



RÉPUBLIQUE
FRANÇAISE

*Liberté
Égalité
Fraternité*

Conseil

de l'IA
et du Numérique

Note

Les intelligences artificielles à l'heure de la **vague agentique** : de quoi parle-t-on?

février 2026

« Et si demain votre collègue était une IA ? », « Comment l'IA réservera bientôt vos vacances à votre place ? », « Réserver un restaurant, payer vos courses avec l'IA... » : ces titres de presse se sont multipliés ces derniers mois. Depuis 2022, les débats sur l'intelligence artificielle (IA) se sont principalement cristallisés sur les modèles de langage et les assistants conversationnels, comme ChatGPT (OpenAI), Gemini (Google), Claude (Anthropic) ou Le Chat (Mistral). Mais la profondeur de ce qui peut être confié à des IA s'étend d'ores et déjà sous la bannière des « agents IA » ou de « l'IA agentique ». La promesse est la suivante : déléguer une succession de tâches à une suite logicielle intégrant de l'IA, le plus souvent avec une interface conversationnelle.

Les agents IA suscitent de grandes attentes et autant d'interrogations : y aura-t-il un remplacement progressif du travail humain ? Quelle protection des données des utilisateurs ? Quelle responsabilité en cas d'erreurs ? Quelle transparence de ces systèmes logiciels ? Les agents IA vont-ils vraiment permettre d'accroître la productivité de façon significative ?

Malgré l'effet d'annonce ambiant, l'IA agentique ne constitue pas une rupture technologique unique mais correspond plutôt à une évolution des capacités logicielles et des infrastructures, portée par le perfectionnement des IA génératives et l'émergence d'infrastructures à grande échelle.

Par cette note, le CIANum inaugure une série de travaux sur les enjeux autour de l'IA agentique. Aborder ces questions implique au préalable de poser des bases communes d'analyse : qu'est-ce qu'un agent IA ? Comment ces systèmes fonctionnent-ils ? Que peut-on réellement en attendre ? Cette première note vise à apporter de premiers éléments de définition et de vulgarisation pour en comprendre le fonctionnement, et les questions qui en découlent.

Sommaire

Qu'est-ce que l'IA agentique ?	2
Comment fonctionnent les IA agentiques ?	4
Quelles sont les accélérations technologiques qui sous-tendent l'IA agentique ?	4
Premier aperçu des limites et des risques des IA agentiques	7
Des limites techniques	7
Des limites éthiques, sociétales et juridiques	7
La crainte du remplacement du travail humain	10
Quelques références pour aller plus loin	11

Qu'est-ce que l'IA agentique ?

Un agent IA est un programme informatique capable de prendre des décisions ou de réaliser des actions, y compris de les coordonner entre elles, en s'appuyant sur des modèles d'intelligence artificielle¹. Intégré à une suite logicielle (on parle alors de « logiciel agentique »), un agent IA est doté d'une capacité d'action propre. Il peut décomposer un processus, analyser son environnement, proposer une solution, la comparer avec d'autres, voire prendre des décisions ou enchaîner des flux d'opérations, avec ou sans validation humaine.

L'un des exemples les plus souvent cités est la réservation d'un voyage en ligne. À partir d'une requête, l'utilisateur peut s'attendre à ce que l'IA :

- consulte son agenda pour trouver des dates disponibles ;
- étudie les différents modes de transport et logements existants ;
- compare les prix ;
- propose une gamme d'options adaptées ;
- paye en ligne avec – ou sans – demande de validation intermédiaire.

Accoler le mot « agentique » aux systèmes informatiques n'est cependant pas nouveau.

Dans les années 1960, le robot « Shakey », développé par le *Stanford Research Institute* avec le soutien de la DARPA, fut considéré comme le premier robot mobile autonome doté d'une intelligence artificielle rudimentaire. Innovant pour son époque, il combinait vision par ordinateur, planification automatique (via le système STRIPS), traitement du langage naturel et navigation autonome, lui permettant d'exécuter des tâches complexes comme déplacer des objets ou éviter des obstacles.

Shakey préfigurait les agents IA modernes en intégrant perception, raisonnement et action, trois piliers de l'autonomie décisionnelle. Le qualificatif « agentique » s'est par la suite endormi jusqu'au milieu de l'année 2024². Andrew Ng, un des pionniers des réseaux de neurones, l'a remis au goût du jour après l'explosion des modèles d'IA génératives à la fin de l'année 2022.

Si le terme « Agents IA » s'est imposé dans le débat public, il nous paraît néanmoins préférable de privilégier celui d'« IA agentique ». Cela évite de réduire la discussion à une opposition binaire entre automatisation (les agents) et contrôle humain : l'intervention de systèmes agentiques n'exclut en rien des interventions humaines, bien au contraire. L'IA agentique bénéficie de validations intermédiaires par des experts humains. Son degré d'autonomie est d'ailleurs variable, allant de la simple assistance à l'automatisation de bout en bout³.

¹ Ces modèles peuvent être de différents types en fonction des objectifs et fonctionnalités de l'IA agentique, par exemple des grands modèles de langage, de traitement de la vision, de raisonnement, etc.

² Une recherche des tendances sur les réseaux sociaux montre cette accélération subite à partir d'août 2024.

³ Une source très pédagogique sur ces questions est : SAPKOTA, Ranjan, ROUMELIOTIS, Konstantinos I., et KARKEE, Manoj. Ai agents vs. agentic ai: A conceptual taxonomy, applications and challenges. *arXiv preprint arXiv:2505.10468*, 2025, <https://arxiv.org/abs/2505.10468>.

L'IA agentique en cinq degrés d'autonomie

La littérature converge autour d'une échelle à **cinq degrés pour décrire les niveaux d'autonomies des IA agentiques**⁴ :

1. Des agents d'**automatisation simple fondée sur la reproduction des règles** : ils exécutent des règles prédéfinies et des flux de travail fixes.
Exemples : les technologies dites RPA (Robotic Process Automation) qui saisissent des données de factures ; ou un régulateur de vitesse qui maintient une vitesse constante.
2. Des agents à **automatisation intelligente**, capables d'extraire des données ou d'analyser des documents.
Exemples : la reconnaissance optique de caractères (OCR) pour extraire des données de factures ou de reçus ; l'assistant de maintien dans la voie ou le freinage d'urgence utilisant la vision par ordinateur embarqués dans certains véhicules.
3. Des **flux de travail agentiques** : les agents à automatisation intelligente et/ou des agents d'IA génératives enchaînent les étapes dans des processus alternant automatisation de tâches et validation humaine.
Exemples : une IA agentique collecte l'identification du client bancaire (« KYC » pour « Know Your Customer ») et une autre vérifie l'identité, une troisième prépare des décisions de crédit. Le tout est ensuite validé par un humain.
4. Des **flux d'agents semi-autonomes** : le logiciel devient proactif, sensible au contexte et capable de préparer puis de prendre des décisions soumises à validation humaine.
Exemple : un agent de soutien à l'organisation hospitalière qui prédit l'afflux de patients et propose des règles d'affectation du personnel.
5. Des **flux d'agents entièrement autonomes** : auto-dirigés, ils apprennent en continu et s'exécutent de bout en bout sans orientation ni validation humaine. Sur le plan technique, ce degré d'autonomie nécessite l'apprentissage perpétuel (*continual learning*) et la capacité à décider de ses objectifs propres. Ce niveau n'est pas atteint aujourd'hui et fait encore l'objet de travaux de recherche.
Exemple : une entreprise agentique « accélératrice » de start-ups qui identifierait des opportunités de marché, construirait des prototypes, proposerait un modèle de vente et recruterait des travailleurs pour développer ses produits.

⁴ Voir par exemple BORNET, Pascal et WIRTZ, Jochen. *Agentic artificial intelligence: harnessing AI agents to reinvent business, work, and life*. 2025.

Comment fonctionnent les IA agentiques ?

Les architectures logicielles agentiques s'organisent en plusieurs couches qui composent le système et concourent à son fonctionnement d'ensemble :

- **La couche applicative** est l'interface avec laquelle l'utilisateur interagit, par exemple via un *chatbot* qui permet de tirer parti des capacités conversationnelles de l'IA générative. Elle peut faire appel à des bases de données ou à des sources d'information externes pour comprendre le contexte.
Exemple : un assistant virtuel intégré à votre boîte mail, qui affiche une interface simple d'où vous pouvez formuler des requêtes comme : « planifie une réunion de suivi sur ce chantier ».
- **La couche d'orchestration** séquence les actions nécessaires à l'accomplissement d'une tâche. Elle sollicite différents micro-services logiciels pour exécuter chaque étape d'une tâche. Elle est le cœur des suites agentiques et, sur le plan économique, l'enjeu central du déploiement des plateformes permettant de déployer l'IA agentique.
Exemple : lire l'e-mail > analyser le contenu > rédiger une réponse.
- **La couche logique (de raisonnement)** fonctionne comme un guide intelligent : elle connaît des scénarios types d'enchaînement d'actions (un peu comme des recettes de cuisine ou des modes d'emploi) qu'elle utilise pour accomplir un objectif. Cette couche calcule aussi le coût de chaque alternative et choisit la meilleure en termes de d'objectifs et de coût associé. Elle constitue le cadre et la logique de décision de l'IA.
Exemple : si l'e-mail contient le mot « urgent », répondre dans l'heure et m'alerter.
- **La couche d'évaluation** permet de vérifier la qualité du résultat et de le corriger si nécessaire. Les mécanismes correctifs peuvent être :
 - Automatisés (vérification du résultat par l'IA elle-même) ;
 - Supervisés (l'utilisateur corrige la réponse ou l'action proposée) ;
 - Améliorés (l'IA peut apprendre à partir des retours, pour optimiser ses futures actions, on parle aussi d'apprentissage par renforcement).*Exemple : l'IA génère une réponse à un e-mail, puis demande confirmation avant envoi : « Cette réponse vous convient-elle ? ». Si l'utilisateur modifie le texte, l'IA en tient compte par la suite.*

Quelles sont les accélérations technologiques qui sous-tendent l'IA agentique ?

Comme affirmé précédemment, l'IA agentique n'est pas une rupture technologique unique, mais plutôt la combinaison de cinq accélérations, dont les progrès ont été spectaculaires au cours des dernières années :

1. L'émergence d'une nouvelle génération de modèles d'IA générative capables d'élaborer des **raisonnements itératifs et de plus en plus complexes, même s'il est à noter que leur performance décline au-delà d'un**

certain niveau de complexité⁵. Ces modèles, dits de raisonnement⁶, existent depuis septembre 2024. Ils sont capables de suivre des raisonnements logiques, de répondre à des questions, de générer des données synthétiques et de décomposer un processus de décision.

2. L'émergence d'une **science de l'ingénierie des modèles d'IA ou du « post-entraînement »** (2023 – 2026) : les progrès de l'IA dépendent de moins en moins de la croissance de la taille des modèles, mais de plus en plus des opérations réalisées après leur entraînement. Une fois le modèle créé, ce dernier est affiné pour le rendre plus efficace et plus précis en optimisant son comportement en fonction d'objectifs mieux circonscrits. Cela peut par exemple consister à réentraîner le modèle sur un jeu de données plus précis et plus spécifique, souvent lié à un domaine particulier (médical, juridique, technique...). On parle de *fine-tuning* ou d'affinement. Il peut s'agir d'une optimisation grâce à des retours d'annotateurs humains (*reinforcement learning from human feedback*) qui évaluent les réponses du modèle, lesquelles sont ensuite ajustées pour privilégier les réponses jugées meilleures. Des opérations d'élagage peuvent aussi avoir lieu en supprimant certaines parties du modèle jugées non essentielles afin de réduire la taille du modèle et d'accélérer son exécution sans perdre en performance.
3. **L'allongement des fenêtres de contexte et de la mémoire** allouées aux modèles d'IA ainsi qu'aux logiciels agentiques, ce qui permet d'enrichir leur capacité à tenir compte d'un contexte relationnel ou d'un historique de contacts : les systèmes retiennent des informations sur une plus longue période, relient mieux les idées entre elles en tenant compte du passé et du contexte et peuvent s'adapter à des échéances complexes (par exemple sur plusieurs jours ou avec un grand nombre de facteurs). C'est dans ce domaine que les architectures agentiques progressent le plus actuellement.
4. **L'innovation logicielle**, considérable en 2025, avec notamment :
 - De nouvelles méthodes de dialogue avec l'IA comme c'est le cas avec les agents conversationnels, ce que l'on appelle le *prompting* (que l'on peut traduire par « formulation de requête ») ;
 - De nouvelles fonctionnalités permettant aux systèmes d'IA de lancer des actions grâce aux « appels de fonction », par exemple envoyer un e-mail, réserver un billet, payer ou chercher une information dans une base de données ;
 - De nouvelles boîtes à outils pour combiner plusieurs IA, les connecter à d'autres logiciels⁷, favoriser l'interopérabilité et l'accès

⁵ VALMEEKAM, Karthik, STECHLY, Kaya, GUNDAWAR, Atharva, et al. A systematic evaluation of the planning and scheduling abilities of the reasoning model o1. *Transactions on Machine Learning Research*, 2025.

SHOJAEI, Parshin, MIRZADEH, Iman, ALIZADEH, Keivan, et al. The illusion of thinking: Understanding the strengths and limitations of reasoning models via the lens of problem complexity. *arXiv preprint arXiv:2506.06941*, 2025, <https://arxiv.org/abs/2506.06941>.

⁶ Par exemple o3 d'OpenAI, R1 de DeepSeek, Claude Sonnet 3.7 d'Anthropic.

⁷ CrewAI, AutoGen, LlamaIndex, Langchain en sont des exemples.

à des bases de données externes⁸, intégrer des fonctionnalités d'IA agentiques dans des logiciels classiques⁹, proposer des applications clés en main¹⁰, développer *via* des plateformes *no-code*¹¹, des kits de développement logiciel, des studios de conception ou des plateformes d'intégration de bout en bout¹².

5. La mise à disposition, par des fournisseurs d'infrastructure *cloud*, de capacités de calcul « scalables » à grande échelle, capables de faire face aux besoins croissants en puissance de calcul des modèles d'IA. On peut parler ici **d'un nouveau segment de marché : l'inference-as-a-service**¹³. L'inférence correspond à la phase où un modèle d'IA traite une requête pour produire un résultat (réponse, décision, analyse, etc.). Ces nouveaux services *cloud* permettent aux organisations, quelles que soient leur taille ou leurs ressources, d'accéder de façon virtualisée à une puissance de calcul adaptée à leurs besoins, sans investissement propre en matériel. Même si les données précises sont très variables selon les modèles sollicités et peu transparentes, on observe une baisse drastique de ces coûts d'inférence : les données du *think tank* EPOCH AI indiquent que le coût moyen de l'inférence, mesuré par jeton (*token*), a été divisé par 90 en 18 mois¹⁴. Cette baisse rapide des coûts est le carburant principal de la croissance des usages de logiciels agentiques.

Portée par ces cinq accélérations, l'IA agentique devient un phénomène de marché. Les investissements réalisés sont comparables à ceux ayant permis, en leur temps, l'émergence des infrastructures ferroviaires, énergétiques ou des réseaux de télécommunications¹⁵. La grande différence de ces investissements dans les centres de données est qu'ils se déprécient rapidement (trois à cinq ans) comparés aux infrastructures historiques citées. Les attentes sur l'IA agentique sont donc conséquentes, à la fois du point de vue des utilisateurs et de celui des entreprises du secteur de l'IA¹⁶, qui cherchent à générer des revenus croissants et réguliers au regard des montants investis ces dernières années.

Pour résumer : Ces quatre couches fonctionnent en synergie pour permettre à l'IA agentique de (i) comprendre, (ii) décider, (iii) agir, (iv) s'améliorer, de manière autonome ou en collaboration avec l'humain.

⁸ LangChain and Semantic Kernel en sont des exemples.

⁹ À titre d'exemple, Salesforce, ServiceNow, Adobe, Office 365, UiPath ou encore SAP ont rapidement proposé des services agentiques complémentaires à leur licence.

¹⁰ Eidos, Auxo Dynamics, AutoGPT, BabyAGI, Manus ou Hippocratic AI en sont des exemples.

¹¹ StackAI, Glean, Beam, Relevance.ai, Lovable en sont des exemples.

¹² Mistral Studio, Google Vertex, Azure Studio, Cohere North en sont des exemples.

¹³ Aussi appelé GPU-as-a-service du nom des semi-conducteurs utilisés – les GPUs- dans les infrastructures dédiées à l'IA.

¹⁴ "LLM inference prices have fallen rapidly but unequally across tasks", EPOCH AI, 12 mars 2025.

¹⁵ AZHAR, Azeem, WARREN, Nathan. "Is AI a bubble?". *Exponential View*, 17 septembre 2025.

¹⁶ Greg Brockman, co-fondateur d'OpenAI prédit à ce titre que 2026 sera l'année de l'adoption de l'IA agentique par les entreprises (<https://x.com/gdb/status/2006584251521839141?s=43&t=NOvXIUVo0GBu4sy6swo41w>).

Premier aperçu des limites et des risques des IA agentiques

La route vers une mise en place généralisée des modèles d'IA agentique est longue et implique une transformation en profondeur des processus internes aux entreprises. Leur adoption soulève d'ores et déjà des interrogations de divers ordres.

Des limites techniques

Une partie des limites et des risques sont liés à l'ingénierie propre des IA agentiques. On peut relever en particulier :

- **Les effets de cascade** (*compositional generalization*) : les logiciels agentiques, mobilisant plusieurs opérations consécutives, voient leurs erreurs de raisonnement se cumuler ; se multiplier et s'aggraver mutuellement.
- **La complexité systémique** : à mesure que le nombre et la diversité des agents en interaction augmentent, la probabilité de comportements imprévus augmente.
- **La « dérive de l'orchestration »** : on désigne ce phénomène lorsque les agents sont connectés à d'autres agents sans cadre commun ni logique de coordination claire, les processus pouvant devenir imprévisibles.
- **Le « désalignement sémantique »** : deux agents peuvent interpréter différemment une même consigne et prendre des décisions contradictoires ou répéter inutilement les mêmes actions. Ces désalignements sont très fréquents et impliquent de tester régulièrement les logiciels pour vérifier leur performance.
- **Les risques de sécurité des données** : les suites agentiques peuvent exposer des données sensibles ou être exploités par des acteurs malintentionnés. L'organisme américain NIST¹⁷ a documenté ces nouveaux risques de sécurité qu'il convient de prendre en compte.
- **La vulnérabilité des IA agentiques** qui peuvent, en raison de leur capacité à interagir avec d'autres logiciels ou infrastructures, devenir des cibles privilégiées pour des cyberattaques, des manipulations malveillantes ou des usages détournés.

Des enjeux éthiques, sociétaux et juridiques

Tout d'abord, la **question de la protection des données personnelles** se pose avec une acuité particulière, dans la mesure où ces agents, pour fonctionner efficacement, accèdent, analysent et parfois stockent des volumes considérables d'informations. Selon Meredith Whittaker, fondatrice de Signal, il n'existe actuellement aucune solution technique permettant de concilier l'autonomie des agents IA et la protection effective des données¹⁸.

La question de la responsabilité vis-à-vis des utilisateurs finaux, usagers, clients ou patients, en cas d'erreur, de préjudice ou de décision dommageable obéit aux principes habituels de l'automatisation, avec l'obligation de rattacher la solution agentique à l'entreprise qui en fait usage. La capacité pour cette entreprise à se retourner vers son éditeur est garantie de façon variable selon les acteurs. Qui est responsable si un agent prend une décision erronée ayant des conséquences financières, sanitaires ou sociales par exemple ? Le développeur qui a conçu le modèle, l'éditeur du logiciel, l'utilisateur qui l'a déployé, ou encore l'agent lui-même ?

¹⁷ National Institute of standards and technology. [Artificial Intelligence Risk Management Framework \(AI RMF 1.0\)](#), janvier 2023.

¹⁸ [AI agents are coming for your privacy, warns Meredith Whittaker.](#)

Le règlement européen sur l'IA (RIA ou AI Act) encadre, partiellement, ces enjeux (voir encadré ci-dessous).

En outre, les risques de biais algorithmiques et de discrimination automatique sont souvent relevés. En effet, les agents IA apprennent à partir de données réelles, qui peuvent contenir des biais de différentes natures¹⁹. Programmer une IA implique donc un arbitrage entre la réduction de ces biais, et la reproduction de la réalité vécue (on parle de réglage de la « température » des modèles). Cet arbitrage n'est jamais parfait et il n'existe pas de modèles d'IA parfaitement « débiaisés ». Lorsqu'une IA est utilisée dans une entreprise ou avec des clients, il est à cet égard essentiel de tester et vérifier le caractère potentiellement discriminatoire de ses résultats et de mettre en œuvre toutes les mesures possibles pour le corriger et assurer une parfaite transparence sur ces effets.

La transparence et l'intelligibilité de ces systèmes constituent un autre enjeu central au regard de l'ensemble des points de vigilance précédemment listés. À mesure que ces modèles d'IA agentique se complexifient, la compréhension de leurs mécanismes décisionnels devient de plus en plus difficile. Pourtant, cette opacité, cet effet « boîte noire » peut avoir des conséquences graves, notamment dans des domaines critiques comme la santé, la justice ou la finance.

Ces risques ne tombent néanmoins pas dans un vide juridique. Au niveau européen, le législateur est intervenu pour encadrer ces développements technologiques, notamment dans le règlement sur l'intelligence artificielle.

Le règlement européen sur l'intelligence artificielle

Ce texte européen adopté en 2024 et entré en vigueur depuis 2025, régle les usages des systèmes d'IA en fonction de leurs risques :

- **Systèmes à risque minimal** : pas d'obligation, code de conduite.
Exemple : systèmes IA dans les jeux vidéo
- **Systèmes à risque limité** : obligations de transparence.
Exemple : les utilisateurs doivent être informés qu'ils interagissent avec une intelligence artificielle ou qu'un contenu a été généré ou modifié par IA.
- **Systèmes à haut risque** : obligations renforcées d'autorisation préalable avant leur mise sur le marché. Exigences fortes de documentation, de transparence, de qualité, de supervision humaine et de sécurité.
Exemple : dans la santé, l'éducation, le recrutement, le maintien de l'ordre ou la justice.
- **Systèmes à risque inacceptable** : interdits dans l'UE.
Exemple : systèmes de notation sociale, de reconnaissance des émotions, d'identification biométrique à distance en temps réel, ou de justice prédictive.

¹⁹ Sur la notion de biais cognitifs et les effets de l'IA : GIGERENZER, Gerd. [What Is Bias? And Why Are We Biased?](#). *Proceedings of the Paris Institute for Advanced Study*, 2025, vol. 21.

Les systèmes d'IA à haut risque constituent l'essentiel du texte. Ils sont soumis à des obligations d'identification des risques qu'ils présentent tout au long de leur cycle de vie et de mise en place de mesures de gestion et d'atténuation de ces risques. Les fournisseurs doivent également conserver les *logs* pour permettre une traçabilité des décisions et des actions du système d'IA et établir une documentation technique détaillée sur les données d'entraînement, l'architecture du système, ses performances et ses limites. L'utilisateur doit être explicitement informé qu'il interagit avec un système d'IA et des limites du système.

Les modèles d'IA à usage général (permettant d'accomplir une grande variété de tâches) présentant des risques systémiques sont soumis à des obligations renforcées d'évaluation et d'atténuation de ces risques. Par exemple, s'ils présentent des risques d'accidents, des risques pour les processus démocratiques ou la diffusion de contenus illégaux, faux ou discriminatoires ou encore des risques de conséquences graves pour la santé et la sécurité publiques.

Le texte ne prévoit pas explicitement de dispositions relatives aux IA agentiques. Toutefois, les règles et obligations précédemment détaillées peuvent trouver à s'appliquer à ces systèmes, par exemple s'ils présentent des risques limités ou élevés. Ils sont également concernés par les obligations relatives aux modèles d'IA à usage général dans le cas où ils s'appuient sur ce type de modèles.

En cas de non-conformité, les sanctions peuvent aller jusqu'à 30 millions d'euros ou 7 % du chiffre d'affaires annuel.

En complément, en fonction des usages et de la qualification de leurs fournisseurs, **les IA agentiques peuvent être soumises à d'autres textes européens** comme le règlement sur les services numériques (RSN ou DSA) et le règlement sur les marchés numériques (RMN ou DMA). Dans le premier cas par exemple, si une IA agentique est utilisée au sein d'une plateforme (par exemple, pour modérer des contenus, recommander des produits, ou interagir avec des utilisateurs) et que celle-ci est soumise au RSN, alors le système d'IA doit respecter les obligations qui en découlent, en termes notamment de transparence, de gestion des risques, de modération des contenus, etc. Dans le second cas, si l'IA agentique est développée ou utilisée par un « contrôleur d'accès » (ou *gatekeeper*) au sens du RMN, alors elle peut être soumise à des obligations d'interopérabilité, de portabilité des données, d'interdiction des pratiques anticoncurrentielles, etc.

Enfin, **l'émergence rapide des agents IA risque d'introduire une rupture d'échelle en matière d'empreinte environnementale de l'IA**. Si l'utilisation mondiale de ChatGPT pour le seul mois de janvier 2023 équivalait déjà à l'émission moyenne annuelle de 1264 foyers français²⁰, la charge de calcul permanente sur les serveurs, induite par l'IA agentique, pourrait multiplier ce coût de façon exponentielle. Selon les analyses prospectives de l'ADEME (2026), la généralisation de ces agents autonomes pourrait potentiellement faire basculer l'IA de 20% à près de 49% de la consommation totale des centres de données d'ici la fin de l'année.

²⁰ Data for Good. *Livre blanc sur l'IA générative*. 2023.

La crainte du remplacement du travail humain

En rendant automatisables des tâches qui jusqu'à présent ne l'étaient pas, l'IA agentique emporte un potentiel bouleversement de monde du travail et de son organisation. Le médecin augmenté, le soignant augmenté, l'ingénieur augmenté : on ne compte plus le nombre des professions sur le point d'être « augmentées ». Ce discours de l'« augmentation » ne permet toutefois pas de penser de façon exhaustive la redéfinition des professions à laquelle l'IA agentique peut conduire.

Dès lors, la question de l'acceptation et de la désirabilité de ces modèles se pose : l'économiste Erik Brynjolfsson²¹ a notamment montré l'écart qui existe entre le potentiel des IA agentiques et la réalité des activités pour lesquels les salariés se projettent avec elles. En outre, Daron Acemoglu²² a souligné aussi la lenteur dans la pénétration de l'IA dans les entreprises chaque fois que la part des tâches automatisées correspond à une petite proportion (moins de 20%) des tâches d'un métier. C'est le plus souvent le cas.

Comme la sociologie du travail l'a montré²³, le travail recouvre à la fois une dimension prescrite (ce qui figure dans votre fiche de poste) et une dimension réelle qui recouvre l'ensemble des missions effectuées au quotidien, dans toute leur diversité. En d'autres termes, le travail n'est pas qu'une succession de tâches, mais aussi une place, un rôle ou une fonction dans l'organisation de l'entreprise où les interactions humaines sont indispensables. Dans ce contexte, les annonces de réductions massives d'emplois doivent être prises avec précaution, l'adoption de l'IA agentique pouvant servir de prétexte aux entreprises pour justifier des licenciements économiques.

L'irruption de l'IA et, demain, de l'IA agentique, devrait d'abord inviter à repenser plus largement l'exercice des tâches – en particulier les plus répétitives – redessiner les contours d'un métier, réorganiser les processus de travail. Il s'agit pour le travailleur de s'extraire des travaux les plus aliénants, les plus pénibles, les moins émancipateurs. Le vœu n'est pas neuf ; Oscar Wilde l'exprimait déjà en 1891²⁴, au sujet de la « machine » qui devait cesser de « concurrencer l'Homme » et « faire les besognes d'assainissement », « pénibles et ennuyeuses », « faire les courses quand il pleut ».

Le CIANum étudiera dans une prochaine note les conditions dans lesquelles l'intégration rapide, massive et pas toujours maîtrisée (par les employeurs comme par les travailleurs) de l'IA agentique dans le monde du travail pourrait conduire à une transformation de l'organisation du travail (lieu, horaires, recomposition des fonctions, modalités de coopération, de relation client/usager, d'évaluation...), de l'identification et la reconnaissance des compétences professionnelles (identification, acquisition et reconnaissance des compétences techniques, des « *soft skills* », recomposition des métiers, etc.). Il existe une préoccupation selon laquelle

²¹ BRYNJOLFSSON, Erik et MCAFEE, Andrew. *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. WW Norton & company, 2014.

Sur l'IA spécifiquement : ["How can AI augment rather than dictate human action? This expert explains"](#). *World Economic Forum*, 18 octobre 2024.

²² JOHNSON, Simon et ACEMOGLU, Daron. *Power and Progress: Our Thousand-Year Struggle Over Technology and Prosperity* | *Winners of the 2024 Nobel Prize for Economics*. Hachette UK, 2023.

Sur l'IA spécifiquement : ACEMOGLU, Daron. ["Don't Believe the AI Hype"](#). *Project Syndicate*. 21 mai 2024.

²³ Voir par exemple : DANIELLOU, François, GRALL, Jérôme, MARTIN, Christian, et al. Prescriptions, injonctions et pressions. *Actes des journées de Bordeaux. Bordeaux : Editions de l'université Paul Segalen*, 2000 ; DUJARIER, Marie-Anne. Chapitre 2. Travail et organisation : quelques définitions. *Éthique et philosophie morale*, 2012, p. 43-54.

²⁴ WILDE, Oscar. *L'Âme humaine sous le régime socialiste*. 1981.

ces modèles d'IA pourraient entraîner des conséquences majeures en termes de création et destruction d'emploi, de qualité de travail (possibilité d'amélioration des conditions de travail mais aussi risque d'intensification du travail) et d'accès à la formation pour tous. Ils peuvent favoriser un marché du travail plus inclusif (outil de travail mieux adapté aux situations de handicap, par exemple), ou – au contraire – plus excluant (notamment pour l'entrée des jeunes dans certaines professions, par exemple dans le droit ou le conseil, où l'IA remplace les fonctions traditionnellement attribuées aux postes d'entrée.

Les décisions qui seront prises aujourd'hui en matière d'accompagnement du changement, de formation, de reconversion professionnelle, d'adaptation de l'organisation et des conditions de travail, sont les déterminants de l'impact de l'IA agentique sur le travail et sur l'emploi demain. Il est donc essentiel que le dialogue social et la négociation collective jouent pleinement leur rôle pour anticiper, prévenir, accompagner, ralentir si nécessaire et évaluer les transformations en cours.

Quelques références pour aller plus loin

- BORNET, Pascal et al. *Agentic Artificial Intelligence: Harnessing AI Agents to Reinvent Business, Work and Life*. Irreplaceable Publishing, 2024.
- NG, Andrew. [Cours sur l'IA agentique](#) [en ligne]. DeepLearning.AI, 2024.
- HENDRYCKS, Dan, BURNS, Collin, BASART, Steven, ZOU, Andy, MAZEIKA, Mantas, SONG, Dawn et STEINHARDT, Jacob. [Measuring massive multitask language understanding](#), 2020, arXiv:2009.03300.
- SRIVASTAVA, Ankit et al. [Beyond the imitation game: Quantifying and extrapolating the capabilities of language models](#). *Transactions on Machine Learning Research*, 2023.
- LIU, Xiaoyuan, YU, Hao, ZHANG, Hao, XU, Yusu, LEI, Xinyi, LAI, Huan, TANG, Jie. [AgentBench: Evaluating LLMs as agents](#), 2023, arXiv:2308.03688.