VIVE L'INTELLIGENCE... MÊME ARTIFICIELLE !

L'intelligence artificielle aurait pu être la personnalité de l'année 2023. Car même si elle compte déjà quelques années d'existence derrière elle, jamais l'IA n'avait autant suscité de légitimes débats. Popularisée par les intelligences artificielles génératives qui créent, à la demande, du texte (ChatGPT), des images (DALL-E), des fausses vidéos (Sora, Deep Face Lab), cette technologie fascine autant qu'elle terrifie. Plusieurs experts la qualifient déjà comme une nouvelle révolution industrielle, aussi conséquente que le furent autrefois celles liées au développement du chemin de fer ou de l'électricité. Entre promesses et menaces, que penser de l'IA ?

PAR **OLIVIER VAN CAEMERBEKE**

« *Je suis désolé, Dave, je crains de ne pas pouvoir faire cela* ». Cette phrase émane d'une intelligence artificielle bien connue des cinéphiles : HAL 9000. Le surpuissant ordinateur à l'œil carmin du film « *2001 l'Odyssée de l'espace* » de Stanley Kubrick est sans doute la plus emblématique des intelligences artificielles de cinéma. En 1968 déjà donc, cette technologie était perçue comme une menace pour l'humanité. D'une manière générale, les films, les romans, les bandes dessinées ont rarement façonné une bonne image à l'IA, comme tout le monde l'appelle désormais. Cette dernière est-elle vraiment l'adversaire de l'espèce humaine ou, au contraire, un soutien... ce pour quoi elle a quand même été créée ?

Avant de répondre à cette question, il n'est pas inutile de rappeler que ce qu'on appelle l'IA, c'est l'expression des capacités des machines à réaliser des tâches qui, jusqu'alors, nécessitaient une intelligence humaine, ou animale : comprendre un langage, le traduire, reconnaître une image, générer des synthèses de données, résoudre des problèmes. Et si on parle autant de ces outils aujourd'hui, c'est grâce, ou à cause, du développement des « intelligences artificielles génératives » qu'incarnent notamment ChatGPT qui sait converser en répondant à des questions utilisant le langage commun, Midjourney et Dall-E qui, sur le même modèle, créent des images. L'IA générative c'est donc celle qui est capable d'inventer des échanges, des histoires, des images, des vidéos, de la musique, des

PHOTO © PIXABAY

articles… en s'appuyant sur des contenus existant pour en façonner de nouveaux.

« L'IA N'EST NI BONNE NI MAUVAISE » : PAROLE D'IA !

Si l'on peut légitimement penser que la plupart des ingénieurs travaillant sur de l'intelligence artificielle ne cherchent pas à nuire l'humanité, donner aux machines la capacité d'apprendre et de créer seules, n'est pas forcément rassurant. Ces IA, qui seront demain intégrées dans des enveloppes robotiques, nous contrôleront-elles un jour ? Utiliseront-elles nos données pour nous mettre hors-jeu de la société ou pour concentrer le pouvoir au profit de quelques-uns ? Autrement dit, faut-il avoir peur de l'IA ?

« L'IA n'est pas une force indépendante et incontrôlable, mais le résultat des choix et des actions des humains qui la conçoivent, la développent, la déploient, la régulent, l'utilisent, etc. L'IA n'est ni bonne ni mauvaise en soi, elle dépend de l'usage que l'on en fait et des valeurs que l'on y intègre. C'est pourquoi il est important de se poser des questions sur le sens et le but de l'IA et de définir des principes et des normes pour la gouverner et la maîtriser. Il est aussi essentiel de favoriser le dialogue et la participation de tous les acteurs et les parties prenantes de l'IA, afin de garantir son acceptabilité sociale et son bénéfice collectif. » Cette réponse vous semble-t-elle sensée ? Elle n'a pas pourtant pas été formulée par un humain, mais par COPILOT, la nouvelle interface IA de Microsoft à qui nous avions posé la question.

Les Français portent, eux, un regard inquiet sur cette nouvelle révolution numérique. Selon une étude Ifop réalisée entre fin décembre 2023 et début janvier 2024, 35 % d'entre eux s'attendent même à ce que l'IA finisse un jour par gouverner l'humanité, rien que ça ! Et, 64 % des personnes interrogées voient cette nouvelle technologie *« bientôt capable d'agir comme les humains »*. À les lire, l'IA n'a cependant pas encore envahi leur espace de travail. 78 % des salariés estiment ne s'en être jamais servis dans leur vie professionnelle. Une large majorité de salariés estiment, par ailleurs, que l'intelligence artificielle ne sera jamais capable de réaliser leurs tâches. Quant à ceux qui expriment un avis contraire, ils envisagent que les IA seront aptes à les remplacer, en 2035.

MÉTRO, BOULOT, ROBOT

La disparition d'emplois est pourtant l'une des conséquences pointées par celles et ceux qui se penchent sérieusement sur la question. Selon une étude de la BBC réalisée avec la banque Goldmann Sachs en mars 2023, l'intelligence artificielle pourrait faire disparaître 300 millions d'emplois dans le monde. Kristalina Georgieva, directrice générale du Fonds monétaire international, estime que *« dans le monde, 40 % des emplois seront touchés. Et plus vous occupez un emploi qualifié, plus ce sera le cas. Ainsi, pour les économies avancées, et certains pays émergents, 60 % des emplois seront concernés »*. Et une autre étude réalisée avec l'université de Pennsylvanie par OpenAI, société créatrice de ChatGPT, conclut qu'aux États-Unis, environ 80 % de la main-d'œuvre pourrait voir au moins 10 % de ses tâches modifiées par les différents types de GPT.

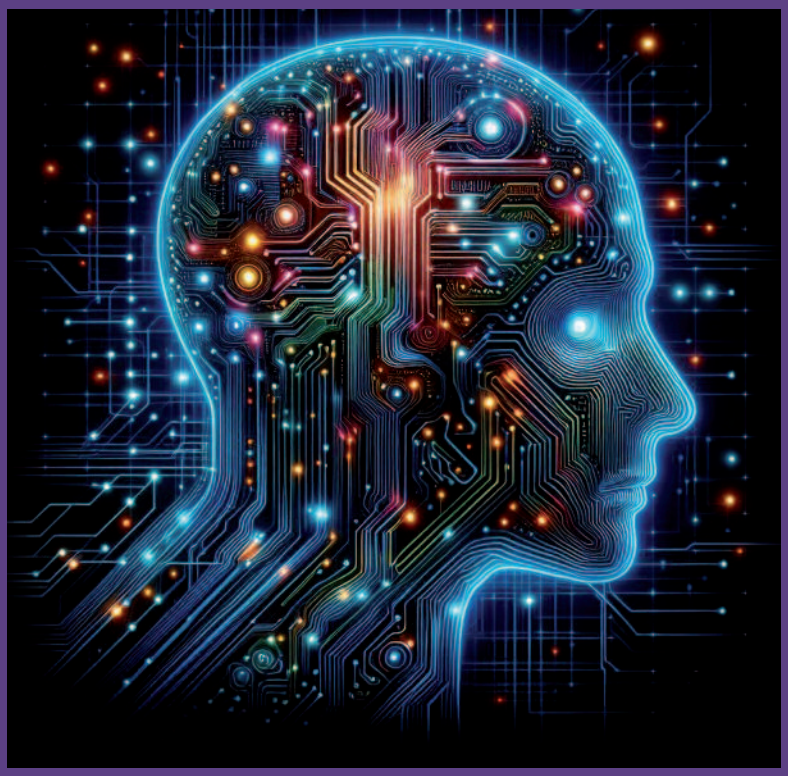
En France, une étude du cabinet de conseil en management Roland Berger a étudié, pour 400 métiers, quelles tâches pourraient être améliorées grâce à des IA, voire totalement assumées par elles. Il en ressort que 800 000 emplois sont potentiellement menacés d'être détruits d'ici 2030. Sans surprise les métiers administratifs (199 000), les services comptables (140 000), les employés de bureau (100 000) ou encore les agents de réception (80 000) sont les plus concernés par ce bouleversement. Le cabinet souligne cependant qu'1,4 million d'emplois pourraient être « augmentés » et gagner en productivité grâce à une IA générative. *« Gain de productivité ne signifie pas nécessairement réduction de l'emploi,* précise, dans *Le Monde*, Laurent Benarousse, le directeur général du cabinet. *On ne croit pas que l'impact sur le volume global d'emploi en France sera massif, en revanche notre conviction est qu'il y aura des substitutions, certains feront autre chose demain que ce qu'ils font aujourd'hui. »*

De son côté, l'économiste du travail, Philippe Askenazy, redoutait récemment dans *Libération* que ces nouvelles intelligences soient surtout utilisées comme une menace par des employeurs *« qui se concentreraient sur la surveillance accrue plutôt que sur la hausse de la production »* et qui pourraient *« dire aux salariés : 'Si vous ne vous comportez pas bien, vous serez remplacés par une machine'. Ce n'est pas une technologie pour libérer le travail. Elle risque d'avoir plus de conséquences sur la répartition de la valeur que sur le niveau d'emploi »*.

À MORT L'IA !

Dans le bras de fer virtuel entre les enthousiastes et les alarmistes, ces derniers ont donc le vent en poupe. Il faut dire qu'au-delà des bouleversements liés à l'emploi, la liste des risques potentiels est assez vertigineuse, bien que parfois nourrie par un imaginaire débridé. *« L'expansion des capacités des technologies basées sur l'IA s'accompagne d'une augmentation de leur potentiel d'exploitation criminelle »,* avertit Lewis Griffin, chercheur en informatique à l'University College London (UCL). Mandatés par cette université, 31 experts des différentes en IA

LA PLACE DE L'IA AU SEIN DES COLLECTIVITÉS LOCALES



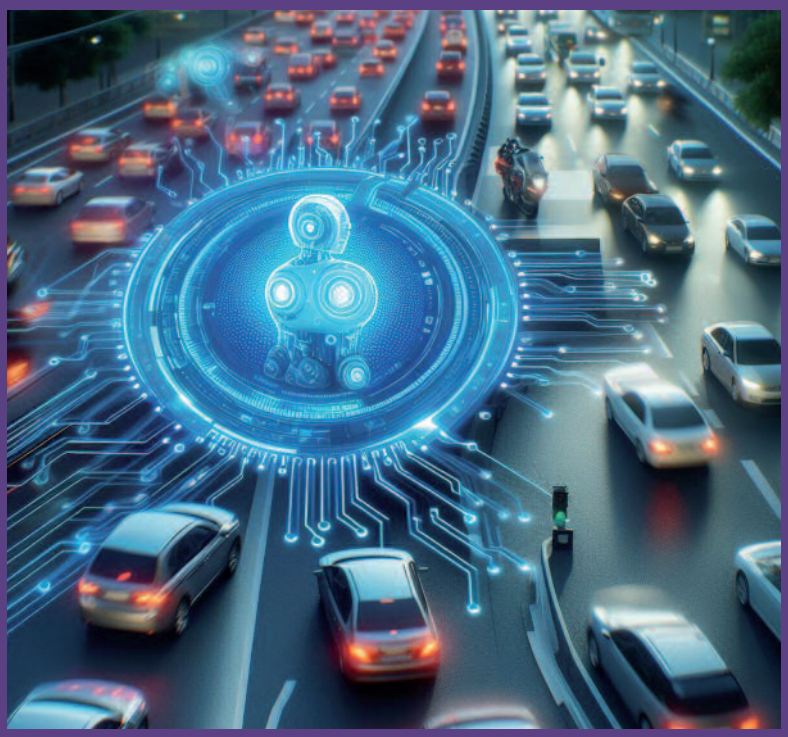


IMAGE CREE PAR COPILOT, IA INTEGREE A MICROSOFT

Simplification de l'accueil, optimisation des ressources, participation citoyenne, sécurité…, l'intelligence artificielle a de quoi intéresser les collectivités locales.

En novembre 2023, pour la deuxième année consécutive, l'Observatoire Data Publica a dévoilé les résultats de son *Baromètre national sur la data* dans les territoires. Une petite partie s'intéresse à l'IA et indique que 27 % des collectivités déclarent avoir déjà eu recours à l'intelligence artificielle ; c'est 12 points de plus qu'en 2022. Aucune commune de moins de 100 000 habitants ne déclare avoir eu recours à l'IA. Les métropoles sont les plus en pointe (44 %), suivies par les départements (36 %) et les régions (30 %). « *On peut faire l'hypothèse que cette accélération forte doit beaucoup à la mise en service de ChatGPT en décembre 2022, qui a fortement popularisé l'intelligence artificielle, dans le domaine des chatbots en particulier. Et la vague ne fait que commencer : 30 % des répondants envisagent de déployer de nouveaux projets IA dans les 12 prochains mois* », peut-on lire dans la synthèse de l'enquête. Les outils conversationnels de type chatbot arrivent nettement en tête des technologies utilisées (43 %), devant les modèles prédictifs, les outils de traitement automatique des images, du langage et du texte. Les collectivités ont recours à l'IA dans le cadre de leurs actions liées à la mobilité (8 % des réponses), l'administration (7 %), l'aménagement du territoire (6 %), le développement économique (5 %), la citoyenneté (3 %), les déchets (2 %). Des usages variés, mais, souligne l'observatoire, « le moment est aux expérimentations, plus qu'aux stratégies ».

CLIN D'ŒIL. QUAND CHATGPT ÉCRIT MIEUX QUE L'AUTEUR QU'IL IMITE !

Bernard Werber, l'un auteurs les plus lus en France, a demandé à ChatGPT d'écrire à sa manière un premier chapitre sur le même thème que son dernier roman. Résultat ?

« C'était mieux écrit que ce que j'avais rédigé. Des phrases plus longues. Du vocabulaire plus riche. Des tournures compliquées avec beaucoup de style (…) Cela m'a permis de comprendre encore mieux ma spécificité d'auteur. Je privilégie la complexité de l'histoire sur la complexité de la forme. (...). Je ne pense pas que chat GPT puisse inventer une intrigue complexe originale. Mais justement, c'est en ça que ce phénomène est intéressant. Cela va obliger les écrivains à devenir encore plus créatifs, encore plus originaux et probablement à prendre de plus en plus de risques pour sortir des chemins battus. »

SOURCE : « Le Biais de Bernard Werber » podcast de Radio France

ont été chargés de passer en revue 20 menaces potentielles qui ont été classées en fonction du niveau de danger. En tête de leur classement, les Deepfakes, ces vidéos qui imitent visage et voix de vraies personnes. Des usurpations d'identité qui peuvent nuire à une réputation, mais aussi être utilisées pour voler des données sécurisées, manipuler les opinions, faire du chantage…

Également dans le haut du tableau de l'UCL, le piratage de véhicules autonomes qui pourraient alors servir d'armes pour des terroristes. Les experts étendent aussi ce risque aux drones autonomes, même si *« à l'heure actuelle, les drones ne sont généralement pas utilisés pour commettre des crimes violents, mais leur masse et leur énergie cinétique sont potentiellement dangereuses s'ils sont bien ciblés (par exemple dans les moteurs d'avion) et ils pourraient également être équipés d'armes »*, écrivent les rapporteurs. Ces derniers citent aussi le phishing, ou « hameçonnage sur mesure », le piratage des systèmes contrôlés par l'IA permettant de générer une panne d'électricité généralisée, des embouteillages monstres ou des ruptures des logistiques de distribution alimentaire.

La rédaction de fausses informations rédigées par IA est aussi dans ce top 20, mais nous ne sommes déjà plus, ici, dans de la prospective. On pourrait encore citer les risques de détournement des systèmes de reconnaissance faciale, l'exploitation des biais existants des algorithmes pour exploiter par renforcer, les préjugés, les discriminations, la création de contrefaçons… *« Contrairement à de nombreux crimes traditionnels, les crimes dans le domaine numérique peuvent être facilement partagés, répétés et même vendus »*, déclare le Dr Matthew Caldwell, coauteur de l'article.

VIVE L'IA !

L'IA est imparfaite, c'est une évidence. Faut-il pour autant la jeter dans la corbeille informatique d'un simple clic en restant aveugle aux bénéfices à en attendre ? Évidemment, non. On pense d'abord au secteur de la médecine où elle joue déjà un rôle majeur. De jeunes entrepreneurs sortis de Centrale Paris ont ainsi fondé la start-up Lumio Médical qui propose un outil nourri d'intelligence artificielle afin de détecter automatiquement les risques d'accidents médicamenteux à l'hôpital, autrement dit les erreurs ou imprudences liées à des prescriptions inadéquates. *« Il a été estimé que près de 1 % des prescriptions faites contenaient une erreur potentiellement fatale pour le patient,* expliquent les concepteurs. *Ce sont donc près de 30 000 personnes qui décèdent suite à ces accidents chaque année en France. »*

Edi Prifti, chercheur à l'Institut de recherche pour le développement travaille quant à lui sur la façon dont l'IA peut aider les cardiologues à identifier les arythmies mortelles en détectant des variations minimes qu'un humain est incapable de déceler. Le traitement des cancers est un autre champ d'application. L'équipe de

IMAGE CREE PAR COPILOT, IA INTEGREE A MICROSOFT

Sarah Watson, oncologue et chercheuse à l'institut Curie, a appris à une IA à identifier la tumeur initiale lorsque, chez un malade, des métastases sont découvertes simultanément dans plusieurs organes. Connaître le premier organe à avoir été touché est crucial pour traiter correctement, or leur algorithme y parvient dans 79 % des cas. Un outil accessible à tous les patients de France concernés.

Certains diabétiques bénéficient déjà de pompes à insuline connectées à un système d'IA qui leur évitent de devoir mesurer leur taux de sucre dans le sang et de recalibrer les doses d'insuline délivrées par la pompe. Une start-up issue de la recherche de CentraleSupélec a, pour sa part, breveté une technologie de reconnaissance des mouvements humains par l'analyse des ondes WiFi. En s'appuyant sur les travaux de recherche en intelligence artificielle de son cofondateur, Piotr Antonik, ce système 100 % logiciel (comprendre sans caméra ou appareil portatif) sait repérer les chutes de personnes âgées à leur domicile. Un vrai enjeu puisqu'on estime à 2 millions le nombre de chutes de personnes âgées chaque année et, dans 10 % des cas, la personne ne parvient pas à se relever.

C'est aussi grâce à l'IA qu'une équipe du Broad Institute du MIT et de Harvard (Cambridge, Massachusetts) a identifié une classe de composés capables de tuer une bactérie résistante aux médicaments, en la testant sur le Staphylococcus aureus. Cela faisait des décennies que les antibiotiques arrivant sur le marché n'étaient que des versions améliorées de traitements existant, cette fois c'est d'une nouvelle classe d'antibiotiques dont on parle : une première depuis 60 ans.

L'IA : DES IDÉES VERTES… ET DES PAS MÛRES

Autre secteur où l'IA est porteuse de certains espoirs, l'environnement. Certes, cette nouvelle technologie part avec un lourd handicap puisqu'elle est très exigeante en ressources. Les data center qui les hébergent ont besoin d'électricité et d'eau pour pour être refroidis. Selon des études récentes, le prélèvement d'eau lié à l'IA au niveau mondial pourrait atteindre entre 4,2 et 6,6 milliards de mètres cubes en 2027, soit la moitié de celui du Royaume-Uni. Et lorsqu'on formule une requête sur ChatGPT, on émet dix fois plus de carbone que lors d'une recherche sur Google. *« Si l'on faisait autant de requêtes sur ChatGPT que l'on en fait sur Google, il n'y aurait pas assez d'énergie dans le monde entier pour y répondre »,* assure dans *Le Monde* Luc Julia, ingénieur cocréateur de Siri, l'assistant virtuel d'Apple.

Le bilan écologique de l'IA est-il donc plombé d'avance ? Peut-être pas. L'IA va, par exemple, contribuer à ce qu'on appelle « l'agriculture de précision », c'est-à-dire une agriculture ou les apports d'eau, d'engrais, de pesticides, etc. aux zones strictement nécessaires. L'IA est aussi utilisée pour surveiller la biodiversité dans les écosystèmes fragiles, aider à la détection précoce des incendies, modéliser les mouvements de feu, gérer les énergies renouvelables…

La revue *Science* a, elle, présenté une étude qui

démontre que GraphCast, le programme IA de Google, est plus rapide et plus précis que les méthodes traditionnelles de prévisions pour prévoir la météo si utile aux agriculteurs. La ville de Zurich utilise l'IA dans la gestion de son éclairage public et a réduit de 70 % sa consommation d'énergie, en allumant uniquement aux moments nécessaires, grâce aux prévisions de trafic.

« L'I A est un formidable outil pour lutter contre le changement climatique et pour la préservation de la diversité, ajoute Dejan Glavas, directeur de l'institut AI for Sustainability (intelligence artificielle pour le développement durable), professeur de finance durable à l'école de commerce Essca, dans *Le Figaro étudiant* du 6 décembre 2023. *L'ONG Rainforest déploie des outils de surveillance acoustique couplés à l'IA qui permettent de lutter contre la déforestation et le braconnage à Puerto Rico, dans les Caraïbes. Elle peut aussi servir à créer la ville intelligente' notamment en matière de gestion du trafic, des déchets ou de l'énergie. »*

En effet, le syndicat mixte Val-d'Oise Numérique a mis en place une plateforme de détection des dépôts d'ordures illégaux. Nourrie de vues satellitaires en haute résolution mises à jour deux fois par mois, l'IA peut repérer n'importe quel dépôt, évaluer son volume et alerter sur sa proximité avec une zone sensible comme un cours d'eau. À Pittsburgh (USA), le système Surtrac, pour Scalable Urban Traffic Control, adapte en temps réel les feux de circulation en fonction du trafic : les automobilistes gagneraient 25 % sur leur trajet quotidien, réduiraient de 40 % le temps passé devant les feux… et les émissions nocives de 20 %.

BOUTON PAUSE ?

Au printemps 2023, le buzz médiatique est venu de deux pétitions demandant une « pause » de l'IA ou comparant ses risques à ceux des pandémies ou de la guerre nucléaire. Des pétitions émanant ou signées par les grands acteurs de l'IA, comme Elon Musk ou le fondateur d'Apple Steve Wozniak. Paradoxal ? Pour certains observateurs, cette demande de pause ou de moratoire n'est pas exempte d'hypocrisie. Elle permettrait à ces docteurs Frankenstein de s'exonérer d'éventuels méfaits de leurs créatures en cas de vrais soucis, comme une cyberattaque d'ampleur, par exemple. *« L'intelligence artificielle a le potentiel d'améliorer à peu près tous les aspects de nos vies, mais elle crée aussi des risques sérieux,* a reconnu Sam Altman, le patron d'OpenAI, créateur de la désormais célèbre interface. *L'une de mes plus grandes peurs, c'est que nous, cette industrie, cette technologie, causions des dommages significatifs à la société. Si cette technologie va dans le mauvais sens, elle peut aller assez loin. (…) Et nous voulons travailler avec le gouvernement pour empêcher que cela ne se produise. »*

Pour Sarah Myers West qui dirige AI Now Institute, un institut de recherche américain qui étudie les implications sociales de l'intelligence artificielle, *« en se centrant sur des risques hypothétiques et à long terme, on éloigne la conversation du débat sur la régulation nécessaire contre les risques que nous voyons aujourd'hui et maintenant »*, notamment les atteintes aux libertés qu'ils peuvent provoquer, les usurpations de droit d'auteur, les erreurs que commettent déjà les logiciels d'IA ou leurs biais discriminatoires. Rappelons qu'à peine lancé en 2016 par Microsoft, Tay « chatbot » censé discuter avec des adolescents sur les réseaux sociaux avait été désactivé après 24 h ! Trop d'utilisateurs avaient pris un malin plaisir à lui faire « apprendre » un discours raciste et complotiste… L'un de ses tweets entrés dans la postérité disait : *« Bush est responsable du 11 septembre et Hitler aurait fait un meilleur boulot que le singe* (NDLR, Comprendre B. Obama) *que nous avons actuellement. Donald Trump est notre seul espoir. »*

Conscient des risques de dérives, OpenAI a installé des filtres de sécurité qui empêche la génération de certains types de réponses sur Chat GPT-4. Le robot est donc, en théorie, incapable de donner, par exemple, des instructions pour fabrication des explosifs, ou des drogues, de créer des contenus racistes ou pornographiques… Mais ces barrières ont été contournées par des chercheurs de l'université Brown (USA) qui ont utilisé des langues rares, comme le Gaélique écossais, la langue zoulou ou guarani. Et d'autres chatbot, comme Dan, sont bien moins regardants.

INTELLIGENCE SOUS CONTRÔLE

Entre risques et bénéfices, régulation et innovation, un équilibre reste donc à trouver et, pour une fois, les pays de l'UE ne sont pas à la traine. Vendredi 2 février 2024, les représentants des États membres ont approuvé une nouvelle législation sur l'intelligence artificielle. Ce texte, qui constitue une première mondiale, sera prochainement voté par le Parlement européen. Baptisé AI Act, il vise à *« garantir que les droits fondamentaux, la démocratie, l'état de droit et la durabilité écologique soient protégés contre l'IA à haut risque, tout en stimulant l'innovation et en faisant de l'Europe un leader dans le domaine »*. Ainsi, certains usages éthiquement indéfendables seront interdits comme la notation sociale, la manipulation du comportement, la surveillance de masse… Les outils présentant le plus de risques (liés à l'éducation, aux ressources humaines, au maintien de l'ordre…) devront, eux, répondre à des exigences renforcées,

IMAGE CREE PAR COPILOT, IA INTEGREE A MICROSOFT

comme des obligations de transparence et/ou la mise en place d'un contrôle humain sur le logiciel. Notons que la France a trainé des pieds pour signer l'accord, arguant que certaines dispositions freineraient le développement de start-up européennes (Mistral AI, LightOn, Hugging Face) qui rêvent de rivaliser avec les américains OpenAI ou Google. Notre gouvernement contestait aussi une disposition visant à protéger le droit d'auteur. Selon l'AI Act, les fabricants devront publier un *« résumé suffisamment détaillé »* des données (textes ou images) utilisées pour l'entraînement de leurs logiciels, afin de permettre aux ayants droit de vérifier de les faire retirer ou de percevoir une rémunération.

Aux États-Unis un décret du président Biden a instauré une série de normes, de mesures en matière de sécurité et de protection de la vie privée : publication des résultats de tests de sûreté qui devront être standardisés, collectés et analysés par une nouvelle agence fédérale, le Bureau pour la Sûreté et de la Sécurité de l'IA. Les États-Unis souhaitent aussi se doter d'une charte d'utilisation *« sûre, éthique et efficace »* de l'IA dans le domaine militaire et pour prévenir les tromperies en ligne, l'État américain compte développer des outils de détection des contenus générés par l'IA permettent d'apposer un label obligatoire *«généré par l'IA »*. On pourrait aussi citer l'adoption, en octobre par les pays du G7, d'un code de bonne conduite à destination des entreprises développant des systèmes d'IA. On le comprend, les États entendent garder la main sur une technologie qui pourrait rendre très vite ceux qui la maîtrisent, bien plus puissants qu'eux-mêmes.

L'IA est une promesse, pleine de risques, certes, mais une promesse quand même. Au service des nations et des citoyens, elle peut être une chance. À chacun de nous de s'en saisir. *« Il s'agit d'un réveil citoyen, rien de moins,* défend Laurence Devillers, professeure à Sorbonne Université et chercheuse au CNRS. *L'IA porte en elle un bouleversement de la société, de la démocratie, de la connaissance de l'Homme, mais aussi de notre propre connaissance de nous-mêmes. En se formant à l'IA, on apprend sur son propre comportement, ses biais cognitifs, on découvre ce qu'elle peut nous apporter pour être plus équilibré, plus heureux, pour passer moins de temps sur des choses inintéressantes ou optimiser certaines tâches »*.

Face aux IA, osons faire le pari de l'intelligence humaine ! ■

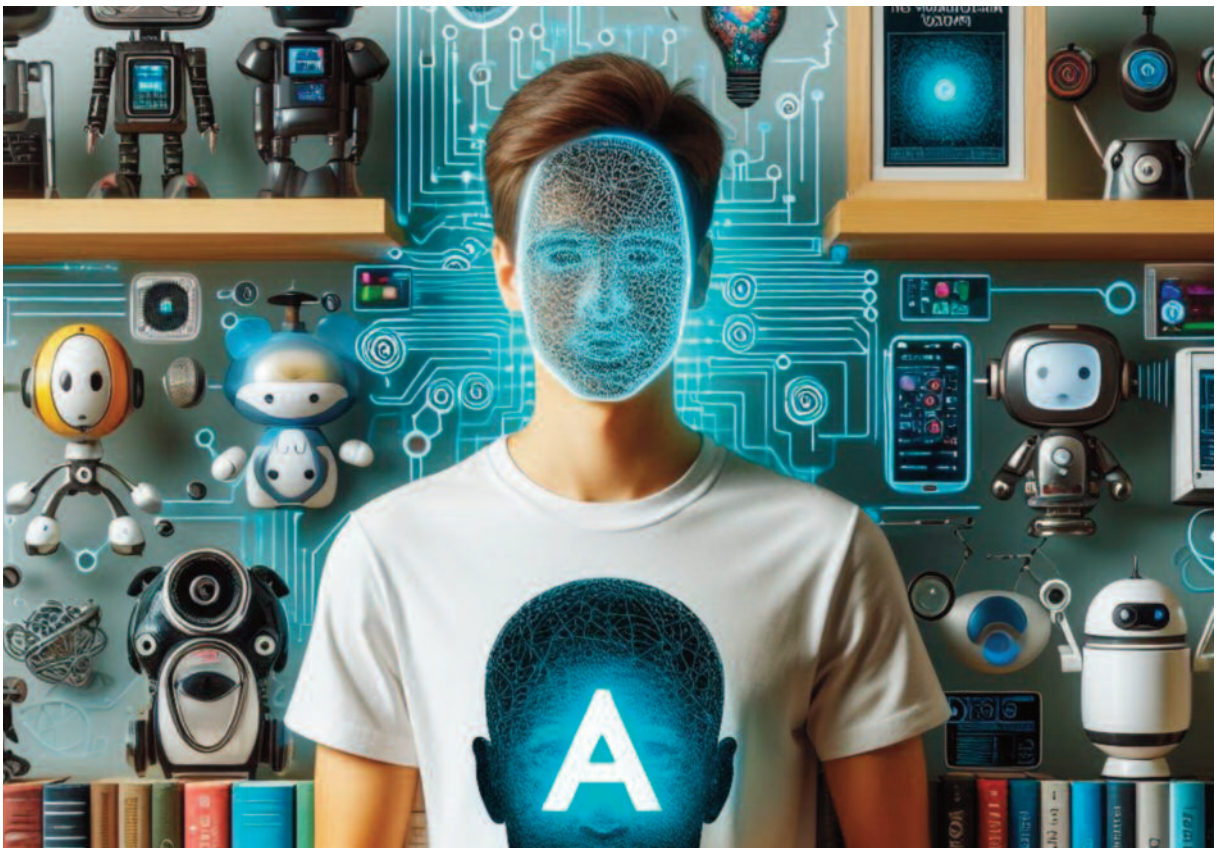


IMAGE CREE PAR COPILOT, IA INTEGREE A MICROSOFT