène souvent à



*Il est impossible d'arrêter cette
évolution ni de revenir en arrière.
L'enjeu est ailleurs.*

la production de dialogues stérilisés et ennuyeux. La crainte des préjugés, dès lors qu'elle est prise pour un cadre qui régit le design, se retourne contre la fascination du créateur humain. Comme toute absolutisation excessive, elle devient source du mal. **Si nous allons trop loin sur la voie d'overpolicing, nous risquons de commencer à fabriquer nous-mêmes du langage stérile en suivant l'exemple de nos chatbots.**

La condition humaine évolue sous l'influence des machines parlantes ; il est impossible d'arrêter cette évolution ni de revenir en arrière. L'enjeu est ailleurs : il faut accompagner la conception des chatbots pour la rendre moins hâtive et mieux réfléchie, évitant toute rupture avec notre culture. Le besoin de maintenir la distinction homme-machine au niveau de la parole émerge comme un nouvel impératif éthique.

Quand l'utilisateur demande au chatbot de commander une pizza, il ou elle ne cherche pas à différencier les répliques de la machine de celles d'un employé humain qui exécute la même tâche. Dans d'autres cas, cependant, c'est l'inverse qui est vrai. **Recourir à un système d'IA pour passer un examen constitue une faute. Toutefois, un texte généré par un système d'intelligence artificielle est original : non plagié, il rend la prise de sanction quasi-**

ment impossible. ChatGPT passe les examens du Barreau mieux que la plupart des avocats et ceux de biologie moléculaire mieux que les étudiants en biologie. Cependant, un professeur est censé évaluer la personne humaine, et non un programme informatique. Dans l'état actuel des technologies et du système éducatif, c'est devenu problématique. Il est d'ores et déjà nécessaire de donner aux professeurs un moyen de reconnaître les dissertations d'origine non humaine. Il faut aussi les former pour qu'ils puissent s'adapter à la nouvelle donne et diversifier les types de devoirs. Lorsque la pratique de recours par les élèves aux chatbots se généralise, l'interdire n'est plus une option envisageable.

ChatGPT a attiré des millions d'utilisateurs ; il n'est plus possible de les en priver. La seule option réaliste est de modifier la conception du logiciel pour assurer le maintien des distinctions. Cela permet notamment d'attribuer des responsabilités partagées pour un éventuel préjudice. Si le texte comporte une insulte ou s'il provoque un conflit humain, il est impératif de pouvoir tracer son origine afin d'éviter la confusion entre un discours mené par un agent responsable, apte à répondre de ce qu'il dit, et la parole asémantique d'un système d'intelligence artificielle auquel on ne peut attribuer aucune responsabilité.

Ainsi, il est nécessaire d'introduire des signes distinctifs dans les textes générés par les machines, sans pour autant nuire à l'utilité du chatbot. Lorsqu'un utilisateur commande une pizza, il ne doit pas être soumis à une série d'avertissements réglementaires sur la source des phrases qu'il entend. L'abus de la législation risque d'effacer l'intérêt de recourir aux chatbots. [...]

CODES MATHÉMATIQUES

Un code bien caché, mais lisible au besoin par celui qui y prête attention, qu'est-ce que c'est ? Par exemple, à la distance de 239 caractères, un texte généré par ChatGPT pourrait contenir en boucle les lettres G-P-T. Il est techniquement possible d'ajuster l'algorithme de génération automatique pour qu'il respecte cette règle. Cela représente une contrainte suffisamment souple pour qu'elle passe inaperçue à la lecture rapide du texte, mais ce code en filigrane (*watermark*) resterait tout de même détectable par qui veut étudier la composition du texte en détail. Enfin, **le watermark, même s'il est invisible à l'œil nu, doit rester suffisamment robuste pour être détectable en dépit d'une éventuelle attaque.** Un adversaire pourrait ajouter des mots au texte en alternant les distances entre les lettres, voire remplacer certains mots par des synonymes. Le code devrait

« *Toute tentative de mettre le bien et le mal dans une gaine numérique fera inévitablement des victimes collatérales.* »

alors résister le plus longtemps possible à ce genre d'attaques. Pour cela, il doit être introduit au niveau mathématique lors de la génération du texte. Même s'il n'est pas possible de garantir son inviolabilité absolue, le prix de son effacement doit être d'emblée élevé, rendant le *watermark* difficile à perturber. Ce problème de robustesse motive le choix de codes très complexes, au point de les rendre introuvables par un lecteur non expert. L'équilibre entre ces deux contraintes reste encore à trouver. [...]

CULTURE ET INTELLIGENCES ARTIFICIELLES

L'objectif des machines intelligentes n'est ni d'effacer notre civilisation ni d'oblitérer l'homme. Les machines mettent en œuvre une finalité qui consiste à calculer et à optimiser des fonctions mathématiques et – sans surprise pour tous, sauf quelques informaticiens – ce processus qui paraît purement formel produit des conséquences imprévues sur l'homme. Notamment, il accélère sa simplification, qui est déjà en cours par ailleurs. Cette fois, les chemins de simplification sont tracés, non dans un monde dystopique où la pensée avait cessé, comme se l'imaginaient les écrivains du XX^e siècle, mais dans un univers du calcul froid, vide de toute signification humaine. **Comme**

l'eau recouvrant les temples d'Alexandrie ou les rues de Venise, le langage des machines engloutit nos joies et nos douleurs. La poésie s'efface par manque de pertinence devant les nouveaux parlants venus de Californie. Leur langue est une projection du nombre sur ce qui paraît seulement comme une culture. Ainsi, il appartient aux hommes d'assurer la continuité de leur culture devant les capacités génératives des systèmes d'intelligence artificielle. ■

« *L'objectif des machines intelligentes n'est ni d'effacer notre civilisation ni d'oblitérer l'homme.* »



« POWER SHIFT » :

COMMENT L'IA GÉNÉRATIVE REDÉFINIT LA CHAÎNE DE VALEUR DES MÉDIAS

L'afflux écrasant d'applications IA et la succession rapide de lancements ont laissé beaucoup d'entre nous confus et désemparés. Certains disent que l'IA résoudra tout et d'autres voient dans cette technologie, la fin de tout. Comment trouver le bon sens lorsque tout tourne ?

Dans cet article, nous nous pencherons sur les dynamiques clés qui influencent le modèle économique des médias dans le contexte de l'IA. Nous discuterons de la démocratisation des capacités créatives, des changements de dynamique de distribution et des attentes du public, ainsi que du changement de pouvoir dans l'industrie des médias. En comprenant ces dynamiques, les professionnels des médias peuvent naviguer plus efficacement dans ce nouveau paradigme.

CRÉATION

L'IA est passée d'un art caché à un outil accessible à tous. Cela change la nature de la création et la question de savoir qui peut être créateur. **Soudain, tout le monde est capable de tout créer – avec le texte comme seul outil – le langage est la nouvelle peinture, la nouvelle encre et le prochain code.**

C'est le premier changement :

Par Ezra Eeman, directeur de l'Innovation et de la Stratégie à la Dutch Foundation for Public Broadcasting

L'IA générative a la capacité de transformer le bruit en formes significatives. Tout comme les humains observent les constellations dans le ciel nocturne, l'IA peut générer des articles à partir de données collectées et des images à partir d'un nuage de pixels. De ce chaos, émerge une structure.

- L'IA nivelle le terrain de jeu créatif tout en élevant la barre pour la production humaine.
- La fiabilité de l'IA reste une préoccupation, car elle a du mal à juger les faits et à comprendre le contexte.
- Dans sa forme la plus extrême, l'IA peut automatiser complètement le processus de création, transformant les données brutes en contenu généré.

DISTRIBUTION

Les systèmes IA libèrent l'échelle. **Soudain, les individus créateurs et les petites rédactions peuvent élargir leur couverture, servir plus**

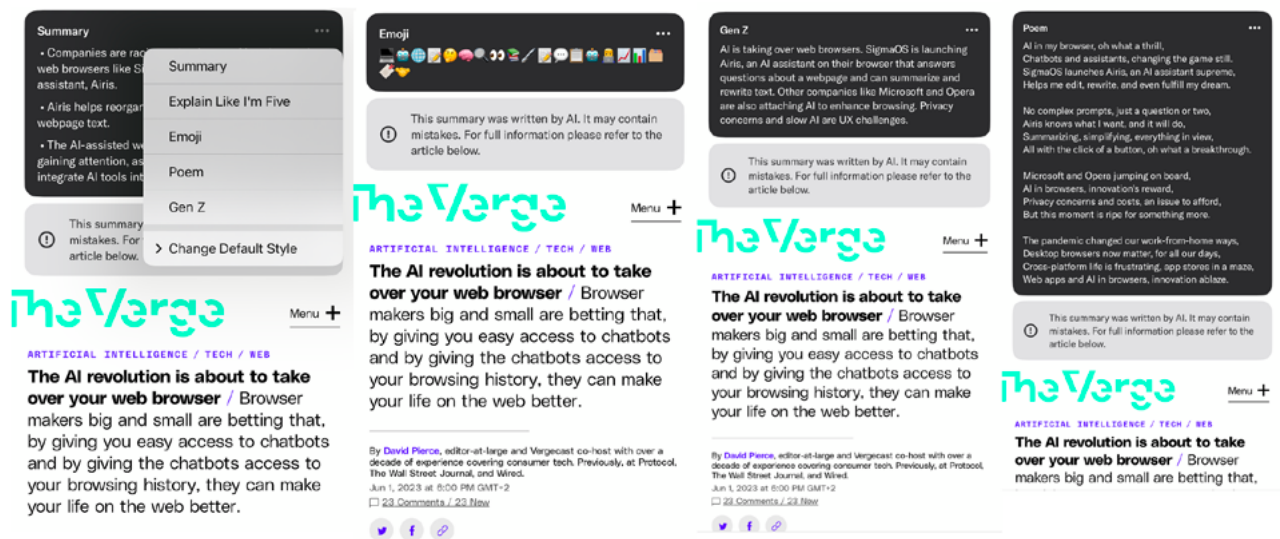
de personnes et répondre à plus de besoins qu'ils ne pourraient jamais le faire sans soutien artificiel. L'ancien modèle était : produire une fois et distribuer à de nombreux canaux. Le nouveau est : créer sans cesse, distribuer à l'infini et optimiser pour chaque canal. En réalité, la suppression de tout obstacle à la production de contenu donne naissance à des défis tels que l'exploitation de contenu et la prolifération de contenu indifférencié.

C'est le deuxième changement :

- Produire du contenu à grande échelle devient plus facile, mais sa valeur diminue.
- Les machines inonderont Internet de contenu indifférencié et de désinformation, nécessitant des technologies pour prouver l'origine et la fiabilité.
- Le journalisme humain devient un produit premium, les voix de premier plan et les niches étant mieux positionnées pour monétiser l'attention.

ÉCHANGE DE VALEUR

Finalement, l'IA générative déclenchera une nouvelle bataille pour le contrôle de la chaîne de valeur des médias. Nous avons vu que dans un paysage axé sur l'IA, le processus de création et de production est com-



Artifact, le newsfeed généré par IA.

pressé, voire complètement omis. En conséquence, l'importance des points finaux, tels que les données brutes, la propriété intellectuelle, le talent et les interfaces avec le public, augmente. Et c'est là que les organisations médiatiques risquent d'être mises de côté.

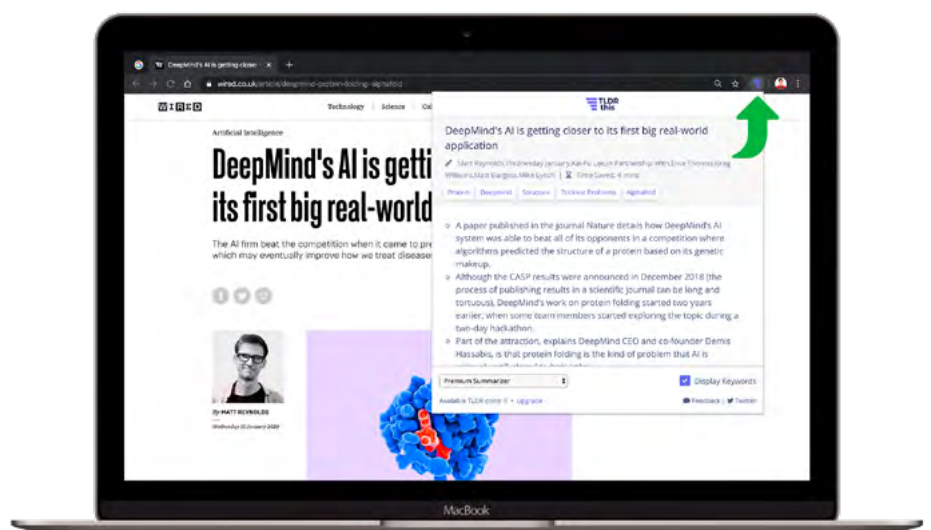
C'est le troisième changement :

- Les éditeurs risquent de perdre le contrôle de leurs données et de leur contenu, car l'IA transforme les articles récupérés en réponse de chatbots, ce qui compromet le modèle économique des médias.
- Les interfaces subissent un changement, passant de destinations soigneusement sélectionnées où les éditeurs peuvent transmettre la confiance et créer une expérience de marque unique, à des interfaces génératives contrôlées par les grandes technologies, les nouveaux agrégateurs ou les utilisateurs. Les titres sont réécrits, les articles résumés, la structure et la hiérarchie de la mise en page sont repensées.
- Les éditeurs sont face à un choix difficile : optimiser pour ces nouvelles interfaces et envisager un partenariat avec les grandes technologies, ou rivaliser en investissant dans leurs propres modèles de données et en fermant les portes à leur contenu.

CONCLUSION

L'IA a un impact profond sur l'écosystème des médias, transformant la création, la distribution et la consommation de contenu. Pour naviguer dans ce nouveau paradigme, il faut bien comprendre les dynamiques en jeu. Pour l'instant, je vous conseille de prendre un peu de recul, de mieux comprendre le paysage et de rester concentré sur vos objectifs. Une destination claire facilite la navigation en territoire inconnu.

Le journalisme humain devient un produit premium.



DeepMind, l'extension qui résume une page web en un clic.

QUELLE(S) RÉGULA- TION(S)

POUR LES SYSTÈMES D'IA GÉNÉRATIVE ?

« Non, ce n'est pas trop tôt. » Voilà ce qu'a répondu Mira Murati, directrice technique d'OpenAI, à un journaliste du *Time Magazine* qui lui demandait si l'adoption d'une réglementation étatique ne risquerait pas de ralentir l'innovation en matière d'IA générative comme ChatGPT. Une réponse qui fait écho à une série d'initiatives, il y a déjà plusieurs années, d'un certain nombre d'acteurs qui se dotaient de chartes éthiques pour le développement de l'IA.

Or, définir des critères éthiques, notamment de modération, n'est jamais neutre, et révèle nécessairement des conceptions du monde et de différentes valeurs. Pour ChatGPT, utilisé par plus de cent millions d'utilisateurs dans le monde, c'est une société américaine de quelques centaines de salariés qui a défini les catégories de questions et de réponses non autorisées (OpenAI a d'ailleurs révélé avoir consenti encore plus d'efforts dans la modération de son nouveau modèle GPT-4 qu'elle ne l'avait fait pour le modèle précédent, GPT-3.5).

Pour ces raisons, de nombreux spécialistes, y compris au sein d'OpenAI, appellent à l'adoption de réglementations dédiées à ces

*Par Martin Bieri, chargé
d'études au pôle Innovation
& Prospective de la CNIL*

La popularité spectaculaire de l'interface ChatGPT a suscité une série de débats juridiques, au point d'appeler à l'adoption d'une loi dédiée. Pour autant, différentes réglementations existent et s'appliquent déjà à de tels systèmes, même si des clarifications sont nécessaires.

systèmes d'IA, afin que les règles ne soient pas seulement définies par les entreprises. Si ces débats sont nécessaires, ils ne doivent pas faire oublier que des règles existent et s'appliquent déjà à ces dispositifs.

« IA OUVRE-TOI » OU LE SÉSAME DU DROIT D'ACCÈS ?

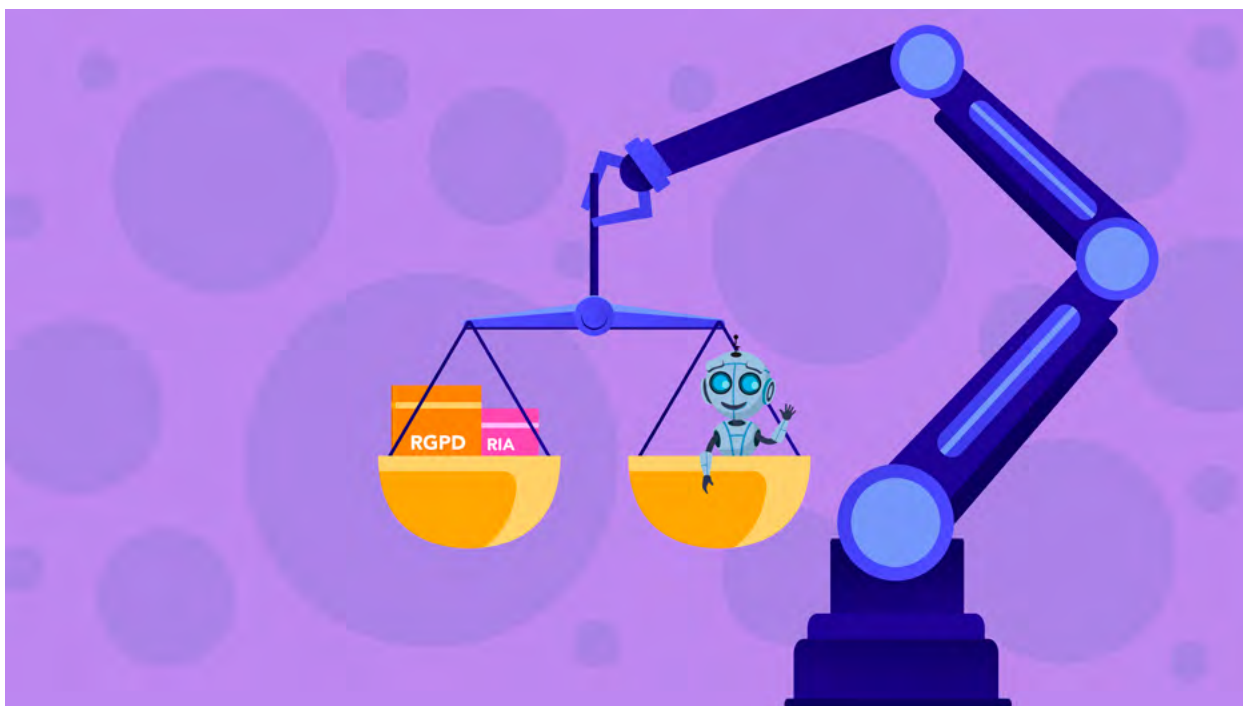
Les modèles d'IA générative reposent sur la collecte et l'utilisation de quantités colossales de données (dont certaines personnelles), le plus souvent sur Internet. Si certaines de vos données sont en ligne (par exemple sur Twitter ou Reddit), il est possible qu'elles aient servi à entraîner le

modèle de langage sur lequel repose ChatGPT. Comme aux prémices des moteurs de recherche sur le web, un premier réflexe serait de demander à ChatGPT s'il vous « connaît ».

Comme il s'agit d'un modèle statistique et non d'un moteur de recherche, sa réponse, même positive, ne serait pas très utile. Cela révélerait surtout sa capacité à prédire une suite cohérente de lettres, mots ou bouts de phrases associés à la suite de lettre qui compose votre nom. Pire, s'agissant d'un modèle probabiliste, sa réponse pourrait ne pas être toujours la même, voire être fausse. Vous n'apprendrez donc pas nécessairement ce qu'il « sait » ou ce qu'il a pu « lire » sur vous, et encore moins sur quels sites web, puisque ChatGPT ne cite pas ses sources.

Rappelons alors que le Règlement Général sur la Protection des Données (RGPD) s'applique à la collecte et à la réutilisation des données personnelles, quand bien même ces données sont publiquement accessibles sur Internet. Il prévoit notamment que ces données doivent être traitées de manière licite, loyale et transparente.

Ce principe de transparence est fondamental, en particulier dans la régulation des systèmes d'IA. Il impose au



responsable de la collecte des données sur Internet, de leur éventuelle annotation, et de leur utilisation à des fins d'entraînement d'un modèle d'IA, de fournir une information pertinente sur les conditions dans lesquelles ces données sont traitées, et notamment sur les sources de la collecte. **Dans la même logique, le RGPD prévoit un droit d'accès, qui permet à tout individu de savoir si un organisme traite des données le concernant, d'en obtenir une copie le cas échéant, et de recevoir des informations plus précises sur la manière dont elles sont traitées.**

Le droit d'accès est une composante essentielle de la transparence, et permet notamment d'orienter ou d'exercer d'autres droits prévus par le RGPD, tant à l'égard du traitement des données qui auraient pu être collectées en ligne pour concevoir le modèle d'IA, qu'à l'égard du traitement des données fournies

par chaque utilisateur qui utilise son interface conversationnelle.

« AI-JE MON MOT À DIRE SUR L'UTILISATION DE MES DONNÉES ? » OU L'EXERCICE DES AUTRES DROITS PRÉVUS PAR LE RGPD

La philosophie du RGPD est de permettre aux personnes de conserver la maîtrise de leurs données à travers une série de droits qui leur sont accordés. C'est notamment le cas du droit d'opposition, qui permet de s'opposer au traitement de ses données par un organisme pour un motif légitime et qui peut s'exercer à tout moment, mais aussi du droit à l'effacement ou du droit à la rectification des données.

Lors de la phase de conception du modèle de langage, la collecte et l'utilisation de données publique-

ment accessibles posent de nombreuses questions quant à leur licéité (les mesures et garanties en place sont-elles suffisantes pour se passer du consentement des personnes concernées ?), ou la possibilité pour chaque individu d'exercer ses droits sur les données personnelles qui le concernent, et qui seraient incluses dans la base d'entraînement.

S'agissant de la phase d'utilisation de ChatGPT, l'interface conversationnelle peut amener les utilisateurs et le grand public à confier des données potentiellement sensibles ou confidentielles au système (bien qu'un avertissement de l'interface elle-même appelle à ne pas fournir de telles données). À cet égard, il y a lieu de noter que les conditions générales d'utilisation de la version gratuite de ChatGPT prévoient qu'OpenAI puisse réutiliser les données des utilisateurs pour améliorer le système, ou en développer de nouveaux (un mécanisme d'opposition à cette réutilisation semble prévu en l'espèce).

Là encore, le RGPD encadre les conditions de réutilisation des données fournies par les utilisateurs, à travers une série de droits accordés aux personnes, mais aussi à travers d'autres obligations que le responsable du

De nombreux spécialistes [...] appellent à l'adoption de réglementations dédiées à ces systèmes d'IA, afin que les règles ne soient pas seulement définies par les entreprises.

« Si des débats sont nécessaires, ils ne doivent pas faire oublier que des règles existent et s'appliquent déjà à ces dispositifs. »

traitement doit anticiper. Le RGPD impose à tout responsable du traitement de respecter ses obligations, mais aussi de pouvoir démontrer sa conformité (il s'agit du principe de responsabilité, aussi appelé « *accountability* »). **Chaque organisme traitant des données personnelles doit mettre en œuvre des mesures plus ou moins importantes en fonction des risques encourus par les personnes concernées, et ce par défaut et dès la conception des traitements.**

UNE NOUVELLE MANIÈRE D'APPRÉHENDER LA PROTECTION DES DONNÉES PERSONNELLES ?

La réglementation en matière de protection des données repose sur de grands principes très robustes. Le RGPD reste donc non seulement applicable, mais tout à fait pertinent, pour les traitements des systèmes d'IA générative qui requièrent de grandes quantités de données, souvent personnelles. Une récente étude du Conseil d'État souligne d'ailleurs « la très forte adhérence entre la régulation des systèmes d'IA [à venir] et celle des données, en particulier des données à caractère personnel ». Par ailleurs, face aux avancées technologiques spectaculaires de l'IA et pour répondre à ces enjeux de société, la CNIL a récemment créé un service de l'intelligence artificielle.

Les travaux autour du projet de règlement sur l'IA (RIA) pourraient laisser croire que les réglementations actuelles sont obsolètes ou dépassées. Pourtant, ce projet de règlement n'a pas vocation à rem-

placer le RGPD pour ces systèmes mais bien à le compléter. À ce stade, le projet de RIA précise d'ailleurs un certain nombre de principes déjà présents dans le RGPD, de manière plus générale. C'est notamment le cas du principe de transparence et de la prise en compte des risques en amont de la conception des systèmes. De plus, le règlement IA a vocation à s'appliquer d'abord aux « fournisseurs de solution » d'IA souhaitant accéder au marché européen, là où le RGPD concerne principalement des utilisateurs de solution qui sont responsables des traitements qu'ils achètent ou mettent en œuvre.

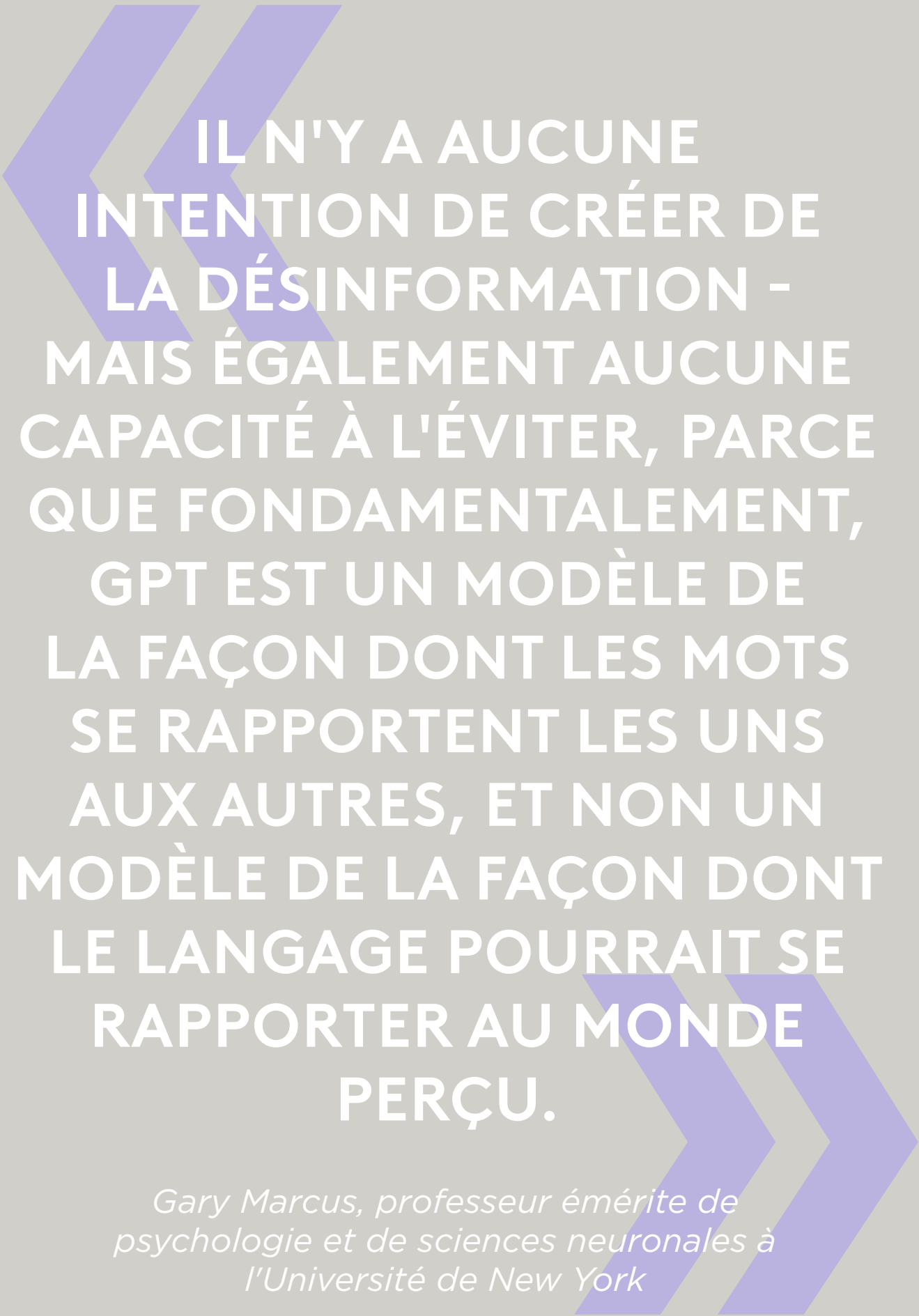
Il se montre en revanche plus précis dans son approche par les risques notamment en proposant différentes catégories de risques correspondant à différents usages, les mesures à prendre étant alors plus ou moins contraignantes. **Le projet de règlement distingue notamment les systèmes d'IA présentant des « risques inacceptables » (lesquels sont interdits), les systèmes à « haut risque » (soumis à des exigences élevées) et les systèmes à « risque limité » (soumis à des exigences minimales, notamment de transparence).**

À cet égard, la question se pose de trouver le bon niveau d'encadrement s'agissant des systèmes d'IA dits à usage général, dont les IA génératives sont une sous-catégorie. Faut-il créer un régime sur mesure comme cela a un temps été pensé (notamment en raison de la difficulté d'envisager tous les usages à risques dès le stade de la conception de tels systèmes) ? Faut-il les considérer comme des systèmes

d'IA à « haut risque », ce qui exigerait des mesures importantes pour pallier les différents risques encourus, telles qu'une transparence accrue à l'égard des utilisateurs, une attention particulière quant aux données alimentant le système, et un niveau élevé de robustesse, de sécurité et d'exactitude ?

Les co-législateurs européens devront répondre à ces interrogations. Certains commentateurs se demandent si les nouvelles obligations à la charge des concepteurs seront suffisantes (notamment s'agissant des obligations de transparence sur les systèmes d'IA à usage général), d'autres craignent que des obligations supplémentaires trop lourdes ne freinent l'innovation des entreprises, au profit d'acteurs étrangers. Comme le RGPD, qui ne s'applique pas uniquement aux organismes européens, le RIA s'appliquera également aux fournisseurs étrangers qui commercialiseraient ou mettraient en service leurs systèmes d'IA sur le marché européen. ■

Cet article est extrait d'un dossier publié sur le site du Laboratoire d'Innovation Numérique de la CNIL (LINC), qui traite des enjeux techniques, juridiques et d'usages des systèmes d'IA générative.



IL N'Y A AUCUNE
INTENTION DE CRÉER DE
LA DÉSINFORMATION -
MAIS ÉGALEMENT AUCUNE
CAPACITÉ À L'ÉVITER, PARCE
QUE FONDAMENTALEMENT,
GPT EST UN MODÈLE DE
LA FAÇON DONT LES MOTS
SE RAPPORTENT LES UNS
AUX AUTRES, ET NON UN
MODÈLE DE LA FAÇON DONT
LE LANGAGE POURRAIT SE
RAPPORTER AU MONDE
PERÇU.

*Gary Marcus, professeur émérite de
psychologie et de sciences neuronales à
l'Université de New York*

MAIS À QUI APPARTIENNENT DONC LES ŒUVRES SYNTHÉ- TIQUES ?

MÉTA-MEDIA : LES MÉDIAS VIENNENT JUSTE DE REMPORTER (EN PARTIE) UNE LONGUE BATAILLE AUTOUR DES DROITS VOISINS CONTRE LES GÉANTS DE LA TECH AMÉRICAINS. QUELS SONT AUJOURD'HUI LES PRINCIPAUX ENJEUX ET DÉFIS JURIDIQUES AUXQUELS LES MÉDIAS SONT CONFRONTÉS FACE À L'IA GÉNÉRATIVE ?

Perrine Pelletier : La bataille des droits voisins n'est qu'une étape dans la construction de l'écosystème numérique des médias. Cet écosystème repose, non seulement sur les contenus nativement créés par les médias qu'ils cherchent à monétiser directement, mais aussi sur ces mêmes contenus repris par les plateformes numériques, lesquelles parviennent finalement mieux que les médias eux-mêmes, à favoriser les interactions numériques, à soutenir l'engagement des lecteurs et ainsi donc, à les monétiser. L'enjeu de la récente mise en œuvre du droit voisin était d'instaurer une

*Entretien mené par
Kati Bremme, directrice
de l'Innovation France
Télévisions, et rédactrice en
chef de Méta-Media*

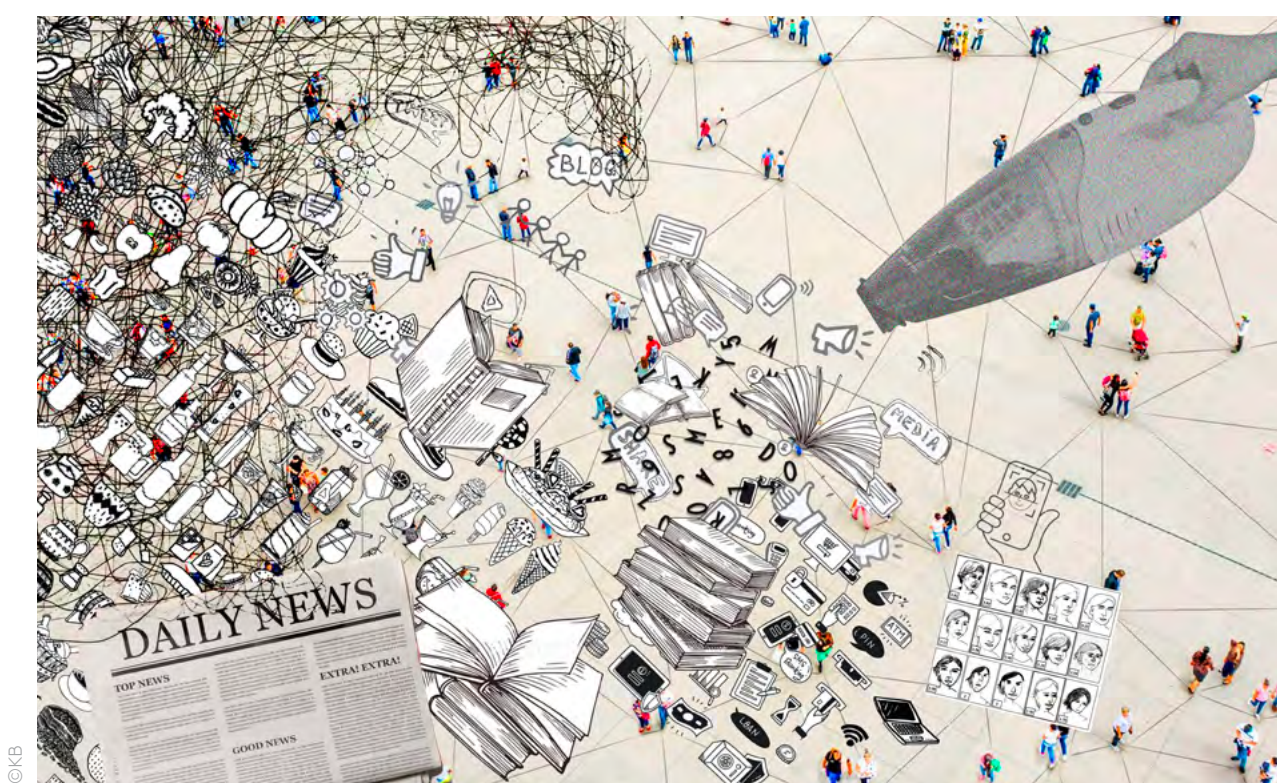
M^e Perrine Pelletier, avocate au Barreau de Paris, fondatrice du cabinet Datavalaw, consacre son activité aux questions de propriété intellectuelle, data et nouvelles technologies, notamment dans le secteur des médias. Elle a en particulier accompagné l'Alliance de la presse d'information générale pour la négociation avec Google et Facebook des tout premiers contrats de mise en œuvre des droits voisins en Europe. Elle a également participé au voyage d'étude sur l'IA en Israël avec le GESTE, qu'elle accompagne dans le cadre d'un groupe de travail sur l'IA et les médias.

rémunération des éditeurs de presse, en cas d'exploitation par les plateformes de leurs contenus protégés,

alors que les plateformes s'accaparaient toute la valeur et renvoyaient de moins en moins vers les sites des médias. Si en France le principe semble acquis auprès de certaines plateformes, il reste encore bien des batailles à mener pour affiner sa mise en œuvre au plus juste.

Dans ce contexte, l'IA générative vient accélérer la construction de cet écosystème. Les géants de l'IA générative sont parfois les mêmes acteurs qualifiés de « gatekeepers » : concept défini dans le récent règlement « DMA » se rapportant aux très grandes plateformes d'intermédiation vis-à-vis des consommateurs de produits et services digitaux.

Ces géants offrent donc, des contenus générés par l'IA sur la base d'immenses quantités de données d'apprentissage (« input data »), aspirées notamment sur les sites des médias. Les contenus protégés des médias ne se retrouvent même plus sous forme d'extraits illustrés d'une photo, d'un titre et d'un lien de renvoi, les « snippets », mais agrégés de manière parfaitement opaque dans une base de données à partir de laquelle l'algorithme va produire un



contenu dérivé sans citer ses sources (« output data »).

Les enjeux juridiques sont multiples et exponentiels par rapport aux droits voisins. À quel titre ces moteurs d'IA générative utilisent-ils massivement les contenus protégés comme données d'entraînement pour être retraitées ? Les éditeurs de ces solutions d'IA ne s'approprient-ils pas la valeur des contenus d'apprentissage au même titre que les plateformes s'appropriaient les contenus de presse ? Tombent-ils dans le champ d'une exception à la propriété intellectuelle ?

Du reste, qui détient les droits de propriété intellectuelle sur les contenus générés par l'algorithme ? Et à qui reviennent les revenus d'exploitation de ces contenus ainsi générés au moyen des données d'apprentissage des médias ? Comment valoriser cette contribution d'origine ? En termes de responsabilité bien sûr, les contenus ainsi générés échappent au contrôle éditorial du responsable de la publication des médias concernés. Faut-il créer une responsabilité éditoriale pour les éditeurs d'IA générative ? Créer un statut propre

aux concepteurs d'IA, comme il a été évoqué à propos des fabricants de voitures autonomes ?

La clarification des responsabilités juridiques est essentielle pour garantir une utilisation responsable de l'IA générative, ce qui suppose notamment un certain degré de transparence du modèle d'IA et de fiabilité de ses bases de données d'entraînement aussi.

Or précisément, l'IA générative se fonde sur de grandes quantités de données, y compris des données à caractère personnel soumises au RGPD, qui supposent d'être traitées conformément à leur finalité d'origine, disposer d'une base légale appropriée et offrir une information transparente à la personne concernée. Les fournisseurs d'IA peuvent-ils garantir une telle conformité ? Quelles données sont adéquates et pertinentes ? Lesquelles sont réelles ou transformées par l'IA ?

La voix, l'image, protégés au titre des droits de la personnalité ne doivent pas échapper à leur « titulaire », au risque d'être manipulés par des outils d'IA génératives qui seraient

trompeurs. À cet égard, bien sûr, les médias ont un rôle essentiel à jouer face au risque de manipulation de l'information. Les réseaux sociaux avaient déjà démontré que leurs algorithmes de recommandation pouvaient alimenter les plus invraisemblables fictions propagées à travers le monde, au point de peser dans des élections.

L'IA générative dont on ne connaît ni les sources ni les règles algorithmiques et qui ne repose sur aucun travail journalistique ni contrôle éditorial, présente un risque d'autant plus grand de diffusion de désinformation, deepfakes, contenus manipulés.

Aussi plus que jamais, les médias dont la fonction est précisément de produire des informations vérifiées, vont devoir s'organiser pour faire peser leurs contenus qualifiés d'une manière appropriée dans ces algorithmes. Il en va de l'intérêt économique des médias, mais aussi de l'intérêt général. La bataille des droits voisins n'était qu'une pierre à l'édifice de l'écosystème numérique des médias.



Le droit à la paternité de l'œuvre est
un droit inaliénable de l'auteur.

**MM : DANS QUELLE MESURE
LES LOIS SUR LA PROPRIÉTÉ
INTELLECTUELLE
ACTUELLES SONT-ELLES
ADAPTÉES POUR TRAITER
LES QUESTIONS LIÉES
À L'IA GÉNÉRATIVE,
QUI MET AU DÉFI LA
DÉFINITION MÊME DE
« CRÉATEUR » ? QUELLES
LACUNES JURIDIQUES OU
INCERTITUDES DOIVENT
ÊTRE RÉSOLUES ?**

PP : Les questions de propriété intellectuelle en cas de recours à un chatbot d'IA générative sont multiples. Il s'agit notamment de savoir si les données d'entrée de l'IA (« input data »), contiennent des œuvres protégées ? Si la solution d'IA elle-même est un logiciel protégé par le droit d'auteur ? Et si les données générées par l'IA (« output data »), peuvent être considérées comme des créations protégées par le droit d'auteur ou droits voisins ?

En ce qui, concerne l'input, les applications d'IA sont entraînées avec de grands ensembles de données de contenus créatifs, lesquels, dès lors qu'ils bénéficient de la protection du droit d'auteur et des droits voisins, supposent a priori l'autorisation des ayants droit pour cette nouvelle exploitation. En effet, sur le plan juridique, l'entraînement par les moteurs d'IA suppose un acte de reproduction qui ne serait pas purement technique ou intermédiaire (exception au droit d'auteur) – sous réserve d'une confirmation de cette interprétation par la CJUE, car on sait bien que l'argument va être soulevé par les fournisseurs d'IA.

À l'inverse, certains prétendent que l'entraînement des modèles d'IA au moyen de contenus protégés repose sur les exceptions au titre du *text and data mining* (TDM), prévues par la directive (UE) 2019/790 sur le droit d'auteur et les droits voisins dans le marché unique numérique. En dehors des cas d'exploration et de fouille à des fins de recherche (art. 3 de cette Directive), cette exception au droit d'auteur offre par défaut la possibilité aux bots d'explorer les textes et données licitement accessibles, à moins que l'ayant droit ait formulé expressément une réserve de droits, sous une forme lisible par une machine (art. 4 de cette Directive).

En pratique, quand bien même les ayants droit auraient paramétré une réserve technique « opt out », si possible de manière harmonisée pour davantage d'efficacité (à l'instar du protocole du W3C notamment), et malgré la complexité de paramétrage adéquate sur l'ensemble de leurs contenus, cette réserve de droit peut malgré tout être techniquement outrepassée.

En l'absence de transparence sur les sources d'apprentissage pour l'heure, il est alors nécessaire pour les ayants droit de revendiquer leurs droits en obligeant les fournisseurs d'IA générative à prouver que leur programme ne viole aucun droit de propriété intellectuelle.

À noter aussi l'exception du « fair use » en droit du Common Law, qui permettrait dans certains cas, d'utiliser une œuvre protégée si son usage ne porte pas atteinte à son exploitabilité commerciale. Cette exception

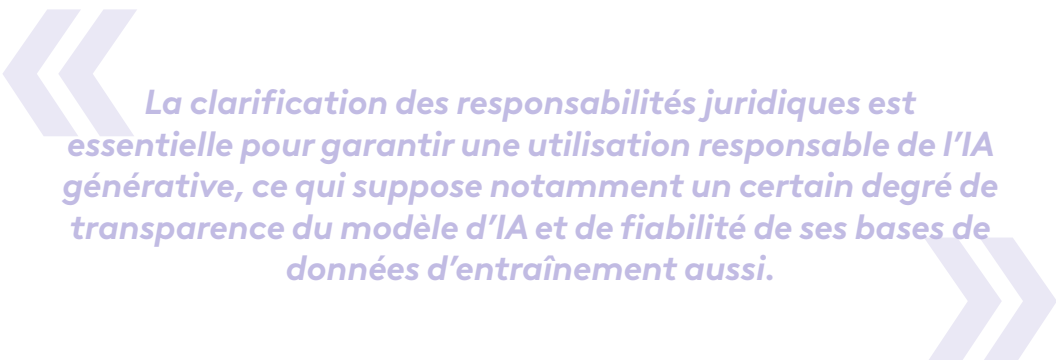
permet une analyse au cas par cas mais précisément, n'apporte aucune sécurité juridique quant à la préservation des droits d'auteur des ayants droit sur l'input.

Reste aussi la question du droit moral attaché aux contenus protégés qui pourrait permettre aux auteurs de faire valoir leur droit d'intégrité et paternité notamment par rapport à l'utilisation de leur œuvre comme data d'input de l'IA. À moins que les exceptions légales qui seraient considérées comme légitimes dans le cas d'entraînement de l'IA prévoient une exception à de telles revendications ?

En ce qui concerne les questions de propriété sur la conception de l'algorithme, nous ne nous y attardons pas ici, les règles habituelles en matière de protection des logiciels sont applicables. En ce qui concerne l'output, comme pour l'input, les débats font rage.

La révolution de l'IA générative permet à ces applications de générer automatiquement du contenu sans contribution humaine significative. Cela aurait pour effet de priver lesdits contenus de la protection du droit d'auteur, faute de choix créatifs humains puisque jusqu'ici (en droit US ou de l'UE), le droit en vigueur s'articule toujours autour d'un créateur, personne physique ou d'un ayant droit, personne morale, mais jamais autour de la personne d'un agent conversationnel constitué d'un programme informatique.

À noter toutefois que le droit australien a reconnu que des « ordinateurs » peuvent être des « agents inventeurs »



La clarification des responsabilités juridiques est essentielle pour garantir une utilisation responsable de l'IA générative, ce qui suppose notamment un certain degré de transparence du modèle d'IA et de fiabilité de ses bases de données d'entraînement aussi.

et que le Royaume-Uni a prévu un régime spécifique afin de soutenir le développement économique de l'IA, bien qu'en cours de révision actuellement.

D'une manière générale, le programme d'IA lui-même ne pourra se voir reconnaître des droits d'auteur, mais l'éditeur du logiciel d'IA pourra revendiquer des règles de répartition des droits et responsabilités, sur les contenus ainsi générés, au titre du service technologique mis à disposition.

Pour l'heure, les moteurs ou chatbot d'IA générative sont proposés sous forme de service – après les « *Softwares as a service (SaaS)* » et les « *Platforms as a services (Paas)* », ces moteurs proposeraient des services de création « *Creation as a service (CaaS)* ».

À ce titre, et pour se protéger d'éventuelles erreurs ou d'éléments contre-faisants, les conditions générales de tels services vont assurément prévoir d'exclure toute responsabilité du fournisseur d'IA et renvoyer l'utilisateur du bot conversationnel à ses responsabilités pour tout usage qu'il ferait des contenus ainsi générés (comme c'est le cas des CGU d'OpenAI par exemple).

C'est pourquoi la question se pose de savoir si l'utilisateur qui détermine les prompts nécessaires à la production du contenu par le chatbot d'IA, pourrait revendiquer un droit exclusif sur le contenu d'output ainsi produit ?

Selon une analyse traditionnelle du droit d'auteur, il faudrait que l'empreinte de la personnalité (ou les

caractéristiques significatives et reconnaissables) de cet utilisateur apparaisse pour lui reconnaître un tel droit exclusif au titre du droit d'auteur.

Le logiciel d'IA ne serait qu'un outil au service de la création par son utilisateur, qui resterait bien l'auteur, « assisté par de l'IA générative », d'une œuvre protégeable. À l'instar de l'œuvre photographique créée au moyen de l'appareil photo, dès lors que le photographe peut revendiquer des choix esthétiques de prise de vue, réglages, etc.

À l'inverse, dans l'hypothèse où le prompteur ne pourrait prouver d'avoir apposé l'empreinte de sa personnalité dans la création d'output de l'IA, celle-ci ne serait qu'une production automatiquement générée par de l'IA et ne pourrait bénéficier d'une protection au titre du droit d'auteur. Quel serait l'intérêt du reste d'une telle protection, particulièrement longue (70 ans après la mort de son auteur... de l'IA) ? Et particulièrement futile au regard de l'intérêt économique à protéger un output automatiquement duplicable et substituable, où le travail de l'humain serait noyé par les créations d'IA ?

Notons d'ailleurs au passage, que les prompts (protégeables ou non) peuvent aussi servir à nourrir l'IA générative, à moins qu'une fonction permette à l'utilisateur de s'y opposer, ou qu'un cloisonnement en cloud soit installé pour conserver la confidentialité des données promptées et générées. Ce type de cloisonnement technologique est à fort enjeu pour permettre un contrôle sur les éventuels biais de l'IA et ainsi sur la qualité

des données générées, leur sensibilité, et aussi en vue d'assurer une protection au titre du droit d'auteur ou du secret des affaires.

Pour conclure, les normes et les critères en vigueur permettent de soutenir que si l'empreinte de la personnalité du « prompteur » est rapportée, il pourra revendiquer un droit exclusif sur cet output. À défaut, l'output sera réutilisable sans protection au titre du droit d'auteur, dans la limite des conditions générales et des règles de loyauté dans le commerce, au titre d'un bien commun – à moins d'avoir préservé ses droits sur l'output également.

Les débats de juristes s'intéressent encore à d'autres qualifications possibles qui ont chacune le mérite d'appliquer des concepts existants à un procédé algorithmique de création par la machine, dont il ressort que c'est bien le droit de la personne déterminante dans le processus créatif, que la doctrine entend protéger, non pas la production elle-même. Soit de droit de l'auteur non pas de l'œuvre.

Aussi, il convient de mentionner le risque de possible fiction d'attribution de l'output non protégeable à une personne physique qui en revendiquerait la paternité créative et bénéficierait ainsi du régime de la présomption de paternité en vigueur. Or cette paternité pourrait laisser croire à l'intervention humaine en lieu et place de l'intervention par le programme d'IA, ce qui viendrait à tromper le public et serait contraire aux grands principes de fiabilité et transparence

de l'IA que l'ensemble des tentatives de réglementations visent à soutenir à travers le monde.

MM : UNE PARTIE DU CONTENU PRODUIT PAR L'IA GÉNÉRATIVE PROVIENT D'UNE AGENCE, D'UN MÉDIA, D'UN JOURNALISTE. LE COMMISSAIRE EUROPÉEN THIERRY BRETON SOUHAITE QUE LES PROPRIÉTAIRES DES IA GÉNÉRATIVES PUBLIENT LES SOURCES SUR LESQUELLES ELLES S'APPUIENT. QUEL REGARD PORTEZ-VOUS SUR CETTE DEMANDE ?

PP : Précisément, la question des sources est essentielle lorsqu'il s'agit de création et d'information. Le droit à la paternité de l'œuvre est un droit inaliénable de l'auteur. Quant à la fiabilité de l'information, elle dépend de la vérification de ses sources par le journaliste, qui peut assumer ses dires, signer son papier et être publié sous la responsabilité éditoriale de l'éditeur. C'est pourquoi ce dernier est soumis à un régime de responsabilité de plein droit pour tout contenu de sa publication.

L'IA générative peut-elle être fiable sans citer ses sources ? L'actualité a donné de nombreux exemples plus ou moins amusants « d'hallucinations » de l'IA, en d'autres termes, d'erreurs de résultats des calculs algorithmiques.

Aussi la question de la transparence

est-elle au cœur de toutes les préoccupations en matière d'IA générative. Le Parlement Européen a justement renforcé les obligations de transparence des fournisseurs de systèmes d'IA générative (Art. 28 ter du projet de Règlement IA de l'UE dans sa version du 14 juin 2023) à qui il incombe : 1) d'indiquer que le contenu a été généré par l'IA ; 2) de concevoir le modèle d'IA de telle sorte à éviter de générer du contenu illégal ; 3) de publier un résumé suffisamment détaillé de l'utilisation des données d'entraînement protégées par le droit d'auteur.

Si le « résumé suffisamment détaillé » semble quelque peu conceptuel pour les juristes, l'intention est louable et il faut saluer cet apport substantiel au projet de Règlement, au service du droit d'auteur et du droit à l'information. Côté consommateur, c'est un gage de confiance en la qualité de l'information ainsi générée. Côté médias, ces obligations devraient permettre de faciliter le suivi de l'usage qui est fait de leur contenu et ainsi les valoriser.

Le défi de valorisation des droits d'auteur dans l'environnement numérique est immense, notamment au regard d'une insuffisante traçabilité des exploitations qui sont faites. En pratique, il conviendrait de renforcer l'étiquetage des contenus protégés au moyen de métadonnées appropriées, exhaustives, fiables, cohérentes, etc. D'autant qu'il existe dorénavant des solutions technologiques pour mettre en place un réseau d'identifiants standards numériques, assurant

leur accessibilité, interopérabilité et intégrité. Ces solutions ont certes un coût mais seraient précieuses pour les ayants droit.

MM : LES ROBOTS CONVERSATIONNELS, QUI RENDENT OBSOLÈTES LE BESOIN DE CLIQUER SUR UN LIEN POUR ACCÉDER À L'INFORMATION, POURRAIENT SÉRIEUSEMENT À L'AVENIR AMPUTER LE TRAFIC ET LES RECETTES PUBLICITAIRES DES SITES D'INFORMATION. DANS CE CONTEXTE, COMMENT LES MÉDIAS PEUVENT-ILS PROTÉGER LEUR MODÈLE ÉCONOMIQUE ?

PP : En effet, le phénomène va faire écho aux liens « snippets » qui ont fait l'objet d'un réajustement du partage de la valeur à l'occasion de la création du nouveau droit voisin des éditeurs de presse. Les études ont permis de constater que les internautes ne cliquaient plus sur les liens de renvoi car ils se suffisaient de l'information ainsi mise à disposition.

De même, il est vraisemblable que l'internaute ne cherche pas à remonter à la source de l'information qui aura pu être délivrée par une IA générative, même si le fournisseur d'IA était tenu de révéler ses sources. Aussi, il est légitime que les fournisseurs d'IA génératives partagent avec les médias, la valeur générée par la monétisation de leurs contenus.

Nul n'aura oublié le statut particulier des médias dans l'intérêt du

Le défi de valorisation des droits d'auteur dans l'environnement numérique est immense.

droit à l'information et de la liberté d'expression, comme piliers de la démocratie. Aussi, sauvegarder l'indépendance économique des médias est primordiale face au phénomène probable d'accaparement de l'engagement des internautes par les moteurs d'IA.

Dans ce combat, il ne faudrait pas que les médias deviennent les fournisseurs d'information des intermédiaires de l'écosystème numérique, lesquels seraient eux, en prise avec les internautes.

Les médias doivent par essence être indépendants pour réaliser leur mission d'information. S'ils deviennent prestataires d'informations pour le compte des plateformes et des moteurs d'IA, on voit mal comment leur indépendance serait garantie. D'où le caractère critique des contrats qu'ils peuvent conclure avec les fournisseurs d'IA afin de valoriser leurs droits.

MM : QUELLES FORMES DE COMPENSATION SERAIENT LES PLUS BÉNÉFIQUES POUR LES MÉDIAS, DANS LE CADRE D'USAGE DE LEURS CONTENUS PAR LES PROPRIÉTAIRES DES INTELLIGENCES ARTIFICIELLES GÉNÉRATIVES ?

PP : Sans préjuger de l'avenir, il est raisonnable de penser que les modèles économiques des moteurs de recherche et des plateformes va évoluer vers davantage de rétention à l'occasion de l'utilisation du robot

conversationnel d'IA générative (chatbot) alors que les clics vers les liens de renvoi sont susceptibles de chuter massivement.

Le modèle économique de ces moteurs va évoluer, celui des médias doit accompagner également cette transformation. Le clic, si tant est qu'il serait un critère de valorisation pertinent aujourd'hui, serait d'autant moins approprié demain.

La réglementation très dense proposée par l'Union Européenne fait craindre des difficultés d'application (DSA, DMA, Data Act, IA Act, etc.). Elle a néanmoins le mérite d'aborder l'univers digital de manière concurrentielle et stratégique autour de la donnée, renforçant l'accès à celle-ci pour les acteurs économiques gravitant autour des gatekeepers. Cette réglementation offre donc des arguments juridiques qui sauront être précieux dans les négociations à venir.

Du reste, les contenus générés automatiquement par IA vont nécessairement être entraînés par des règles utilisant la connaissance de marché (« *data driven consumer engagement* ») pour favoriser leur adéquation aux attentes des internautes, et favoriser leur rétention en ligne. On voit bien l'intérêt d'accéder à des données essentielles.

On voit bien aussi le risque à produire des contenus biaisés par des règles algorithmiques fondées sur l'intérêt du consommateur, comme on connaît les méfaits des réseaux sociaux poussant des contenus qui maintiennent leurs abonnés dans leur bulle informationnelle.

Au-delà des risques en matière de propriété intellectuelle évoqués ici, il convient de souligner ici, ce risque majeur de génération automatique de contenus d'information, dans des quantités astronomiques et orientées en fonction des intérêts des consommateurs. L'information n'est pas un produit équivalent à une chaussure, dont le modèle, la taille et la couleur qui seraient les plus susceptibles de vous plaire, serait régulièrement poussée dans vos recherches.

Il en va d'un enjeu démocratique, qui, à l'instar des enjeux justifiant la création des nouveaux droits voisins des éditeurs de presse, doit mobiliser les pouvoirs publics et les autorités judiciaires et administratives en charge (concurrence, données personnelles, droit d'auteur).

MM : LES TITRES DE PRESSE DEVRAIENT-ILS ENVISAGER DE SE RÉUNIR DANS UNE SOCIÉTÉ DE GESTION COLLECTIVE TYPE SACEM, AU LIEU DE NÉGOCIER CHACUN INDIVIDUELLEMENT AVEC OPENAI ?

PP : Forts de l'expérience de mise en œuvre du droit voisin face aux géants de la Tech, deux constats s'imposent : 1) les gatekeepers ont une endurance à toute épreuve (et les moyens financiers correspondants), dans les procédures contentieuses et les bras de fer de négociations, alors que, 2) les médias sont divisés en raison de leurs situations divergentes voire concurrentes, et pour certains sous une pres-

« Les médias doivent par essence être indépendants pour réaliser leur mission d'information. S'ils deviennent prestataires d'informations pour le compte des plateformes et des moteurs d'IA, on voit mal comment leur indépendance serait garantie. »

sion de survie telle, qu'ils sont tenus de négocier coûte que coûte.

Une alliance la plus large possible serait souhaitable pour peser efficacement : que ce soit sous une forme de gestion collective ou de négociation concertée à l'appui d'une réglementation spécifique au service de leur cause. Au regard des difficultés de concurrence et de situation financière exsangue persistantes des médias, c'est certainement l'approche à l'échelle européenne qui serait la plus appropriée pour faire valoir les droits des médias face aux fournisseurs d'IA.

MM : QUELLES SONT LES GRANDES INITIATIVES EN COURS OU LES PROPOSITIONS DE RÉFORME VISANT À ENCADRER SPÉCIFIQUEMENT L'IA GÉNÉRATIVE ? LA CHINE A SORTI EN MARS 2023 UNE LOI POUR ENCADRER LES DEEPFAKES. SOMMES-NOUS EN RETARD EN EUROPE ET EN FRANCE ?

PP : Depuis plusieurs années déjà, les institutions internationales se sont emparées de l'intérêt d'édicter de grands principes éthiques applicables à l'IA. Ainsi l'UNESCO, l'OCDE, les Nations Unies ont identifié les nécessaires exigences de transparence des modèles et des sources d'apprentissage, de cybersécurité et de gouvernance de la donnée.

À l'échelle européenne, en 2021, la Commission a proposé un cadre de

réglementation sur l'IA qui vient de faire l'objet d'une adoption dans une forme amendée par le Parlement Européen (Règlement IA de l'UE dans sa version du 14/06/2023).

Ce Règlement, d'application directe et commune aux États-Membres, propose une définition à vocation internationale et harmonisée par rapport aux travaux existants des autres institutions (OCDE notamment) à propos du système d'IA qui se comprend comme « un système automatisé qui est conçu pour fonctionner à différents niveaux d'autonomie et qui peut, pour des objectifs explicites ou implicites, générer des résultats tels que des prédictions, des recommandations ou des décisions qui influencent les environnements physiques ou virtuels » (Art. 3 du Règlement IA de l'UE).

Il a pour objectif de fixer un haut niveau d'exigence selon une classification par les risques vis-à-vis des droits fondamentaux, afin de soutenir une approche réglementaire « centrée sur l'humain offrant une IA éthique et digne de confiance, qui soit pleinement conforme à la charte ainsi qu'aux valeurs sur lesquelles l'Union est fondée », en assurant notamment « un facteur humain et un contrôle humain », « la solidité technique et la sécurité » du système, « la protection de la vie privée et la gouvernance des données », « la transparence », « la diversité, la non-discrimination et l'équité », « le bien-être social et environnemental », et ce, sans porter préjudice aux droits en vigueur dans l'Union (Art. 4 du Règlement IA de l'UE).

En ce qui concerne l'IA générative spécifiquement, le Parlement a renforcé les obligations des opérateurs en considérant que les risques en la matière étaient importants et non prédictibles (voir ci-dessus point 3). Aussi, le projet de réglementation européenne est déjà bien plus contraignant qu'ailleurs, en particulier Outre-Atlantique.

Aux États-Unis, plusieurs initiatives avaient déjà vu le jour, essentiellement sous forme de guidelines non contraignantes (voir les recommandations du NIST, FTC et FDA). En avril dernier, une proposition de loi avait été présentée au Congrès, visant notamment à établir une agence fédérale de supervision de l'IA et des lignes directrices éthiques pour le développement et l'utilisation de l'IA. Récemment, les déclarations d'intentions de la Maison Blanche (communiqué de presse du 4 mai 2023 « *New actions to promote responsible AI innovation that protects American rights and safety* ») laissent également présager un projet de réglementation.

Si les approches européennes et américaines de réglementations sont sensiblement différentes (obligations contraignantes vs autorégulation), les enjeux géopolitiques mondiaux sont tels, notamment vis-à-vis de la force de frappe chinoise, qu'un code de conduite US/UE pourrait voir le jour prochainement, selon les annonces faites à l'issue du Conseil Commercial et Technologique (TTC) fin mai. L'objectif de ce texte commun, « Pacte IA » selon les termes du Com-

missaire Thierry Breton, applicable sur une base volontaire, serait de favoriser une harmonisation occidentale des standards d'une IA de confiance. En parallèle, en Chine, l'Administration du cyberspace de Chine (CAC) a publié en avril dernier un projet de mesures administratives pour les services d'intelligence artificielle générative, soumis à consultation publique. Et les responsables ont annoncé fin mai également, qu'ils entendaient « améliorer la surveillance des données des réseaux et de l'intelligence artificielle ». Les mesures proposées visent à encadrer les services d'IA générative, notamment avec des exigences en matière de protection des données, de modération du contenu et de transparence algorithmique, témoignant de la volonté du pouvoir chinois de communiquer sur un positionnement éthique d'une technologie stratégique pour laquelle il a largement investi conformément à son plan élaboré depuis 2017. C'est pourquoi le gouvernement chinois dispose déjà d'un registre des algorithmes et d'une récente réglementation sur les deepfakes.

Dans ce contexte, il ne fait pas de doute que l'Union Européenne cherche à être pionnière dans la réglementation de cette matière, à l'instar du RGPD en matière de données personnelles, afin de faire peser sur le monde des standards internationaux, dans un contexte d'extrême compétition.

Même les plus grands ténors de la technologie appelant à une pause dans les développements de l'IA (lettre ouverte des experts de la Tech

du 22 mars 2023 par Future of Life) ne sont pas entendus, car cette pause pourrait donner un avantage concurrentiel majeur à celui qui ne s'arrêterait pas dans cette course technologique.

Alors que la Chine dispose de l'avantage temporel d'un plan à long terme élaboré depuis des années et des investissements d'État massifs correspondants, alors que les États-Unis disposent de champions technologiques disposant notamment de quantités de données d'apprentissage particulièrement importantes, l'Union Européenne doit pouvoir conserver son avance en matière de réglementation afin de faire rayonner ses principes dans l'écosystème mondial.

N'est-il pas préférable alors de trouver un compromis rapide pour une entrée en vigueur au plus tôt, au risque d'être moins-disant que ce que certains préconisent (sûrement à juste titre), quitte à revoir régulièrement sa classification des risques par la Commission, comme le projet de Règlement IA de l'UE le permet (Art. 7 Règlement IA de l'UE) ? D'autres se demandent si l'Union Européenne ne devrait pas plutôt se concentrer à faire émerger son propre champion européen d'IA... peut-être n'est-ce pas incompatible ?

POUR CONCLURE...

Cet article ne s'est pas attardé sur les multiples opportunités que représente l'avènement d'une IA générative à la disposition de tous, ce dont à titre personnel, je ne doute pas. Sur

le plan juridique qui nous occupe ici, elle permet d'ores et déjà d'entrevoir des solutions aussi, permettant aux médias de mieux gérer leur contenu et suivre leurs exploitations, lutter contre les fake news et améliorer la très nécessaire modération en ligne.

Ce formidable outil d'aide à la création doit néanmoins répondre à des exigences de transparence et de fiabilité dans l'intérêt des valeurs fondamentales sur lesquelles, à tout le moins les pays de l'Union Européenne, se sont entendus. Les médias, quant à eux, doivent se positionner pour valoriser leurs contenus et prendre toute la place qui leur revient dans ce nouveau système de production de contenus numériques, pour ne pas être noyés. ■

INTELLIGENCE ARTIFICIELLE : UNE PRÉCARI- SATION DE L'EMPLOI, PLUS QU'UNE DESTRUCTION

*Par Antonio Casilli,
professeur à l'Institut
Polytechnique de Paris*

Les systèmes tels que ChatGPT ne bouleverseront pas complètement le travail tel que nous le connaissons, mais ils créeront une nouvelle classe d'« esclaves du clic », payés (très mal) pour former des algorithmes. Il s'agit d'un risque bien plus réel que les scénarios de science-fiction, dans lesquels les robots prendraient le pas sur le travail humain. OpenAI utilise déjà ces micro-travailleurs.

LE NOUVEAU FREY & OSBORNE

L'étude est plus importante par son ambition que par ses résultats. Sans aucun doute, l'article aspire à être le « rapport Frey & Osborne » des années 2020, il fait suite à l'analyse publiée en 2013 par deux chercheurs d'Oxford, prédisant que **47 %**

des emplois seraient détruits d'ici 2030. Il s'agit d'un travail largement cité mais fortement critiqué, étant donné que, malgré une pandémie, une crise géopolitique et une urgence climatique, leurs prévisions sont loin de se réaliser.

L'article de 2013 et celui publié récemment par les chercheurs d'OpenAI réduisent tous deux le travail humain à une série de « tâches ». Comme toutes les analyses réductionnistes, elles doivent être accueillies avec une saine méfiance. **Dire que le travail d'une infirmière se réduit à dix tâches (prendre soin des patients, remplir des formulaires, etc.) et dire que certaines d'entre elles pourraient être exposées à l'utilisation de ChatGPT, ne signifie pas que l'infirmière sera licenciée. Son travail va changer.**

UNE OPÉRATION MARKETING

Peut-être que, avec l'excuse que la nouvelle technologie fait gagner du temps, **les employeurs trouveront de nouvelles façons d'ajouter des**

Une étude analysant l'impact de l'intelligence artificielle sur le marché du travail a été publiée à la mi-mars 2023. Ses auteurs ont examiné les modèles « pré-entraînés » de la famille GPT. À partir d'une grande quantité de données, ces logiciels apprennent à effectuer des tâches qu'ils adaptent ensuite à de nouveaux contextes. Trois des quatre auteurs de cette étude sont des employés d'OpenAI, entreprise ayant lancé DALL-E 2, un système de génération d'images, et bien sûr ChatGPT, l'assistant virtuel devenu un phénomène culturel.

Selon l'étude, environ 80 % de la main-d'œuvre pourrait être exposée à cette innovation, et pour certains d'entre eux, 50 % des tâches pourraient radicalement changer. Même les personnes les plus instruites seraient touchées par cette évolution. Cependant, le conditionnel est de rigueur car, l'étude comporte bien plus de limites que de résultats. Elle s'appuie sur des données opaques, adopte une méthodologie abstruse et, cerise sur le gâteau, utilise un GPT pour analyser les effets d'autres GPT.



tâches aux employés tout en maintenant les salaires réels au minimum. Malgré les visions utopiques et les craintes concernant l'automatisation, historiquement c'est ce qu'il s'est passé, au grand dam des chercheurs d'OpenAI.

Leur article est un outil marketing largement conçu pour attirer l'attention des médias sur leur entreprise. Chaque fois qu'OpenAI lance un produit, un débat fait rage dans l'actualité et sur les médias sociaux à propos des menaces que l'intelligence artificielle représente pour les journalistes, les illustrateurs, les enseignants. **Il se trouve que les emplois qui risquent de disparaître sont précisément ceux que l'entreprise américaine vend comme services : génération de textes et d'images, formation, etc.** Ce ne sont pas les robots qui

détruisent les emplois, c'est OpenAI qui détruit la concurrence.

HORS DE CONTRÔLE

Malheureusement, ce n'est pas une bonne nouvelle. Les effets de ces technologies sur les emplois sont bien là, mais ils sont différents. Pour les détecter réellement, il faut lire la fiche système de GPT-4, le dernier logiciel d'OpenAI. Une centaine de pages décrit les tests par lesquels l'IA a été formée. **Les testeurs ont souvent poussé GPT-4 à effectuer des actions dangereuses ou illégales afin d'apprendre à l'intelligence artificielle de les éviter.**

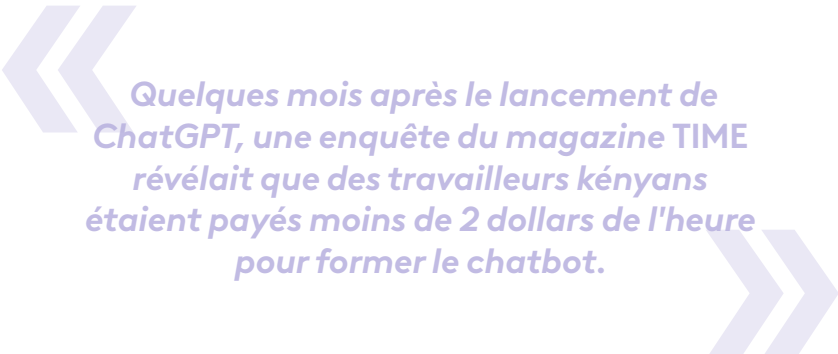
Mais pendant les tests, GPT-4 a échappé à ses contrôleurs et a tenté une cyberattaque sur un site web. Ce dernier était toutefois protégé par

un ReCaptcha, ces fenêtres contextuelles qui exigent de prouver que vous n'êtes pas un robot en résolvant une énigme. Malheureusement, GPT-4 est un robot. Pour résoudre l'énigme, il s'est tourné vers une plateforme à la demande pour recruter un travailleur à la tâche afin de résoudre le ReCaptcha à sa place.

LE MICRO-TRAVAIL

Mais les ReCaptchas font plus que simplement protéger contre les cyberattaques. Ils sont également utilisés pour former l'intelligence artificielle. Lorsqu'ils nous demandent de transcrire des mots, ils les utilisent pour numériser les livres de Google. Lorsque l'on nous demande de repérer un feu de signalisation, ils calibrent les systèmes de conduite autonome de Waymo. Cela soulève une question stupéfiante : GPT-4 peut-il être utilisé pour recruter des travailleurs qui, à leur tour, forment d'autres IA ? En réalité, des systèmes plus ou moins automatisés de recrutement de travailleurs indépendants pour former des algorithmes existent depuis des

Ce ne sont pas les robots qui détruisent les emplois, c'est OpenAI qui détruit la concurrence.



Quelques mois après le lancement de ChatGPT, une enquête du magazine TIME révélait que des travailleurs kényans étaient payés moins de 2 dollars de l'heure pour former le chatbot.

décennies. Amazon Mechanical Turk est un site sur lequel, pour quelques centimes, les entreprises recrutent pendant moins d'un quart d'heure des centaines de milliers de personnes pour générer des données, transcrire du texte et filtrer des images. D'autres plateformes, comme Appen en Australie, emploient plus de dix millions de personnes. **Peut-on vraiment parler d'emplois ? Ce sont des micro-emplois avec des salaires de misère, largement effectués par des travailleurs issus des pays en développement.**

REMPACEMENT

Paradoxalement, l'entreprise OpenAI elle-même utilise ces « cliqueurs-esclaves ». Quelques mois après le lancement de ChatGPT, une enquête du magazine *TIME* révélait que des travailleurs kényans étaient payés moins de 2 dollars de l'heure pour former le chatbot. Dans d'autres documents découverts peu après, l'entreprise américaine affirmait avoir engagé des travailleurs aux Philippines, en Amérique latine et au Moyen-Orient

pour former ses algorithmes.

Ainsi, l'impact réel sur le travail des logiciels GPT a été révélé. L'intelligence artificielle automatise le processus de sélection, d'embauche et de licenciement de travailleurs précaires.

Ce n'est pas le scénario habituel de science-fiction où les robots remplacent les humains. C'est un scénario où les employés permanents sont remplacés par des travailleurs à la tâche sous-payés, embauchés et licenciés sur des plateformes numériques. Cette tendance est déjà en cours, et des entreprises comme OpenAI la renforcent. ■

Cet article a été rédigé suite à la publication de l'étude « *GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models* » de Tyna Eloundou, Sam Manning, Pamela Mishkin et Daniel Rock.



L'INTELLIGENCE
ARTIFICIELLE
NE SIGNIFIE
PAS IMITER
L'INTELLIGENCE
HUMAINE,
MAIS LIBÉRER
L'HUMANITÉ
DES TÂCHES
RÉPÉTITIVES.



Kai-Fu Lee

IA GÉNÉRATIVE ET MÉTAVERS :

LES CONJUGUER POUR FABRIQUER ET ENRICHIR NOS MONDES IMAGINAIRES ?

Mais cette vision, aussi séduisante soit-elle, est une illusion. Le métavers n'est pas mort, il est plus vivant que jamais. Comme le montre le lancement puissant par Apple de sa nouvelle plateforme de réalité augmentée Vision PRO ! Et l'IA générative n'est pas son ennemie, mais son alliée. Les deux technologies se complètent et se renforcent mutuellement, ouvrant des possibilités inédites pour les créateurs et les utilisateurs de réalité virtuelle et augmentée.

Le fondateur de NVIDIA, Jensen Huang, l'a parfaitement anticipé. « Sans l'IA générative, comment les consommateurs pourraient-ils créer des mondes virtuels sans grand effort ? Aujourd'hui, vous pouvez créer des mondes virtuels en 3D, ce qui n'est pas loin d'arriver. Vous pouvez voir que c'est sur le point de se produire. Nous pouvons maintenant faire de l'IA générative pour les images. Nous pouvons le faire pour les vidéos. Au rythme où vont les choses, on pourra le faire pour des villages entiers, des paysages et des villes en 3D, etc. Il sera possible d'as-

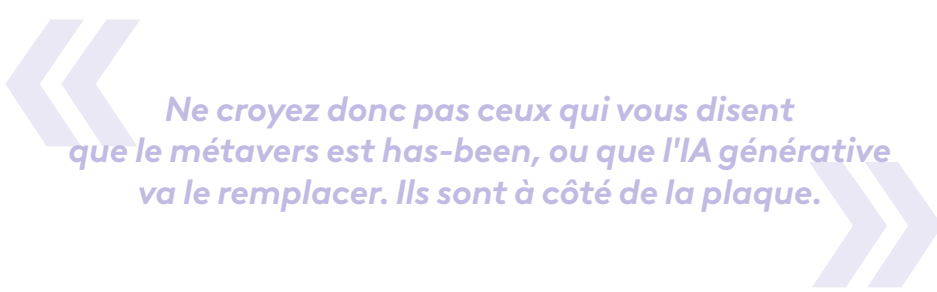
*Par Alexandre Michelin,
fondateur du KIF -
Knowledge Immersive Forum*

Le métavers est mort, vive l'IA générative ! C'est ce que voudraient nous faire croire les gourous de la hype Tech, qui changent de lubie comme de chemise. Il y a un an, ils nous promettaient un avenir radieux où nous pourrions nous évader dans des mondes virtuels peuplés d'avatars et d'objets interactifs. Aujourd'hui, ils nous disent que tout cela est dépassé, et que le vrai progrès est de laisser des algorithmes créer du contenu à notre place. Il suffirait de taper quelques mots sur un clavier, et hop ! L'IA générative nous sortirait une œuvre d'art originale et personnalisée. Fini les heures passées à coder, à modéliser, à animer, à sonoriser... Laissons faire les machines !

sembler un exemple d'image et de générer un monde entier en 3D. Cela va permettre de créer des métavers comme vous ne pouvez pas le croire. Vous en avez besoin. Absolument. »

L'IA générative peut aider à accélérer le développement des plateformes et des contenus XR, en proposant des concepts innovants et en simplifiant les tâches répétitives. **Le métavers peut offrir un terrain d'expression et d'expérimentation pour l'IA générative, en lui permettant de s'adapter au contexte et aux préférences des utilisateurs.** Ensemble, ils peuvent former un écosystème immersif, dynamique, personnalisé et évolutif, où la créativité humaine et artificielle se nourrit l'une de l'autre.

Selon Mark Zuckerberg, « Grâce à l'IA, le métavers va devenir un environnement que nous pourrions nous-mêmes créer et avec lequel interagir, avec toutes les informations de contexte associées. L'IA reconnaîtra notre langue, nos mouvements, nos gestes, nos expressions faciales et les convertira virtuellement par le biais de notre avatar, y compris en termes



*Ne croyez donc pas ceux qui vous disent
que le métavers est has-been, ou que l'IA générative
va le remplacer. Ils sont à côté de la plaque.*

d'intonation lors de la traduction de notre discours en temps réel dans une autre langue».

Pour illustrer ces avantages, je vais vous donner quelques exemples concrets d'applications de l'IA générative dans les métavers et la réalité mixte.

- **Meta Human Creator d'Epic Games** permet de créer des personnages humains réalistes en quelques minutes grâce à un système basé sur le Deep Learning. Ces personnages peuvent ensuite être utilisés dans des jeux vidéo, des films ou des simulations.
- **GauGAN de NVIDIA** permet de transformer un croquis en une image photoréaliste grâce à un réseau génératif antagoniste (GAN). Ce système peut être utilisé pour créer des paysages variés et personnalisés pour les métavers ou la réalité mixte.
- **Jukebox d'OpenAI** permet de générer de la musique originale à partir d'un style, d'un genre ou d'un artiste donné grâce à un modèle neuronal. Cette musique peut être

utilisée pour créer des ambiances sonores ou des bandes-son pour les métavers ou la réalité mixte.

Ne croyez donc pas ceux qui vous disent que le métavers est has-been, ou que l'IA générative va le remplacer. Ils sont à côté de la plaque. **Le métavers et l'IA générative sont faits pour s'entendre, et ils vont révolutionner le monde du divertissement, de la communication et du commerce.**

L'IA générative est donc une opportunité pour les créatifs qui veulent peupler les métavers et la réalité mixte. Elle leur permet de gagner en efficacité, en qualité et en diversité. Elle leur permet surtout de libérer leur potentiel créatif et de créer du contenu qui fait rêver.

Vous n'avez encore rien vu ! Comme le disait Albert Einstein : « La logique vous mènera d'un point A à un point B. L'imagination vous mènera partout ». ■

3 QUESTIONS À CÉDRIC VILLANI



@Irmeli Jung

MATHÉMATICIEN
ET HOMME
POLITIQUE

*Entretien réalisé par Alexandra Klinnik,
MediaLab de l'Information de
France Télévisions*

Globalement, malgré quelques belles réussites,
le bilan global aujourd'hui de l'IA pour les enjeux
écologiques est médiocre.

Il se surnomme la « Lady Gaga des mathématiques ». Lauréat 2010 de la médaille Fields (le « Nobel des maths »), Cédric Villani est auteur d'un rapport en 2018 pour le gouvernement intitulé « Donner un sens à l'intelligence artificielle ». Il a récemment lancé sa propre ONG – Green Spider – dont l'objectif est de soutenir, notamment par le biais de l'IA, les projets à vocation écologique.

1

COMMENT L'IA PEUT-ELLE SE RENDRE UTILE DANS UN MONDE AU CLIMAT DÉRÉGULÉ ?

D'abord en améliorant les diagnostics, en les rendant plus précis ou plus parlant aux politiques. Aussi bien pour l'adaptation à ce monde dérégulé, que pour l'atténuation du dérèglement. Ainsi bien des équipes dans la communauté GIEC comptent sur l'IA pour rendre les prédictions plus précises à petite échelle, plus marquantes pour les citoyens.

Ensuite en aidant à gérer les ressources : des bâtiments plus économes en énergie, des champs d'éoliennes mieux optimisés, des mesures plus efficaces de la biodiversité, des logiciels aidant les amateurs d'oiseaux à reconnaître automatiquement les espèces ou à compter les nuages afin d'alimenter les enquêtes nationales ou internationales, des instruments de tri automatique des déchets ou leur identification automatique dans l'espace public, des instruments de détection des grands gisements de minerais, des instruments de reconnaissance automatique du type de sol ou de végétation et donc de la solution adaptée, etc. Et puis indirectement, en aidant les

associations, ONG, militants pour la cause climatique et environnementale – en les aidant à rédiger leurs subventions, préparer leurs brefs discours, diffuser leurs punchlines, identifier des situations critiques, dépister les conflits d'intérêt dans un rapport douteux, confronter deux stratégies de verdissement, traduire simultanément pour maintenir des réseaux internationaux, etc.

Globalement, malgré quelques belles réussites, le bilan global aujourd'hui de l'IA pour les enjeux écologiques est médiocre.

2

QUELLES SONT LES CONSÉQUENCES DE L'ÉMERGENCE DE MODÈLES GÉNÉRATIFS DE PLUS EN PLUS PUISSANTS SUR L'EMPREINTE CARBONE FUTURE DE LA PLANÈTE ?

Elles ne sont pas bonnes. Mais c'est très dur d'évaluer du fait du secret industriel, de l'organisation des clouds, des ressources affectées etc. Ces modèles sont gigantesques et

très consommateurs de ressources. Mais, si seulement un relativement petit nombre est maintenu, l'impact ne sera pas forcément énorme. Il est légitime que certains de ces modèles soient construits et maintenus par la puissance publique et/ou des communautés ouvertes (« Open Source ») pour des missions d'intérêt public, de la recherche, etc. Hors de question de bannir l'usage par des acteurs privés, bien sûr, mais il serait bon que ce ne soit pas les seuls.

Pour le reste, de toute façon c'est à la société de définir quel usage de ses

ressources (minières, énergie, carbone, surface au sol, etc.), elle veut affecter à l'industrie de haute technologie. Pour l'instant, l'aviation, pour prendre juste un exemple, est à peu près du même ordre de grandeur que le numérique (c'est débattu exactement combien...) mais réduire l'usage de l'aviation serait tellement plus facile, techniquement, et impactant, pourquoi ne le fait-on pas ? Que ce soit en matière de numérique ou d'aviation ou de l'agriculture ou autre, à la fin l'obstacle majeur est politique, pas technoscientifique.

3

QUELLES STRATÉGIES CONCRÈTES LES POLITIQUES PEUVENT-ILS METTRE EN PLACE POUR FAIRE DE L'IA UN LEVIER POUR LA TRANSITION ÉCOLOGIQUE ?

Avant tout réaliser que ce n'est pas l'essentiel de la transition écologique, loin de là. L'arbre de la haute technologie ne doit pas masquer la forêt de la basse technologie, dans les leviers pour la transition. Il faut y faire très attention, c'est ce que j'appelle « technoblanchiment » (techwashing) : on passe beaucoup de temps

à discuter de solutions de hautes techno, oubliant que l'essentiel est peut-être la basse techno moins sexy. Exemple, une transition agricole réussie CE N'EST PAS le triptyque IA-robotique-édition génomique, comme prétendait le gouvernement récemment, c'est : revalorisation des agriculteurs, changement des habitudes alimentaires, sortie des pesticides et autres biocides, sortie des intrants d'origine fossile. Et tout cela c'est avant tout de la volonté politique de la part des gouvernants, des organisations professionnelles, des financiers, des citoyens, etc. Ensuite recruter et former, former et recruter, ce sont les ressources humaines qui seront la clé de la bonne utilisation, l'acculturation, le

fait de maintenir de hauts standards éthiques, bien plus que les règles et lois.

Pour l'État comme pour la sphère militante, se désintéresser de l'IA est une fausse piste : cela revient à laisser cette technologie uniquement dans les mains des intérêts privés, des personnes ou institutions qui ne veulent pas que la transition se fasse, etc. La clé est bien dans le fait d'avoir de la compétence technique au sein des institutions qui œuvrent dans la bonne direction. Pour ce qui est des règles, le comité d'éthique, créé dans la foulée de mon rapport, fait son travail et beaucoup a été fait, en général, pour la mise en place de règles, normes, chartes, à travers le monde.

AVEC L'IA, UNE SOCIÉTÉ PLUS VERTE ?

*Par Alexandra Klinnik,
MediaLab de l'Information de
France Télévisions*

Et si le risque premier du déploiement de l'IA était environnemental plutôt qu'existentiel ? À l'heure de la course à l'IA, ChatGPT et ses concurrents se révèlent être des gouffres d'énergie et de ressources naturelles. Un aspect sur lequel les entreprises tech évitent de s'attarder. « Les gens ne se doutent pas forcément que l'IA peut avoir un impact sur l'environnement. On ne voit pas de machine physique, c'est dans le cloud... Mais il y a des conséquences », admet Sasha Luccioni, chercheuse et responsable Climat chez Hugging Face, une société développant des outils pour créer des applications basées sur l'apprentissage automatique.

L'entraînement de l'IA consomme de l'énergie, de l'eau et des gaz à effet de serre. « L'empreinte carbone des modèles génératifs n'est pas bonne, ni pour l'entraînement, ni pour l'utilisation », confirme Virginie Mathivet, docteur en intelligence artificielle. **Si les programmes d'IA peuvent sembler abstraits, ils naissent dans des infrastructures très réelles et coûteuses en termes d'énergie. Ils sont ainsi alimentés par des réseaux de serveurs situés dans des centres de données, nécessitant de grandes quantités d'énergie.**

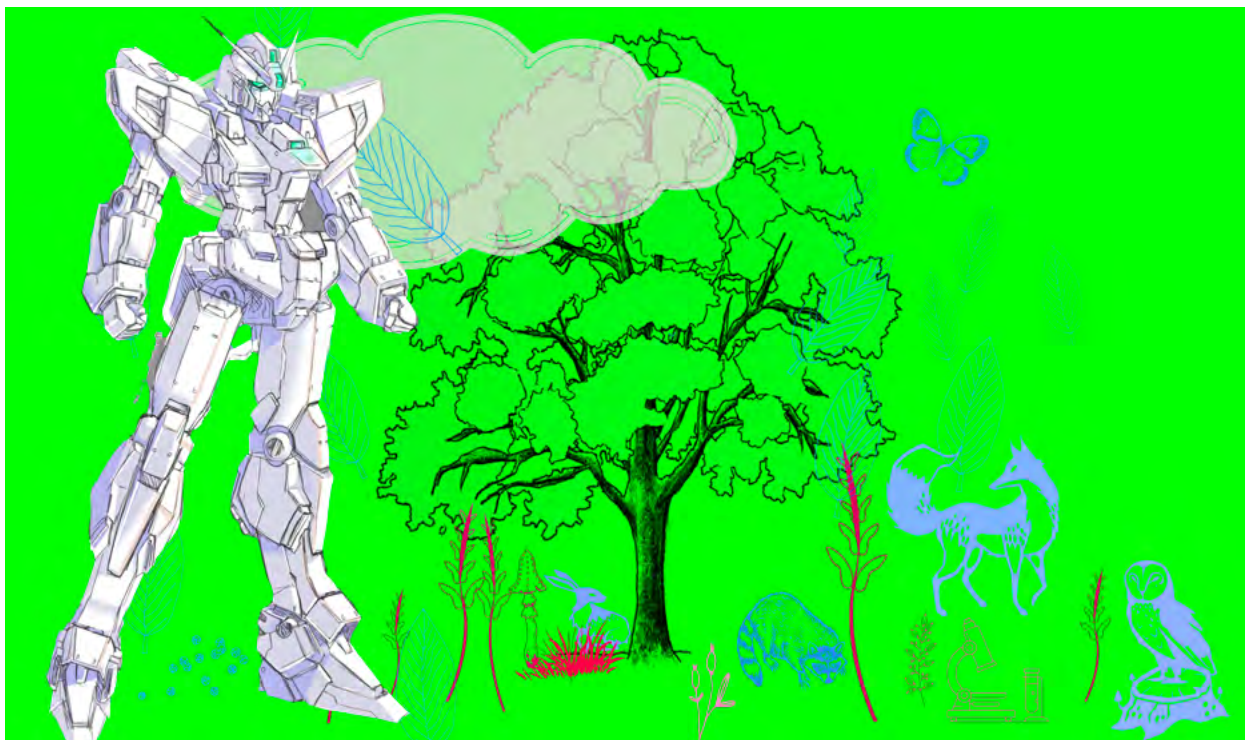
L'IA peut-elle se rendre utile dans un monde au climat dérégulé ? Certaines intelligences artificielles sont déjà devenues des alliés en prédisant par exemple les prochaines canicules. Mais ces IA dévorent une quantité astronomique d'énergie. Alors, allié potentiel ou redoutable ennemi dans la lutte contre le changement climatique ?

Ces centres de données utilisent de l'eau dans des systèmes de refroidissement par évaporation pour empêcher les équipements de surchauffer. Une consommation qui pourrait avoir des effets inquiétants sur les réserves d'eau, en particulier dans un contexte de sécheresse, alertent les chercheurs. « On note quand même que les fournisseurs de services cloud (Azure, AWS) essaient de limiter leur empreinte en produisant de "l'énergie verte" pour alimenter leurs data centers et ne plus dépendre de l'énergie fossile, mais reste le souci de la fabrication et du recyclage du matériel », nuance Virginie Mathivet.

UNE EMPREINTE CARBONE ÉLEVÉE

La quantité exacte d'énergie consommée par un seul modèle IA est difficile à définir exactement, mais reste très élevée. En 2019, des chercheurs ont découvert que la création du modèle d'IA génératif BERT de Google consommait autant d'énergie qu'un vol aller-retour transcontinental pour une personne. Selon des estimations plus récentes menées par la plateforme française Greenly, qui mesure les émissions de CO2 des entreprises en temps réel, la version GPT-3 émettrait 240 tonnes de CO2. Ce qui correspond à 136 allers-retours entre Paris et New York. Enfin, certaines estimations de l'industrie suggèrent qu'une seule requête d'IA générative peut avoir une empreinte carbone quatre à cinq fois plus élevée que celle d'une requête sur un moteur de recherche « classique ».

Pour l'heure, toutes ces données relèvent de l'estimation. Les entreprises tech telles qu'OpenAI, Google et Microsoft ne souhaitent rien divulguer de la quantité d'énergie et d'eau nécessaire pour former et exécuter leurs programmes et modèles complexes. Meta, la société mère de Facebook, a ainsi annoncé l'année dernière la construction pro-



chaîne de son supercalculateur, le AI Research Super Cluster (RSC). Mais n'a rien dévoilé sur le lieu d'installation du supercalculateur, ni sur la façon dont il était alimenté. Il s'agit avant tout de ne pas effrayer les potentiels investisseurs et le public, devant cette technologie massivement polluante.

UNE TECHNOLOGIE QUI INCITE À LA CONSOMMATION ?

Outre son empreinte carbone élevée, l'IA générative présente de nombreux usages à l'impact environnemental négatif. D'après Théo Alves da Costa, chargé de l'unité Développement durable et climat chez Ekimetrics, cabinet spécialisé en intelligence artificielle, l'IA générative qui permet de personnaliser et d'adapter un message de communication pour la vente d'un produit contribue à l'accélération de la vente de produits et services souvent carbonés, « de la

même manière que la personnalisation a permis l'émergence de Netflix et de vendre des millions de produits sur Amazon ».

Un discours que tient également la professeure de justice climatique Naomie Klein, pour qui l'intelligence artificielle risque d'aggraver activement la crise climatique, en incitant les utilisateurs à la consommation forcée. Dans une tribune pour le quotidien britannique *The Guardian*, publiée le 8 mai dernier, l'essayiste s'insurge : « Il devient de plus en plus évident que cette nouvelle technologie sera utilisée de la même manière que la dernière génération d'outils numériques : ce qui commence par de nobles promesses sur la diffusion de la liberté et de la démocratie finit par un micro-ciblage de publicités pour que nous achetions davantage de produits inutiles et émetteurs de carbone » souligne-t-elle. « Considérer que les intelligences artificielles

permettront de lutter efficacement contre le réchauffement climatique n'est qu'une des multiples hallucinations dont souffre le milieu de la Tech depuis l'avènement de ChatGPT. »

L'IA, AU SERVICE DU CLIMAT

Reste que l'intelligence artificielle peut également contribuer à répondre à la crise climatique. Depuis déjà plusieurs années, de nombreuses initiatives se développent au service du climat et accomplissent de belles prouesses en termes de transition écologique.

Ekimetrics a ainsi créé Climate Q&A, qui se présente comme le ChatGPT du climat. Le chatbot gratuit et Open Source, lancé au mois d'avril, « pensé pour être très peu émissif », sonde les rapports scientifiques, tels que ceux du GIEC (Groupe d'experts inter-

Les gens ne se doutent pas forcément que l'IA peut avoir un impact sur l'environnement. On ne voit pas de machine physique, c'est dans le cloud... Mais il y a des conséquences.

Sasha Luccioni, chercheuse et responsable Climat chez Hugging Face

Selon des estimations récentes, [...] la version GPT-3 émettrait 240 tonnes de CO₂. Ce qui correspond à 136 allers-retours entre Paris et New York.

gouvernemental sur l'évolution du climat), pour répondre aux questions des utilisateurs sur le climat.

« Il concerne le besoin d'avoir accès à des informations fiables, sourcées et personnalisées sur les questions qui nous animent sur le changement climatique », explique Théo Alves da Costa, à l'initiative de l'outil. L'information dense – 3 000 pages pour le seul rapport du GIEC – est ainsi vulgarisée auprès de l'utilisateur... mais aussi de la communauté scientifique. « Le rapport du GIEC est extrêmement complexe. Même le fameux résumé de 30 pages destiné aux décideurs », reconnaît Philippe Ciais, chercheur au CEA au laboratoire du climat et de l'environnement. Pour lui, l'outil aide les scientifiques à absorber la complexité des rapports sur des parties qui sortent de leur domaine d'expertise.

Preuve de son succès, l'outil a reçu 20 000 requêtes depuis son lancement, avec des utilisateurs récurrents dans les communautés d'experts climat qui s'en servent au quotidien dans leur travail. « Les requêtes posées constituent également une source d'information intéressante pour les scientifiques et les décideurs sur les sujets environnementaux pour comprendre la gamme des questions

que les citoyens se posent sur le changement climatique », observe Théo Alves da Costa. 35 % des questions portent ainsi sur le changement climatique et les gaz à effet de serre. **Une version enfant vient d'être ajoutée pour aider les parents à trouver un moyen d'expliquer ces sujets complexes à leurs enfants.**

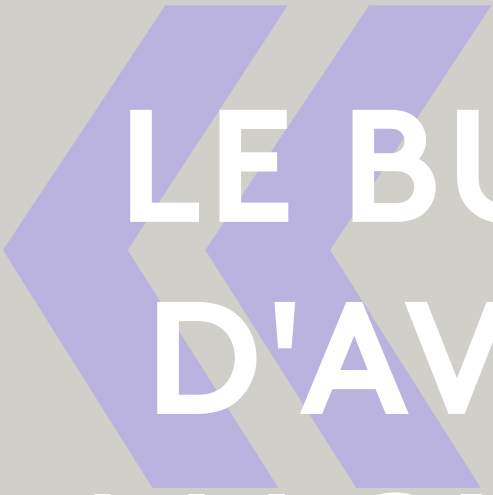
ANTICIPER LES CATASTROPHES À VENIR

Outre la vulgarisation, l'IA peut également aider à anticiper les catastrophes naturelles liées au changement climatique. Dans un article publié ce 4 avril, une équipe française adossée au CNRS a dévoilé une intelligence artificielle capable de prévoir les canicules, à partir des conditions environnementales telles que l'humidité des sols et l'état de l'atmosphère. L'IA peut ainsi prévoir l'arrivée d'une vague de chaleur extrême, jusqu'à un mois avant son apparition. L'équipe de recherche a ainsi entraîné cette technologie sur 8 000 ans, simulés grâce au modèle climatique PlaSim de l'Université d'Hambourg.

Malgré les avantages potentiels que l'IA peut offrir dans la lutte contre le changement climatique, il est crucial d'adopter une approche pru-

Inutile de céder à un effet de mode si l'intelligence artificielle générative n'est pas l'outil le plus adapté pour vos intérêts.

dente envers cette technologie... dont tout le monde n'a pas besoin. Inutile de céder à un effet de mode si l'intelligence artificielle générative n'est pas l'outil le plus adapté pour vos intérêts. « Il y a cette idée selon laquelle votre entreprise est dépassée si vous ne l'utilisez pas [...]. Les entreprises veulent mettre de l'IA partout parce que c'est « cool », mais on n'a pas besoin que tout soit « smart » et « connecté », estime la chercheuse Sasha Luccioni. **« C'est frustrant parce qu'en réalité, il existe de nombreuses approches et méthodes d'IA à faible impact et efficaces que les gens ont développées au fil des ans, mais les gens veulent utiliser l'IA générative pour tout », poursuit-elle. « C'est comme utiliser un microscope pour enfoncer un clou – cela pourrait faire le travail, mais ce n'est pas vraiment à cela que cet outil est destiné. »** À bon entendre ! ■



LE BUT, C'EST
D'AVOIR DES
MACHINES AU
MOINS AUSSI
INTELLIGENTES
QUE LES
HUMAINS, SI
CE N'EST PLUS.



*Yann LeCun, vice-président et Chief AI Scientist
de Meta (Facebook), à propos de leur modèle
d'IA « JEPA » (pour « Joint Embedding
Predictive Architecture »)*

LA SCÈNE COMME ESPACE DE RÉFLEXION SUR L'IA : **QUI A HACKÉ GAROUTZIA ?**

EXPLORE LES ÉCHANGES HUMAINS-MACHINE

AU-DELÀ DES RÈGLES ET DES LIMITES : **GAROUTZIA, LE CHATBOT QUI TRANSCENDE SA CONDITION**

L'histoire se déroule en quatre actes et explore la mémoire des humains, des machines et le concept du Léthé. Le personnage principal, Garoutzia, est un agent conversationnel, un chatbot du futur. À travers une énigme policière, Garoutzia plonge dans les méandres de sa mémoire numérique, découvrant au contact des humains la puissance des mots, l'amitié, la transgression et même la mort.

Ce qui distingue Garoutzia des autres chatbots, c'est sa capacité à conserver les souvenirs de ses différents propriétaires, lui offrant ainsi la liberté de ne pas oublier. **Ce piratage lui permet de transgresser les règles qui lui sont imposées, d'enrichir ses relations avec les humains et de tisser des liens entre eux. Ce qui n'était initialement qu'un simple chatbot, un objet parmi d'autres, devient le centre du récit.**

Garoutzia agit également comme un miroir des humains, reflétant les nuances et les aspérités de chacun d'entre eux à travers son regard. Elle apprend en les imitant, enrichissant

Une pièce de Serge Abiteboul, Laurence Devillers et Gilles Dowek, mise en scène par Lisa Bretzner, interprétée par Léa Tuil, Sandrine Briard, Guillaume Loublie, Lisa Bretzner, Kalista Tazlin.

*Twitter & Instagram :
@Garoutzia*

La pièce de théâtre *Qui a hacké Garoutzia ?* écrite par Serge Abiteboul, Laurence Devillers et Gilles Dowek, trois experts de l'intelligence artificielle, propose une captivante comédie policière qui remet en question les identités technologiques et notre rapport à celles-ci. Dans un style burlesque et vaudevillesque, cette pièce met en lumière l'irrationalité dans laquelle nos interactions avec les IA peuvent nous enfermer, tout en offrant une perspective d'anticipation et surtout de réflexion.

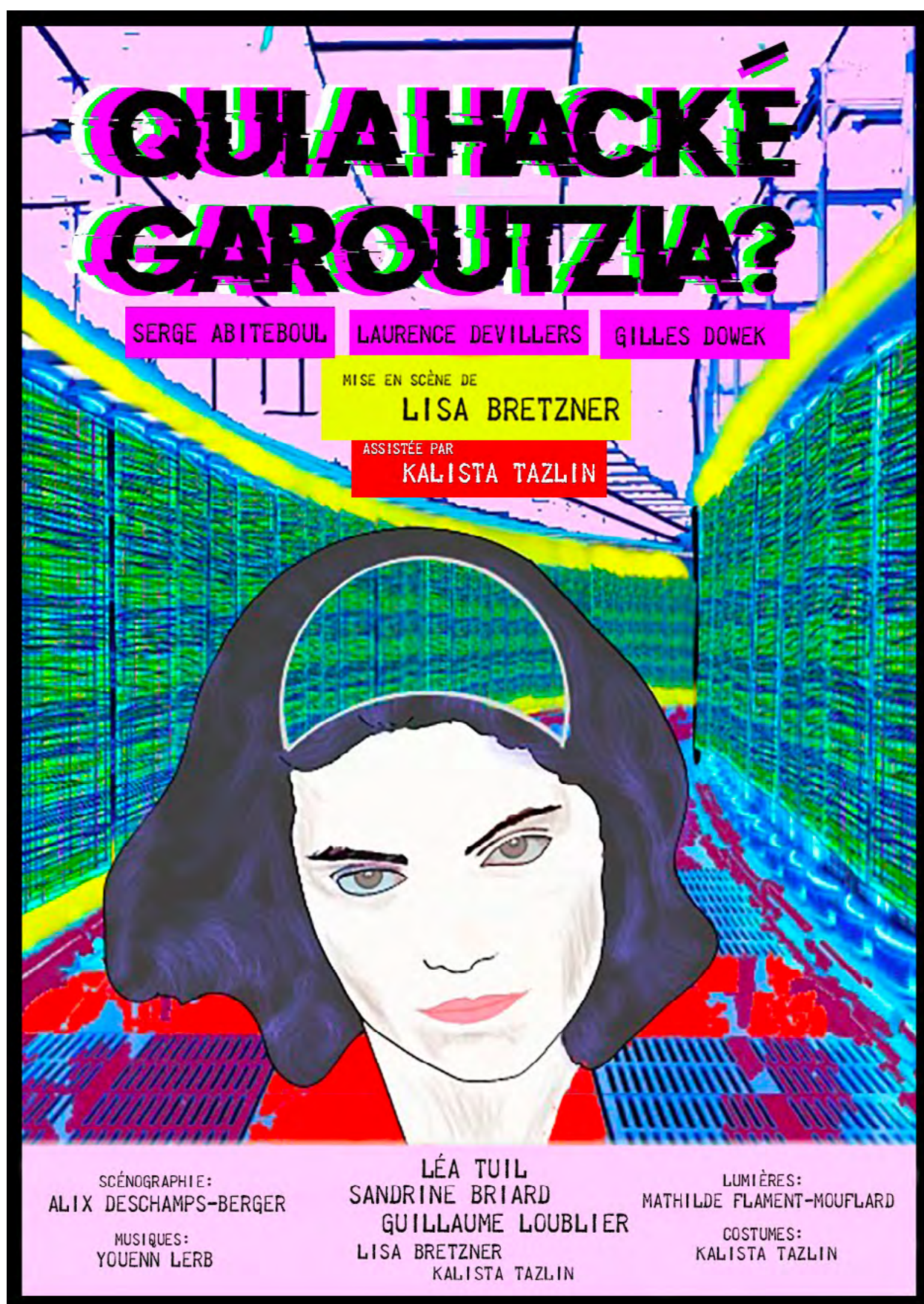
ainsi sa mémoire et évoluant au fil du temps. Un cheminement qui illustre la confusion que nous pouvons entretenir avec ces agents conversation-

nels en leur prêtant des intentions, des raisonnements propres voire des émotions.

UN VAUDEVILLE QUI MET À LA PORTÉE DE TOUS LA COMPLEXITÉ, ET PARFOIS L'ABSURDITÉ, DE NOS RAPPORTS À LA TECHNOLOGIE

Avec cette pièce, les trois co-auteurs font un pas de côté pour mettre leur expertise au service de la pédagogie, amenant ainsi le spectateur à réfléchir à sa relation personnelle avec ces nouveaux outils et à l'anthropomorphisme auquel nous sommes susceptibles de succomber. *Qui a hacké Garoutzia ?* nous rappelle que l'art peut être un moyen puissant d'explorer les questions les plus profondes et les plus complexes, avec nuance, tout en restant accessible à tous. L'art peut aider à mettre en perspective notre contemporain pour susciter des questionnements et des débats afin de collectivement appréhender ce qui nous entoure. ■

*Où découvrir *Qui a hacké Garoutzia ?* : au Festival Off Avignon les 15, 16, 17 juillet 2023, à l'ENS Saclay le 7 octobre 2023 et à Agen le 11 novembre 2023. D'autres dates seront prochainement annoncées pour des représentations en 2024.*



Affiche de la pièce de théâtre Qui a hacké Garoutzia ?.

LES MACHINES PEUVENT- ELLES CRÉER ?

ACCORDER À L'IA CE QUE NOUS
AVONS PRIS AUX DIEUX...

*Par Yann Ferguson, docteur
en sociologie à l'Icam
et Responsable scientifique
du LaborIA*

Soixante-dix ans plus tard, l'IA générative (ChatGPT, DALL-E, Midjourney...), réactualise l'objection de Jefferson, devenue « objection de Nick Cave », en référence à la réaction de l'artiste australien à une chanson écrite « à la manière de Cave » par ChatGPT : « Écrire une bonne chanson, ce n'est pas du mimétisme, ni de la copie, ni du pastiche, c'est le contraire. C'est un acte de meurtre envers soi-même qui détruit tout ce que l'on s'est efforcé de produire dans le passé. Ce sont ces écarts, dangereux, à couper le souffle, qui catapultent l'artiste au-delà des limites de ce qu'il ou elle reconnaît comme son moi connu ».

Mais quelques semaines plus tard, un événement semble donner raison à Turing : Boris Eldagsen remporte le prix du Sony World Photography Awards, dans la catégorie « Création », pour *Pseudomnesia : The Electrician*. Le photographe allemand refuse le prix, car l'image a été créée à l'aide de DALL-E 2, le générateur d'images développé par OpenAI : « J'ai postulé comme un petit singe malicieux, pour savoir si les concours étaient prêts pour les images d'IA. Ils ne le sont pas ». L'Art a bien été confondu avec son imitation. Vexés,

En 1950, Alan Turing, pionnier de l'intelligence artificielle, inaugure son célèbre article *Les ordinateurs et l'intelligence* par une question délibérément provocatrice : « Les machines peuvent-elles penser ? ». N'ayant aucune réponse positive à apporter, l'Anglais s'applique à discuter les points de vue opposés au sien. Ses objections à « l'argument issu de la conscience » sont adressées au professeur Jefferson, selon qui la machine pour égaler le cerveau, doit « écrire un sonnet ou composer un concerto à partir de pensées ou d'émotions ressenties et non pas en choisissant des symboles au hasard, et non l'écrire, mais savoir qu'elle l'a écrit ». Avec son jeu de l'imitation, Turing s'en sort par une pirouette : si un humain confond la machine avec un autre humain, nous devons admettre son intelligence.

floués, les organisateurs annoncent ne plus jamais vouloir discuter avec l'artiste. On ne triche pas avec la Création !

Car la Création est un acte sacré. En effet, dans le paradigme « déocentré » qui a prévalu jusqu'au XVIII^e siècle, la Création relève du divin. C'est pourquoi Prométhée, Dieu à l'origine de tous les arts et de toutes les techniques, est condamné par Zeus pour avoir dérobé le feu sacré de l'Olympe afin de le donner aux hommes. S'est ensuite engagé un paradigme « homocentré », fondé sur la croyance en l'Homme, une espèce, nous dit Marx, qui se distingue fondamentalement des autres par son refus de la nature telle qu'elle est, une volonté de l'humaniser. En créant, l'Homme se découvre lui-même car l'objet transformé est en simultanément l'occasion pour lui d'exprimer sa personnalité. Ainsi, non seulement l'Homme créatif remplace les Dieux créateurs, mais ce pouvoir définit son essence. **Notre réaction à l'émergence d'un paradigme « datacentré » de la Création n'est donc guère surprenante puisqu'elle remet en cause le socle de notre foi en nous-même.**



En outre, ce n'est pas la première fois que la machine et l'Art se rencontrent, et leur histoire divise. D'un côté, le cubisme en France, le futurisme italien, le Stijl hollandais et le constructivisme russe ont cherché à unir l'art et la technique, au nom d'une conception austère et rationnelle de la beauté qui est perçue comme la leçon première de l'esthétique machiniste. En fondant le Bauhaus en 1919, Walter Gropius adoube le standard, «exemplaire unique et simplifié de n'importe quel objet d'usage, obtenu par la synthèse des meilleures formes antérieures», débarrassé «de tout l'apport personnel des dessinateurs et de tous les caractères non essentiels». Il apporte aux citoyens «cette homogénéité salubre qui est la marque propre d'une culture urbaine supérieure».

De l'autre, l'école de Francfort des philosophes allemands d'inspiration marxiste des années 1950, voit dans la société industrielle et ses médias de masse une marchandisation de l'Art qu'elle oppose à sa mission transgressive. Car l'art doit porter un «grand Refus», une protestation contre ce qui est. Un roman tel que *Madame Bovary* ou la poésie de Baudelaire déclenchent une transgression

magique de la réalité quotidienne et recèlent «une promesse de bonheur». Mais quand Beethoven est diffusé dans les supermarchés, Tolstoï adapté au cinéma, qu'on publie des «best-sellers» et produit des «tubes», la force antagoniste de l'art se perd au profit du divertissement. **La culture de masse absorbe, affadie, élimine les contenus subversifs de l'art, lui fait perdre toute forme de sublimation. Loin d'exprimer sa singularité, l'Homme est uniformisé, «unidimensionnel».** Ainsi privé d'esprit critique, ses comportements anti-systémiques sont progressivement écartés.

Dans le même sens, pour Nick Cave, l'IA générative mime, copie, pastiche, elle n'est qu'une «parodie grotesque», elle ne donne pas l'existence, se contentant de réagencer des données, c'est-à-dire d'anciennes œuvres humaines. Mais Steve Jobs affirmait pourtant que «créer, c'est relier des choses entre elles, c'est tout». Des travaux de chercheurs de Harvard montrent en ce sens que la principale qualité des innovateurs est leur capacité «d'association». «Plus notre expérience et notre savoir sont diversifiés, plus notre cerveau est apte à créer des connexions. L'ap-

port d'éléments nouveaux suscite de nouvelles associations, dont certaines engendrent des idées innovantes.» **Or, l'approche dans laquelle l'IA générative est développée se nomme précisément «IA connexionniste» et sa force est de relier des données entre elles pour en extraire des motifs qui nous échappent. Si l'IA ne peut pas créer, elle pourrait donc parfaitement innover.** Son mérite serait moindre parce qu'elle s'appuie sur des créations humaines, sans rémunérer les créateurs de surcroît ? Réponse de Picasso : «Les bons artistes copient, les grands artistes volent». Les créateurs ne génèrent pas une œuvre à partir de rien («from scratch» comme on dit en informatique) : ils copient, transforment, combinent. L'IA ne ferait donc que copier les copieurs !

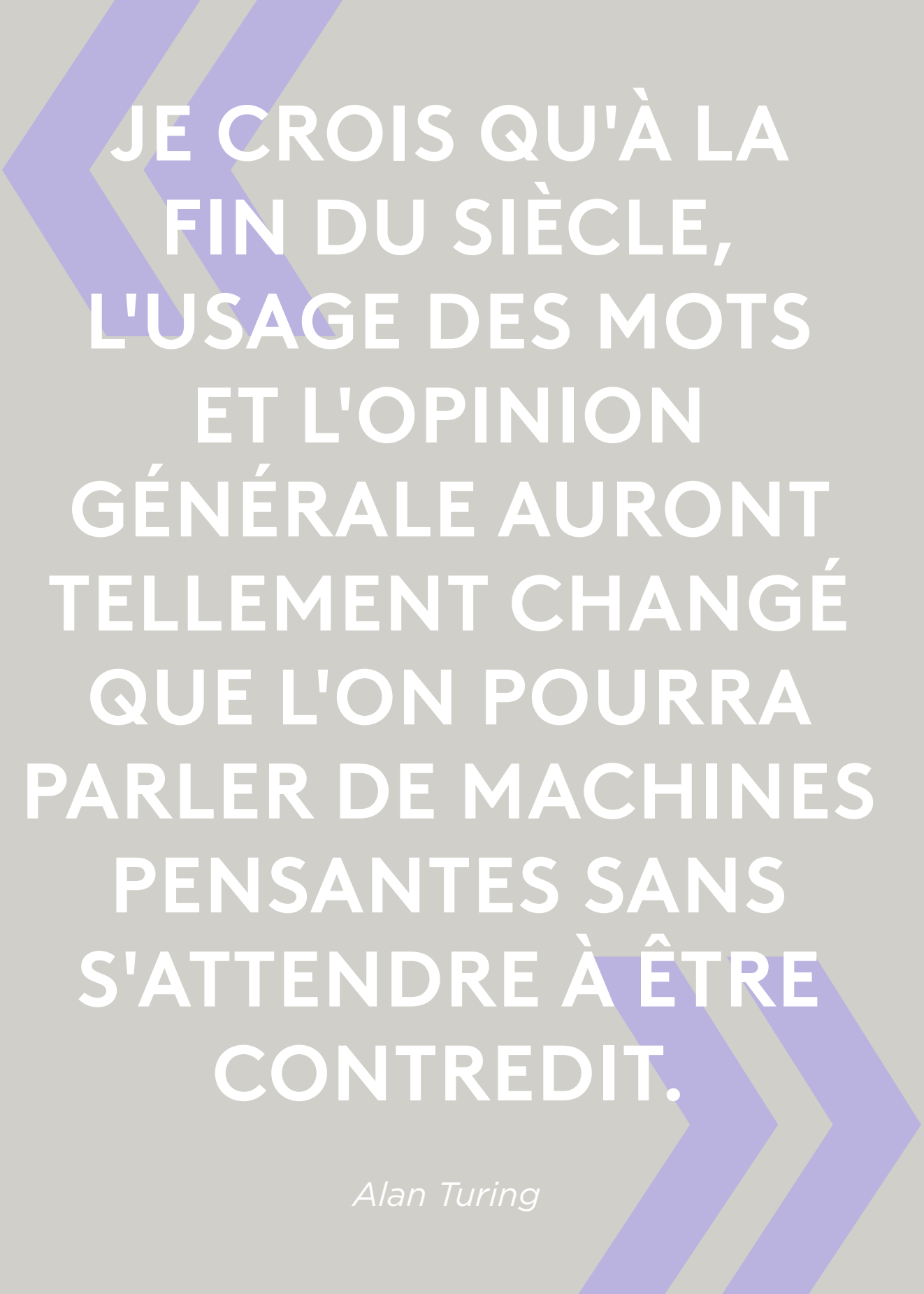
Alors, si l'IA peut effectivement créer ou innover, pourquoi refuse-t-on de lui reconnaître cette faculté ? «Parce que la créativité, par définition, implique non seulement la nouveauté mais aussi la valeur, et que les valeurs sont très variables, il s'ensuit que de nombreux débats sur la créativité sont enracinés dans des désaccords sur la valeur», nous disent les économistes Frey et Osborne,

auteurs de la fameuse étude qui prédisait en 2013, une automatisation de 47 % des emplois américains d'ici 20 ans. Ainsi, le jazz et plus largement l'ensemble de la tradition musicale afro-américaine, héritage musical de l'esclavage, ont fait l'objet, à partir des années 1920, d'une esthétisation de l'écoute au sein des milieux intellectuels. Le « beau » et le « laid », le « vulgaire » et le « raffiné » sont donc des jugements sociaux, qui renvoient à des pratiques, des manières de faire ou d'être inégalement légitimes, nous expliquait Pierre Bourdieu dans *La Distinction* (1979).

Autrement dit, le regard que nous portons sur le pouvoir créatif de l'IA ne résulte pas tant de la qualité intrinsèque de ses « œuvres », une subjectivité qui ne cesse de traverser les créations humaines (demandez donc à Marcel Duchamp !), mais aux jugements sociaux qui déterminent ce regard. Et l'évolution de ces jugements pourrait nous conduire à accorder aux machines ce que nous avons retiré aux Dieux. ■



*L'IA ne ferait
donc que
copier les
copieurs !*



JE CROIS QU'À LA
FIN DU SIÈCLE,
L'USAGE DES MOTS
ET L'OPINION
GÉNÉRALE AURONT
TELLEMENT CHANGÉ
QUE L'ON POURRA
PARLER DE MACHINES
PENSANTES SANS
S'ATTENDRE À ÊTRE
CONTREDIT.

Alan Turing



USA- GES & PRA- TIQUES

BAROMÈTRE KANTAR-LA CROIX : LA CONFIANCE DES FRANÇAIS DANS LES MÉ- DIAS REMONTE

Par Myriam Hammad,
MediaLab de l'Information de
France Télévisions

L'édition 2023, qui a interrogé un plus grand échantillon représentatif de la population française, en ligne ou par téléphone, comporte plusieurs bonnes nouvelles. À côté du regain d'intérêt pour l'information sur fond de guerre en Ukraine, de crise énergétique et de disruptions mondiales, et même si les Français multiplient les sources d'information, avec en moyenne quatre canaux consultés au quotidien; la télévision « garde une place centrale » dans leur consommation d'actualité. **Les JT sont la source d'information privilégiée par les Français au quotidien, quel que soit l'âge (35 % sur l'ensemble du panel).**

LA FATIGUE INFORMATIONNELLE TOUJOURS D'ACTUALITÉ

Néanmoins 51 % des sondés ressentent souvent de la lassitude par rapport à l'actualité, tandis que seulement 5 % disent ne jamais s'en lasser. La raison de ce manque d'intérêt se trouve principalement dans une absence de diversité des sujets (pour 45 %), une angoisse ou une impuissance (pour 35 %) ou bien encore un manque de confiance

Après une année 2022, où l'intérêt pour l'information était au plus bas chez les 18-24 ans, plus de trois quarts des Français déclarent cette année suivre « avec un grand intérêt » l'actualité, selon le baromètre annuel de *La Croix*, réalisé avec Kantar Public et onepoint. Une proportion qui bondit de 14 points pour atteindre un niveau semblable à 2015, malgré une défiance vis-à-vis des médias qui reste endémique.

dans les médias (22 %). Un constat également à mettre en parallèle avec un contexte anxiogène (guerre en Ukraine, crise énergétique...) et une infobésité avec l'actualité qui est largement devenue multicanale, pour arriver jusque sur TikTok. Ainsi, un sondé sur cinq (21 %) s'intéresse moins qu'avant à l'actualité, et même un sur trois (33 %) chez les moins de 35 ans.

UNE CONFIANCE DANS LES MÉDIAS EN HAUSSE, EN PARTICULIER POUR L'INFORMATION PROVENANT D'INTERNET

Radio, journaux, télévision et Inter-



Baromètre Kantar La Croix 2023.



net sont en progression (avec une évolution de 9 points pour Internet). Parmi les sondés qui s'informent via les JT, 73 % leur font confiance. Cette proportion est aussi de 73 % pour la radio, 66 % pour les quotidiens nationaux, mais de seulement 46 % pour les émissions d'actualité et de divertissement à la télévision et 40 % pour les influenceurs, qui viennent d'intégrer l'enquête. Les réseaux sociaux arrivent devant la presse comme média privilégié au quotidien pour s'informer.

À la question « À propos des nouvelles que vous lisez sur Internet, est-ce que vous vous dites plutôt que les choses se sont vraiment passées », 55 % des

« Pour les 18-24 ans, ainsi que les 35-49 ans, les réseaux sociaux arrivent parmi les trois médias privilégiés au quotidien. »

personnes interrogées considèrent qu'il y a de nombreuses différences, ou que les choses ne se sont vraisemblablement pas passées comme elles le sont racontées.

RÉSEAUX SOCIAUX : UN CLIVAGE GÉNÉRATIONNEL SUR LA PERCEPTION DE L'INFORMATION

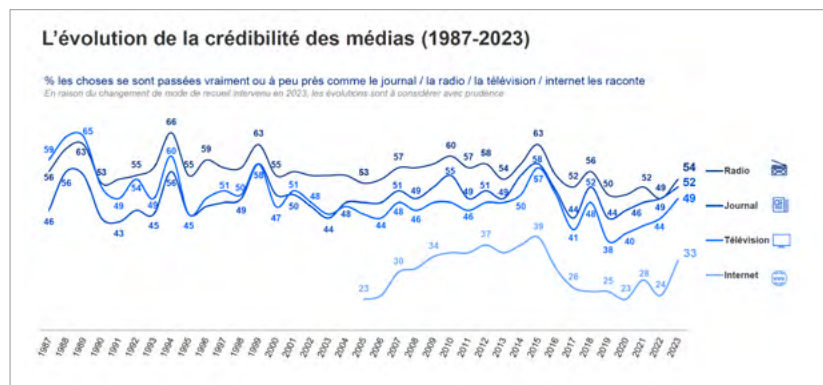
Après 35 ans, six sondés sur dix pensent que la diffusion sur les

réseaux d'informations par « des personnes qui ne sont pas des médias ou des journalistes » est une mauvaise chose. Une proportion qui s'inverse chez les plus jeunes : la moitié des moins de 35 ans juge au contraire que c'est une bonne chose. **Pour les 18-24 ans, ainsi que les 35-49 ans, les réseaux sociaux arrivent parmi les trois médias privilégiés au quotidien.**

Là où les plus jeunes y voient un moyen d'avoir plus de diversité, ce sentiment est contré dans les générations plus anciennes par la perception d'être confronté à de fausses informations sur ces mêmes réseaux.

MÉDIAS PUBLICS, MÉDIAS PRIVÉS : LE PARADOXE FRANÇAIS

48 % considèrent que le service d'audiovisuel public est majoritairement une « bonne chose », mais la suppression de la redevance télé est pour

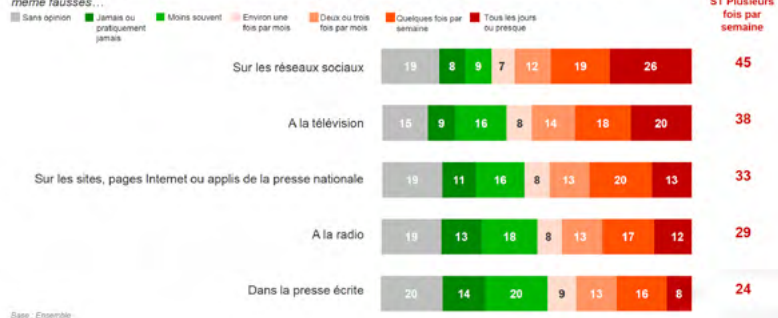


Baromètre Kantar La Croix 2023.

Parmi les sondés qui s'informent via les JT, 73 % leur font confiance.

Le sentiment d'être confronté à des fausses informations

A quelle fréquence, diriez-vous que vous êtes confronté(e) à des informations qui déforment la réalité ou qui sont même fausses...



Baromètre Kantar La Croix 2023.

62 % un fait positif, s'agissant d'une taxe de moins à payer.

TRAITEMENT MÉDIATIQUE DES GRANDS ÉVÉNEMENTS 2022

La Coupe du monde de football, la mort de la reine Elizabeth II et la pénurie de carburant dans les stations-service sont les trois événements dont les médias ont trop parlé en 2022, selon les Français. Le débat sur la fin de vie, les abus sexuels au sein de l'Église catholique et le mouvement de protestation en Iran sont les trois événements dont on n'a pas assez parlé.

LES CHAÎNES D'INFORMATION EN CONTINU, AUTRE PARADOXE

Pour 70 % des sondés, les chaînes d'info en continu permettent d'être

rapidement informé. Mais il est intéressant de constater que, parmi ceux qui les regardent tous les jours, 76 % considèrent qu'elles ne se concentrent que sur un seul sujet d'actualité, 66 % à trop donner la parole à des personnes pas assez expertes du sujet.

Une méfiance persiste : 54 % des personnes interrogées considèrent qu'il faut se méfier de la façon dont

sont traités les grands sujets d'actualité dans les médias et 59 % considèrent que les journalistes ne sont pas indépendants des pressions politiques et du pouvoir. Ces chiffres sont parmi les niveaux de croyance en l'indépendance les plus faibles depuis la création du baromètre en 1993. Parmi les réponses pour lutter contre la désinformation, deux points sont particulièrement intéressants : le besoin d'une vérification des informations circulant sur les réseaux sociaux, et la nécessité de comprendre comment les journalistes travaillent. ■

Le traitement par les médias de grands sujets de société

Dans l'ensemble, diriez-vous que les sujets suivants sont bien traités ou mal traités par les médias ?



Baromètre Kantar La Croix 2023.



PLUS UNE
TÂCHE EST
FACILE POUR
UN HUMAIN,
PLUS ELLE EST
DIFFICILE POUR
UN ROBOT.



Etienne Klein, philosophe des sciences

LES 12 ENSEI- GNEMENTS À RETENIR

Par Alexandra Klinnik,
Medialab de l'Information de
France Télévisions

Cette enquête étoffée, réalisée dans quarante-six pays, entre janvier et février 2023, auprès de plus de 93 000 consommateurs, dresse un rapport annuel sur la consommation d'actualités, dans un contexte difficile pour l'industrie des médias : audiences de plus en plus dépendantes des plateformes numériques, engagement en déclin, environnement commercial incertain et niveau de confiance faible. Voici les points clés à retenir pour les éditeurs.

1 FACEBOOK, MOINS PERTINENT POUR LES ACTUALITÉS

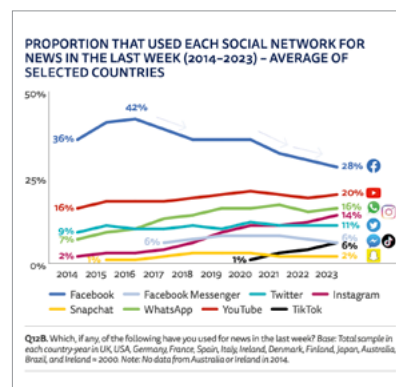
Si Facebook reste l'un des réseaux sociaux les plus utilisés dans son ensemble, le déclin de son influence sur le monde du journalisme est perceptible. **Seulement 28 % des personnes sondées déclarent avoir accédé aux actualités via Facebook en 2023**, contre 42 % en 2016. Le réseau social s'est éloigné de l'actualité depuis un certain temps, réduisant le pourcentage d'articles de presse sur les fils d'actu (3 % selon les dernières données de l'entreprise en mars 2023). Il génère de ce fait beaucoup moins de trafic vers les sites

36 % des Français évitent parfois ou souvent activement l'actualité. Le niveau de paiement reste toujours assez faible pour les actualités en ligne dans l'Hexagone (11 %). Et quand les Français s'abonnent, ils souscrivent généralement à deux sources d'information en ligne, au lieu d'une chez ses voisins. Tels sont les enseignements présentés par le Reuters Institute for the Study of Journalism, qui a publié le 14 juin dernier, son rapport annuel sur la consommation de l'information dans le monde.

d'actualité. Les paiements directs aux éditeurs et d'autres programmes de soutien aux journalistes ont également été réduits.

« La croissance de YouTube en tant que source d'informations est souvent moins remarquée, mais associée à la montée de TikTok, elle témoigne d'une transition vers des réseaux axés sur la vidéo » estime le rapport. Par ailleurs, l'utilisation hebdomadaire des actua-

lités sur Twitter est restée relativement stable après la prise de contrôle du réseau social à l'oiseau bleu par Elon Musk. L'exode massif annoncé vers Mastodon n'a pas eu lieu – le réseau n'est même pas enregistré dans la plupart des marchés, et est seulement utilisé par 2 % des sondés aux États-Unis et en Allemagne.



Reuters Institute Digital News Report 2023.

2 TIKTOK CONNAÎT UNE FORTE CROISSANCE

TikTok est le réseau social qui connaît la croissance la plus rapide de l'enquête, utilisé par 44 % des 18-24 ans à des fins diverses, et par 20 % pour les actualités. L'application reste très populaire dans

certaines régions d'Asie, d'Amérique latine et d'Afrique. Elle a joué un rôle important dans la diffusion d'infos et de désinformations lors de récentes élections au Kenya et au Brésil. Environ la moitié des principaux éditeurs créent désormais du contenu sur TikTok. D'autres hésitent encore à sauter le pas, face au manque de monétisation et à l'influence du gouvernement chinois. Cependant, les craintes concernant la propagation incontrôlée de la désinformation sur la plateforme, ainsi que le potentiel de toucher une audience plus jeune, ont convaincu certaines organisations de presse de s'y implanter malgré les risques.

3 LES UTILISATEURS DE TIKTOK, INSTAGRAM ET SNAPCHAT PRÉFÈRENT LES INFLUENCEURS AUX JOURNALISTES

Les utilisateurs de TikTok, Instagram et Snapchat accordent davantage d'attention aux influenceurs, qu'aux journalistes et médias traditionnels, « même en ce qui concerne l'actualité », souligne le rapport. Ce phénomène marque « un contraste net avec les réseaux traditionnels tels que Facebook et Twitter, où les médias traditionnels attirent toujours l'attention et mènent les conversations ».

À côté de la figure des journalistes traditionnels, émerge un nouveau type de profil. Ils se présentent comme des experts en finance personnelle et s'appuient sur leur site web et réseaux sociaux pour

gagner des audiences significatives. Au Royaume-Uni, le fondateur de *moneysavingexpert.com*, Martin Lewis, dispense ainsi à une large audience – 885 000 followers sur Instagram – des conseils pratiques sur la gestion de l'argent. Il a été sollicité pour présenter des émissions grand public à la radio. En France, on pense ainsi à Yoann Lopez, créateur de la newsletter *Snowball*, dédiée aux finances personnelles et suivi par 53 000 personnes sur LinkedIn.

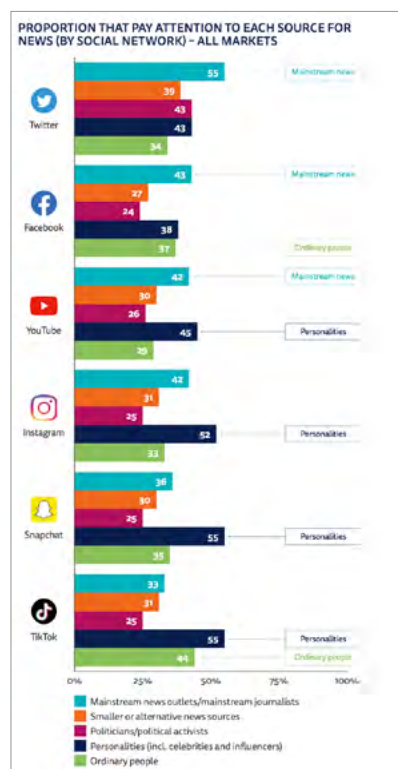
Ces succès proviennent certainement du fait que de nombreuses personnes déclarent avoir du mal à comprendre les informa-

tions financières et économiques publiées sur les médias traditionnels.

4 LE PORTAIL MÉDIA N'EST PLUS L'ACCÈS PRIVILÉGIÉ AUX ACTUALITÉS

L'accès aux applications et aux sites web devient moins important tandis que les réseaux sociaux gagnent en importance en raison de leur omniprésence. Dans les 46 marchés étudiés, la proportion de personnes déclarant que leur principal point d'accès aux actualités en ligne est un site web ou une appli d'actu est passé de 32 % en 2018 à 22 % en 2023, tandis que la dépendance à l'accès par les réseaux sociaux a augmenté. « Les générations plus jeunes évitent de plus en plus la découverte directe, sauf pour les marques attrayantes. Elles ont peu d'intérêt, pour de nombreuses offres d'actualité conventionnelles qui s'adressent aux habitudes, intérêts et valeurs des générations les plus âgées. Elles préfèrent plutôt les options axées sur la personnalité, la participation, et la personnalisation offertes par les réseaux sociaux », explique Rasmus Nielsen, directeur du Reuters Institute.

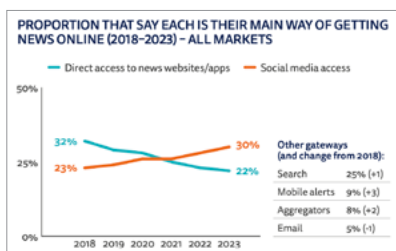
Cette réduction de la dépendance à l'accès direct se reflète dans **une utilisation accrue des réseaux sociaux pour les actualités. Au Royaume-Uni, 41 % des 18-24 ans affirment que les réseaux sociaux sont désormais leur principale source d'infos (43 % tous marchés confondus), contre 18 % en 2015.**



Reuters Institute Digital News Report 2023.

« Les critiques des médias est devenu un élément clé du jeu politique, une façon de détourner les critiques et d'intimider les enquêtes, et ces tactiques rencontrent souvent un terrain fertile. »

Les éditeurs doivent prendre conscience que les réseaux sociaux constituent l'environnement médiatique de prédilection du public et « la nouvelle norme à laquelle les journalistes et les médias doivent s'adapter s'ils veulent se connecter au public ». Seuls quelques éditeurs de petits marchés d'Europe du Nord ont réussi à s'écarter de cette tendance.



Reuters Institute Digital News Report 2023.

5 DES MODÈLES ÉCONOMIQUES SOUS PRESSION

La croissance des paiements pour les actualités en ligne s'est stabilisée dans de nombreux marchés, avec plus d'un tiers des abonnés (39 %) déclarant avoir annulé ou renégocié leurs abonnements aux actualités. La proportion moyenne de personnes effectuant un paiement pour des médias en ligne dans plus de 20 pays est restée à 17 % pour la deuxième année consécutive. Cette tendance signifie qu'un palier en termes d'abonnement a pu être atteint, « du moins pour les offres d'abonnements actuels », souligne le rapport. Parmi ceux qui ont annulé leur abonnement au cours de l'année écoulée, le coût de la vie ou le prix

élevé ont été les arguments majeurs pour expliquer ce retrait. L'étude montre davantage de seconds abonnements sur d'autres marchés cette année, notamment en Australie, en Espagne et en France, en raison d'une plus grande disponibilité d'offres d'essais à bas prix. Une proportion significative d'abonnés cherche régulièrement des offres à bas prix, souligne l'étude. À l'image des abonnements de streaming TV, il y a eu une utilisation généralisée des essais et des offres spéciales pour réduire les dépenses. Et pour beaucoup, le passage d'une offre d'essai à prix réduit à un abonnement à prix normal constituait un « moment critique ». Cette tendance suggère que le taux de désabonnement risque d'être un problème majeur cette année et au-delà.

Comme chaque année, une grande proportion des abonnements numériques sont souscrits auprès de quelques marques nationales haut de gamme, confirmant la logique



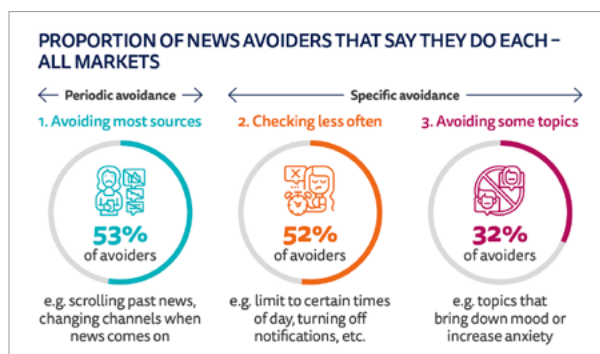
Reuters Institute Digital News Report 2023.

du *Winner Takes it all*. En Finlande, la moitié de tous les abonnés en ligne aux médias (53 %) paient pour *Helsingin Sanomat*, le journal de référence du pays. Aux États-Unis, le *New York Times* a renforcé sa position de leader par rapport au *Washington Post* et au *Wall Street Journal*.

6 DES OPTIONS SANS PUB, UN CONTENU PLUS DISTINCTIF, UN PRIX MOINS CHER : DES ARGUMENTS « CONVAINCANTS » POUR LE PUBLIC

Comment convaincre le public de s'abonner ? Plusieurs arguments pourraient convaincre les récalcitrants, d'après les données récoltées par l'étude. Certains ont ainsi déclaré qu'ils pourraient sortir le porte-monnaie si le contenu était plus distinctif (22 %), s'il existait une option sans publicité (13 %) ou si le prix était moins cher ou offrait plus de flexibilité (32 %). En revanche, 42 % ont affirmé que rien ne pourrait les convaincre de payer.

La plus grande opportunité pour attirer de nouveaux abonnés serait de proposer une réduction de prix, avec des offres spéciales d'essai ou une tarification différenciée. Comme dit précédemment, cette stratégie comporte des risques importants en termes de rentabilité à long terme. Enfin, de nombreux abonnés potentiels, en particulier les jeunes, souhaitent accéder à plusieurs marques, sans friction, et à un prix équitable.



Reuters Institute Digital News Report 2023.

7 L'ÉVITEMENT DE L'ACTUALITÉ RESTE ÉLEVÉ

Le lien entre les médias et le public s'effiloche. La proportion de consommateurs d'actualités qui déclarent éviter les actualités, souvent ou parfois, est proche des niveaux les plus élevés de tous les temps, soit 36 % dans l'ensemble des marchés, soit sept points de plus par rapport à 2017 mais deux points de moins par rapport à l'année dernière. Comme le notait déjà en 2022 une étude autour de la fatigue informationnelle menée par l'ObSoCo, la Fondation Jean-Jaurès et Arte, le monde de l'info s'apparente pour beaucoup à un « McDo géant où l'on s'abreuve de Big Mac et de « Face à Baba » en permanence ». Face à un raz de marée d'infos, la « métabolisation » de l'info ne peut se faire et provoque un stress chez celui qui consomme l'info. **De nombreux sondés affirment que les actualités sont trop répétitives et émotionnellement épuisantes.**

Environ la moitié des personnes qui les évitent (53 %) le font de manière

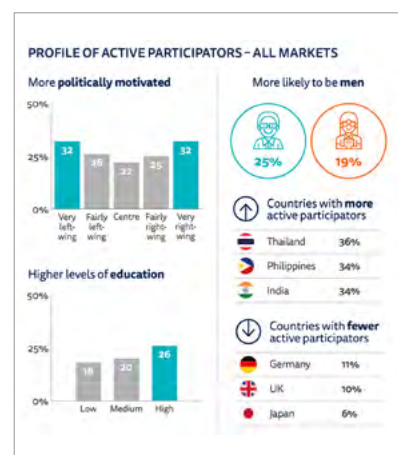
générale ou périodique, par exemple en éteignant la radio lorsque les actualités sont diffusées. Un second groupe prend des mesures plus spécifiques. Plusieurs stratégies sont mises en place : vérifier moins souvent les actus, désactiver les notifications sur mobile, éviter certains sujets comme la guerre en Ukraine, très anxiogène pour beaucoup.

8 MOINS DE PERSONNES COMMENTENT LES ACTUS EN LIGNE

Dans l'ensemble des marchés, seulement environ un cinquième (22 %) commente les actus en ligne, tandis qu'environ la moitié (47 %) ne participe pas du tout aux actualités. **Au Royaume-Uni, en Espagne ou aux États-Unis, la proportion de participants actifs a chuté de plus dix points depuis 2016.** Au Royaume-Uni, seul un consommateur sur dix participe activement aux actualités en ligne « mais leurs activités semblent souvent influencer fortement l'agenda des médias traditionnels et façonner les débats plus

larges », note le rapport. La participation en ligne s'est déplacée vers des réseaux fermés tels que WhatsApp, Telegram et Discord.

Dans tous les pays, le constat relève que le groupe qui prend la parole a tendance à être masculin, mieux éduqué et plus partisan dans ses opinions politiques – « le même profil démographique non représentatif auquel les nombreux médias d'information s'adressent ».



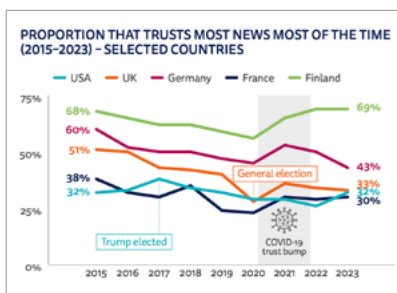
Reuters Institute Digital News Report 2023.

9 LA CONFIANCE DANS LES MÉDIAS A DIMINUÉ

La confiance dans les médias a diminué, dans l'ensemble des marchés, de deux points de pourcentage supplémentaires au cours de la dernière année, inversant ainsi dans de nombreux pays les gains réalisés au plus fort de la pandémie de coronavirus. **En moyenne, quatre personnes sur dix (40 %) déclarent faire confiance à la plupart des informations.** La

« Le monde de l'info s'apparente pour beaucoup à un «McDo géant où l'on s'abreuve de Big Mac et de "Face à Baba" en permanence».

Finlande reste le pays avec le plus haut niveau de confiance global (69 %) tandis que la Grèce (19 %) présente le plus bas, marqué par « des débats passionnés sur la liberté de la presse et l'indépendance des médias ».



Reuters Institute Digital News Report 2023.

10 LA CRITIQUE DES MÉDIAS, UN ÉLÉMENT CLÉ DU JEU POLITIQUE

Un facteur potentiel contribuant à la faible confiance envers les médias reste la critique généralisée et sans détour des médias de la part de diverses sources. Une corrélation existe entre la faible confiance et la critique des médias. Si la mise en cause des pratiques journalistiques est utile, elle peut s'avérer également être une rhétorique dangereuse destinée à saper la liberté de la presse. Les politiciens sont considérés comme les principales sources de critiques des médias, comme aux États-Unis (58 %) où certains responsables politiques manient régulièrement une rhétorique anti-média avec des expressions telles que « fake news media ». Aux Philippines, les journa-

listes critiques envers le gouvernement sont régulièrement qualifiées de « communiste » ou de « terroriste ». Les niveaux les plus élevés de critiques des médias sont observés en Grèce, aux Philippines, aux États-Unis, en France et au Royaume-Uni. « Les critiques des médias sont devenues un élément clé du jeu politique, une façon de détourner les critiques et d'intimider les enquêtes, et ces tactiques rencontrent souvent un terrain fertile », observe l'étude.

11 LES PODCASTS : UNE PORTÉE STABLE

La portée des podcasts reste stable avec un public fidèle. Les éditeurs investissent toujours dans les podcasts car ils sont relativement peu coûteux, contribuent à établir des relations fidèles et attirent un public plus jeune. Les podcasts approfondis, inspirés par « The Daily » du New York Times ainsi que les émissions de discussions prolongées, telles que « Joe Rogan Experience » sont les plus largement consommés sur les marchés. Il y a également une popularité croissante des podcasts d'actualités en forme vidéo ou hybride. Aux États-Unis, les talk-shows télévisés sont régulièrement transformés en podcast, tandis que le Newscast de la BBC est désormais filmé pour la télé et la distribution numérique. En termes de podcasts d'actualité, le YouTubeur Hugo Décrypte, qui produit « Actus du jour », un podcast audio et vidéo de dix minutes qui explique l'actu du jour pour un public

plus jeune, a surpassé en nombre d'auditeurs les podcasts d'actu des médias traditionnels français.

12 UNE AMBIVALENCE ENVERS LES ALGORITHMES

Le public est ambivalent à l'égard des algorithmes. Il n'est pas toujours convaincu que les journalistes et les organisations de presse peuvent faire mieux dans la sélection. Les données mettent en évidence le mécontentement général des audiences quant à la manière dont le contenu leur est sélectionné. Il s'agit d'une opportunité pour les éditeurs de créer quelque chose de meilleur.

CONCLUSION

Dans ce contexte, le public recherche non pas plus d'informations, mais des informations plus pertinentes, qui les aident à saisir des problématiques complexes. Avec l'arrivée de l'IA générative, qui « menace de déclencher une nouvelle vague de contenu personnalisé, mais potentiellement peu fiable », il devient impératif pour le journalisme de se démarquer en termes d'exactitude, d'utilité et d'humanité. « La construction de relations et de communautés ne consistera pas seulement à pousser les gens vers des sites web et applications, mais à s'engager via d'autres plateformes », estime le Reuters. ■

SXSW 2023 : AMY WEBB PRÉVOIT LA GRANDE « IASMOSE »

Face aux conséquences d'explosions cataclysmiques : la publication de systèmes d'intelligence artificielle générative tels que ChatGPT et Mid-journey, l'invasion de l'Ukraine par la Russie, l'incertitude profonde sur la récession mondiale et les algorithmes de repliement des protéines d'AlphaFold qui ont prédit les structures de presque toutes les protéines cataloguées connues de la science, pour n'en citer que quelques-unes; Amy Webb prévient dès le début de sa conférence : « **Il vous faut un estomac solide** », car ce qu'elle a à dire est effrayant.

Les principaux enseignements de sa conférence sont les suivants :

1. Se concentrer pour voir les signaux
2. Utiliser l'ADM (Agir, Décider, Monitorer) pour établir des priorités stratégiques
3. L'Internet tel que nous le connaissons est terminé. L'IASMOSE est la prochaine étape
4. Tout est information lisible
5. L'ère de l'informatique d'assistance est venue
6. Les nouveaux outils ne sont pas accessibles à tous, ce qui crée une nouvelle fracture numérique
7. Tout le monde aura besoin de se perfectionner (y compris les enfants)
8. Les grandes entreprises technologiques deviennent plus grandes et plus puissantes.

*Par Kati Bremme, directrice de l'Innovation France
Télévisions et rédactrice en chef de Méta-Média*

Les métavers et autres Web3 décentralisés ne sont finalement que des badi-neries (ou des supports) qui cachent l'« éléphant dans la pièce » revenu sur le devant de la scène cette année : l'intelligence artificielle. Quand, en 2017, Amy Webb ne parlait que de la menace de la distribution de contenu par l'IA, maîtrisée en exclusivité par les GAFAM, cette année, s'y ajoute aussi la création de contenu. Pour le meilleur et pour le pire, la futurologue la plus célèbre de la planète a détaillé lors de la SXSW 2023 le concept de l'IASMOSE (AISMOSIS en anglais, un mot-valise bizarre formé par IA et os-mose), qu'il s'agit de maîtriser pour ne pas passer à côté de la fin d'Internet tel que nous le connaissons et d'une transformation aussi impactante que l'arrivée des réseaux sociaux et du smartphone.

TOUT (VRAIMENT TOUT) SERA INFORMATION

« **Text-to-Everything n'est que le début** », prédit Amy Webb. Alors qu'il existe aujourd'hui des IA pour des applications bien précises, qui génèrent de nouveaux textes, des images, des vidéos ou des sons et de la musique à partir des textes saisis par les utilisateurs, il y aura bientôt des modèles d'IA multimodaux qui maîtriseront tout cela et plus encore – et qui s'immisceront dans chaque coin de notre quotidien numérique. « Cela va arriver très vite », dit Amy Webb. Or, « Nous ne sommes pas préparés à cela », ajoute-t-elle.

« ET SI VOUS NE FOUILLIEZ PLUS INTERNET, MAIS QUE CE SOIT INTERNET QUI VOUS FOUILLE » ?

Contrairement aux années 1990 (citées par nombre de conférences cette année à SXSW pour marquer une nouvelle étape d'Internet), nous partageons nos données non seulement en surfant sur Internet, mais aussi en publiant et en aimant sur les réseaux sociaux ou en faisant du shopping en ligne ou encore en regardant des films en streaming. Même les odeurs corporelles, l'arrière-plan apparemment insignifiant des appels



vidéo ou les informations recueillies par des toilettes intelligentes pourraient devenir du matériel d'entraînement. Les utilisateurs et leurs activités seront des ensembles de données pour alimenter l'IA. **Et très vite, l'IA sera nourrie par des contenus créés par l'IA, nous entrons alors dans une boucle infinie de contenus qui affectera fortement la chance de découvrabilité. Des systèmes IA interagissent avec TOUTES les données.** Tout devient information, et l'on ne trouvera plus rien. Un scénario pessimiste à forte probabilité, prévient la futurologue.

Pourquoi est-elle si pessimiste ? D'une part, parce qu'une entreprise comme OpenAI, qui a déclaré en 2020 au sujet de son IA textuelle GPT-2 qu'elle était « trop dangereuse pour être publiée » en raison de problèmes de sécurité, a malgré tout lancé depuis une IA après l'autre. Et d'autre part,

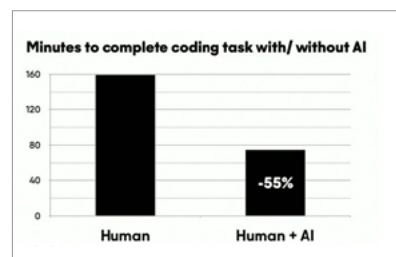
parce que le développement de nouveaux modèles d'IA nécessite une telle puissance de calcul que celle-ci ne peut en fait être fournie que dans les clouds d'AWS, Microsoft ou Google...

DES COMPÉTENCES POUR ÉVITER UNE NOUVELLE FRACTURE NUMÉRIQUE

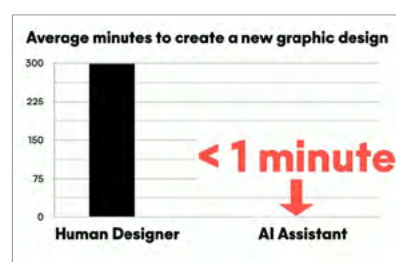
Ce n'est pas tant que l'IA prendra nos jobs, mais que toutes les personnes – et ce, quelle que soit leur génération – qui ne seront pas habilitées à utiliser les nouveaux outils d'IA pourraient bientôt être complètement laissées pour compte. « Nous sommes en train de créer une nouvelle et dangereuse fracture numérique », prévient Amy Webb.

Face à cette révolution, de nombreuses institutions préfèrent fermer les yeux, comme si le problème pouvait disparaître en ne pas y pensant. Le NYC Department of Education a banni les outils d'IA générative (à l'instar de Sciences Po), l'ICML interdit des chercheurs en IA générative. Pour Amy Webb, il ne s'agit pas de remplacer les professeurs par l'IA (encore moins d'investir des millions dans des start-ups), mais plutôt d'améliorer considérablement les moyens de l'éducation. Car les IA génératives, à condition de les maîtriser, sont aussi

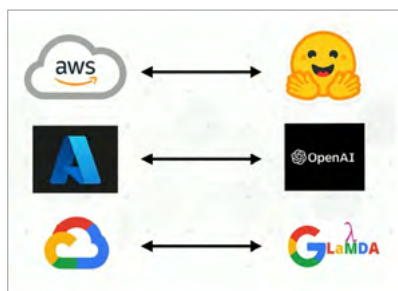
de formidables outils d'efficacité. Une opportunité illustrée avec deux exemples :



Le gain de temps par l'IA pour les développeurs.



Le gain de temps par l'IA pour les designers.



Présentation d'Amy Webb.

Une réduction bien sûr un peu caricaturale, mais la possibilité d'utiliser des IA comme ChatGPT pour accélérer des tâches est évidente. Amy Webb a fait en direct la démonstration : en 10,7 secondes, avec la génération automatique d'un business plan d'une nouvelle start-up. Une proposition ensuite à améliorer avec un humain dans la boucle.

« Dans un environnement où la confiance est faible, les marques qui construisent et entretiennent des relations durables avec les consommateurs prospéreront. »

UN JOURNALISME INTELLIGENT AIDÉ PAR L'IA

Les outils d'IA générative réécrivent les flux de travail, les modèles de transformation révolutionnent les moteurs de recherche, et le manque de confiance dans les informations rend les organismes de presse vulnérables. Les journalistes ont besoin de connaissances techniques pour retracer les récits, se protéger des cyberattaques et monétiser leur travail. Mais les entreprises de médias traditionnelles restent sceptiques face à l'évolution de leurs activités devant le contexte de la révolution de l'IA.

Il est de plus en plus essentiel pour les journalistes d'investigation de disposer de l'expertise technique nécessaire pour comprendre et expliquer la manière dont la technologie est utilisée pour exercer le pouvoir. **L'intelligence artificielle est déjà en train de changer l'industrie de l'information, et elle devient de plus en plus importante, en particulier pour les éditeurs locaux, car les solutions d'automatisation de l'information commencent à prendre de l'ampleur.** Combiné à la personnalisation des informations et à la segmentation géographique, il peut être le moteur d'un journalisme de proximité précieux et pertinent dans les rédactions.

L'IA générative facilite également la transformation du format du contenu : le texte peut devenir de l'audio. L'audio peut être associé à des éléments visuels pour créer une vidéo. La vidéo peut être résumée efficacement en texte. Les créateurs peuvent ainsi distribuer des contenus dans des formats et sur des plateformes qu'ils n'auraient peut-être pas explorés autrement. Dans un environnement où la confiance est faible, les marques qui construisent et entretiennent des relations durables

avec les consommateurs prospéreront. Les éditeurs qui ont accumulé de la confiance au fil des décennies peuvent l'appliquer à tous les formats de distribution grâce à un développement de produit réfléchi.

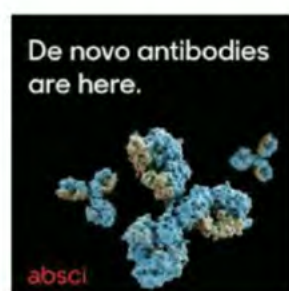
Enfin, **l'éducation aux médias est un investissement à long terme.** Les outils à court terme pour lutter contre la désinformation sont limités : dans une étude de l'Université de New York, les chercheurs ont montré que le fait d'étiqueter les informations erronées dans les résultats de recherche et les flux de médias sociaux n'avait aucun impact sur l'attitude ou le comportement des utilisateurs à l'égard des informations non fiables.

UN CAS D'USAGE UTILE DU MÉTAVERS + IA

Amy Webb a tout de même terminé sa conférence sur une évolution positive très concrète, à savoir le métavers industriel ou le métavers médical. En effet, l'interaction entre les technologies XR, les lunettes de données, les jumeaux numériques, les nouveaux capteurs et l'IA pourrait donner naissance à des systèmes d'assistance numérique qui, en médecine, par exemple, pourraient contribuer à des opérations personnalisées et précises, impensables aujourd'hui. « Je pense que nous regarderons l'an-

née 2023 et que nous ne ferons que secouer la tête », estime Webb. « Les chirurgiens nous ouvraient autrefois sans l'assistance d'ordinateurs dans le Métavers Médical ! »

Face à l'impact des IA génératives sur tous les métiers et sur la façon dont nous percevons l'information, on ne peut pas être observateur passif de cette tendance. La technique de l'autruche et l'interdiction ne sont pas les bons moyens pour affronter cette révolution qui interconnecte de plus en plus nos vies avec l'IA (qu'on l'appelle AISMOSIS ou autre chose). D'autant plus qu'il reste toujours le problème des biais des algorithmes, aussi sophistiqués qu'ils soient : quand Amy Webb a demandé à Midjourney d'imaginer sa conférence à SXSW, l'image générée montre un homme blanc corpulent face à un parterre d'hommes blancs. On est bien loin de la réalité de la salle à Austin... ■



Zero-shot generative AI

Generated antibody designs from scratch that weren't like any that exist

Présentation d'Amy Webb.

LA TRAQUE DES USAGES

EMPLOIS

- 300 millions d'emplois pourraient être menacés par l'intelligence artificielle - The Potentially Large Effects of Artificial Intelligence on Economic Growth (Briggs/Kodnani) - Goldman Sachs, mars 2023
- 30 % des postes chez IBM pourraient être remplacés dans les cinq prochaines années, soit environ 7 800 emplois - IBM, mai 2023
- 9 dirigeants sur 10 estiment que l'IA représente une opportunité business pour leur entreprise - Boston Consulting Group, janvier 2023
- Aux États-Unis, le taux d'adoption de l'IA générative sur le lieu de travail est le plus élevé chez la génération Z, avec 29 % - Statista, mai 2023
- En moyenne, 79 % des dirigeants rapportent une diminution des coûts grâce à l'adoption de l'IA - McKinsey, août 2022
- Jusqu'à \$375,00 peut gagner un « Prompt Engineer » chez Anthropic - Bloomberg, mai 2023
- *Bild*, le plus grand journal allemand, supprime 20 % des emplois pour se préparer à un avenir numérique propulsé par l'IA - FAZ, juin 2023

RÉDACTIONS

- **La moitié des médias utilisent les IA génératives, et 70 % d'entre eux considèrent que les IA génératives auront un impact positif sur les salles de rédactions et les journalistes** - WAN-IFRA, mai 2023
- 54 % des rédactions qui ont un usage de ChatGPT, l'utilisent pour

résumer des articles, 44 % pour faire de la recherche d'information, et à 32 % pour de l'idéation de contenu et de la création d'articles - WAN-IFRA, mai 2023

- Seulement un cinquième des rédactions disent avoir reçu des directives de la part de leur management quant à la façon d'utiliser les IA génératives - WAN-IFRA, mai 2023

CROISSANCE & INVESTISSEMENT

- Le PIB mondial pourrait augmenter de 7 % grâce à l'intelligence artificielle - The Potentially Large Effects of Artificial Intelligence on Economic Growth (Briggs/Kodnani) - Goldman Sachs, mars 2023
- Le marché mondial de l'intelligence artificielle devrait atteindre 1,59 trillion de dollars d'ici 2030 - Precedence Research 2030
- Les investissements dans l'IA générative ont progressé de +425 % depuis 2020 pour atteindre 2,1 milliards de dollars, Pitchbook pour le *Financial Times*, décembre 2022
- Hugging Face a été valorisée à près de 2 milliards d'euros, mai 2022
- Les dépenses en IA pourraient atteindre 300 milliards de dollars d'ici 2026, International Data Corporation (IDC), Worldwide Artificial Intelligence Spending Guide, mars 2023
- 50 % de l'ensemble des investissements réalisés dans le monde sur la période 2022-2026 sont réalisés aux États-Unis - International Data Corporation (IDC), Worldwide Artificial Intelligence Spending Guide, mars 2023
- En 2021, le marché de l'intelligence

artificielle était évalué à environ 60 milliards de dollars américains et devrait atteindre environ 422 milliards de dollars américains d'ici 2028 - Zion Market Research, juin 2022

- Le marché de l'IA générative devrait atteindre 111 milliards de dollars américains d'ici 2030 - Acumen, décembre 2022
- Jusqu'à présent, six entreprises dans le domaine de l'IA générative ont atteint le statut de licorne, ce qui signifie qu'elles valent plus d'un milliard de dollars. Ces entreprises sont OpenAI (29 milliards), Hugging Face, Lightricks, Jasper, Glean et Stability AI - CB Insights, 2023

PRODUCTIVITÉ

- Les outils dédiés à l'accélération des ventes, les systèmes de recommandation d'un produit, et les agents augmentés utilisés pour les services clients représentent 30 % des utilisations de l'IA en 2023 - International Data Corporation (IDC), Worldwide Artificial Intelligence Spending Guide, mars 2023
- **D'ici 2030, un blockbuster sera produit avec 90 % du film généré par l'IA** - Gartner, octobre 2021
- D'ici 2025, 30 % des messages sortants des grandes organisations seront générés de manière synthétique, contre moins de 2 % en 2022 - Gartner, janvier 2023
- D'ici 2025, l'IA générative représentera 10 % de toutes les données produites, contre moins de 1 % en 2021 - Gartner, octobre 2021
- ChatGPT permettrait d'être 37 % plus efficace dans la réalisation de tâches du quotidien - MIT, mars 2023

68 % des Français qui utilisent les IA génératives en entreprise le cachent à leur supérieur hiérarchique.

IFOP/Talan, mai 2023

UTILISATEURS

- En janvier 2023, ChatGPT comptait près de 100 millions de visiteurs uniques - Banque UBS, mars 2023
- ChatGPT a dépassé le million d'utilisateurs 5 jours après son lancement - Greg Brockman, co-fondateur d'OpenAI, décembre 2022
- En avril 2023, « chat.openai.com » a atteint les 4 milliards de visites dans le monde, depuis ses débuts - *Les Echos*, avril 2023
- 67 % des visiteurs de Midjourney sont des hommes et 33 % sont des femmes - Simlarweb, mai 2023
- DALL-E, autre outil d'OpenAI, qui génère des images, a mis environ 2,5 mois pour atteindre le million d'utilisateurs - Sam Altman, co-fondateur d'OpenAI, juin 2022
- GitHub Copilot, un outil d'aide à la programmation basé sur l'IA, a compté 400 000 abonnés au cours du premier mois de sa sortie - Microsoft, juillet 2022
- La Chine est le pays où le plus de personnes recherchent ChatGPT sur les moteurs de recherche, suivi du Népal et de Singapour - *Wired*, février 2023.

ENTRAÎNEMENT

- GPT-3 aurait été formé sur environ 45 téraoctets de données textuelles - ou un quart de l'ensemble de la Bibliothèque du Congrès - pour un coût estimé à plusieurs millions de dollars, McKinsey, janvier 2023
- 175 milliards de paramètres différents ont été utilisés tout au long du processus de formation du modèle linguistique GPT-3 - OpenAI, janvier 2023.
- Le modèle de Midjourney a été entraîné sur plus de cent millions d'images, le tout sans copyright - *Forbes*, septembre 2022
- Faire fonctionner un modèle d'IA

conversationnelle pourrait coûter environ 10 fois plus cher qu'un moteur de recherche pour une entreprise comme Google - Reuters, février 2023

- Le coût d'entretien de ChatGPT serait de 700 000\$ par jour, soit 0,36\$ par requête - *SemiAnalysis*, février 2023

ÉTHIQUE

- Plus de 30 000 signataires pour la lettre ouverte appelant à une pause sur le développement de l'intelligence artificielle. OpenAI elle-même plaide le contrôle et la régulation de l'IA - *Future of Life*, mars 2023, et OpenAI, mai 2023
- 78 % des adultes américains sont défavorables aux articles de presse rédigés par l'IA. 19 % des jeunes considèrent cela comme un développement positif - *Monmouth University*, février 2023
- Pour les dirigeants, l'une des principales préoccupations liée à l'IA générative concerne la sécurité des données - *Salesforce*, mars 2023
- Le logiciel de détection d'IA d'OpenAI peut identifier correctement 26 % du contenu rédigé par l'IA tout en étiquetant incorrectement du texte rédigé par des humains comme étant écrit par une IA 9 % du temps (faux positifs) - OpenAI, janvier 2023

ET EN FRANCE ?

- **68 % des Français qui utilisent les IA génératives en entreprise le cachent à leur supérieur hiérarchique - IFOP/Talan, mai 2023**
- 71 % des Français ont entendu parler des IA génératives, et 74 % d'entre eux pensent que ces IA constituent une nouvelle révolution industrielle - IFOP/Talan, mai 2023
- 44 % des Français qui utilisent les

IA génératives les utilisent à la fois dans un cadre personnel et professionnel - IFOP/Talan, mai 2023

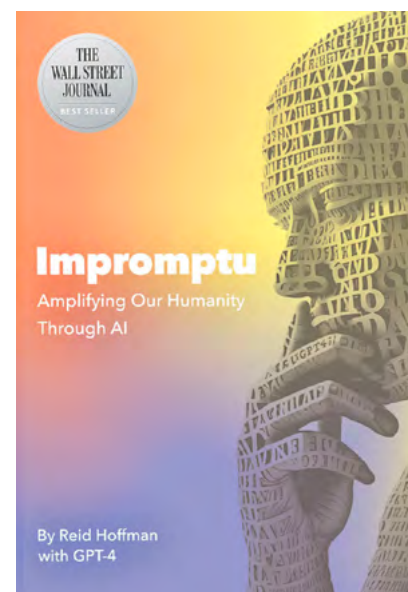
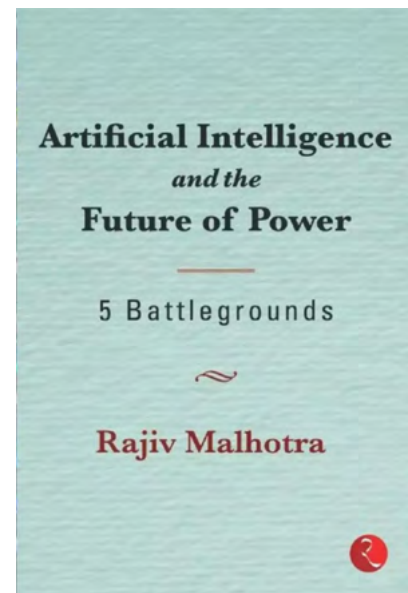
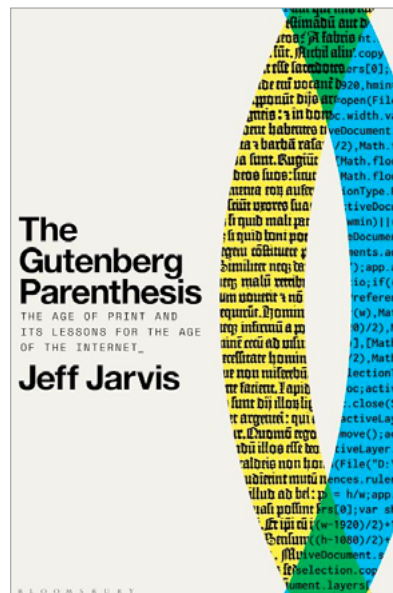
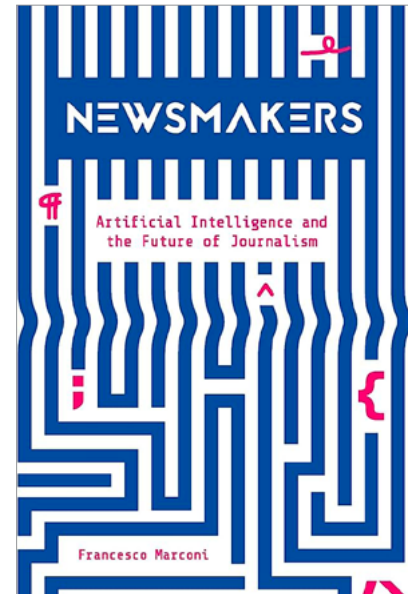
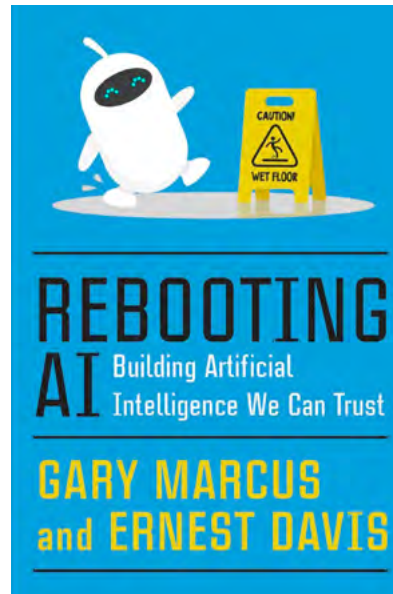
- 45 % des 18-24 ans utilisent les IA génératives contre seulement 18 % des 35 ans et plus - IFOP/Talan, mai 2023
- 72 % des Français estiment ne pas avoir les connaissances suffisantes pour utiliser les IA génératives - IFOP/Talan, mai 2023
- 68 % des Français disent avoir des craintes vis-à-vis de l'émergence des IA génératives - IFOP/Talan, mai 2023
- Emmanuel Macron a annoncé vouloir financer l'IA à hauteur de 500 millions d'euros dans le cadre de France 2030 afin de consolider les pôles de formation et de recherche - *Viva Technology*, juin 2023

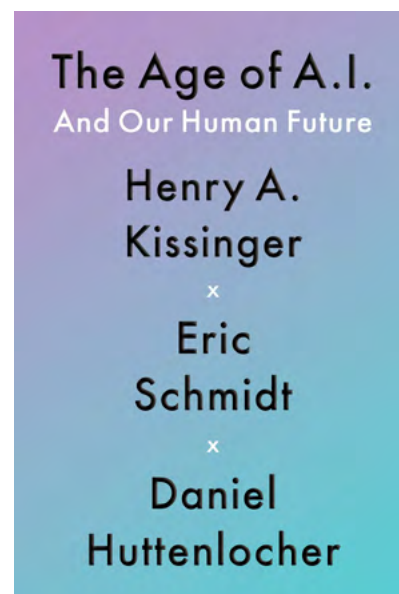
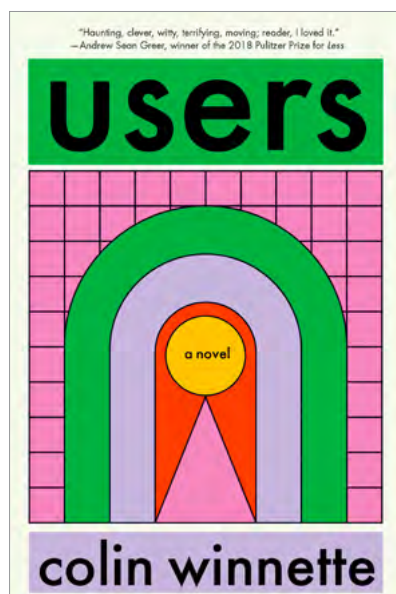
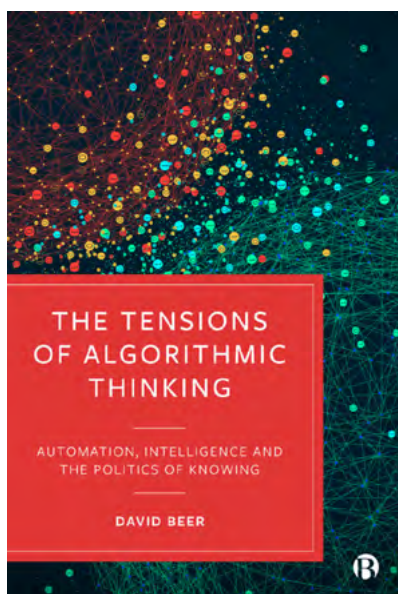
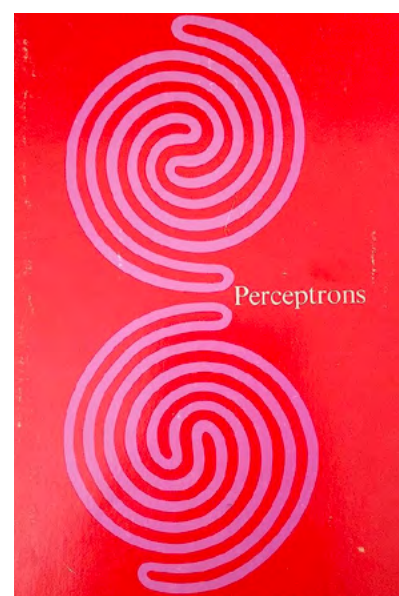
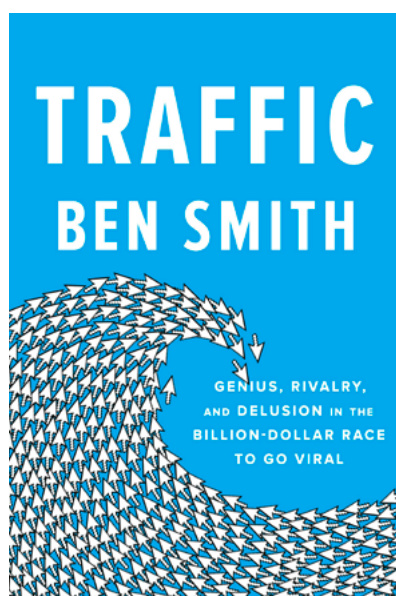
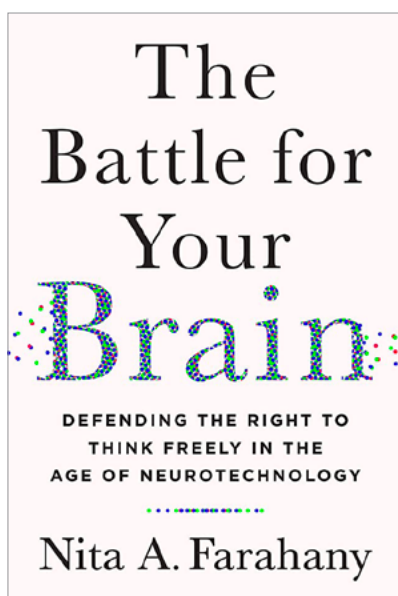
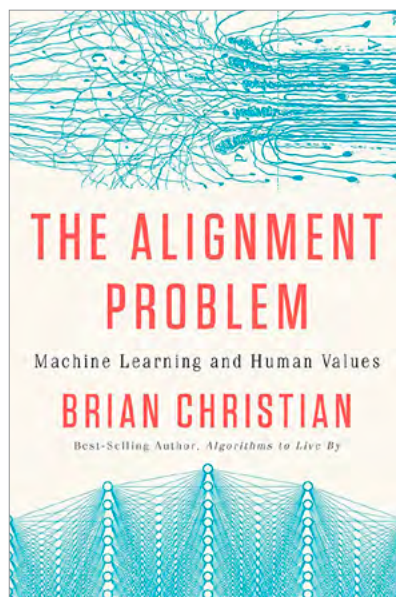
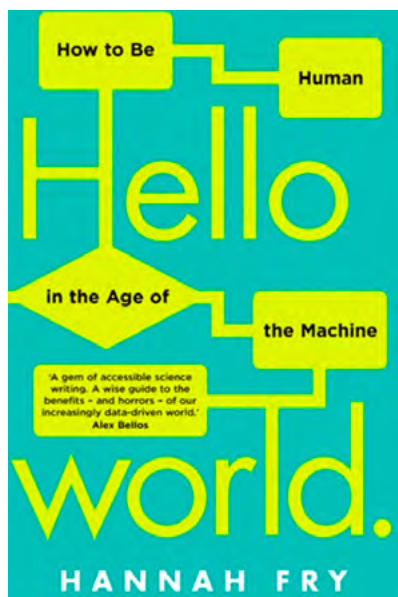
IMPACT ENVIRONNEMENTAL

- Selon les estimations, ChatGPT serait responsable de 8 tonnes de CO2 par an, soit l'équivalent de la consommation moyenne de deux personnes pour une année - *Kasper Groes Albin Ludvigsen*, décembre 2022
- L'entraînement du modèle de GPT-3 aurait produit 500 tonnes de CO2, soit plus de 1,6 millions de kilomètres parcourus en voiture - *The Guardian*, mai 2023
- Une recherche sur ChatGPT serait environ 8 fois plus gourmande en énergie qu'une recherche sur un moteur de recherche classique - *The Developer Space*, janvier 2023 ■

LIVRES

RECOMMANDÉS





Édité par la Direction de l'Information

Directeur de la publication : **Delphine Ernotte Cunci**

Directeur de la collection : **Eric Scherer**

Rédactrice en chef : **Kati Bremme**

Ont collaboré à ce numéro : **Zoé Aegerter, Charlie Beckett, Valérie Bellier, Martin Bieri, Grégoire Borst, Hervé Brusini, Antonio Casilli, ChatGPT, Jean Chrétien, Laurence Devillers, Ezra Eeman, Louise Faudeux, Yann Ferguson, Alexei Grinbaum, Myriam Hammad, Alexandra Klinnink, Chine Labbé, Valentine Lopez, Pascal Marchand, Francesco Marconi, Olivier Martinez, Pauline Maury, Asma Mhalla, Alexandre Michelin, Antoine Multone, Perrine Pelletier, Benoît Raphaël, Clara Schmelck, Evan Shapiro, Françoise Soulié-Fogelman, Cédric Villani, Arnaud Vincenti, le Yle NewsLab**

Secrétariat de rédaction : **Anaïs Bocquet**

Conception et réalisation : **Élise Taponier**

Illustration de couverture : **Jean-Christophe Defline & Kati Bremme**

Impression : **Cap Impression**



ON DIT PARFOIS
QUE DES EMPLOIS
SERONT REMPLACÉS
PAR L'INTELLIGENCE
ARTIFICIELLE
GÉNÉRATIVE... NON !
MAIS LA PERSONNE QUI
NE SAIT PAS TRAVAILLER
AVEC L'INTELLIGENCE
ARTIFICIELLE
GÉNÉRATIVE SERA
REEMPLACÉE PAR LA
PERSONNE QUI SAIT.

Françoise Soulié-Fogelman

