

Intelligence artificielle et devoir de vigilance

Il y a intérêt à agir

INTÉRÊT
À AGIR



Intelligence artificielle et devoir de vigilance : il y a intérêt à agir !

Impossible de passer une journée sans entendre parler d'intelligence artificielle. Les prouesses de l'IA semblent chaque jour plus spectaculaires que la veille. On peut aujourd'hui « discuter » avec un robot conversationnel, obtenir en temps réel la traduction d'un document ou générer une image ou une vidéo à partir d'un texte. Sur cette partie émergée de l'iceberg s'échafaudent diverses théories : le robot remplacera-t-il l'Homme ? Les emplois vont-ils disparaître ?, etc.

Loin de ces fantasmes, les succès de l'IA se fondent sur la mobilisation intensive de main d'œuvre et de matières premières. Les ruptures technologiques ont été rendues possibles, notamment, par la captation d'une quantité massive de données que les individus ont produites et que des travailleurs doivent aujourd'hui classer pour entraîner les machines. Les puces grâce auxquelles l'IA fonctionne n'existeraient pas sans métaux rares, sans mineurs, sans marins pour les transporter, ni les supercalculateurs sans électricité.

Il faut donc (re)mettre l'IA à son juste niveau : derrière ces deux lettres ne se cache nulle magie mais, comme hier pour le textile ou la manufacture, une chaîne globale de valeur, ce que le débat public a tendance à occulter. La principale réglementation sur le sujet est le Règlement européen « AI Act ». Il permet aux destinataires des services d'IA – consommateurs et travailleurs européens – de s'assurer que les entreprises ne les exposent pas à des risques disproportionnés. Mais *quid* des risques sociaux et environnementaux, notamment pour les populations « du Sud » qui contribuent au fonctionnement de l'IA ?

La loi française sur le devoir de vigilance impose aux plus grandes sociétés françaises d'identifier chaque année les risques sociaux et environnementaux dans leur chaîne de valeur et de prendre des mesures effectives pour les éviter. *Intérêt à Agir* a confronté les plans de vigilance de 11 sociétés françaises aux solutions d'IA qu'elles déclarent publiquement déployer. Il en ressort deux constats. D'une part, l'IA est désormais présente partout, quel que soit le secteur d'activité : commerce, banque, défense, cosmétique, etc. D'autre part, il n'est jamais fait mention de la chaîne de valeur se trouvant en amont de cette technologie. Les discours publics sont plutôt rassurants et même exaltants : grâce à ces sociétés, la population gagnerait en confort, en sécurité, en optimisation. Cela peut être vrai, tout dépendant, *in fine*, de l'utilisation faite de l'IA. Mais à quel prix ?

La première partie du rapport relate ces discours ; la deuxième revient sur les risques sociaux et environnementaux liés à l'élaboration et au fonctionnement de l'IA ; et la dernière porte sur l'application du devoir de vigilance. Le défaut de prise en compte des risques liés à la confection de l'IA dans les plans de vigilance expose les sociétés françaises à des risques d'actions en justice.

Résumé et principales conclusions

Intérêt à agir a analysé la presse généraliste et spécialisée pour découvrir quels sont, précisément, les services d'intelligence artificielle utilisés par les plus grandes entreprises françaises. Il en ressort le constat suivant : toutes les plus grandes entreprises françaises ont recours à l'IA pour des cas d'utilisation d'une diversité insoupçonnée : automatisation des caisses dans la restauration collective, robot conseil en ligne pour faire ses courses dans la grande distribution, diagnostic dermatologique dans les cosmétiques, recherche et développement de nouveaux médicaments, passages piétons "intelligents", etc. Du secteur bancaire au BTP en passant par l'aéronautique et la défense, toutes les sociétés observées ont investi dans l'Intelligence Artificielle.

C'est le premier apport de ce rapport : centraliser dans un document unique les solutions d'IA utilisées à ce jour par plusieurs sociétés soumises au devoir de vigilance. Les informations fournies sont parfaitement publiques puisque diffusées dans la presse.

Mais il y a plus : les conséquences sociales et environnementales de l'élaboration des services d'IA qu'achètent les plus grandes entreprises françaises sont la plupart du temps éludées. Seule la société Téléperformance en fait mention dans son plan de vigilance, pourtant obligatoire depuis 2017 pour les sociétés employant au moins 5 000 salariés en France.

Sur la vingtaine d'entreprises ayant fait l'objet de recherches, 11 sont présentées dans ce rapport, chacune représentative d'un secteur d'activité. Dans leur plan de vigilance, les sociétés Sodexo, Carrefour, L'Oréal, Sanofi, Société Générale, Stellantis, Capgemini, Bouygues construction, Airbus et Thales n'abordent pas du tout les conséquences sociales et environnementales des services d'IA qu'elles achètent, en contrariété apparente avec le devoir de vigilance que la loi française leur impose depuis 2017. Téléperformance a mis en place des mesures de protection de la santé mentale des modérateurs de contenu après avoir fait l'objet d'une enquête en Colombie.

Il faut dire que tout est fait pour occulter ce que l'IA fait aux territoires et aux habitants des pays dans lesquels cette rupture technologique laisse des traces. Le rapport remis par la Commission de l'IA au Président de la République au printemps 2024 réalise l'exploit de présenter la chaîne de valeur de l'IA sans jamais parler des sous-traitants recrutant des millions de travailleurs dans le monde qui contribuent à son élaboration dans des conditions parfois extrêmement difficiles.

C'est là le second apport de ce rapport : rappeler que l'IA cause des dégâts sociaux et environnementaux dans le monde, que les entreprises soumises au devoir de vigilance doivent identifier et évaluer pour chercher à les éviter. Les sociétés soumises au devoir de vigilance portent une responsabilité particulière dans les conséquences sociales et environnementales du développement de l'IA.

A ce jour, les sociétés Sodexo, Carrefour, L'Oréal, Sanofi, Société Générale, Stellantis, Capgemini, Bouygues construction, Airbus, Thales et Téléperformance présentées dans ce rapport s'exposent à des actions en justice pour ne pas prendre suffisamment en considération les risques que les services d'IA qu'elles acquièrent et déploient font peser sur l'environnement et les droits humains de millions de travailleurs.

Il y a donc intérêt à agir, sur le plan juridique pour les sociétés elles-mêmes, sur le plan judiciaire pour les organisations de la société civile et sur le plan politique pour que la "révolution de l'IA" ne se fasse pas au détriment de la planète et des territoires "du Sud".

Table des matières

1

Le recours à l'IA : une pratique banalisée

Les solutions d'IA développées et/ou acquises par les plus grandes sociétés françaises sont systématiquement présentées dans la presse comme synonymes de progrès, bien-être, efficacité voire sécurité pour les ménages et entreprises (françaises). Tous les secteurs sont concernés : commerce, chimie, construction, métallurgie, banque, bureaux d'études, télécommunications et même défense. L'IA est donc bien massivement présente dans le quotidien des plus grandes entreprises françaises.

2

L'élaboration de l'IA : un processus risqué

L'IA est pourtant le résultat d'une chaîne de valeur qui s'appuie en définitive sur de la matière première et de la main d'œuvre. Qu'il s'agisse de la collecte de données ou de l'extraction de métaux rares pour faire fonctionner les outils d'IA, les risques de pollution environnementale et d'atteinte aux droits des travailleurs de la donnée et des mineurs sont réels et documentés. Des propos de spécialistes du sujet, recueillis exclusivement pour ce rapport, l'expliquent.

3

Encadrer l'IA : le devoir de vigilance

Dès lors que l'IA est appréhendée pour ce qu'elle est, soit une nouvelle chaîne mondiale de valeur, la question de sa réglementation se pose. Parmi les outils existants, le devoir de vigilance, qui impose aux plus grandes sociétés françaises de s'assurer des conditions de production, est parfaitement applicable aux solutions d'IA. Or, ces risques ne sont pas ou sont à peine mentionnés par les entreprises observées, ce qui les expose à des demandes juridiques.

1

Le recours à l'IA : une pratique banalisée



Fréquence d'utilisation des restaurants

« Avec le télétravail, nous connaissons de grandes fluctuations de fréquentation qui peuvent passer du simple au double entre un mardi et un vendredi, assure Stéphane Roger, le directeur général de Sodexo services. Nous avons développé une solution d'intelligence artificielle qui analyse les données pour faire des prévisions de fréquentations à quinze jours sur une cinquantaine de restaurants d'entreprise. Nous observons un taux de fiabilité de 90 % ».

Aujourd'hui en France, samedi 17 septembre 2022

Livraison de repas à robot sur les campus américains

« L'intelligence artificielle ouvre sur l'automatisation, explique Aurélia Valot, vice-présidente en charge de la transformation digitale chez Sodexo. En partenariat avec la start-up Kiwibot, une solution de véhicules autonomes qui circulent dans un périmètre de 2 km autour d'un site, a été testée au sein de l'université de Denver (États-Unis). L'idée est d'offrir aux étudiants, aux enseignants et au personnel du campus un service de livraison de repas à l'endroit de leur choix. Un service facturé 2 \$ (1,69 E). Nous observons un taux de fiabilité de 90 % ».

Néorestaurant, mercredi 1 septembre 2021

Automatisation des caisses

« Après avoir choisi les plats de son déjeuner, le client se rend sur une borne équipée d'une webcam qui en capture une image. Le système analyse le contenu du plateau, puis l'envoie au logiciel de caisse qui calcule instantanément le prix des aliments choisis. Le client règle ensuite son repas. C'est grâce à une technologie de deep learning que le dispositif est aussi performant. Les produits sont reconnus car la base de données contient déjà des photos de ces produits ».

L'Usine Nouvelle, vendredi 28 février 2020

Robot nettoyeur dans les hôpitaux anglais

« L'équipe de nettoyage de Sodexo à l'hôpital Stoke Mandeville a une nouvelle recrue de la variété robot qui a été nommée Chloé par les jeunes patients du service pédiatrique.

Chloé aidera l'équipe qui veille à ce que l'hôpital Stoke Mandeville reste un espace propre et sûr pour tous en s'assurant que les sols des zones communes, des couloirs, des entrées et des services pédiatriques restent impeccables, tandis que ses collègues humains se concentrent sur l'assainissement des points de contact dans l'hôpital, ce qui est essentiel pour prévenir la propagation d'infections telles que Covid-19 ».

<https://uk.sodexo.com/media/news-room/chloe-cleaning-robot-stoke-mandeville.html>

(traduction)



Magasin 100% autonome

« Carrefour ouvre aujourd'hui en Belgique un nouveau concept de magasin connecté. Baptisé BuyBye, il est entièrement autonome grâce à l'intelligence artificielle ».

Linéaires (site web), 12 décembre 2023

Plateforme « Marketplace »

« ChatGPT a décidément la cote au sein du secteur de l'e-commerce. Ce mardi, c'est au tour de la start-up Mirakl, qui édite des sites de ventes pour le compte de marques comme Carrefour de s'appuyer sur le célèbre logiciel d'intelligence artificielle générative du groupe de Sam Altman. Ce dernier va permettre, entre autres, d'enrichir de façon automatisée les "fiches produits" vendues sur les places de marché de marques tierces créées par Mirakl ».

Le Figaro Économie, mardi 13 juin 2023

Vérification de la disponibilité des produits en magasin

« Le français Belive.ai installe des caméras en hauteur afin de vérifier en temps réel si les produits sont disponibles et étiquetés, s'il y a du désordre, si les fruits ou les légumes sont abîmés... "Nous sommes là pour corriger tous les irritants pour le consommateur", explique Grégoire Provost, qui travaille dans la petite entreprise créée il y a deux ans et s'apprête à déployer sa technologie dans 80 supermarchés Carrefour. "Il suffit de trois ou quatre semaines pour entraîner l'intelligence artificielle", poursuit-il, pour effacer la corvée quotidienne du relevé manuel des rayonnages [...]. ».

Les Echos, vendredi 20 janvier 2023

Développement de solutions de machine learning

« Carrefour ne se contente pas d'être le premier distributeur étranger à implanter des magasins en Israël. Le groupe annonce le 10 mai 2023 avoir construit des partenariats avec six start-up. Le retailer va effectuer dans les prochains mois des tests avec ces jeunes pousses, dont LSA vous tire le portrait. AI21Labs : spécialisée dans le traitement du langage naturel, cette société fondée en 2017 a développé Jurassic, un outil d'intelligence artificielle capable de générer automatiquement du texte, comme le fait ChatGPT d'Open AI. La version française de cet outil est disponible depuis mars 2023. "Un test sera mené en France sur l'analyse et la labellisation des verbatims clients relatifs à la navigation sur Carrefour.fr et en magasin, ainsi que sur l'analyse de réponses apportées aux requêtes des clients Carrefour", indique la compagnie dans un communiqué ».

LSA, jeudi 11 mai 2023

Surveillance en temps réel des magasins

« Carrefour [...] se sont lancés fièrement dans l'expérience : pour lutter contre les vols en magasin, une intelligence artificielle scrute les comportements de leurs clients et salariés, via la vidéosurveillance, et lance l'alerte auprès du service de sécurité en cas d'agissement équivoque. En tout cas, telle est la promesse de start-up françaises - elles sont au moins trois - qui leur vendent ces solutions de "repérage des gestes suspects en temps réel", y compris ceux du personnel, car 35 % des larcins seraient commis en interne, précise l'une d'elles ».

I'Humanité, vendredi 9 décembre 2022

Liste de courses

« L'intelligence artificielle vient simplifier la vie des clients. Carrefour s'est associé le 8 juin dernier à ChatGPT pour proposer un service d'assistance aux courses. Un chatbot doté d'une intelligence artificielle est désormais disponible sur le site Carrefour.fr, pour proposer des recettes et composer des paniers de courses en fonction des contraintes alimentaires (végan, absence de porc...), des goûts et du budget des clients. Le robot est encore en phase d'apprentissage, et demande pour l'instant "l'indulgence des clients" ».

Le Figaro, jeudi 22 juin 2023

Campagne marketing ciblée

« Carrefour expérimente depuis une semaine un outil de Google Cloud baptisé Vertex AI, dopé à l'intelligence artificielle générative, qui permet de créer en quelques clics des campagnes de marketing ciblées et omnicanales. Les scénarios développés peuvent être utilisés sur les réseaux sociaux, mais aussi en presse écrite, à la télévision ou à la radio. Pour le moment, ces pubs n'ont pas été réellement diffusées ».

LSA, lundi 19 juin 2023

L'ORÉAL

Diagnostic dermatologique

« [...] La Roche Posay [Groupe L'Oréal] pour sa part propose Spotscan, un outil de diagnostic dermatologique grâce à l'intelligence artificielle gratuit pour les personnes atteintes d'acné, et disponible sur le site de la marque. L'acné touche 80 % des adolescents et 40 % des adultes. À partir de trois selfies, Spotscan évalue les imperfections, grâce à une banque de 50 000 photos de tous types de peaux et sur la base d'une échelle de sévérité, le score GEA (Global Acne Severity scale), reconnu par les dermatologues [...] ». www.vital.com/loreal-accelere-sur-les-

<https://www.larevuedudigital.com/oreal-accelere-sur-les-terminaux-dedies-pour-la-beaute-personnalisee/>

Diagnostic capillaire

« [...] L'Oréal propose l'outil K-Scan avec Kerastase. Il s'agit d'une caméra intelligente utilisant l'intelligence artificielle afin de scanner les cheveux et le cuir chevelu, établir un diagnostic précis et en suivre l'évolution ». L'Oréal sur-
www.loreal-paris.com

<https://www.larevuedudigital.com/oreal-accelere-sa-les-terminaux-dedies-pour-la-beaute-personnalisee/>

Analyse des signes de l'âge

« "Tout commence également par un selfie. Ensuite, les algorithmes analysent les sept signes de l'âge déterminants (éclat, pores, ridules, fermeté, rides sous les yeux, taches brunes, rides profondes) pour mettre en lumière les forces et faiblesses de la peau et déterminer un programme ciblé avec les références de la marque. vichy.fr" [...] Pour établir le diagnostic, l'algorithme au cœur de la technologie SkinConsult a été entraîné à l'aide de 6 000 images cliniques issues des évaluations et des études menées par les équipes de L'Oréal Recherche et Innovation et publié dans les Atlas du Vieillissement Cutané [...] ». *LSA, mercredi 11*

LSA, mercredi 20 février 2019

IA générative

« L'Oréal lance également son "Beauty Content Lab, CreAITech", un laboratoire interne doté d'outils avancés basés sur l'intelligence artificielle générative dédié à la production du contenu du groupe. Ce "laboratoire" permet à L'Oréal de faciliter la création d'images en interne pour ses 37 marques de beauté, et en parallèle, de former ses équipes marketing aux dernières technologies créatives. Au cours des huit derniers mois, le groupe a testé plus de 20 technologies d'intelligence artificielle générative et mené des dizaines d'ateliers avec ses marques pour créer plus de 1 000 images beauté ».

Les Echos, 24 mai 2024

Essayaages virtuels

« Tester en temps réel et en 3D une coloration ou une coupe de cheveux avec un effet bluffant de réalité ? C'est possible via un simple selfie sur l'application Style My Hair. Développée par la start-up ModiFace, pilier de la Digital Services Factory de L'Oréal, la plateforme a élaboré son intelligence artificielle sur le tracking et le diagnostic capillaire de 220 000 images faciales. A la clé : une expérience beauté innovante et personnalisée, déjà expérimentée avec le maquillage [...] ».

Paris Match, mardi 17 juillet 2018

Recherche de nouveaux médicaments

« L'israélien CytoReason a annoncé hier soir la signature d'un accord de licence avec Sanofi portant sur l'utilisation de ses outils d'intelligence artificielle en vue de découvrir de nouveaux médicaments. Ce projet vise à exploiter les modèles cellulaires d'intelligence artificielle de CytoReason afin d'identifier des sous-types de patients dans le traitement des maladies inflammatoires de l'intestin. Selon les termes de cet accord pluriannuel, CytoReason percevra plusieurs millions de dollars de la part de Sanofi [...] ».

ABC Bourse, mardi 24 janvier 2023

Fabrication et approvisionnement de médicaments

« [...] L'an passé, le groupe a mis la main sur Amunix Pharmaceuticals, jeune pousse de San Francisco spécialisée dans l'immuno-oncologie, qui s'appuie sur l'IA pour fournir des médicaments ciblés, activables uniquement dans le microenvironnement tumoral de manière à épargner les dommages collatéraux sur les tissus sains.

La même année, Sanofi s'est associée à la biotech britannique Exscientia, qui s'appuie sur l'IA pour découvrir de nouveaux traitements contre le cancer et les maladies auto-immunes.

Depuis 2021, Sanofi a également conclu trois partenariats : avec Owkin, dont la plateforme pilotée par l'IA s'appuie sur les données issues de partenariats avec de multiples centres hospitaliers universitaires pour prédire les réactions des patients aux traitements ; ainsi qu'avec Insilico Medicine et Atomwise, afin d'accélérer le développement de médicaments à l'aide de leurs plateformes respectives basées sur l'IA ».

La Tribune, mercredi 12 juillet 2023

Gestion de données

« SANOFI a également annoncé le déploiement à grande échelle de son application plai, développée en collaboration avec la plateforme d'intelligence artificielle Aily Labs. Une couche supplémentaire en matière de gestion des données qui s'adresse aux cadres de l'entreprise pour leur permettre d'optimiser le pilotage de cette dernière et de prendre des décisions éclairées [...] ».

L'Usine Nouvelle, jeudi 10 novembre 2022

Développement de médicaments

« [...] le groupe s'appuie également sur le traitement des données pour déployer des algorithmes d'intelligence artificielle afin de réduire les temps de recherche. Sanofi recourt ainsi à des algorithmes d'apprentissage automatique et à de grands modèles linguistiques pour gagner en efficacité très en amont de la chaîne de développement des médicaments ».

La Tribune, mercredi 12 juillet 2023

Collaboration avec OpenAI à l'origine de ChatGPT

« Sanofi se renforce dans l'intelligence artificielle. Le laboratoire pharmaceutique français a annoncé, mardi, une collaboration avec la société californienne OpenAI, à l'origine du célèbre robot conversationnel ChatGPT ».

Les Echos, 22 mai 2024

Découverts

« [...] l'IA est très efficace dans la prédiction des dépassements de découverts autorisés. Ainsi, plutôt que d'appeler chaque client concerné pour lui demander de régulariser sa situation, et au passage risquer des frictions dans la relation, l'algorithme sélectionne les cas où les clients ont un réel besoin d'accompagnement, ce qui fait gagner un temps considérable aux conseillers [...] ». *L'AGEFI Hebdo, jeudi 22 septembre 2022*

Gestion de patrimoine

« [...] Parmi les cas d'application de l'IA [de la Société Générale], figure notamment le programme Synoé, développé pour les clients de la banque privée en France. Il offre une solution automatique de conseil en placement, qui se base sur les objectifs et les connaissances financières des clients [...] ». *Les Echos, jeudi 25 novembre 2021*

Chatbots et Callbots

« Les techniques d'IA sont également mises à profit [au sein de la Société Générale] pour faire fonctionner les "chatbots" et les "callbots" qui permettent de répondre aux questions basiques des clients et qui traitent environ 15.000 conversations par jour. L'IA vient aussi épauler les équipes de risque et de conformité, avec des logiciels d'analyse qui peuvent traiter des volumes très importants de données et repérer les éventuelles situations anormales [...] ». *Les Echos, jeudi 25 novembre 2021*

Détection de malversation

« Du côté des banques françaises, Société Générale s'est illustrée en lançant dès 2015 un vaste projet d'IA visant à détecter les malversations touchant tout type de virement : instantané, de masse, d'open banking (via l'API DSP2), Paylib, etc. Baptisée Mosaic, pour More Security with Artificial Intelligence, la plateforme en question est mise au service des clients particuliers, professionnels et entreprises, à la fois côté Société Générale et côté Crédit du Nord [...] ». *Le Journal du Net, lundi 2 mai 2022*

Gestion de la fraude

« [...] Mais les projets de Société Générale ne sont pas uniquement à destination directe de ses clients puisqu'une large partie (29%) est dédiée à l'efficacité opérationnelle. "Le service CAST (compliance analytics surveillance tool, ndlr) analyse des communications des employés régulés dans 16 pays, les retranscrit et peut les traduire en 26 langues", décrit Claire Calmejane. L'intelligence artificielle est également utilisée pour lutter contre les escroqueries : "les solutions Filae et Defi aident à gérer la fraude dans le monde digital", qu'il s'agisse de lutte anti-blanchiment pour la première, ou de paiements illicites pour la seconde (chèques falsifiés, virements frauduleux...) [...] ». *mind Fintech, mercredi 24 novembre 2021*



ChatGPT

« DS Automobiles, la marque premium du groupe Stellantis, a intégré le système d'intelligence artificielle générative ChatGPT sur l'ensemble de sa gamme, DS 3, DS 4, DS 7 et la très luxueuse DS 9. La voiture devient capable de vous répondre. Et même de tenir la conversation ».

Aujourd'hui en France, 20 décembre 2023

Conduite autonome

« Stellantis accélère ses recherches en conduite autonome avec l'acquisition d'aiMotive, une start-up spécialisée dans l'intelligence artificielle et la conduite autonome. Cette opération permet au constructeur de consolider ses technologies fondamentales d'intelligence artificielle et de conduite autonome, d'élargir son vivier de talents internationaux et de dynamiser le développement à moyen terme de sa nouvelle plateforme STLA AutoDrive [...] ».

Boursier, jeudi 17 novembre 2022

Aide à la conduite

« L'annonce de l'acquisition de CloudMade par Stellantis, confirme l'arrivée en force des technologies d'intelligence artificielle dans l'analyse des données des conducteurs et la mise en place de services personnalisés embarqués ».

Auto Infos (site web), 30 janvier 2024

Maintenance assistée

« Dans la technologie, Stellantis a investi dans Nauto, une société qui conjugue intelligence artificielle, vision par ordinateur et analyses de données de risques. Ses autres investissements concernent Viaduct, une plateforme d'intelligence artificielle conçue pour améliorer l'analyse de la qualité et de la maintenance des véhicules [...] ».

ABC Bourse, jeudi 15 juin 2023

Service après-vente

« Stellantis introduit le marketing prédictif dans l'après-vente. L'intelligence artificielle indique au concessionnaire le moment opportun pour inviter son client à passer à l'atelier ou à renouveler son auto. Un savoir-faire 100 % français, qui vise à lutter contre le gaspillage publicitaire [...] ».

Challenges, vendredi 15 avril 2022

Services innovants et personnalisés au sein des véhicules

« [...] STLA SmartCockpit, supporté par la coentreprise Mobile Drive entre Stellantis et Foxconn utilisera des applications basées sur l'intelligence artificielle pour délivrer des fonctions et des services comme la navigation, l'assistance vocale, l'e-commerce et des services de paiement. La plateforme utilisera des produits et des solutions Amazon qui seront spécialement conçus pour les véhicules de Stellantis afin de créer des expériences à bord personnalisées, intuitives et différencierées selon les marques [...] ».

<https://www.larevuedudigital.com/stellantis-sur-les-rails-pour-20-milliards-deuros-de-chiffre-d'affaires-en-logiciel/>.

Conduite de changement

« [...] l'éditeur de la 1ère plateforme SaaS de conduite du changement basée sur l'intelligence artificielle InsideBoard annonce une deuxième augmentation de capital de plus de 20 M. [...], sa solution a déjà convaincu une trentaine de grands noms tricolores tels que PSA [groupe Stellantis] ».

Capital Finance, mercredi 29 avril 2020

Cartographier la végétation

« Le groupe français Eramet utilise des drones et des technologies d'IA fournies par Capgemini pour cartographier la végétation initialement présente sur la zone minière qu'elle exploite au Sénégal. Cela devrait lui permettre de réhabiliter plus facilement l'environnement qu'elle devra restituer intact au gouvernement local à la fin de son contrat d'exploitation [...] ».

L'Usine Nouvelle, mardi 18 avril 2023

Repérer la fraude

« La direction générale des finances publiques (DGFIP) semble particulièrement fière de son IA, commandée à Capgemini et Google, capable de repérer les piscines non déclarées sur des photos satellites. Elle devrait même permettre de récupérer plusieurs millions d'euros en impôt foncier, selon les affirmations du directeur de l'administration, la semaine passée, dans Ouest France. [...] Bercy souhaite aussi que, à terme, une autre IA, en cours d'expérimentation, mette automatiquement à jour les plans cadastraux. Plutôt que d'aller sur le terrain, le géomètre se retrouve à choisir sur son écran, parmi six propositions, celle qui correspond le mieux à la forme de l'édifice nouvellement bâti [...] ».

L'Humanité, mardi 25 avril 2023

Création d'une intelligence artificielle générative

« [...] l'annonce par Capgemini de l'extension de son partenariat stratégique avec Google Cloud – la division de Google spécialisée dans l'informatique dématérialisée (cloud), l'analyse de données et l'intelligence artificielle (IA) – a été reçue cinq sur cinq par le

Marché. [...]

Concrètement, les deux partenaires vont créer un centre d'excellence mondial Google Cloud dédié à l'IA générative afin d' « aider les entreprises à exploiter pleinement le potentiel des technologies d'IA », ont-ils indiqué dans un communiqué. Capgemini entend utiliser l'IA générative de Google Cloud pour développer dans les vingt-quatre prochains mois une bibliothèque riche de plus de 500 cas pratiques d'utilisation, dont les 100 premiers se concentreront sur les secteurs des services financiers, de l'assurance, de la vente au détail et de l'automobile ».

L'AGEFI Quotidien, jeudi 1 juin 2023

Offre de technologies d'intelligence artificielle

« Capgemini et le ministère de la Défense (MOD) britannique annoncent la signature d'un contrat de cinq ans pour la gestion du centre de services IT du ministère [...]

Le centre reposera sur les capacités de pointe de Capgemini en matière d'intelligence artificielle et d'analyse intelligente, pour offrir plus de fonctionnalités en libre-service ainsi qu'une meilleure prise en main pour ses usagers ».

L'AGEFI Quotidien, vendredi 19 juin 2020

Offre d'annotation de données

« Les services d'étiquetage des données de Capgemini permettent des opérations d'étiquetage et d'annotation des données évolutives et de bout en bout qui améliorent l'agilité et les résultats commerciaux de votre organisation. [...]

Pourquoi Capgemini ?

Capgemini, l'un des principaux fournisseurs de technologies au monde, [...]

Cette combinaison de technologies, de méthodologies et de ressources mondiales évolutives d'annotateurs de données dédiés nous permet de travailler avec vous de manière véritablement collaborative afin d'apporter une valeur accélérée et de minimiser les risques pour votre entreprise à l'ère du numérique [...] ».

Signalisation intelligente

« Le passage piéton à détection de présence fonctionne grâce à la mise en place de nombreux capteurs intelligents de notre [établissement Flowell de la société Colas*] partenaire Starling Technologies. A travers une Intelligence Artificielle et une analyse automatique des données, Starling Technologies mesure :

- o Le nombre de véhicules et de piétons
- o Le pourcentage de piétons respectant le marquage (traversée sur les bandes de passage piétons)
- o La distance d'arrêt des véhicules par rapport au passage piéton
- o Le pourcentage de véhicules qui respectent la priorité piétonne [...] ».

<https://www.flowellbycolas.com/realisations/passage-pieton-portsmouth.html>

*Colas SA est une entreprise du groupe Bouygues

Maintenance de l'infrastructure routière

« Cibler et prioriser les opérations de maintenance pour mieux prévenir les accidents et optimiser les investissements, tel est l'objectif que se fixe Mobility by Colas au travers du développement de deux solutions complémentaires. La plus

aboutie, Anaïs, a été lancée il y a quatre ans. Cette plateforme, adoptée par les départements du Cher et d'Eure-et-Loir, utilise les données relatives aux comportements de conduite pour surveiller l'état des routes. [...] Ce dispositif peut désormais être enrichi du système Infracare. Il s'agit cette fois d'utiliser les images captées par une caméra GoPro installée dans un véhicule. "Traitée dans le cloud par nos algorithmes et une intelligence artificielle, le film sert à produire un tableau de bord dans lequel non seulement les dégradations sont identifiées, mais aussi leurs évolutions futures", souligne Laurent Le Boulc'h, directeur général adjoint infrastructures durables et mobilités intelligentes chez Colas [...] ».

Le Moniteur des Travaux Publics et du Bâtiment, vendredi 10 juin 2022

Détection de réseaux enterrés

« [...] Spécialisée dans l'imagerie de sub-surface non intrusive, Exodigo utilise une technologie basée sur les multicapteurs et l'intelligence artificielle pour identifier les réseaux enterrés. Identifier 30 ouvrages enterrés contre 19 sur les plans [...] Identifier 30 ouvrages enterrés contre 19 sur les plans [...] Exodigo affiche également Colas Rail parmi ses références françaises [...] ».

Le Moniteur, mercredi 12 juillet 2023

Analyse des sols pollués

« Utiliser l'intelligence artificielle pour accélérer la dépollution des friches industrielles. C'est la solution innovante proposée par la start-up Tellux, [qui] permet d'identifier des objets mais aussi d'obtenir des indications sur la quantité (ou la concentration) des constituants présents.

[...] Intervention chez Colas : [la solution] permet d'obtenir l'information en temps réel sur site et de caractériser des centaines d'échantillons ».

L'Usine Nouvelle, lundi 10 juillet 2023

Retards de pièces dans la supply chain

« La filiale d'Airbus, Stelia Aerospace, vient d'annoncer un plan d'investissement de 110 millions d'euros sur quatre ans. Stelia anticipe un usage accru de l'intelligence artificielle et du big data pour optimiser la supply chain. "Nous faisons déjà appel à l'IA et la data pour contrôler les éventuelles pièces manquantes avec nos sous-traitants. Nous comparons le rythme de production en temps réel avec nos données passées sur ces mêmes sous-traitants. Il est possible de prédire les pièces manquantes avant même la fin de la production" ».

La Tribune, mardi 16 mars 2021

Vols autonomes

« En janvier 2020, le géant européen de l'aéronautique a réussi à faire décoller et atterrir un A350 automatiquement. Puis en janvier 2023, il a dévoilé le démonstrateur DragonFly, intégré dans le cockpit et développé par sa filiale Airbus UpNext. Truffé d'algorithmes intelligents et de capteurs, DragonFly a été testé en conditions réelles plusieurs mois en 2022. Le logiciel a piloté un A350 sur le tarmac jusqu'au décollage. Puis a dérouté l'avion, après simulation d'un accident, en générant un nouvel itinéraire communiqué au contrôle aérien, et s'est posé tout seul ».

Le Figaro, 22 octobre 2023

Traitements des appels d'urgence

« HighWind a développé une solution qui vise à réduire le temps [des appels d'urgence] en étant plus efficace dans la communication. [...] à travers le déploiement d'une intelligence artificielle, brevetée, capable de fournir un diagnostic de la situation d'urgence sur la base des photos envoyées. Après avoir signé un partenariat avec Airbus Développement [...], la deeptech entame une série de négociations techniques et commerciales pour intégrer sa solution chez Airbus dans le cadre de leur politique contre les accidents du travail [...] ».

La Tribune, mardi 1 novembre 2022

Routeurs

« Collins Aerospace s'est imposé comme le fournisseur leader des routeurs Fomax, une sorte de box internet pour les avions qui permet de collecter les données de vol utilisées par la plateforme de maintenance prédictive Skywise d'Airbus. [...] Les routeurs Fomax sont utilisés également pour collecter les informations lors du vol et les envoyer à des serveurs gérés, soit par les constructeurs (Airbus) soit par la compagnie, qui vont pouvoir faire dérouler des algorithmes d'intelligence artificielle en vue d'offrir tout un tas de services pour améliorer les vols et les coûts d'opération [...] ».

La Tribune, lundi 27 septembre 2021

Ravitaillement

« Le constructeur aéronautique européen Airbus cherche à automatiser le ravitaillement de drones en plein vol grâce à l'intelligence artificielle. Les essais qui viennent d'être menés, lors desquels le guidage et le contrôle autonome d'un drone ont été réalisés, s'inscrivent dans le cadre du programme Auto'Mat visant à tester en vol les technologies clés de demain [...] ».

L'Usine Nouvelle, mardi 28 mars 2023

Gestion de situations militaires

« Si l'IA est déjà utilisée par les armées, l'Otan s'apprête à tester un "outil" inédit par sa puissance et sa capacité d'analyse, de déduction, et de préconisation lors d'un exercice de grande taille. Cela, avec Anticipate, le produit développé par Thales et ses partenaires, notamment le département de recherche de l'US Air Force et l'École nationale de

Anticipate intègre au sein de la plateforme d'analyse de big data Kast de Thales, la technologie d'IA développée par la start-up française NukkAI, pionnière de l'intelligence artificielle de nouvelle génération. Le groupe de défense a formalisé son partenariat en annonçant, ce jeudi, la signature d'un premier contrat avec la jeune poussée [...] ».

Le Figaro, vendredi 13 janvier 2023

Cartographie militaire

« Thales et LuxCarta, entreprise spécialisée dans les produits géospatiaux dérivés de l'imagerie spatiale, s'associent pour proposer des solutions de renseignement et de cartographie militaire optimisées grâce à l'Intelligence Artificielle (IA). Cet accord renforce le partenariat informel déjà en place entre les deux sociétés depuis 2020.

En planification militaire, il est capital d'obtenir rapidement des données géospatiales précises, qui facilitent la prise de décisions opérationnelles [...]. Pour le renseignement et la cartographie militaire, l'IA est utile en particulier pour la photo-interprétation. Elle détecte, interprète et numérise automatiquement les bâtiments, les réseaux routiers, les obstacles, les véhicules, etc. [...] ».

<https://www.zonebourse.com/cours/action/THALES-4715/actualite/Thales-et-LuxCarta-signent-un-partenariat-pour-proposer-des-solutions-de-cartographie-et-de-reseignement-43068713/>

Plateforme de traitement de données massives

« L'intelligence artificielle pour analyser des quantités massives de données sensibles : les groupes français Atos (services informatiques) et Thales (électronique) ont annoncé hier la création d'une société commune, Athea, qui développera une plateforme pour les secteurs de la défense, du renseignement et de la sécurité intérieure. Cette plateforme s'adressera "tant aux acteurs publics que privés" et se développera d'abord sur "le marché français puis, à terme, les marchés européens", précise Thales. Appelée à devenir "le champion européen du Big Data et de l'Intelligence Artificielle pour la défense et la sécurité", elle associera donc "traitement de données massives et intelligence artificielle" [...] ».

La Correspondance économique, vendredi 28 mai 2021

Solution de détection d'attaques informatiques

« Le Groupe a poursuivi la commercialisation de Cybel Security, sa plateforme de détection d'attaques avancée basée sur l'Intelligence Artificielle [...] ».

Document d'Enregistrement Universel de 2022 de Thales, p. 27

Modération de contenu pour TikTok

« Leader mondial des centres d'appels, son chiffre d'affaires était de 7,1 milliards d'euros en 2021. Ces dernières années, la société a diversifié ses activités. Elle opère également la modération de certains réseaux sociaux, dont TikTok, qui affirme employer 687 personnes pour analyser et censurer au besoin les contenus francophones publiés sur la plateforme ». *Liberation, 15 janvier 2024*

Enquête pour violation des droits des modérateurs de contenu en Colombie

« Le gestionnaire de centres d'appels fait l'objet d'une enquête de la part du ministère colombien du Travail pour violation de droits des syndicats et sur les conditions de travail des employés en charge de la modération du réseau TikTok ». *Investir, 10 novembre 2022*

Modération de contenu pour Meta (Facebook)

« Certains fournisseurs externes qui gèrent le contenu des plateformes de Meta, propriétaire de Facebook, ont été affectés par la panne technologique mondiale qui a paralysé les aéroports, les banques et les hôpitaux vendredi. Comme la plupart des entreprises de médias sociaux, Meta s'appuie sur un mélange d'intelligence artificielle et d'examen humain pour modérer les milliards de messages postés sur ses plateformes, qui comprennent également Instagram, WhatsApp et Threads.

Une partie de cet examen humain est effectué par le personnel de Meta, mais la plus grande partie est sous-traitée à des fournisseurs de services commerciaux employant des armées de travailleurs mal payés qui évaluent si les messages contiennent des discours haineux, de la violence et d'autres violations des règles de l'entreprise. Les principaux fournisseurs concernés sont Teleperformance ». *Boursorama, 20 juillet 2024*

Leader dans la modération de contenus

« Teleperformance, basée à Paris, est devenue l'un des leaders du marché des services de modération de contenu, avec plus de 7 000 travailleurs de ce type dans le monde ». *Time, 20 octobre 2022 (traduit)*

<https://time.com/6223340/tiktok-content-moderators-latin-america/>

Maintien de l'activité de modération de contenu

« Teleperformance a annoncé qu'il maintenait son offre complète de services de modération de contenu, y compris le segment le plus offensant. En novembre 2022, la société de services aux entreprises avait annoncé son retrait de ce segment, au coeur de "l'affaire colombienne". Le groupe revient sur sa décision à la suite de la revue exhaustive de son activité de modération de contenu avec ses différentes parties prenantes, et notamment ses collaborateurs, ses clients et ses actionnaires ». *Agence Option Finance, 22 mars 2023*

Création d'un ChatGPT interne

« Lors de son point d'activité du premier trimestre, le groupe dirigé par Daniel Julien a assuré que la technologie GPT offrait des «opportunités significatives». La société n'est d'ailleurs pas novice en la matière. Elle travaille avec OpenAI depuis 2021 et a développé son propre produit, «TP GPT», un outil aux capacités saisissantes, selon les données présentées par Teleperformance lors d'une conférence avec des analystes le mois dernier.

Grâce à TP GPT, Teleperformance affirme que la durée d'un appel téléphonique peut être écourtée de près de 40%, tout en réduisant les erreurs de 90% ». *L'AGEFI Quotidien, 31 mai 2023*

2

L'élaboration de l'IA : un processus risqué

La chaîne de valeur de l'IA est structurellement fondée sur trois risques majeurs :



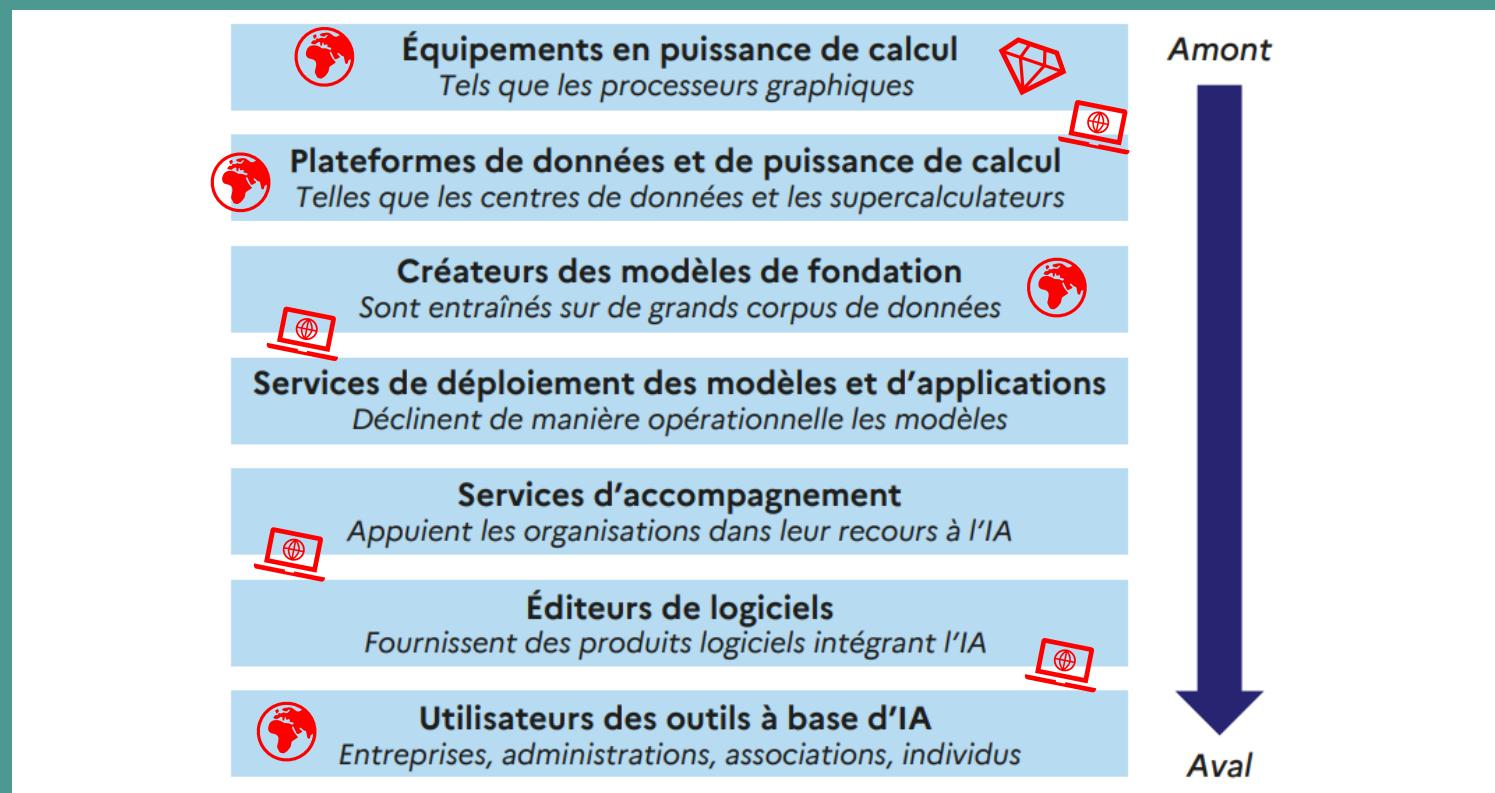
L'atteinte aux droits humains dans le secteur des mines, nécessaire pour extraire les métaux rares utilisés dans la fabrication de puces électroniques puissantes ;



Des pollutions environnementales occasionnées par l'extraction des métaux rares et par la grande consommation d'énergie nécessaire à l'élaboration et au fonctionnement des solutions d'IA.



L'atteinte aux droits des travailleurs de la donnée, nécessaires pour 1) générer les données, 2) annoter les données, 3) vérifier les résultats de l'IA et parfois 4) remplacer l'IA.



Ce schéma est issu du rapport remis au gouvernement en mars 2024 par la Commission de l'Intelligence artificielle, « IA : notre ambition pour la France ». Comme il ne le fait pas, nous y avons ajouté en rouge les risques sociaux et environnementaux liés à l'élaboration de l'IA tout au long de la chaîne de valeur.

Risque 1/3 Métaux rares

L'IA fonctionne grâce à **des puces numériques** qui rendent possible l'apprentissage complexe des machines. Parmi elles, les plus puissantes du marché sont celles vendues par la société américaine NVIDIA. Or, ces puces doivent leur existence à quantité de métaux rares qu'il faut extraire puis transformer. La puce peut alors commencer son cycle de vie : être transportée et utilisée avant d'être détruite.

Tout ceci a un coût environnemental et social documenté depuis des années.

2023, une année record pour Nvidia

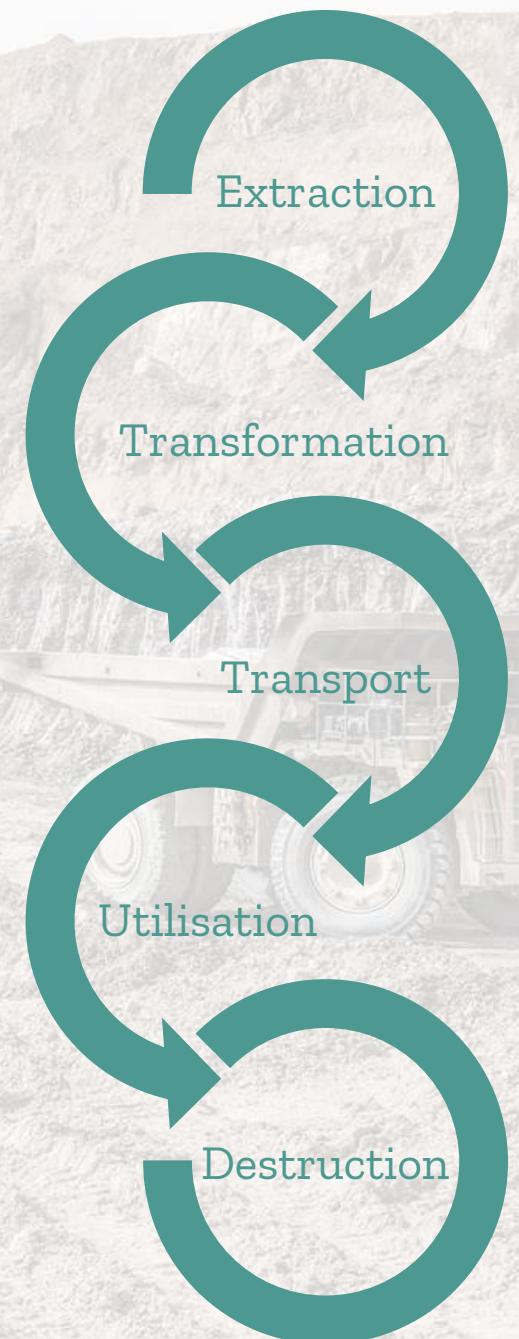
481,68 USD : +447,63 (1314,63%) en 8 ans



Cours de l'action : NVDA (NASDAQ)

Source : ActuIA, 2024, n° 14

Cycle de vie d'une puce



Du Lithium, cobalt, nickel, étain, silicium, pour les batteries, micro-processeurs ou les circuits électriques imprimés..., les minéraux sont indispensables pour accéder à l'intelligence artificielle. Ils sont la source de production des composants électroniques et des technologies de stockage d'énergie nécessaires au fonctionnement des systèmes d'IA, de la collecte des données à leur traitement et à leur utilisation.

Or, l'extraction minière se concentre dans certaines régions du monde. Ainsi, la République Démocratique du Congo (« RDC » ci-après) est le premier lieu, incontesté, de production de cobalt, avec 60% du commerce mondial. De son côté, l'Indonésie représente la moitié de la production mondiale de nickel. Ces pays sont l'objet de fortes pressions internationales, en particulier par la Chine, leader de la production de produits électroniques et de batteries.

Il est important de distinguer deux types de production minière : la production industrielle, caractérisée par ses gros volumes et des machines titaniques d'une part, et la production artisanale caractérisée par sa forte intensité de main d'œuvre d'autre part. Bien que chaque mineur ait des conditions d'exploitation différentes, l'exemple de la RDC est assez représentatif des enjeux humains et environnementaux : les mines industrielles de cuivre et cobalt posent de véritables problèmes d'expropriation des populations locales et de pollution des fleuves environnants ; les mines artisanales, qui embauchent parfois plus de 20.000 « creuseurs », sont caractérisées par des risques humains d'accidents graves et de morts (tunnels sans oxygène, non-étayés, absence d'équipements de protection individuelle comme le casque et les chaussures, etc.).

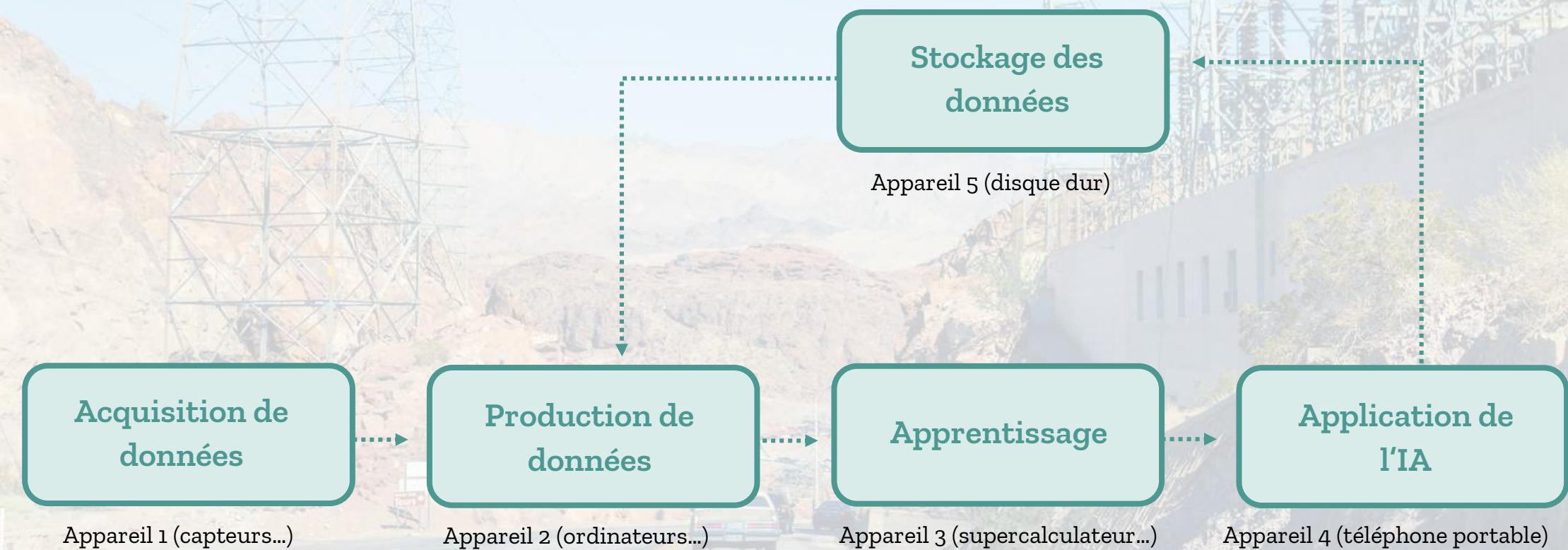
Il existe très peu de statistiques sur les accidents miniers, bien que les médias locaux et internationaux témoignent régulièrement de catastrophes meurtrières. Selon une étude de la Revue Africaine de Médecine et de Santé Publique néanmoins, entre 2015 et 2017, 23% des 3300 mineurs auraient été victimes d'accidents de travail. Les accidents les plus fréquents seraient les chutes et les blessures sur les mains avec des outils contondants. Il s'agit de chiffres qui soulignent l'importance des équipements minimaux de protection, fréquemment manquants, comme les chaussures et les gants dont le coût est souvent à la charge du mineur artisan. Le nombre total de victimes devrait toutefois augmenter au regard de la croissance du secteur.

Fabrice Warneck

Directeur du monitoring à l'ONG Electronics Watch,

organisation de défense des droits des travailleurs de la chaîne de sous-traitance des secteurs de l'électronique et des batteries

L'IA doit son fonctionnement à une **utilisation massive de données**. Or, pour pouvoir être utilisées, ces données doivent être collectées afin d'alimenter un système d'apprentissage machine. L'utilisation d'un service d'IA conduit également à produire des données, qui sont toutes, finalement, stockées dans des *data center* ou centres de données qui doivent être régulièrement refroidis. Selon l'Agence Internationale de l'Energie, la consommation totale d'électricité des centres de données pourrait atteindre plus de 1 000 TWh en 2026. Cette demande équivaut à peu près à la consommation d'électricité du Japon. Microsoft, par exemple, qui est propriétaire de la société OpenAI, créatrice de ChatGPT, déclare détenir 300 data centers dans le monde. Certains se trouvent dans des déserts tels que l'Arizona.



Les modèles d'apprentissage automatique reposent sur des réseaux de neurones profonds composés de plusieurs millions, voire milliards de paramètres. Chacun de ces paramètres doivent être appris. Cela se fait en utilisant des bases de données d'entraînement de très grandes tailles.

D'après les études, l'entraînement du modèle de traitement automatique des langues BLOOM (Big Science Large Open-Science Open-Access Multilingual) a consommé 433 MWh d'électricité tandis que le modèle GPT-3, à l'origine de ChatGPT, en a de son côté consommé 1287 MWh. D'autres études ont également estimé les consommations moyennes par requête lors des tests du modèle et une fois déployé, c'est -à-dire lors de l'inférence. Elles varieraient entre 0.045 et 9 Wh pour des modèles de langage en fonction des hypothèses et consommations considérées. Ces résultats sont dépendants des modèles utilisés, de leur taille, de leur application (texte ou image par exemple), de la fréquence des requêtes ou du matériel sur lequel sont déployés les modèles. Les quantifications précédentes ne prennent en compte que la phase d'usage des équipements informatiques, c'est-à-dire qu'elles ne considèrent que la consommation électrique des serveurs lors de leur utilisation. Il faudrait y ajouter les autres étapes du cycle de vie des équipements (extraction de ressources, fabrication, transport, fin de vie).

Pour Meta, les émissions de gaz à effet de serre liées aux phases d'extraction de ressources et fabrication représenteraient ainsi 30 % des émissions liées au fonctionnement des modèles d'apprentissage automatique, en supposant que les serveurs sont utilisés en moyenne de 30% à 60% de leur capacité maximale et qu'ils ont une durée de vie de 3 à 5 ans. Pour l'entraînement de BLOOM, elles représentent 22 % des émissions.

Ces phases d'entraînement et d'inférence ne représentent en outre qu'une partie des impacts environnementaux du service d'IA. Afin d'en appréhender l'intégralité, il conviendrait d'étudier, d'une part, l'ensemble des impacts sur tout le cycle de vie du service dont le nombre de phases est conséquent : conception, acquisition et collecte des données pour l'entraînement ou le ré-entraînement, traitement et stockage de ces données, évaluation, mise en production, utilisation sur des terminaux utilisateurs, maintenance, fin de vie du programme d'IA ; et d'autre part les impacts liés aux changements profonds, économiques et sociaux permis par ce service tels que des changements dans les habitudes de consommation ou les processus de production.

Enfin, les impacts environnementaux ne concernent pas uniquement la consommation électrique et les émissions de gaz à effet de serre. Une première étude publiée récemment extrapole que l'utilisation de chatGPT pour les recherches Google pourrait nécessiter plus de 500 000 serveurs correspondant à 4 milliards de GPU. La fabrication de ces serveurs et de tous les autres équipements nécessaires au fonctionnement des IA a des impacts notamment en termes de déplétion de ressources, de consommations d'énergie primaires ou d'eau.

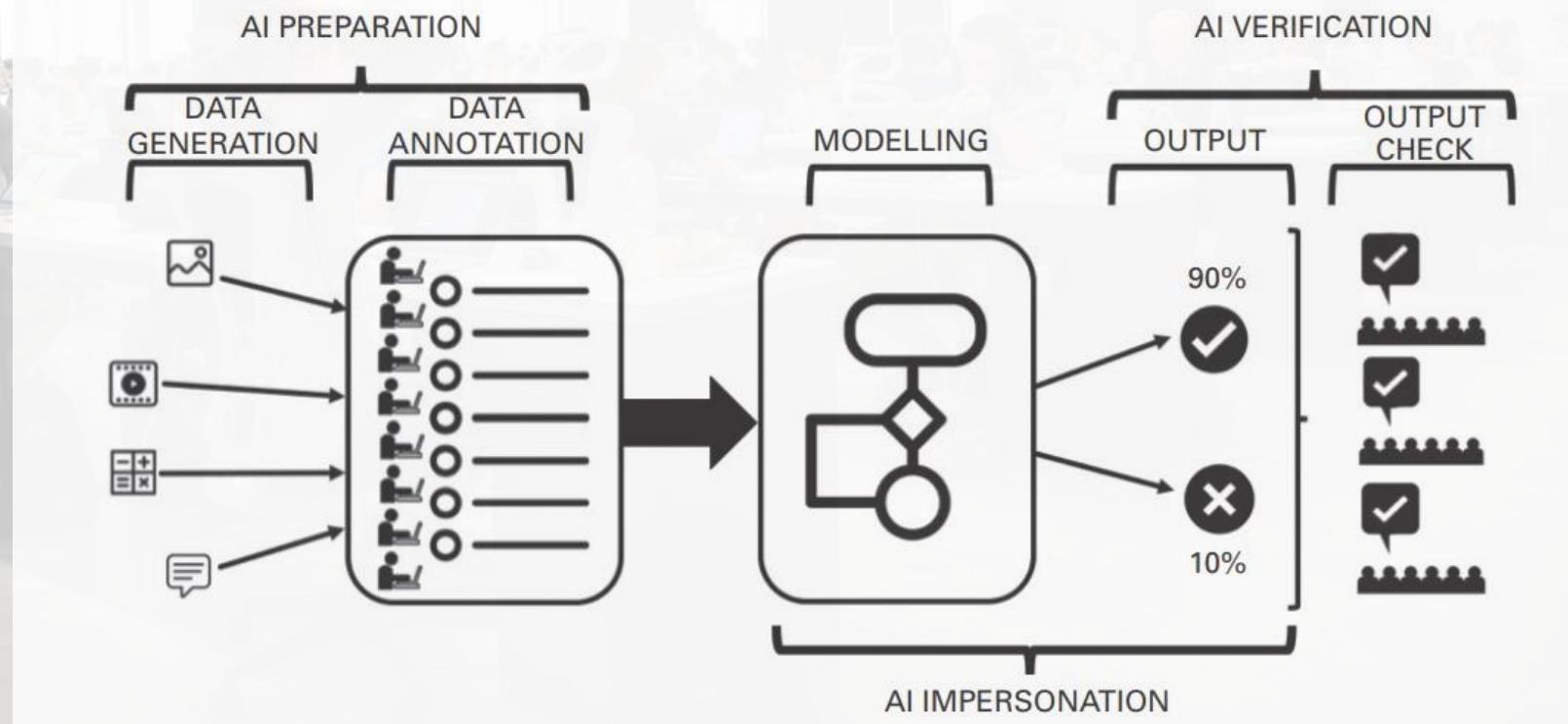
Aurélie Bugeau

**Professeure en informatique à l'Université de Bordeaux
Membre de l'Institut Universitaire de France**

Risque 3/3 Supervision

Le dernier risque lié à la production de l'IA concerne les **travailleurs de la donnée**. Des hommes et des femmes, se trouvant le plus souvent dans des pays dits « du Sud », interviennent à au moins trois moments dans le fonctionnement de l'IA : en amont pour générer et classer la donnée ; en aval pour corriger les erreurs commises par la machine et parfois en direct pour remplacer l'IA. Ces travailleurs peuvent être soumis à des conditions de travail difficiles : longues heures de travail, exposition à des contenus violents, faibles rémunérations et parfois absence de rémunération, discrimination syndicale, éviction des règles protectrices du droit social par le recours au statut du travailleur indépendant voire au secteur informel.

Les travailleurs de la donnée annotent les textes, images, vidéos ou sons pour entraîner un algorithme puis corrigent les erreurs commises par l'IA (l'enceinte qui entend « boisson » au lieu de « poisson » par exemple). Parfois même, ils « font » l'IA en répondant à la place d'un *chatbot* ou en visionnant des vidéos de surveillance en direct.



Source : Paola Tubaro et alii, « The trainer, the verifier, the imitator: Three ways in which human platform workers support artificial intelligence », *Big Data & Society*, 2020

L'intelligence artificielle est fréquemment perçue comme une menace pour l'emploi dans les discussions publiques. Cependant, sa création et son fonctionnement reposent sur une force de travail importante. Si les experts reconnus tels que les ingénieurs et les "data scientists" sont essentiels, le développement de l'IA s'appuie également sur une multitude de travailleurs moins visibles. Ces millions de "data workers" (ouvriers de la donnée), répartis à l'échelle mondiale, jouent un rôle crucial dans l'élaboration et le maintien des systèmes d'intelligence artificielle.

Les entreprises d'IA recrutent ces ouvriers de la donnée pour effectuer des tâches rapides et répétitives comme l'étiquetage d'images ou la transcription de textes. Souvent mal payés et précaires, ils sont majoritairement issus de pays à bas revenus. Cette situation souligne la nécessité d'un devoir de vigilance dans toute la chaîne de valeur de l'IA.

Des enquêtes journalistiques successives ont exposé l'ampleur du travail humain caché derrière l'IA. En 2019, on a appris que des employés distants écoutaient les conversations que des détenteurs de téléphones ou d'enceintes intelligents avaient avec Siri et Alexa. En 2023, le Time Magazine a dévoilé que des centaines de travailleurs au Kenya contribuaient à l'entraînement de ChatGPT. Plus récemment en 2024, il s'est avéré qu'Amazon dissimulait que ses supermarchés "sans caissiers" étaient en fait gérés à distance par des milliers de travailleurs indiens.

Pour donner du sens à la diversité de ces tâches, on peut les catégoriser selon trois phases principales de production de l'IA :

1. Préparation de l'IA : les ouvriers de la donnée génèrent ou filtrent des informations pour constituer les bases de connaissances servant à l'entraînement des IA.
2. Vérification de l'IA : les ouvriers de la donnée contrôlent le fonctionnement des IA.
3. Imitation de l'IA : Dans certains cas, les ouvriers de la donnée interviennent en temps réel pour corriger les erreurs ou effectuer les tâches que l'IA ne parvient pas à réaliser.

L'ampleur de cette main-d'œuvre fait débat. En 2021, l'Oxford Internet Institute estimait à 163 millions le nombre de travailleurs en ligne dans le monde. En 2023, la Banque mondiale a revu ce chiffre à la hausse, évoquant une fourchette haute de 435 millions de personnes, soit potentiellement 12% de la main-d'œuvre mondiale.

Contrairement aux attentes, cette force de travail ne semble pas vouée à disparaître avec l'amélioration des performances de l'IA. La préparation de l'IA nécessite une mise à jour continue pour s'adapter aux évolutions sociales et technologiques. La phase de vérification reste indispensable et ne peut être confiée qu'à des humains comprenant le fonctionnement attendu de l'IA. Quant à l'imitation de l'IA, elle demeure nécessaire pour le développement de prototypes et de nouvelles applications.

Antonio Casilli

**Professeur de sociologie à l'Institut polytechnique de Paris
Co-fondateur de l'International Network on Digital Labor**

3

Encadrer le recours à IA : le devoir de vigilance

Article L. 225-102-4.-I du Code de commerce

Toute société qui emploie, à la clôture de deux exercices consécutifs, au moins cinq mille salariés en son sein et dans ses filiales directes ou indirectes dont le siège social est fixé sur le territoire français, ou au moins dix mille salariés en son sein et dans ses filiales directes ou indirectes dont le siège social est fixé sur le territoire français ou à l'étranger, établit et met en œuvre de manière effective un plan de vigilance.

Le plan comporte les mesures de vigilance raisonnable propres à identifier les risques et à prévenir les atteintes graves envers les droits humains et les libertés fondamentales, la santé et la sécurité des personnes ainsi que l'environnement, résultant des activités de la société et de celles des sociétés qu'elle contrôle au sens du II de l'article L. 233-16, directement ou indirectement, ainsi que des activités des sous-traitants ou fournisseurs avec lesquels est entretenue une relation commerciale établie, lorsque ces activités sont rattachées à cette relation.



Le plan a vocation à être élaboré en association avec les parties prenantes de la société, le cas échéant dans le cadre d'initiatives pluripartites au sein de filières ou à l'échelle territoriale. Il comprend les mesures suivantes :

- 1° Une cartographie des risques destinée à leur identification, leur analyse et leur hiérarchisation ;
- 2° Des procédures d'évaluation régulière de la situation des filiales, des sous-traitants ou fournisseurs avec lesquels est entretenue une relation commerciale établie, au regard de la cartographie des risques ;
- 3° Des actions adaptées d'atténuation des risques ou de prévention des atteintes graves ;
- 4° Un mécanisme d'alerte et de recueil des signalements relatifs à l'existence ou à la réalisation des risques, établi en concertation avec les organisations syndicales représentatives dans ladite société ;
- 5° Un dispositif de suivi des mesures mises en œuvre et d'évaluation de leur efficacité.

Le plan de vigilance et le compte rendu de sa mise en œuvre effective sont rendus publics et inclus dans le rapport mentionné à l'article L. 225-102.

	Publication d'un plan de vigilance	Identification des risques liés au recours à l'IA
Sodexo	Oui	Non
Carrefour	Oui	Risques associés aux droits humains, à la santé et à la sécurité plutôt détaillés mais absence de mention du lien avec le recours à l'intelligence artificielle.
L'Oréal	Oui	Risques associés aux droits humains, à la santé et à la sécurité plutôt détaillés mais absence de mention du lien avec le recours à l'intelligence artificielle. Publication en interne en 2023 des « Principes pour une intelligence artificielle digne de confiance » (non accessible).
Sanofi	Oui	Risques associés aux droits humains, à la santé et à la sécurité plutôt détaillés mais absence de mention du lien avec le recours à l'intelligence artificielle.
Société Générale	Oui	Non
Capgemini	Oui	Non (Mention pour l'intelligence artificielle du risque éthique et juridique relatif à la protection des données à caractère personnel et liés au biais).
Stellantis	Oui	Non
Thales	Oui	Non Publication d'une charte éthique du numérique mentionnant notamment les risques liés au biais, à la confidentialité des données et l'importance du maintien du contrôle par l'humain.
Airbus	Oui	Non
Bouygues	Oui	Non Risques associés aux droits humains, à la santé et à la sécurité plutôt détaillés mais absence de mention du lien avec le recours à l'intelligence artificielle.
Teleperformance	Oui	Mention de mesures particulières dédiées aux modérateurs de contenu (prévention des risques psychologiques).

Conclusion

Les sociétés du panel analysé sont dans une très large mesure en contrariété avec les exigences issues de la loi sur le devoir de vigilance s'agissant des risques liés au développement et/ou au recours à l'IA. L'absence d'une cartographie précise, dans leur plan de vigilance, des risques sociaux et environnementaux liés à la production et/ou aux utilisations de l'IA pourrait justifier leur mise en demeure par des organisations de la société civile afin que des injonctions soient prononcées par un juge de prendre des mesures pour identifier et hiérarchiser ces risques mais également pour les prévenir.

Il y a donc intérêt à agir : à la fois pour les sociétés mentionnées, sur le plan juridique et pour se mettre en conformité ; pour les organisations de la société civile, sur le plan judiciaire et pour les amener à se mettre en conformité si elles ne le font pas volontairement ; et, plus généralement, pour l'ensemble de la société et sur le plan politique pour que la "révolution de l'IA" ne se fasse pas au détriment de la planète et des territoires "du Sud".

Nos contacts

<https://www.interetaagir.org>

contact@interetaagir.fr

Intérêt à agir est une association réunissant des juristes, docteur.es en droit, universitaires et avocat.es qui accompagnent des organisations de la société civile sur la réflexion puis la réalisation de contentieux stratégiques pour faire évoluer le droit ou renforcer son effectivité.