



Renaissance
Numérique



RAPPORT

OCTOBRE 2025

Déployer une littératie en IA pour une société inclusive et émancipatrice

Sommaire



RÉSUMÉ EXÉCUTIF	7
INTRODUCTION	9

01 Appréhender le double visage de l'IA générative pour l'inclusion **12**

Limites du numérique dans la réduction des inégalités	13
Opportunités et risques de l'IA générative	15
Illectronisme : évolution et limites des mesures	19
Inégalités structurelles non numériques	20
Compétences et capacités	21
<i>Compétences</i>	21
<i>Capacités</i>	22
<i>Des approches complémentaires</i>	22
Capital numérique : entre reproduction et émancipation	23
IA et "fracture générationnelle"	24
<i>L'âge comme variable descriptive non explicative</i>	24
<i>Shadow "IA" et fragmentation des collectifs de travail</i>	25

02 Définir et nommer les systèmes d'IA **28**

Repères conceptuels	29
<i>Pluralité et évolutions du concept</i>	29
<i>Risques des amalgames</i>	31
IA générative et IA à usage général	32
<i>IA générative</i>	32
<i>IA à usage général ("GPAI")</i>	33

<i>IA générale (“AGI”) et autres concepts philosophiques</i>	34
Avancées et limites des systèmes actuels	35
<i>Évolution vers une autonomie accrue mais limitée</i>	35
<i>Limites structurelles</i>	37
L’IA est un dispositif sociotechnique	39
<i>L’IA n’est pas un outil</i>	39
<i>Même le marteau n’est pas qu’un simple outil</i>	41
Gouvernance, choix éthiques, sociétaux et civilisationnels	43

03 Concevoir

une littératie en IA 46

Premiers éléments de définition	47
<i>Continuum d’apprentissage</i>	47
<i>Culture, éducation et émancipation</i>	48
<i>L’IA, la porte d’entrée vers la littératie numérique</i>	49
Objectifs politiques et réglementaires	51
<i>Ambitions et angles morts de l’AI Act</i>	51
<i>En France, un cadre législatif en construction...</i>	52
<i>... mais un manque de stratégie globale</i>	53
Finalités démocratiques et économiques	54
<i>Enjeu démocratique : former des citoyens éclairés</i>	55
<i>Enjeu économique : redistribuer les bénéfices de l’innovation</i>	56
<i>Enjeu social : résoudre la tension adoption-émancipation</i>	57

04 Analyser les discours

et usages de l’IA générative 58

Discours et diffusion de l’IA générative	59
<i>Imaginaires structurellement ambivalents</i>	59
<i>Stratégies d’énonciateurs</i>	59
<i>Stratégies de légitimation éthique</i>	61
<i>Mythe de la rupture technologique et bulle de l’IA</i>	63
Diffusion et pratiques de l’IA générative	66

SOMMAIRE

<i>Adoption massive et usages diversifiés</i>	66
<i>Adoption généralisée dans l'enseignement</i>	68
<i>Intégration progressive dans le monde professionnel et défi pour les TPE et PME</i>	69
<i>IA intégrée et lectures des usages</i>	69

05 Renforcer le sens critique 72

Risques cognitifs individuels	73
<i>L'alarmisme technologique et ses limites</i>	73
<i>Béquille versus agilité intellectuelle</i>	73
<i>Effets méconnus de l'IA sur la cognition</i>	75
Risques informationnels amplifiés	77
<i>Désinformation, deepfakes et vulnérabilités cognitives</i>	77
<i>Biais cognitifs amplifiés par l'IA</i>	78
Impacts environnementaux de l'IA	80
<i>Impact de l'IA dans l'écosystème numérique</i>	80
<i>Stratégies de minimisation des impacts</i>	81
<i>Externalités négatives</i>	83
Stratégies d'autodéfense intellectuelle	84
Diversité des interactions et enjeux éducatifs	86
<i>Relation avec les agents conversationnels et anthropomorphisme</i>	86
<i>Compétences éducatives et émotionnelles</i>	87

06 Propositions pour déployer une littératie en IA inclusive et émancipatrice 90

Leviers individuels pour contribuer à une littératie en IA	91
<i>Proposition 1 : Éviter les amalgames</i>	91
<i>Proposition 2 : S'imposer, et imposer, une rigueur évaluative</i>	92
<i>Proposition 3 : Distinguer les usages effectifs des usages promis</i>	93

<i>Proposition 4 : Affirmer la littératie en IA comme un continuum évolutif</i>	93
Leviers collectifs pour déployer une littératie en IA	94
<i>Proposition 5 : Renforcer l'enseignement au numérique et à l'IA tout au long de la formation initiale</i>	94
<i>Proposition 6 : Organiser des États généraux du numérique et de l'IA pour l'éducation et la formation</i>	98
<i>Proposition 7 : Partager un référentiel de compétences au niveau national</i>	99
<i>Proposition 8 : Responsabiliser les concepteurs et opérateurs de formations</i>	101
<i>Proposition 9 : Amplifier la formation professionnelle à l'IA grâce à des mécanismes redistributifs spécifiques</i>	101
<i>Proposition 10 : Renforcer les aides à la parentalité numérique</i>	102
<i>Proposition 11 : Ne pas abandonner la médiation numérique</i>	104
REMERCIEMENTS	106
GROUPE DE TRAVAIL	108
BIBLIOGRAPHIE	110
À PROPOS	126

Résumé exécutif



L'intelligence artificielle (IA) générative transforme notre société à un rythme sans précédent dans l'éducation, la santé, le travail, l'accès aux services publics... Simultanément, 16 millions de Français demeurent éloignés du numérique, révélant un défi majeur : comment faire de l'IA un facteur d'inclusion et d'émancipation plutôt que d'amplification des inégalités existantes ?

Ce rapport s'adresse aux décideurs et acteurs qui façonnent, par leurs actions et leurs discours, les transformations numériques de notre société. Il soutient que la littératie en IA constitue un impératif à la fois démocratique, social et économique. Au-delà des compétences techniques, la littératie en IA porte un projet de société qui place l'inclusion et l'émancipation citoyenne au cœur de la transformation numérique, créant un rempart contre la dépolitisation des choix technologiques.

Le défi central réside dans la résolution de la tension entre l'urgence économique de l'adoption et la nécessité du temps long de l'acculturation critique. Cette tension révèle des rapports de pouvoir où l'intérêt général appelle à préserver l'autonomie intellectuelle des citoyens face aux stratégies de maximisation de l'adoption par les détenteurs de capitaux technologiques.

L'IA générative présente une dualité structurelle. D'une part, son interface en langage naturel démocratise l'accès aux outils numériques et ouvre des perspectives pour les publics

traditionnellement exclus. D'autre part, son usage éclairé exige des connaissances critiques et des compétences variées pour éviter dépendance, manipulation et reproduction des biais. Pour mieux éclairer les débats autour de l'IA, ce rapport clarifie les concepts fondamentaux de l'intelligence artificielle et explore les enjeux critiques (biais cognitifs, impacts environnementaux et dynamiques spéculatives).

Le projet d'une société numérisée plus juste, inclusive, démocratique et soutenable nous rend toutes et tous responsables.

Au niveau individuel, chaque décideur doit adopter une posture de responsabilité intellectuelle : éviter les amalgames technologiques, exiger et s'imposer une rigueur évaluative, distinguer prescription et adoption réelle des usages, et concevoir la littératie comme un continuum évolutif.

Au niveau collectif, l'ampleur des transformations exige une réponse systémique coordonnée. Le rapport préconise un changement de paradigme dans la formation initiale, dépassant le simple ajout d'heures d'éducation aux médias pour repenser globalement les programmes scolaires. Un groupe de travail national sous supervision du Conseil national de l'IA et du numérique peut permettre de définir un socle de réflexion intégrant justice sociale et environnementale, préparant de nouveaux États généraux du numérique et de l'IA pour l'éducation et la formation. Un référentiel national de compétences structurerait de son côté les parcours d'apprentissage en s'appuyant sur une grille d'évaluation mesurant l'impact sur la réduction des inégalités.

La mise en œuvre nécessite une coordination renforcée entre acteurs publics et privés, des financements durables orientés vers l'inclusion, une gouvernance démocratique plaçant les citoyens au cœur des choix technologiques, et des mécanismes redistributifs orientant une part des investissements IA vers l'inclusion de tous les publics. L'accompagnement de la parentalité numérique et le maintien de la médiation de proximité complètent cette approche, s'appuyant sur le maillage des centres sociaux, bibliothèques, fablabs, maisons France Services et autres structures au plus proche des habitants, pour déployer une littératie accessible, avec un accompagnement humain adapté aux nouveaux enjeux de l'IA générative.

Intro | duction



L'intelligence artificielle générative transforme notre société dans de nombreux secteurs (éducation, formation, santé, travail, information, etc.), touchant autant les sphères privées que publiques, individuelles que collectives. Dans le même temps, près de 16 millions de Français restent éloignés du numérique. La capacité des individus à tirer profit des possibilités de l'IA générative et de ses usages apparaît dès lors comme un enjeu de justice sociale et un impératif démocratique.

Pour qu'une technologie soit en capacité de transformer la société, il ne suffit pas de la diffuser, il faut qu'elle soit adoptée. Cela ne se fait pas simplement en "touchant" les publics. Cela suppose un travail d'éducation et de formation tout au long de la vie, ainsi que la construction d'un narratif qui favorise cette appropriation. La manière dont l'IA est décrite, les promesses et les attentes qui sont placées en elle peuvent être à la fois des moteurs ou des pièges, en cas de retournement de conjoncture comme l'éclatement d'une bulle.

Si certains érigent l'IA générative comme la nouvelle solution pour résorber les inégalités (numériques), elle présente pourtant un double visage, avec un apport indéniable, mais aussi un revers de médaille moins reluisant. Les fonctions de traduction automatique, de synthèse vocale ou de simplification lexicale ouvrent des perspectives prometteuses pour les publics les plus éloignés des parcours éducatifs et de formations traditionnels. Mais cela doit

être mis en regard de la nécessité d’être constamment formé à son utilisation, du fait de la rapidité des évolutions, pour en tirer les avantages attendus.

Au regard du rythme effréné des évolutions technologiques, l’IA générative peut apparaître comme un risque d’amplification des inégalités existantes et une source de déceptions à la hauteur des promesses qui ont été formulées. Il y a urgence à agir, car la situation actuelle est porteuse de risques systémiques, sociétaux. Les moyens financiers et humains manquent pour sensibiliser, éduquer, former et accompagner les individus dans ces transformations. Si de nombreuses initiatives se développent malgré ce manque de ressources, elles restent souvent dispersées et non reliées par une vision commune de l’intérêt général.

Ce rapport s’adresse aux décideurs (État, territoires, entreprises, société civile) et toute personne ayant un pouvoir d’agir sur les transformations numériques de la société. Les discours tenus sur l’IA ont un impact, il faut en être conscient. Il est donc indispensable que chacun adopte une posture responsable en employant par exemple un langage précis et en refusant les simplifications abusives. Cela peut sembler bien peu, mais c’est un préalable. Il faut également se donner les moyens pour que la population s’approprie (ou non, par choix) ce dispositif sociotechnique. Cela doit dépasser la simple formation aux usages, afin de fournir à chacune et chacun les compétences et savoirs nécessaires pour comprendre et se faire sa propre opinion de ce que l’IA peut apporter concrètement. C’est encore plus urgent pour les personnes considérées comme éloignées du numérique, qui peuvent vivre comme une relégation définitive, le fait d’être écartées d’un ensemble de technologies qui peut transformer en profondeur la société.

Ce rapport propose des éclairages sur les débats et controverses qui ont lieu autour des promesses et limites de l’IA générative pour les publics les plus en difficulté avec le numérique. Il propose ensuite des recommandations pour faciliter le passage à l’action et pour aider à la prise de décision. Pour y parvenir, ce rapport se structure en trois temps. Une première partie propose des fondements conceptuels nécessaires à une action éclairée. Le premier

chapitre rappelle des concepts, chiffres et réalités vécues par des millions de personnes victimes d'inégalités afin de considérer l'IA comme une solution possible d'émancipation et d'inclusion, sous conditions. Le second chapitre propose une définition de l'IA générative, notamment, afin de partager un cadre commun d'analyse. Ce chapitre tente de satisfaire à la fois une exigence de définition au plus près de la situation de septembre 2025, à l'heure où les technologies évoluent extrêmement vite, et à la fois une exigence de pédagogie. Il fait le choix de ne pas revenir sur l'histoire de l'IA ni de tomber dans une technicité trop complexe, considérant fournir aux décideurs suffisamment d'arguments pour prendre des décisions éclairées ; il existe par ailleurs une vaste littérature à ce sujet. Le troisième chapitre propose une définition de la littératie qui vise à doter chaque citoyenne et citoyen des compétences nécessaires pour comprendre, utiliser, évaluer et produire un discours relatif à un environnement numérique. Il présente également ses objectifs pour l'IA spécifiquement, dans une perspective d'inclusion et d'émancipation.

Une deuxième partie s'intéresse au développement de l'esprit critique nécessaire pour comprendre les transformations en cours. Le chapitre quatre propose d'abord une lecture des discours et représentations qui accompagnent le développement et la diffusion de l'IA générative et de ses usages. Le chapitre cinq s'intéresse spécifiquement au développement du sens critique, la pierre angulaire d'un grand nombre de recommandations évoquées par les personnes auditionnées dans le cadre de ce rapport. Il aborde les enjeux de la littératie en IA à pouvoir identifier les risques cognitifs et informationnels, autant que la capacité à développer un regard critique sur les impacts systémiques des systèmes d'intelligence artificielle (SIA), notamment les impacts environnementaux. Il aborde enfin les capacités éducatives et émotionnelles qui sont nécessaires pour faire face à ces nouveaux enjeux.

La troisième partie, au travers du sixième chapitre, propose des leviers opérationnels pour une action systémique, pour une littératie en IA qui transforme l'IA d'un facteur potentiel d'exclusion en instrument d'inclusion et d'émancipation citoyenne.

Appréhen| der le double visage de l'IA géné| rative pour l'inclusion

01

↳ Limites du numérique dans la réduction des inégalités

En mars 2024, la Commission de l'intelligence artificielle indiquait, dans son rapport "IA : notre ambition pour la France", que l'IA (de manière générale, pas uniquement générative) "peut être mise au profit de la réduction des inégalités sociales, de la prospérité collective et de la qualité du travail" ; bien qu'elle conditionne ces possibles qualités à un "projet politique" et à de "l'engagement politique" (Commission de l'intelligence artificielle, 2024). D'autres acteurs, comme la Commission Supérieure du Numérique et des Postes (CNSP), sont restés plus prudents, estimant que "le développement de l'IA est porteur d'opportunités mais peut également potentiellement creuser la fracture numérique qui existe déjà dans notre pays, en raison de l'illectronisme mais également celle liée à l'insuffisance des infrastructures et des équipements notamment dans la ruralité" (CNSP, 2024).

De nombreux travaux en sciences humaines et sociales montrent depuis des années que "le numérique" n'a jamais été une réponse évidente aux problématiques d'inclusion (dans l'administration, dans le recours aux droits, à l'école...) et à la résorption des inégalités sociales (Plantard, 2015 ; Deville, 2023). Ainsi, en 2025, ce sont au minimum entre 11 et 16 millions de Français qui ont des difficultés avec "le numérique". Puisque que l'on sait par ailleurs que "tous les individus ne sont pas en situation de tirer les mêmes profits des technologies numériques [et que] l'exclusion liée aux technologies se superpose aux facteurs d'exclusion existants"(ANCT et al., 2023), il est possible que les personnes qui sont aujourd'hui victimes d'inégalités sociales et confrontées à des facteurs d'exclusion soient également celles qui rencontrent des difficultés pour s'approprier et tirer bénéfice des technologies numériques et plus précisément de l'IA générative.

Ces observations révèlent des inégalités numériques persistantes qui s'aggravent. Un rapport de l'OCDE confirme qu'au-delà des baisses générales de niveau, les inégalités de compétences s'accroissent dans la plupart des 31 pays participants à leur enquête 2022-2023 (OCDE, 2024a).

Cette situation est particulièrement préoccupante dans un contexte de transformation technologique rapide dans lequel les adultes ayant de faibles compétences cumulent souvent plusieurs vulnérabilités qui les rendent moins aptes à bénéficier des évolutions technologiques et économiques en cours (OCDE, 2024b). Ces constats soulignent que l'IA générative ne peut constituer un levier d'inclusion que si les compétences de base sont d'abord renforcées. À ce titre, dans le cadre de personnes en reconversion professionnelle en informatique, une personne nous indiquait lors d'une audition constater qu'*“ il faut de bonnes bases pour utiliser l'IA générative. Et cela passe d'abord par ne pas faire d'IA ”*.

↳ Opportunités et risques de l'IA générative

Le tableau ci-après présente quelques opportunités et risques de l'IA générative pour les personnes considérées comme éloignées du numérique, ou rencontrant des difficultés dans son accès ou son usage.

	Opportunités	Risques
Accès aux dispositifs	<p>Gratuité = accès pour toutes et tous.</p> <p>Accès immédiat en ligne ou via une application, sans création de compte (identification) = peu de compétences requises pour l'usage.</p>	<p>Gratuité = exploitation des données (l'utilisateur est "le produit").</p> <p>Gratuité = version limitée de la solution, par rapport aux versions payantes (premium) qui ne limitent pas les résultats.</p> <p>Passage possible à la monétisation des offres gratuites à court ou moyen termes. Il y a donc un risque d'accroissement des inégalités d'accès entre celles et ceux qui pourront payer, et les autres.</p>
Accès aux savoirs	<p>Accès immédiat au savoir = "on pose une question, on a la réponse".</p>	<p>Considérer pour vrai les "hallucinations" (création de résultats absurdes ou erronés) et la reproduction des biais, ou encore être victime de manipulations (Serret, 2024).</p> <p>Renforcement des inégalités par un développement des savoirs inégaux.</p>
Assimilation des savoirs	<p>Parcours pédagogique personnalisé et interactif pour une meilleure assimilation des savoirs et savoir-faire.</p>	<p>L'outil devient une "béquille cognitive" (Andler, 2023) au profit d'une délégitimation de l'apprentissage.</p>



	Opportunités	Risques
Interaction	<p>Interaction grâce au langage naturel qui permet de surmonter plus facilement des obstacles liés à la lecture et à l'écriture¹ ; notamment pour les personnes en situation de handicap.</p> <p>Personnalisation des interactions lors d'échanges et dans les réponses apportées.</p> <p>Projection de l'IA comme outil et compagnon, au service de " l'Humain ".</p>	<p>Le langage naturel a des codes qu'il faut acquérir, connaître. Si s'exprimer à l'oral peut lever certaines barrières, il peut également reproduire celles déjà existantes. Ainsi, les outils d'IA génératives pourraient empêcher, du moins ralentir, le développement des compétences langagières, de réflexion, de calcul, nécessaires à l'autonomie des individus.</p> <p>Risque d'attachement et de dépendance affective à des systèmes d'IA générative.</p> <p>Risque de " déshumanisation " de certaines relations, perte du lien social.</p> <p>Facilitation de l'usurpation d'identité (<i>phishing</i> par e-mail, <i>spear phishing</i>, <i>smishing</i>, <i>vishing</i>, <i>pharming</i>, hameçonnage par réseaux sociaux), facilitation de fraude (HUB France IA, 2024).</p>
Création	<p>Création débridée pour des personnes peu enclines au maniement des outils informatiques (génération de textes, d'images animées ou statiques, de code informatique, etc.).</p> <p>Soutien à la rédaction d'écrits demandant une compréhension de codes formels ou spécialisés (administratifs, juridiques, lettres de motivation, rédaction de mails...).</p>	<p>Homogénéisation des productions.</p> <p>Distinction des productions par des personnes qui maîtrisent les outils et qui ont un capital culturel plus grand.</p> <p>Plagiat / triche.</p>



1. En France, 10% des personnes âgées de 18 à 64 ans éprouvent des difficultés dans les domaines fondamentaux de l'écrit. Par ailleurs, les personnes en difficulté à l'écrit utilisent moins Internet et font moins de démarches administratives en ligne. Source (Insee & DEPP, 2024).

	Opportunités	Risques
Productivité	“Gain de productivité les plus importants pour les employés les moins qualifiés et productifs” (Commission de l'intelligence artificielle, 2024).	<p>“Ces études ne permettent pas encore de conclure à un effet sur un horizon de quelques années” (Commission de l'intelligence artificielle, 2024).</p> <p>Difficulté de l'accès à l'emploi, en début de carrière notamment, pour les professions les plus exposées à l'IA générative.</p>

Tableau : Quelques opportunités et risques de l'IA générative pour les publics dits éloignés du numérique

IA et emploi des jeunes : au-delà de l'automatisation

Une récente étude de l'Université de Stanford a mis en avant un repli de l'emploi dans les secteurs où certaines tâches sont “ automatisées ” par l'IA, mais pas nécessairement dans ceux où elles sont “ augmentées ” (Brynjolfsson, Bharat, Chen, 2025). Beaucoup de commentaires ont pourtant déduit un remplacement direct des jeunes actifs par les modèles d'IA. Cependant, comme le souligne le sociologue Antonio Casilli, ce type d'étude économique souffre d'un biais méthodologique majeur puisqu'elle assimile le travail à une addition de tâches, sans prise en compte des dimensions relationnelles, politiques et organisationnelles que l'automatisation ne capte pas (Casilli, 2025). Pour ce dernier, cette réduction conduit à sous-estimer le redéploiement des jeunes vers le micro-travail ou d'autres emplois précaires, souvent invisibles dans les statistiques traditionnelles, et à masquer la transformation qualitative du travail sous l'effet de l'IA.

Dans le cas de personnes en situation de handicap, le numérique peut constituer un levier d'autonomie, en facilitant par exemple l'apprentissage, les interactions sociales, l'accès aux droits ou encore la gestion du quotidien. Toutefois, il peut devenir un "surhandicap" lorsque les services numériques ne sont pas conçus de manière accessible, en particulier pour les personnes en situation de handicap² cognitif ou intellectuel (Emmaüs Connect et al., 2024).

Dans le champ de l'IA, deux grands axes d'usages émergent : l'amélioration de l'accessibilité numérique (contenus, interfaces) et le soutien à la vie quotidienne (mobilité, autonomie domestique). Dans le premier cas, les avancées du traitement automatique du langage ont permis le développement d'outils de transcription ou de communication adaptés aux personnes malentendantes ou non verbales (RogerVoice, TypingTalk AAC...). Les agents conversationnels proposés par les systèmes d'IA (SIA) peuvent également s'avérer utiles pour aider les personnes souffrant de troubles de l'apprentissage, en adaptant les contenus aux besoins individuels, ou encore intéressants pour aider les personnes neurodivergentes (autisme, TDAH...) en tant que compagnons numériques ; bien que des précautions restent de rigueur quant au risque d'une dépendance affective excessive (Richter, 2025). Dans le second cas, l'IA permet d'interagir avec l'environnement physique, par exemple via des applications de description visuelle pour les malvoyants (OOrion).

À plus long terme, les assistants numériques pourraient jouer un rôle majeur dans l'autonomisation. Néanmoins, ces solutions demeurent souvent coûteuses ou complexes à utiliser, ce qui posent également des questions relatives à leur pérennité et à la dépendance des personnes à leur égard. D'autres aspects comme les enjeux éthiques, notamment en matière de confidentialité des données personnelles restent également cruciaux. Il est alors utile de rappeler la nécessité - pour ne pas dire l'évidence - de

2. Soit environ 2 millions parmi les 7,7 millions de Français déclarant une limitation fonctionnelle sévère.

co-construire ces outils et dispositifs avec les personnes concernées (handicap sensoriel, cognitif, moteur...), afin d'éviter que des solutions technologiques soient pensées "à côté" de leurs besoins réels et quotidiens.

↘ Illectronisme : évolution et limites des mesures

Pour l'Insee, l'illectronisme est "la situation d'une personne ne possédant pas les compétences numériques de base (rechercher des informations en ligne, communiquer en ligne, utiliser des logiciels, protéger sa vie privée, résoudre des problèmes en ligne) ou ne se servant pas d'Internet (incapacité ou impossibilité matérielle de l'utiliser dans les trois derniers mois)" (Bendekkiche, Viard-Guillot, 2023). Les chiffres officiels montrent une amélioration apparente : de 19% en 2017 à 15,4% en 2021 (Cousteaux, 2019 ; Bendekkiche, Viard-Guillot, 2023). L'Insee estime ainsi que "l'illectronisme a diminué de 3 points entre 2019 et 2021, dans le contexte de la crise sanitaire". Cependant, 28% des internautes conservent des capacités numériques faibles dans au moins un domaine, la protection de la vie privée étant la compétence la moins maîtrisée (Bendekkiche, Viard-Guillot, 2023).

Ces statistiques masquent une réalité plus complexe, puisque la multiplicité des termes employés tels que "fracture numérique", "illectronisme", "vulnérabilité numérique", "éloignement numérique", etc., reflète des approches politiques différentes. Conséquemment, les politiques publiques d'inclusion ne ciblent pas les mêmes publics ni les mêmes priorités (ex. formation des seniors vs. lutte contre la pauvreté numérique des jeunes) selon qu'il est question "d'illectronisme" ou de "vulnérabilité numérique".

Cette diversité de termes révèle également l'absence de consensus méthodologique. Le chiffre largement repris de "13 millions" de personnes éloignées du numérique constitue ainsi "une

extrapolation fondée sur des statistiques univariées et décontextualisées qui néglige la complexité et la diversité du phénomène” (Clerget, Plantard, Bléhaut, 2024). Face à ces limites, certains proposent “une quantification sous la forme d’un halo rendant compte de l’éloignement plus ou moins marqué des individus, reposant sur leur ressenti subjectif”, qui évalue à 31,5% de la population (16 millions de personnes) le nombre de Français concernés par l’éloignement numérique (ANCT, CREDOC, Université Rennes 2 CREAD-M@rsouin, 2023)³.

↘ Inégalités structurelles non numériques

L’illectronisme touche principalement les 60 ans et plus, les 20% de la population les plus modestes, et les personnes sans diplômes ou CEP (Bendekkiche, Viard-Guillot, 2023). Cette répartition révèle que l’éloignement numérique s’enracine dans des inégalités sociales préexistantes plutôt qu’il ne les crée. Ainsi, “l’âge, la situation géographique, le milieu social et le niveau de revenus, le niveau de certification scolaire, et plus largement les niveaux de capitaux (culturel, économique, social et numérique) [...] ont pour particularité d’être [des variables] interreliées et souvent interdépendantes les unes des autres, voire cumulatives” (CREDOC, 2023). Conséquemment, les usages numériques des classes populaires “restent moins variés et moins fréquents que ceux des classes moyennes et supérieures [...] Les individus non diplômés ont plus de mal à s’adapter à la dématérialisation des services administratifs, font moins de recherches, pratiquent moins les achats, se lancent très rarement dans la production de contenus. Bref, il y aurait en quelque sorte un “internet du pauvre”, moins créatif, moins audacieux, moins utile en quelque sorte...” (Pasquier, 2018).

3. Plus précisément, 22,9% de la population de 18 ans et plus (11,5 millions de personnes) peuvent être considérées comme éloignées du numérique, car “ne se sentant pas compétentes dans l’utilisation d’Internet” ; auxquels il faut ajouter 8,8% (4,5 millions de personnes) qui est “non-internaute”.

↳ Compétences et capacités

Compétences

L'analyse de la littérature révèle une évolution significative de la conception des compétences numériques, qui dépasse la simple maîtrise technique pour intégrer des dimensions sociales, cognitives et émancipatrices plus complexes. La notion de compétence peut intégrer “les savoirs, savoir-faire, savoir-être, mais aussi le devoir, vouloir, savoir et pouvoir agir en situations. Pour être plus précis, être compétent c'est non seulement avoir des ressources (connaissances techniques ou scientifiques, savoir-être, savoir-faire techniques, méthodologiques, capacités cognitives...) mais aussi savoir les sélectionner, les mobiliser et les organiser de façon pertinente en situation (et donc en fonction de cette situation). En somme, une personne peut avoir acquis des connaissances, savoir-faire, etc. mais ne sera pas reconnue compétente si elle ne sait pas les combiner et les mobiliser en situation. Cela implique donc que les étudiants puissent construire et mettre en œuvre concrètement, dans l'action, des combinaisons de ses propres ressources comme de celles de son environnement, appropriées et pertinentes. La compétence renvoie donc également à des dimensions collectives, évolutives et réflexives” (Université Rennes 2, 2025).

Cette approche situe la compétence dans l'action contextualisée plutôt que dans la simple possession de connaissances. Elle reconnaît également que les compétences se construisent et s'actualisent dans l'interaction sociale et professionnelle, dépassant l'individualisme méthodologique traditionnel. L'École Polytechnique Fédérale de Lausanne (EPFL) développe ainsi une approche des compétences transversales qui englobe par exemple “la résolution de conflits, la pensée critique, la communication intra- et interpersonnelle, l'inclusion, la durabilité, le raisonnement éthique”, pour répondre aux exigences d'un monde professionnel où “les trajectoires comportent de multiples facettes” (EPFL, 2022).

Capabilités

La théorie des capabilités d'Amartya Sen offre un cadre d'analyse différent de la conception élargie des compétences, puisqu'elle distingue ce que les individus savent faire (compétences) de ce qu'ils peuvent réellement accomplir compte tenu de leurs ressources et de leur environnement (capabilités) (Sen, 1999). Appliqué au numérique, ce cadre implique l'accès aux équipements, la qualité de la connexion, mais aussi la capacité à mobiliser ces ressources pour les exploiter, le temps disponible, le soutien social, ou encore la confiance en soi. Cette approche est par exemple adoptée par l'Agence nationale de la cohésion des territoires (ANCT) qui propose de redéfinir l'éloignement numérique à partir de la capacité des individus à utiliser le numérique pour accroître leur bien-être et leur pouvoir d'agir (ANCT et al., 2023).

Des approches complémentaires

Bien que comparables, ces deux approches présentent des nuances importantes. D'un côté, la conception élargie des compétences se concentre sur l'enrichissement et la diversification des compétences individuelles et collectives, en intégrant des dimensions techniques, sociales et réflexives dans une logique de développement des capacités d'action. De l'autre, l'approche par les capabilités adopte une perspective plus systémique en analysant les possibilités réelles de développer et mettre en œuvre des compétences selon les opportunités et les moyens d'agir qui vont dépendre des contextes et cadres familiaux ou professionnels (organisation, management, outils à disposition, etc.). Elle s'intéresse donc plus précisément aux conditions sociotechniques d'actualisation des potentialités.

Ces évolutions conceptuelles témoignent d'une transformation profonde dans l'appréhension des compétences numériques, qui ne peuvent plus être réduites à de simples habiletés techniques. Elles s'inscrivent désormais dans une perspective plus large d'émancipation et de développement du pouvoir d'agir des individus dans un environnement numérique complexe et en constante évolution.

Ces approches ne sont toutefois pas exemptes de limites et soulèvent des questions importantes. D'une part, la conception élargie des compétences risque de reproduire une vision normative qui standardise les pratiques au détriment de la diversité des appropriations numériques, tandis que l'approche par les capacités peut conduire à une individualisation excessive des difficultés, masquant les déterminants structurels de l'exclusion numérique. D'autre part, ces deux approches posent la question cruciale de qui détient le pouvoir de définir ce qui constitue une "compétence" ou une "capacité" numérique. Si les institutions et les acteurs économiques dominent largement cette définition, les publics concernés restent souvent peu associés à l'élaboration des référentiels. Enfin, **ces évolutions conceptuelles transforment profondément les politiques publiques de formation, en orientant les dispositifs vers une logique d'évaluation individualisée et de certification standardisée qui peut entrer en tension avec les approches d'éducation populaire privilégiant l'émancipation collective et la critique sociale.**

↳ Capital numérique : entre reproduction et émancipation

Le "capital numérique" articule trois dimensions : l'accès matériel, les compétences d'usage et la capacité de transformation stratégique des pratiques numériques en avantages sociaux (Ragnedda, 2018 ; Ruiu & Ragnedda, 2020). Comme les autres formes de capital analysées par Bourdieu, il se transmet inégalement selon les trajectoires sociales. En d'autres termes, les familles qui possèdent déjà des ressources économiques et culturelles transmettent plus facilement à leurs enfants les compétences numériques avancées. Toutefois, si l'accès initial aux ressources numériques reste conditionné par les situations économiques ou culturelles, certaines dimensions comme les usages stratégiques ou la réflexivité numérique peuvent se construire dans des contextes extrascolaires, professionnels ou communautaires. Concrètement, cela signifie qu'un individu peut développer une expertise numérique par ses

expériences lors d'une formation professionnelle ou en autoformation, au sein d'une association, ou encore grâce à des réseaux structurés mais plus informels (comme une "guilde" pour un jeu vidéo par exemple (Servais, 2020)). Cet argument reste néanmoins discuté, dans la mesure où les enfants issus de milieux favorisés développent plus aisément une littératie numérique critique, bénéficiant d'un accompagnement familial, de réseaux d'acteurs diversifiés, d'équipements de qualité et d'un environnement qui valorise l'usage réflexif du numérique, consolidant ainsi leur position sociale.

Former à l'IA générative ne peut se limiter à l'acquisition de techniques de *prompt* (où comment rédiger une "requête"). Il est nécessaire de développer un véritable capital culturel numérique permettant de transformer l'IA en levier d'autonomie, pour transformer les usages en opportunités concrètes, choisies et non subies, plutôt qu'en instruments de dépendance. Les premières études sur l'adoption de l'IA générative confirment cette analyse en dévoilant que les clivages éducatifs traditionnels se reproduisent dans l'usage de ces nouvelles technologies (Sidoti & McClain, 2025). **Il faut donc concevoir des politiques d'inclusion qui s'attaquent aux inégalités structurelles (sociales et économiques) plutôt qu'à leurs manifestations techniques.**

↘ IA et "fracture générationnelle"

L'âge comme variable descriptive non explicative

Dans leur étude conjointe, l'ANCT, le CREDOC et l'Université Rennes CREAD-M@rsouin indiquent que les personnes âgées de 70 ans ou plus comptent "une part de non-internautes nettement supérieure à la moyenne" (ANCT et al., 2023). Mais, pour les auteurs, c'est avant tout l'année de naissance, plus que l'âge, qui permet de comprendre cet écart relatif à "un effet de génération" puisque "les personnes les plus âgées ont, pour la plupart d'entre elles, découvert Internet tardivement, et n'ont pas bénéficié d'une socialisation scolaire et/ou professionnelle aux outils numériques, ce qui tend à

expliquer l'écart constaté entre les 'jeunes retraités' et les retraités les plus âgés". Conséquemment, poursuivent-ils, "l'âge n'est en rien un facteur explicatif de l'éloignement numérique, tout au plus il s'agit d'un facteur descriptif" (ANCT et al., 2023).

Cela vaut également pour les plus jeunes. Bien que le mythe des digital natives soit régulièrement déconstruit, et cela depuis de nombreuses années, l'âge nourrit toujours la croyance selon laquelle les enfants nés "avec" le numérique se "débrouilleraient" mieux que les adultes. Pourtant, la part de jeunes considérés comme éloignés du numérique reste élevée (20% des moins de 25 ans).

Shadow "IA" et fragmentation des collectifs de travail

L'IA générative marque une rupture dans les modalités d'introduction technologique en entreprise. Contrairement aux logiciels métiers déployés traditionnellement par les directions informatiques, elle se diffuse "par le bas", à l'initiative des salariés eux-mêmes. Cette appropriation spontanée peut générer des usages masqués, non déclarés au management par les collaborateurs. Phénomène qualifié de *shadow IA*, qui correspond à "une constellation d'usages informels plus ou moins visibles par lesquels les salariés s'approprient les outils d'intelligence artificielle générative en dehors des cadres officiels" (Ferguson, 2025).

Cette situation révèle elle aussi des écarts générationnels complexes. Comme une évidence, l'arrivée de l'IA générative ravive l'idée d'un "fossé" entre des jeunes supposés à l'aise avec ces outils et des personnes plus âgées, qui le seraient moins. Pourtant, la familiarité avec l'IA générative dépend avant tout de l'exposition, de la formation et du contexte professionnel, non de l'âge. Une fois de plus.

Le véritable enjeu réside dans le risque d'inégalité collective quand ces différences d'usages ne sont ni discutées ni partagées dans les équipes. Comme le souligne le sociologue Yann Ferguson, "on remarque au sein des entreprises des jeunes qui travaillent de

plus en plus vite, mais qui n'ont pas l'expertise pour s'assurer de la qualité du travail produit, et des anciens qui avancent plus lentement, mais avec suffisamment d'expertise pour délivrer un travail de qualité. Le collectif est donc fracturé et ne communique plus, notamment parce que les jeunes utilisent beaucoup l'IA pour se rassurer. Au lieu de se tourner vers ceux qui détiennent le savoir, ils considèrent l'IA comme une sorte de robot omniscient, capable de leur dire s'ils ont raison ou non. Cela pose un problème majeur : non seulement le collectif n'évolue pas à la même vitesse et n'a pas le même rapport à la connaissance, mais il cesse aussi d'interagir, entraînant une perte de connaissance collective" (Consul, 2025).

Dans ce contexte, les auditions menées dans le cadre de ce rapport convergent vers une recommandation forte : *“avant toute démarche de déploiement généralisé, il faut avoir une démarche d'expérimentation accompagnée par du dialogue”*. Cette approche préalable permet d'identifier les usages pertinents, d'anticiper les résistances et de co-construire des règles d'usage acceptées par tous. Le dialogue social apparaît ainsi comme un prérequis à l'efficacité, comme le mentionnait une autre personne auditionnée : *“dès lors où il n'y a pas de démarches réfléchies et co-construites avec les travailleurs quant à l'introduction des technologies, alors nous sommes à peu près assurés, du côté des représentants des travailleurs, que les technologies seront mal déployées”*. Cette co-construction devient d'autant plus nécessaire que l'IA générative soulève des enjeux inédits de responsabilité, de qualité du travail et de préservation des savoirs expérientiels.

L'élaboration de chartes éthiques du numérique ou IA représente un levier intéressant pour mobiliser les parties prenantes, notamment les collaborateurs (dans un cadre professionnel) ou les citoyens, indépendamment du niveau de compétences techniques. Toutefois, quel que soit le contexte (à l'échelle d'une entreprise ou d'une collectivité), une charte n'a de valeur que si le processus réunit l'ensemble des parties prenantes. En effet, c'est son processus de co-construction plus que son contenu final qui prime souvent le plus (considérant que le contenu est finalement assez prévisible). Enfin, une fois publiées, de telles chartes resteront

insuffisantes et symboliques si elles ne sont pas accompagnées d'actions et de mise en pratique dans les équipes.

Conséquemment, **le défi particulier des TPE-PME réside dans l'inadéquation du dialogue social existant, souvent peu formalisé dans ces structures. Il convient donc de développer des modalités adaptées, s'appuyant sur les réseaux professionnels, les chambres consulaires ou les organisations patronales sectorielles pour créer des espaces d'échange et de formation mutualisés.**

Définir et nommer les systèmes d'IA

02

↳ Repères conceptuels

Pluralité et évolutions du concept

Les premières définitions formelles de l'IA datent de l'été 1956, lorsque John McCarthy, Marvin Minsky, Claude Shannon et Nathan Rochester réunissent, à Dartmouth, une poignée de scientifiques pour tester l'hypothèse selon laquelle “tout aspect de l'apprentissage ou toute autre caractéristique de l'intelligence peut, en principe, être décrit avec suffisamment de précision pour qu'une machine puisse le simuler” (McCarthy et al., 1955). Huit ans plus tard, Minsky proposait une définition plus pragmatique, considérant l'IA comme “la science de faire faire à des machines des choses qui demanderaient de l'intelligence si elles étaient faites par des humains” (Minsky, 1998).

Désormais, l'expression “intelligence artificielle” recouvre une diversité de disciplines et de méthodes informatiques visant à obtenir les mêmes résultats que la cognition humaine : apprentissage, raisonnement, création ou prise de décision automatisée. Comme le rappelait Claude Kirchner dans un entretien réalisé avec Renaissance Numérique, l'IA ne constitue pas une entité homogène, mais “le nom d'un domaine scientifique et technologique du numérique. C'est une petite partie des domaines scientifiques du numérique qui incluent globalement les sciences informatiques, les sciences des réseaux de communication, le traitement du signal, l'automatisation, la robotique, l'apprentissage machine, les sciences des données, ainsi que les domaines des mathématiques, de la physique et des sciences cognitives” (Van de Blankevoort, 2025).

Parler de “l'IA” comme d'un tout homogène est donc trompeur, puisque ce terme recouvre aujourd'hui très majoritairement des systèmes fondés sur l'apprentissage profond (réseaux de neurones), qui se déclinent selon leurs applications et architectures. Ainsi, on peut distinguer les systèmes spécialisés, conçus pour des tâches précises comme le diagnostic médical ou la détection de fraudes, des systèmes génératifs, capables de produire

du contenu original à partir d'instructions en langage naturel, bien que ces frontières tendent à s'estomper avec l'émergence de modèles hybrides.

De la même manière qu'il n'existe pas "un numérique", terme générique qui englobe des infrastructures, des équipements, des services, des pratiques et des usages variés, il n'y a pas UNE intelligence artificielle. Il suffit d'ailleurs d'évoquer plusieurs solutions commerciales de *chatbot* (ou agent conversationnel) d'intelligence artificielle pour se représenter cette diversité : ChatGPT (OpenAI), Copilot (Microsoft), Le Chat (Mistral), DeepSeek (DeepSeek), ou encore des solutions spécialisées en génération d'images statiques ou animées comme Midjourney (Midjourney) ou encore Dream Machine (Luma Labs).

Quelques ressources pour approfondir

Nous proposons ici une définition opérationnelle de l'IA générative adaptée aux besoins des décideurs, privilégiant la pédagogie sur l'exhaustivité technique. Il existe de nombreuses ressources complémentaires qui permettent d'approfondir différents aspects de l'intelligence artificielle selon les besoins spécifiques de chaque lecteur. En voici quelques exemples :

- [L'article "Intelligence artificielle"](#) de Wikipédia constitue une bonne porte d'entrée pour comprendre l'ensemble des notions et des techniques relatives à l'intelligence artificielle, et distinguer notamment intelligence artificielle, apprentissage automatique et apprentissage profond.
- [Le MOOC "Introduction à l'Intelligence Artificielle"](#), gratuit, du Campus numérique public.
- Une série de trois articles de la CNIL sur l'explicabilité de l'IA (Berkouk, Arfaoui, Pialat, 2025) représente une excellente ressource pour le lecteur qui s'intéresse aux problématiques sociotechniques et de littératie qui sont au cœur de ce rapport.
- Enfin, le Pôle d'Expertise de la Régulation Numérique (PERen) propose une collection de publications scientifiques et techniques de grande qualité.

Risques des amalgames

L'Observatoire international sur les impacts sociétaux de l'IA et du numérique considère l'IA comme un ensemble de “technologies informatiques capables de raisonner, créer, apprendre ou décider” (OBVIA, 2024). L'OCDE, quant à elle, définit un système d'IA comme un ensemble d'algorithmes qui, “pour des objectifs définis, infère à partir d'entrées comment générer des sorties” (OCDE, 2023). D'autres organismes, ou encore le récent règlement européen sur l'intelligence artificielle (Union européenne, 2024), peaufinent à leur tour leur propre définition, signe qu'aucun consensus n'a encore réellement émergé. Cette pluralité définitionnelle n'est pas problématique en soi, mais elle exige une vigilance particulière dans les débats publics pour éviter les confusions.

Malheureusement, on observe trop souvent des amalgames problématiques entre différentes technologies étiquetées “IA”. D'une part, certaines comparaisons hâtives mélangent des applications qui, bien qu'utilisant des fondements techniques similaires (apprentissage profond), soulèvent des enjeux sociétaux radicalement différents. Un modèle de prédiction climatique et ChatGPT posent des questions très distinctes en termes de transparence, de biais ou d'impact social. D'autre part, certaines productions créent des amalgames technologiques en traitant ensemble des systèmes aux fondements techniques différents. Par exemple, l'avis du Conseil Économique Social et Environnemental (CESE), “Impacts de l'intelligence artificielle : risques et opportunités pour l'environnement” (Tatot & Vermot Desroches, 2024), mêle des algorithmes d'optimisation logistique (relevant de la recherche opérationnelle), des objets connectés préexistants (capteurs simples sans capacité d'apprentissage) et l'IA générative récente.

Il est donc essentiel de reconnaître la diversité des systèmes d'IA, chacun reposant sur des principes, présentant des limites, générant des impacts et comportant des biais qui lui sont propres. Cette distinction constitue le préalable indispensable pour exercer un jugement éclairé et ne pas confondre les capacités, les risques et les usages de ces technologies. (cf. [proposition 1 - “Éviter les amalgames”](#)).

↘ IA générative et IA à usage général

“ Les efforts que consacre l’IA [ici la “ profession ”], avec une énergie désormais décuplée, à la conception de systèmes toujours plus intelligents, c’est-à-dire, selon elle, toujours plus proche de l’intelligence humaine [ainsi que le fait qu’elle] vise aussi à conférer à ces systèmes une autonomie aussi large que possible, et à la limite de l’autonomie totale [représente un double objectif qui] est à la fois incohérent, dangereux et inutile. Incohérent parce que l’autonomie véritable, celle des humains, est de l’ordre du mystère : un SAI ou un robot autonome en ce sens est un concept inintelligible. Dangereux, parce qu’il ne peut conduire qu’à des systèmes difficiles à comprendre et plus encore, par définition, à contrôler. Inutile, parce que ce dont nous avons vraiment besoin, ce qui nous aiderait de manière peut-être décisive à résoudre certains problèmes pressants, c’est d’outils dociles, puissants et versatiles, et non de pseudo-personnes munies d’une forme inhumaine de cognition ” (Andler, 2023, 17-18).

IA générative

L’intelligence artificielle générative désigne des systèmes capables de créer de nouveaux contenus (textes, images, sons) en s’appuyant sur des modèles entraînés sur de vastes corpus de données. Ces systèmes peuvent générer des contenus originaux et résoudre des tâches inédites, ce qui explique à la fois leur potentiel et leur imprévisibilité.

Les premiers modèles génératifs ont d’abord été critiqués comme de simples “ perroquets stochastiques ” (Bender et al., 2021), capables de reproduire des formes textuelles cohérentes sans compréhension véritable. Cette caractérisation, largement vraie en 2021, demande à être nuancée au regard des évolutions des modèles récents. Les systèmes actuels montrent des capacités remarquables de compréhension textuelle complexe, incluant

l'attribution d'intentions sophistiquées et le raisonnement sur des concepts abstraits. Cette vitesse des évolutions est l'un des traits les plus marquants du domaine de l'intelligence artificielle, et l'une des difficultés à pouvoir la commenter.

Cependant, malgré des évolutions extrêmement rapides depuis une dizaine d'années notamment (Bengio, 2025), les systèmes d'IA conservent des limites significatives dans certains domaines (raisonnement relatif à un déplacement spatial par exemple) et leur fonctionnement interne reste largement opaque. Cette opacité et ces limites expliquent l'un des phénomènes les plus problématiques de l'IA générative, puisqu'elle peut inventer des sources, mal interpréter un document, produire un calcul erroné, tout en présentant ces erreurs comme des informations exactes. Or, malgré les progrès en cours ou annoncés à court et moyen termes, les “grands modèles de langage produiront toujours des résultats plausibles mais faux, même avec des données parfaites, en raison de limites statistiques et informatiques fondamentales” (Kalai et al., 2025).

IA à usage général (“GPAI”)

Un système ou modèle d'intelligence artificielle à usage général (*General-Purpose AI* - GPAI) “peut être utilisé et adapté à un large éventail d'applications pour lesquelles il n'a pas été conçu intentionnellement et spécifiquement”⁴. Contrairement aux systèmes d'“IA étroite” (voir encadré ci-dessous), conçus pour des tâches précises, les systèmes GPAI peuvent être adaptés et spécialisés dans des domaines d'application variés. Un même modèle de base peut ainsi servir à la fois à l'assistance médicale, à la création littéraire, à l'analyse de données financières, ou malheureusement à la génération de désinformation.

4. CNIL, “Système d'IA à usage général”, disponible sur : <https://www.cnil.fr/fr/definition/systeme-dia-usage-general>

IA étroite (“Narrow AI”)

Les systèmes d'IA étroite (Russell & Norvig, 2021) sont spécialisés dans l'accomplissement d'une tâche précise : traduction automatique, reconnaissance d'images, génération de texte, diagnostic médical assisté, ou optimisation logistique. Ces systèmes sont souvent performants dans leur domaine de spécialisation, mais ne peuvent pas s'adapter à des tâches différentes de celles pour lesquelles ils ont été conçus. C'est le niveau de maturité actuel de l'IA, déployé massivement dans l'industrie, les services et le quotidien des individus (algorithmes de recommandation des plateformes de streaming, assistants vocaux, systèmes de navigation GPS, ou encore outils de modération automatique des contenus en ligne).

IA générale (“AGI”) et autres concepts philosophiques

L'IA générale (*Artificial General Intelligence* - AGI) est un concept philosophique polysémique désignant une machine capable d'accomplir toute tâche cognitive humaine avec une flexibilité et une adaptabilité comparables à celles de l'intelligence humaine (Goertzel, 2014).

Certains considèrent ainsi que les systèmes techniques actuels les plus avancés, notamment les grands modèles de langage, se rapprochent de l'AGI, car ils démontreraient des capacités émergentes pouvant résoudre des tâches dans des domaines très variés (diagnostic médical, traduction, programmation) et parfois sur des problèmes absents de leurs données d'entraînement.

Toutefois, quand l'IA à usage général (GPAI) est une réalité réglementaire déjà inscrite dans le règlement européen sur l'intelligence artificielle de 2024 (AI Act), l'AGI demeure un horizon théorique encore débattu, à la définition fluctuante et non unanimement arrêtée. Les débats portent par exemple sur les critères de reconnaissance d'une véritable AGI (compréhension vs. performance, conscience vs. simulation), les délais pour y parvenir,

et la désirabilité d'un tel objectif. En quelque sorte, quand certains voient une IA étroite devenir de plus en plus polyvalente, d'autres y voient l'émergence d'une IA générale.

Bien que les capacités actuelles des systèmes d'IA évoluent, il convient de distinguer les avancées techniques récentes de ces concepts philosophiques. Leur mobilisation excessive dans le débat public peut conduire à des représentations déformées des capacités réelles des systèmes d'IA contemporains, et par là-même à sous-estimer les enjeux présents au profit de spéculations futures.

IA forte et “super-intelligence artificielle (ASI)”

Les expressions d’“IA forte”, de “super-intelligence artificielle (ASI)” (Bostrom, 2014) ou encore de “conscience de la machine” apparaissent fréquemment dans les débats et les médias. L’“IA forte” est un concept philosophique désignant une machine dotée d'états mentaux réels, incluant conscience et compréhension (Searle, 1980). L'ASI, quant à elle, correspond au scénario d'une intelligence artificielle surpassant largement l'humain dans tous les domaines cognitifs.

↳ Avancées et limites des systèmes actuels

Évolution vers une autonomie accrue mais limitée

L'évolution actuelle de l'IA générative vise à surmonter ses principales limitations, notamment les hallucinations, le manque de fiabilité factuelle, et la difficulté à accomplir des tâches complexes nécessitant plusieurs étapes.

Récemment, trois évolutions font évoluer l'IA générative d'un modèle isolé vers un système sociotechnique capable de générer une réponse à partir de données qui sont issues de sources autres que celles avec lesquelles un modèle a été entraîné (RAG), d'utiliser plusieurs outils pour répondre au mieux à la requête d'un utilisateur, sans que ce dernier n'en ait fait explicitement la demande (MCP), ou encore de s'auto-corriger (IA agentique).

	RAG (Retrieval-Augmented Generation)	MCP (Model Context Protocol)	IA agentique
Définition	Le RAG permet à un modèle d'IA d'interroger en temps réel une base documentaire (interne ou externe), puis de construire sa réponse en s'appuyant sur les éléments récupérés, plutôt qu'uniquement sur les données avec lesquelles il a été entraîné (Lewis et al., 2020).	Le MCP décrit comment un modèle peut orchestrer séquentiellement l'usage d'outils externes comme des interfaces de programmation API, bases de données, calculatrices, navigateurs web, ou logiciels spécialisés (OpenAI, 2023). Cette approche transforme l'IA générative d'un simple générateur de texte en un orchestrateur capable d'interagir avec un environnement numérique.	L'IA agentique dote les modèles d'une boucle appelée " <i>Plan-Act-Observe-Evaluate</i> " (PAOE) : l'agent planifie une tâche, sélectionne dynamiquement les outils appropriés (RAG, MCP, etc.), exécute les actions prévues, observe les résultats, puis s'auto-évalue pour ajuster sa stratégie si nécessaire.
Fonction principale	Consultation de bases documentaires en temps réel.	Orchestration d'outils externes.	Planification et auto-évaluation.
Avantage	Réduction des hallucinations.	Automatisation de tâches complexes.	Autonomie dans l'exécution.

	RAG (Retrieval-Augmented Generation)	MCP (Model Context Protocol)	IA agentique
Exemple d'usage	Un utilisateur consulte un chatbot pour calculer le montant de sa retraite. Grâce au RAG, l'IA va consulter les derniers textes réglementaires en vigueur, extraire l'article pertinent, puis générer une réponse personnalisée accompagnée de la citation source.	Un utilisateur demande à un dispositif d'IA : "extraire le texte d'un PDF, calcule la moyenne d'une colonne de données, puis rédige une synthèse juridique". Cette suite d'instructions peut être automatiquement décomposée et exécutée sans intervention humaine, chaque étape mobilisant un outil approprié, sans que l'utilisateur dicte à la machine comment procéder étape par étape.	Un agent de veille scientifique programmé pour "collecter chaque semaine les articles sur la littérature en IA" va interroger plusieurs bases de données via RAG, classer les résultats selon des critères définis, rédiger une synthèse, puis relancer automatiquement la recherche si la couverture thématique est jugée insuffisante.

Tableau : Trois exemples d'avancées de l'IA générative contemporaine

Limites structurelles

Malgré ces avancées, les IAs actuelles, même les plus performantes, restent limitées par plusieurs facteurs structurels :

- **La qualité des données** : Bien que des techniques de post-entraînement (alignement par retour humain) permettent de corriger partiellement les biais initiaux liés aux données d'entraînement, la qualité des données de base reste déterminante. Des données biaisées, incomplètes ou erronées influencent durablement les performances, même après correction. Par ailleurs, certains modèles utilisent des données issues d'articles scientifiques qui n'ont pourtant pas fait l'épreuve de validation par la communauté scientifique ou qui ont été rétractés (Ananya, 2025).

- **Les critères d’alignement** : Si l’entraînement initial d’un modèle vise la prédiction optimale, la phase d’alignement ultérieure (*fine-tuning*, apprentissage par renforcement à partir du *feedback* humain ; ou *Reinforcement Learning from Human Feedback* - RLHF) intègre les valeurs de l’organisation développeuse. Ces choix d’alignement (définition du “comportement souhaitable” d’une IA par exemple, critères éthiques, contraintes commerciales) orientent fondamentalement le comportement final des systèmes. C’est ainsi que des IAs seront plus ou moins orientées politiquement, ou selon certaines valeurs éthiques.

- **Les limites cognitives** : Les IAs actuelles ne possèdent pas d’intentions propres, elles manipulent des représentations symboliques selon des règles mathématiques. Les débats restent toutefois ouverts quant à de possibles capacités de “compréhension”, ou d’“états de conscience”, selon les définitions consacrées.

- **La production d’hallucinations** : Ce phénomène résulte de la fonction première des grands modèles de langage (*Large Language Model* - LLM), c’est-à-dire compléter du texte de manière statistiquement plausible. Par exemple, si un LLM doit fournir une réponse quant à la météo qu’il fera demain, alors qu’il n’a pas accès à des données en temps réel, il générera une réponse à partir de données statistiques notamment. La réponse fournie sera jugée probable ou vraisemblable par le modèle, mais restera “inventée”. Si les modèles récents montrent une réduction notable du taux d’hallucinations, celles-ci, selon des chercheurs d’OpenAI, ne disparaîtront pas du fait de la nature même de ces systèmes (Kalai et al., 2025).

De la même manière, l’IA agentic, présentée depuis quelques mois comme la “nouvelle révolution de l’IA” n’est pas radicalement nouvelle, puisqu’il s’agit techniquement d’agents conversationnels (LLMs) spécifiquement entraînés pour utiliser des outils externes de façon autonome. Or, la possible amélioration de l’autonomie des systèmes IA s’accompagne d’une complexification des tâches qu’ils peuvent accomplir.

Anthropic a par exemple mené une expérience consistant à laisser gérer un distributeur de snacks à leur agent conversationnel Claude (Anthropic, 2025). L'agent était équipé de plusieurs outils lui permettant d'interagir avec les consommateurs, de s'approvisionner auprès de différents fournisseurs et de fixer les prix en fonction de la demande. L'analyse des comportements de l'agent ont permis de montrer son incapacité à réaliser des tâches nécessitant une planification au-delà de quelques heures. Si l'on assiste néanmoins à un accroissement rapide de cette limite (qui était de l'ordre de quelques minutes pour GPT4 et qui atteint entre 1h et 4h30 pour GPT5) (METR, 2025), l'accumulation d'erreurs sur des tâches longues peut créer des effets "boule de neige" rendant les résultats incohérents.

Finalement, la progression spectaculaire de l'IA générative, et de l'IA agentique ici, ne doit pas masquer les limites encore bien présentes : les erreurs se multiplient au fil des étapes, les agents perdent vite le fil sur des projets complexes, et la gestion de tâches étalées dans le temps reste très fragile. Autrement dit, la vitesse de l'innovation technique coexiste avec une fragilité persistante des résultats, ce qui peut créer un décalage entre l'enthousiasme médiatique et la réalité opérationnelle (cf. chapitre 4). Précisons toutefois que les progrès concernant les solutions d'IA spécialisées dans la génération d'images ou de vidéos sont particulièrement spectaculaires.

↘ L'IA est un dispositif sociotechnique

L'IA n'est pas un outil

L'adage voudrait que l'IA générative - le numérique par extension - soit "un outil" au service de "l'Humain". En d'autres termes, les usages positifs ou négatifs de l'IA dépendraient avant tout de son utilisation, donc de ce que chacun d'entre nous en fait.

Or, quand un utilisateur demande à une IA générative de produire un texte ou une image, la qualité de la réponse ne dépend pas simplement de la qualité du *prompt* (la requête). C'est le système, dans son ensemble, qui conditionne cette réponse, soit les données avec lesquelles il a été entraîné, celles qu'il va mobiliser pour générer le contenu, la manière dont a été configurée l'IA, etc. Interagir avec un système d'IA générative, c'est interagir avec une chaîne de médiations (Akrich, Callon, Latour, 2006), voir un système d'innovation (Amable, 2003), faits de décisions et d'actions humaines et non-humaines (relatives à des environnements, des cultures, etc.), qui orientent les choix de l'utilisateur, les conditionnent, les façonnent...

Cette perspective permet de dépasser la vision de l'IA comme simple outil afin de la considérer comme un système sociotechnique, c'est-à-dire comme un agencement complexe d'éléments techniques et sociaux, donc économiques et politiques. Parler d'IA, c'est parler d'un ensemble d'éléments interconnectés, des infrastructures numériques (centres de données, bases de données, algorithmes, interfaces...) aux pratiques sociales (usages et détournement d'usages, construction de sens, représentations collectives...). Loin d'être neutre, un dispositif sociotechnique modifie par exemple les interactions humaines, transforme les processus de communication et restructure les modalités de l'action collective.

Prenons l'exemple d'un *chatbot* reposant sur un système d'intelligence artificielle utilisé dans un service public d'aide sociale, qui permet aux usagers de poser des questions sur leurs droits, les démarches à effectuer ou encore les aides disponibles (Guillaud, 2025). Sa finalité est de pouvoir interagir avec des publics divers (souvent en difficulté), selon des objectifs politiques d'inclusion, et de respecter des normes de clarté, de confidentialité et d'accessibilité. Dans ce cas, **l'IA générative est dispositif sociotechnique qui structure la manière dont les citoyens accèdent à l'information et à leurs droits, qui modifie leurs interactions et leurs rapports aux institutions. Il influence aussi bien les comportements des utilisateurs que les pratiques professionnelles des agents publics. Ainsi, il participe à une recomposition des formes d'action publique et de médiation sociale.**

Le system-prompt et la non-neutralité des agents conversationnels

Le *system-prompt* est un ensemble de consignes préprogrammées qui définit le “comportement” d'un agent conversationnel. Systématiquement pris en compte avant chaque réponse apportée à un utilisateur, il oriente fondamentalement les réponses produites selon les critères définis par le fournisseur d'IA ; que ce soit en termes de restrictions de réponses à apporter ou d'orientations politiques à privilégier. Par exemple, le *system-prompt* de Claude Opus (Anthropic) enjoint le modèle à avoir un ton “naturel, chaleureux et empathique” et à “garder une conversation amicale et appropriée s'il suppose interagir avec un mineur”⁵. L'usage de *system-prompt* soulève plusieurs enjeux critiques, dont l'opacité (la majorité des *system-prompts* ne sont pas publics) et une certaine tendance à l'anthropomorphisme des agents conversationnels (plus le modèle sera apprécié, plus la rétention de l'utilisateur sera forte). L'effet du *system-prompt* sur le comportement des agents conversationnels est souvent mal perçu ou compris, révélant un domaine encore hautement expérimental malgré son influence déterminante sur les interactions.

5. Claude Opus 4.1. System Prompts : <https://docs.claude.com/en/release-notes/system-prompts#august-5-2025>

Même le marteau n'est pas qu'un simple outil

Pour bien comprendre que l'IA n'est pas simplement un outil, il est utile de mobiliser la notion de “pharmakon”, que Bernard Stiegler (2007) reprend à la philosophie grecque (et à Jacques Derrida ; 1968).

En grec, le mot *pharmakon* signifie à la fois remède et poison. Platon l'utilise dans le dialogue *Phèdre*, où il parle de l'écriture alphabétique comme d'un *pharmakon*. Pour le philosophe, c'est d'abord un poison, qui peut devenir un remède sous certaines conditions ; à la fois un outil de mémoire, et un danger pour la mémoire vivante. Le web, lui aussi peut être dit pharmacologique, “ parce qu'il est à la fois un dispositif technologique associé

permettant la participation⁶ et un système industriel dépossédant les internautes de leurs données pour les soumettre à un marketing omniprésent et individuellement tracé et ciblé par les technologies du *user profiling*⁷.

La notion de *pharmakon* exprime donc l'encapsulation des virtualités positives et négatives dans la technique, et non dans l'usage qui en est fait dans un contexte spécifique. Une technique est un remède et un poison, par nature, et non pas par celle de son usage.

Conséquemment, la référence partielle au concept de *pharmakon*, telle que nous l'observons de plus en plus dans les débats autour de l'IA générative, a cela de problématique qu'elle peut laisser penser que la technique serait "neutre" (en cela qu'elle serait à la fois positive et négative), et dépendrait exclusivement de l'usage. Or, en plus de ne jamais être neutre, la technique est variée et plurielle (Simondon, 1958). Ainsi, même ce fameux marteau, qui ne serait "qu'un outil", comme l'IA, n'est ni neutre ni unique. Faire le choix d'un marteau, c'est devoir faire un choix réfléchi en fonction des usages qui peuvent être faits. Par exemple, le marteau de charpentier a ceci d'avantageux qu'il a deux fonctions : la face plate de sa tête permet d'enfoncer des clous, quand sa face à deux crochets séparés par une rainure permet de les retirer d'où ils sont logés. Le marteau d'électricien, lui, est spécialement conçu pour fixer des attaches de conduits ou de fils électriques.

Faire de l'IA un outil, c'est lui donner une unicité qu'elle n'a pas, c'est enlever au débat politique un pluralisme des choix technologiques qui doivent exister et être discutés. D'aucuns ont parfaitement compris l'intérêt politique de ces systèmes technologiques, à l'image de Grok, l'assistant d'IA de xAI, qui se positionne ouvertement comme une IA moins "censurée" que les autres, et défendant les principes idéologiques de son propriétaire, Elon Musk.

6. Notons que le "et" est important (et en italique dans la citation originale), car toute technique n'est pas bien "ou" mal, mais à la fois un "remède" et un "poison".

7. *Pharmakon* (pharmacologie), <https://arsindustrialis.org/pharmakon>

↳ Gouvernance, choix éthiques, sociétaux et civilisationnels

La polyvalence qui caractérise les modèles d'IA à usage général soulève des enjeux particuliers de gouvernance et de responsabilité. Contrairement aux systèmes d'IA étroite, dont les usages sont prédéfinis et circonscrits, les concepteurs d'un modèle GPAI ne peuvent, par définition, anticiper tous les usages qui en seront faits une fois déployé. Cette limitation complique considérablement l'attribution des responsabilités en cas de dommage et pose la question cruciale de savoir qui, du fournisseur du modèle ou de l'utilisateur final, porte la responsabilité des préjudices causés.

L'AI Act tente de répondre à ces défis en établissant un cadre réglementaire spécifique pour les modèles GPAI. Le règlement impose des obligations particulières aux fournisseurs de modèles GPAI, notamment en matière d'évaluation des risques, de documentation des capacités et de transparence sur les limitations du système (Commission européenne, 2024). Ces obligations se renforcent considérablement pour les modèles classés comme présentant un “risque systémique”, généralement ceux dont la puissance de calcul dépasse un seuil de 10^{25} FLOPS pour l'entraînement.

Panorama des risques liés à l'IA

Le développement de l'intelligence artificielle expose les sociétés humaines à trois catégories de risques (CeSIA⁸ ; AI Safety atlas⁹ ; Bengio, 2025) :

- Les risques d'utilisation malveillante, c'est-à-dire les usages de l'IA dans un objectif intentionnel de nuisance (armes biologiques, armes autonomes, attaques d'infrastructures, phishing...).
- Les risques de désalignement et de fiabilité, qui apparaissent, malgré de bonnes intentions humaines, lorsque les systèmes d'IA poursuivent des objectifs différents de ces intentions (mauvaise définition des objectifs par l'humain, systèmes d'entraînements qui génèrent des comportements inattendus...) ou qu'ils produisent de mauvaises informations (hallucinations).
- Les risques systémiques, c'est-à-dire les bouleversements socio-économiques liés au déploiement de cette technologie transformatrice à grande échelle (concentration du pouvoir, hausse du chômage de masse résultant de l'automatisation, hausse de la désinformation...).

8. Centre pour la Sécurité de l'IA. Risques de l'IA. <https://www.securite-ia.fr/panorama#risques-ia>

9. AI Safety Atlas. Misuse Risks. <https://ai-safety-atlas.com/chapters/02/03>

Comme le souligne le rapport de Renaissance Numérique sur la coopération mondiale en matière de sécurité de l'IA, “l'absence de consensus scientifique mondial sur les risques liés à l'intelligence artificielle, ainsi que sur les méthodes d'évaluation et de réduction de ceux-ci, constitue l'un des défis les plus urgents” (Renaissance Numérique, 2025). Cette incertitude scientifique complexifie d'autant l'établissement de règles de gouvernance adaptées et nécessite une approche préventive et adaptative qui intègre pleinement les dimensions éthiques, environnementales et sociales. Dans ce contexte, l'OCDE a identifié dix risques prioritaires, notamment l'incapacité des mécanismes de gouvernance à suivre le rythme d'évolution rapide de l'IA et les courses au développement qui compromettent les investissements dans la sécurité (OCDE, 2024c).

Bien que de nombreux cadres et pratiques soient en développement à l'échelle mondiale, des lacunes significatives persistent dans la validation, la standardisation et la mise en œuvre, particulièrement pour l'identification et l'atténuation de risques sans précédent (Bengio, 2025). Deux défis centraux demeurent ainsi. D'abord, la difficulté de prioriser les risques en raison de l'incertitude concernant leur gravité et probabilité d'occurrence. Ensuite, la complexité de déterminer les rôles et responsabilités appropriés à travers la chaîne de valeur de l'IA pour inciter à une action efficace.

Au-delà des questions de gouvernance technique, l'IA générative soulève des enjeux éthiques multidimensionnels qui touchent aux fondements de nos sociétés démocratiques. Comme le rappelle Claude Kirchner, président du Comité consultatif national d'éthique du numérique, l'éthique constitue “ une réflexion relative aux conduites humaines et aux valeurs qui les fondent ” (Van de Blankevoort, 2025). Cette dimension réflexive devient cruciale face aux nombreux défis de l'IA générative, notamment aux questions de vie privée, de biais algorithmiques ou encore de justice que posent ces systèmes (Data For Good, 2023). Les biais algorithmiques représentent un défi particulier pour les modèles GPAI, car leur capacité de généralisation peut amplifier et diffuser des discriminations préexistantes dans les données d'entraînement. Ces biais peuvent toucher différentes dimensions (ethniques, de genre, d'âge, géopolitiques, etc.) créant des inégalités systémiques dans l'accès aux services et aux opportunités. La justice algorithmique devient ainsi un enjeu central, nécessitant des mécanismes d'audit et de correction des discriminations. Parallèlement, les enjeux environnementaux de l'IA générative sont considérables et méritent une attention particulière ([cf. chapitre 5](#)).

Concevoir une littératie en IA

03



Premiers éléments de définition

Continuum d'apprentissage

La notion de littératie trouve ses racines dans la capacité à lire et à écrire. Aujourd'hui, il est question de littératie numérique, financière, médiatique... Cette évolution traduit une transformation des compétences nécessaires à la citoyenneté dans un monde numérisé.

En France, la Commission d'enrichissement de la langue française a proposé dès 2017 l'expression "habileté numérique" pour traduire le concept anglo-saxon de *digital literacy*¹⁰. Cette expression souligne la dimension pratique et opérationnelle de ces compétences, puisqu'elle définit la qualité d'une personne à accomplir une tâche avec adresse, dextérité, tout en laissant la possibilité d'interpréter ou de réussir cette tâche par différents moyens. L'UNESCO va plus loin en définissant la littératie comme "un continuum d'apprentissage et de maîtrise de la lecture, de l'écriture et de l'utilisation des nombres tout au long de la vie et elle fait partie d'un ensemble plus large de compétences qui inclut les compétences numériques, l'éducation aux médias, l'éducation au développement durable et à la citoyenneté mondiale, ainsi que les compétences spécifiques de l'emploi" (UNESCO, 2025). Cette approche dynamique reconnaît que les compétences évoluent et s'enrichissent constamment.

Il existe une richesse conceptuelle autour de la littératie qui peut toutefois être source d'ambiguïté. Olivier Le Deuff souligne ainsi la coexistence – parfois concurrente – de multiples formes de littératies : médiatique, informationnelle, numérique (Le Deuff, 2012). Leur hiérarchie demeure floue et, pour le chercheur, la prolifération de "nouvelles littératies" parfois éphémères témoigne autant d'une adaptation aux environnements en mutation que d'une dispersion des référentiels et des dispositifs de formation. D'autres recherches soulignent l'intérêt d'une "translittératie", notion

10. Commission d'enrichissement de la langue française. Vocabulaire de l'éducation et de l'enseignement supérieur. https://www.education.gouv.fr/bo/17/Hebdo6/CTNR1636064K.htm?cid_bo=112291

complémentaire, définie comme l’aptitude à naviguer entre différents écosystèmes informationnels en modulant ses stratégies cognitives (Haider & Sundin, 2022). Ce foisonnement a néanmoins ouvert la voie à des structurations conceptuelles durables. Dès 2004, Stuart A. Selber proposait une grille de lecture pour la *computer literacy* articulée autour de trois axes : compétences fonctionnelles (utiliser les outils), critiques (analyser les systèmes) et rhétoriques (produire un discours avec les technologies) (Selber, 2004). Depuis, plusieurs évolutions ont été réalisées dans le cadre du développement de l’intelligence artificielle, avec l’ajout de la dimension éthique notamment (UQAM, 2024). D’autres approches et modèles de littératie existent, certains proposant par exemple de considérer des compétences cognitives, émotionnelles et sociales (Eshet, 2004).

Culture, éducation et émancipation

La littératie ne se limite pas à une série de compétences techniques comme le suggère celle d’habitabilité. Elle implique une capacité plus large à comprendre, évaluer et mobiliser des savoirs dans des contextes complexes. C’est pourquoi elle ne peut être dissociée des concepts d’éducation et de culture.

L’éducation renvoie aux dispositifs de transmission des savoirs, qu’ils soient formels (à l’école), informels (en famille) ou non formels (dans les associations par exemple). La culture, elle, permet d’envisager les dimensions symboliques et sociales de notre rapport aux savoirs. À ce titre, le Conseil national du numérique (devenu depuis Conseil de l’intelligence artificielle et du numérique) rappelait l’impératif “de développer une culture numérique qui ne se limite pas à l’acquisition de compétences techniques, mais qui permette aussi une prise de conscience collective des enjeux historiques, sociaux et anthropologiques des transformations techniques contemporaines” (CNNum, 2023).

L'IA, la porte d'entrée vers la littératie numérique

“Si IA et numérisation ne sont pas équivalentes, il ressort que la première ne saurait se concevoir sans la seconde : les deux doivent être considérées comme deux vagues successives. Ainsi, rattraper la première vague de la numérisation sans considération pour la seconde reviendrait à prendre d’emblée un retard conséquent” - Cédric Villani (2018, 64).

Dès 2013, le Conseil national du numérique soulignait l'importance de la littératie numérique pour permettre à chacun de prendre “l’ascenseur social numérique” (CNNum, 2013). Cette littératie combine plusieurs types de compétences :

- **instrumentales** (utiliser les outils),
- **créatives et productives** (concevoir ou modifier des contenus),
- **informationnelles** (chercher et évaluer l'information),
- **réflexives** (développer un regard critique).

L'UNESCO poursuit cette approche en y ajoutant la capacité d'accéder à l'information, de l'évaluer, de la créer et de la partager en toute sécurité. Il s'agit de développer un esprit critique face aux flux d'information et aux risques comme la désinformation ou les discours de haine (UNESCO, 2025)¹¹.

Il est primordial de considérer que la littératie en IA ne se substitue pas à la littératie numérique, elle en est un prolongement avec lequel elle doit s'articuler.

De fait, son déploiement doit susciter un sursaut collectif, car, comme le déclarait une personne auditionnée, “*la littératie en IA est certainement un prolongement de la littératie numérique. Ce qui change, c'est que la littératie numérique n'était pas un sujet crucial. Or, l'ampleur des outils d'IA générative a remis sur le devant de la scène la nécessité d'une littératie numérique commune*”.

11. Dans ce rapport, nous traduisons “digital literacy” par “littératie numérique” en français ; quand bien même l'UNESCO utilise “culture numérique” dans sa traduction officielle.

En cela, l'IA générative constitue une opportunité pour aborder les enjeux plus vastes de la culture numérique, des données, des interfaces, du rapport à l'information et à la vérification des faits, de l'économie de l'attention, des biais algorithmiques, de la gouvernance des plateformes ; autant de sujets déjà inscrits dans DigComp 2.2, qui classe l'IA parmi les “ technologies transversales ” (Vuorikari, Kluzer, Punie, 2022). Cette opportunité est aussi logique car, comme l'exprimait une autre personne auditionnée, *“ la littératie en IA appelle des compétences en littératie des données, puisque comprendre ce qu'est la donnée, comment elle est utilisée, c'est mieux comprendre ce qu'est une IA générative, qui s'entraîne sur des jeux de données ”*, ou qui en mobilise.

Plus encore, l'émergence de l'IA générative révèle un changement de paradigme dans notre rapport au numérique. Alors que les outils traditionnels nécessitent une maîtrise technique préalable, l'IA générative permet une interaction directe par le langage naturel, rendant visible des enjeux auparavant techniques ou abstraits. L'utilisateur qui interroge un *chatbot* découvre concrètement les questions de véracité des données (d'où viennent ces informations ?), de gouvernance algorithmique (pourquoi cette réponse plutôt qu'une autre ?), de droits fondamentaux (mes données personnelles sont-elles utilisées ?) et de communs numériques (qui contrôle ces modèles ?). Ainsi, **la littératie en IA agit comme un révélateur pédagogique puisqu'elle rend tangible des enjeux numériques fondamentaux et constitue une porte d'entrée intuitive vers des littératies spécialisées (données, algorithmes, média et information, etc.) qui pouvaient jusqu'alors perçues comme trop techniques par le grand public.**

➤ Objectifs politiques et réglementaires

Ambitions et angles morts de l'AI Act

L'adoption, en juin 2024, du Règlement européen sur l'intelligence artificielle (AI Act) marque une étape majeure pour encadrer de manière structurée l'usage de l'IA dans l'Union européenne. Au-delà du classement des systèmes par niveau de risque (inacceptable, élevé, limité, minimal), l'article 4 introduit pour la première fois une obligation de "maîtrise de l'IA" (*AI literacy*) à l'intention de l'ensemble des acteurs intervenant dans le cycle de vie d'un système : fournisseurs, intégrateurs, déployeurs et utilisateurs finaux.

L'article 4 indique que "les fournisseurs et les déployeurs de systèmes d'IA prennent des mesures pour garantir, dans toute la mesure du possible, un niveau suffisant de maîtrise de l'IA pour leur personnel et les autres personnes s'occupant du fonctionnement et de l'utilisation des systèmes d'IA pour leur compte, en prenant en considération leurs connaissances techniques, leur expérience, leur éducation et leur formation, ainsi que le contexte dans lequel les systèmes d'IA sont destinés à être utilisés, et en tenant compte des personnes ou des groupes de personnes à l'égard desquels les systèmes d'IA sont destinés à être utilisés" (Union européenne, 2024). Le considérant 20 précise que cet objectif poursuit des finalités claires : garantir des décisions éclairées, protéger les droits fondamentaux, la santé, la sécurité, et favoriser un contrôle démocratique des technologies d'IA (Union européenne, 2024).

Pour faciliter la mise en œuvre, la Commission européenne a mis en ligne un *Living Repository* (Commission européenne, 2025a), recensant des exemples d'initiatives volontaires en matière de littératie en IA. Siemens, SAP ou encore Orange publient par exemple des modules internes et une entreprise comme Workday intègre une formation annuelle à l'IA responsable dans son code de conduite. Toutefois, ce référentiel n'est pas contraignant et ne garantit pas, en l'état, la conformité aux exigences de l'AI Act.

En complément, l'article 66 du règlement confère au Comité IA un rôle de soutien dans la promotion de la littératie, en particulier via un "guichet européen", destiné à éclairer les acteurs économiques sur les obligations réglementaires (Récital 20 de l'AI Act). Enfin, l'AI Act inscrit la littératie en IA dans une perspective plus large que le seul cadre professionnel. Le Digital Education Action Plan 2030 souligne la nécessité de construire une culture de l'IA dans l'éducation, en particulier au sein de l'enseignement primaire et secondaire. Cette approche vise à soutenir la montée en compétence des enseignants et à encourager une adoption éthique de l'IA dès le plus jeune âge (Commission européenne, 2025b).

Malgré ces nombreuses ambitions, plusieurs zones d'ombre subsistent :

- Aucun indicateur chiffré ni seuil minimal de compétences ne sont définis pour qualifier une "littératie de base".
- Les modalités pédagogiques (durée, format, contenus) restent à l'initiative des acteurs concernés.
- L'obligation vise formellement fournisseurs et déployeurs. Il n'est pas précisé comment l'application s'étend aux sous-traitants, intégrateurs ou plateformes open source.

Cette ambition mérite désormais de se traduire en référentiels contraignants, en indicateurs de suivi et en financements dédiés, afin que "maîtriser l'IA" cesse d'être un simple vœu réglementaire pour devenir une réalité partagée par tous les citoyens européens.

En France, un cadre législatif en construction...

La loi n° 2024-449 du 21 mai 2024 visant à sécuriser et réguler l'espace numérique (SREN) marque une première étape dans la reconnaissance institutionnelle de la littératie en IA. En complétant, l'article L.312-9 du Code de l'éducation impose qu'à l'issue de l'école primaire et du collège, chaque élève reçoive une attestation certifiant sa sensibilisation au "bon usage des outils numériques et de l'intelligence artificielle, (...) aux dérives et aux risques liés à

ces outils” ainsi qu’aux mécanismes de désinformation (Journal officiel de la République française, 2024).

Pour traduire cette obligation légale, Élisabeth Borne a annoncé le 4 juin 2025 le déploiement du parcours “Pix IA¹²”, obligatoire dès la rentrée 2025 pour les élèves de 4^e et de 2^{de} des voies générales, technologiques et professionnelles, et mis à disposition de tous les collégiens, lycéens et professeurs volontaires. Ce dispositif de 30 minutes à 1h30 “évaluera les connaissances et les compétences en IA de l’élève, puis lui proposera un programme personnalisé de formation avec des modules sur des thématiques précises comme les bases du prompting (comment poser des questions à l’IA), le fonctionnement des IA génératives, ou encore la gestion des données et les impacts environnementaux” (Ministère de l’Éducation nationale, 2025). Si cette initiative constitue une avancée notable dans l’intégration curriculaire de la littératie en IA, plusieurs défis opérationnels demeurent : formation des enseignants, adaptation aux inégalités territoriales d’équipement, et articulation avec les programmes disciplinaires existants.

... mais un manque de stratégie globale.

Parallèlement, le plan national “Osez l’IA”, lancé en juillet 2025, vise à “accélérer la diffusion de l’intelligence artificielle (IA) dans toutes les entreprises françaises, et en particulier dans les petites et moyennes entreprises et les entreprises de taille intermédiaire” (MEF, 2025a). Bien qu’il prévoit une plateforme de formation “accessible à tous” (apprentis, demandeurs d’emploi, salariés et dirigeants), l’objectif affiché reste prioritairement celui de “convaincre” les entreprises d’adopter l’IA.

12. Le groupement d’intérêt public (GIP) PIX opère une plateforme publique d’évaluation des compétences numériques fondée sur le Cadre de référence des compétences numériques (CRCN) et constitue un levier important pour renforcer la littératie numérique. Si elle couvre 16 domaines, dont certains en lien direct avec les enjeux de l’intelligence artificielle, son usage reste largement concentré dans le secteur éducatif (4,7 millions d’élèves, étudiants et enseignants accompagnés chaque année).

L'ensemble des actions précédemment énumérées révèlent une tension intéressante : peut-on simultanément développer une littératie en IA et promouvoir l'adoption généralisée de l'IA ? Le risque est que les enjeux économiques priment sur la formation d'un regard distancié, réduisant la littératie à une simple appropriation technique au détriment de la dimension évaluative et éthique pourtant centrale dans une démarche éducative émancipatrice.

Pour autant, des initiatives témoignent d'une volonté d'associer les citoyens à cette réflexion. Les Cafés IA¹³, initiés par le Conseil national du numérique depuis 2024 et désormais coordonnés par la Mission IA (MEF, 2025b), proposent une démarche de mobilisation citoyenne avec des kits de formation destinés aux médiateurs numériques pour “engager les publics dans une réflexion constructive sur ces technologies”.

Ces initiatives et projets demeurent toutefois fragmentés et manquent d'une vision stratégique globale cohérente qui ne soit pas exclusivement dictée par les impératifs économiques. Cette absence de stratégie nationale pour l'IA se manifeste par une juxtaposition d'actions de sensibilisation, d'éducation et de formation qui peinent à s'articuler autour d'objectifs partagés (et évalués) de littératie pour l'ensemble de la population (cf. proposition 6 - “Organiser des États généraux du numérique et de l'IA pour l'éducation et la formation”).

↘ Finalités démocratiques et économiques

La littératie en intelligence artificielle poursuit des finalités qui dépassent l'acquisition de simples compétences techniques pour embrasser un projet d'émancipation démocratique. Cette ambition s'articule autour de trois dimensions complémentaires qui révèlent la complexité des enjeux contemporains.

13. À propos des Cafés IA : <https://cafeia.org/a-propos-de-cafe-ia>

Enjeu démocratique : former des citoyens éclairés

Au-delà des dimensions techniques et économiques, la littératie en IA porte un enjeu plus profond de préservation de l'agir démocratique dans des sociétés où les décisions collectives risquent d'être de plus en plus déléguées à des systèmes algorithmiques opaques. En formant des citoyens capables de comprendre, questionner et orienter ces transformations technologiques, elle constitue un rempart contre la dépolitisation des choix de société.

Cette ambition transforme la littératie d'un simple enjeu d'acquisition de compétences et de connaissances en impératif démocratique, condition nécessaire pour que l'IA serve l'émancipation humaine plutôt qu'elle ne renforce les inégalités existantes et les mécanismes de domination.

Au cœur de cette approche réside la volonté de doter chaque citoyen des connaissances et compétences nécessaires pour naviguer de manière autonome dans un environnement technologique en mutation constante. Nous considérons ainsi **la littératie en IA comme la capacité de comprendre, d'utiliser et d'évaluer de façon critique des dispositifs d'IA**. Cette finalité démocratique se décline en cinq objectifs opérationnels :

1. **Comprendre les principes de base de l'IA**, qui déterminent ses capacités et ses limites, sans nécessairement maîtriser les aspects techniques plus complexes. Il s'agit de démystifier ces technologies pour avoir une représentation réaliste de ce que les systèmes d'IA peuvent ou non faire, ainsi que de distinguer les évolutions probables des promesses spéculatives.
2. **Utiliser ces dispositifs de manière efficace, adaptée et sécurisée**, en développant une maîtrise pratique qui permet de tirer parti des opportunités et de limiter les risques.
3. **Évaluer de façon critique les productions de l'IA et ses effets sociétaux**, en comprenant en premier lieu que les réponses apportées par l'IA peuvent être fausses (hallucinations), biaisées, limitées, et qu'elles peuvent avoir des effets sur la désinformation et la manipulation des publics et utilisateurs.

4. **Prendre des décisions éclairées** quant à ses usages de l'IA.
5. **Contribuer aux débats démocratiques**, ne serait-ce que par le fait de décrire et expliciter ses usages, et pouvoir exprimer des besoins ou des attentes, au regard des précédents points.

***Enjeu économique :
redistribuer les bénéfices de l'innovation***

L'histoire des technologies à usage général (machine à vapeur, électricité, numérique) révèle qu'un pays capte durablement les surplus de productivité lorsqu'il élargit massivement la base de compétences capables d'implanter ces innovations dans chaque secteur d'activité (Ding, 2024). Cette leçon historique, observée successivement avec la machine à vapeur, l'électricité et la numérisation, suggère que **l'investissement dans l'éducation et la formation généralisée offre un rendement macroéconomique supérieur aux seuls soutiens à la recherche-développement.**

Ainsi, l'article 4 de l'AI Act, qui pose l'exigence d'une "maîtrise de l'IA" pour les fournisseurs et déployeurs, pourrait inspirer une "clause de diffusion" pour favoriser le développement d'une littératie en IA à grande échelle. Par ailleurs, orienter une part des investissements en IA vers des programmes favorisant l'inclusion de toutes et tous, par la littératie, au-delà des seuls salariés et "talents de demain" et autour d'un projet pour une société apprenante, permettrait de démontrer l'utilité sociale de l'IA (accès aux droits, traduction automatique, simplification des démarches), et du numérique de manière plus générale, afin de créer un marché de diffusion où l'innovation se mesure à l'impact auprès de publics éloignés.

Concrètement, une réallocation d'un petit pourcentage des 644 milliards de dollars d'investissements qui seront réalisés en 2025 (Gartner, 2025) vers la montée en compétences de l'ensemble des acteurs économiques salariés, formateurs et utilisateurs finaux pourrait transformer structurellement les rapports de force économiques et réduire les inégalités d'accès aux bénéfices de l'innovation ; pour un développement économique généralisé profitable à l'ensemble de la société.

Enjeu social : résoudre la tension adoption-émancipation

Comme nous l'évoquions précédemment, la précédente proposition révèle une tension fondamentale entre deux logiques contradictoires. D'un côté, les impératifs économiques et technologiques poussent à une adoption massive et rapide des outils d'IA pour ne pas "manquer le train" de l'innovation. De l'autre, l'exigence démocratique impose de développer préalablement un regard critique et distancié sur ces mêmes technologies. Or, comme le rappelait une personne auditionnée, la formation, et en cela la littératie en IA, c'est aussi "*être capable de perdre du temps pour apprendre. Il faut réussir à trouver un équilibre entre le gain de temps que représente l'IA et la nécessaire perte de temps pour se former initialement*". D'un côté, l'urgence à être sensibilisé, éduqué ou encore formé à l'IA et, de l'autre, la nécessité du temps long pour des processus effectifs.

Cette tension reflète des rapports de pouvoir structurels où les détenteurs de capitaux technologiques ont intérêt à maximiser l'adoption rapide, tandis que l'intérêt général appelle à préserver l'autonomie intellectuelle et la capacité de choix des citoyens. La littératie en IA devient alors l'instrument de résolution de cette tension, en permettant une adoption éclairée plutôt que subie. Précisons toutefois que la littératie n'a pas pour but l'adoption : elle a pour objectif d'éclairer les choix et les décisions. Dans cette perspective, la littératie en IA peut amener des individus à faire le choix, éclairé, de ne plus utiliser de systèmes d'intelligences artificielles, du moins d'en limiter leur recours.

Dans une optique démocratique, d'inclusion, mais aussi de respect des droits et libertés fondamentaux, il convient également d'inscrire dans la loi l'obligation d'alternatives aux démarches dématérialisées, avec un accès à un point de contact physique et téléphonique pour accompagner les citoyens dans leurs démarches administratives, conformément à l'article L.112-8 du Code des relations entre usagers et administration (Commission Supérieure du Numérique et des Postes, 2024). Certaines communes, comme Villeurbanne, ont déjà voté un engagement en faveur d'un droit au non-numérique dans l'accès aux services publics municipaux (Barnes, 2023).

Analyser les discours et usages de l'IA générative

04

↳ Discours et diffusion de l'IA générative

Imaginaires structurellement ambivalents

Tout objet technique porte en lui une double identité : fonctionnelle et fictionnelle. L'intelligence artificielle ne fait pas exception à cette règle fondamentale. La dimension fictionnelle d'un "objet" est constituée d'imaginaires ambivalents, qui oscillent naturellement entre représentations techno-messianiques et techno-pessimistes (Musso, Coiffier, Lucas, 2014). C'est ainsi que les promesses de libération par la technologie côtoient systématiquement les craintes d'asservissement, par exemple. Cette ambivalence s'inscrit dans une tradition cyclique des imaginaires technologiques que l'IA réactualise au travers de mythes anciens (Régnauld, 2025), qui circulent pourtant depuis les années 1960 avec la numérisation de la société (Moles, 1990).

Les discours contemporains sur l'IA illustrent parfaitement cette dynamique. D'un côté, les déclarations grandiloquentes se multiplient. Sundar Pichai (Alphabet) affirme que "l'IA sera le changement le plus profond de notre époque. Elle est plus profonde que le feu ou l'électricité" (Pichai, 2025). Sam Altman (OpenAI) prédisait récemment que "dans une décennie, peut-être que chaque être humain sur Terre sera capable d'accomplir plus que la personne la plus influente aujourd'hui" (Altman, 2025). De l'autre côté, les discours catastrophistes évoquent par exemple une "superintelligence artificielle" susceptible de dépasser les capacités humaines et de menacer l'avenir de l'humanité (quand les premiers en font la promotion).

Stratégies d'énonciateurs

Dans ce contexte, il est important de comprendre que l'IA n'est pas une technologie autonome qui s'impose à nos sociétés, mais le produit de choix humains et organisationnels spécifiques.

L'identification des “littérateurs” de l'IA, pour reprendre l'expression de Patrice Flichy, révèle la multiplicité des acteurs qui contribuent à construire l'imaginaire social attaché à cette technologie (Flichy, 2010). Ces acteurs (entrepreneurs, ingénieurs, journalistes, chercheurs, vulgarisateurs, etc.) ne visent pas nécessairement une analyse critique ou une mise en perspective, mais cherchent souvent à maximiser la visibilité de leur discours dans une logique économique (notamment depuis l'avènement des réseaux sociaux). Ainsi, “la vague de l'IA” dans l'espace public se produit à la faveur de facteurs convergents (technologiques, économiques, médiatiques et politiques) sans être le reflet d'une véritable “révolution scientifique”, promue par les médias et les entreprises mais contestée dans les récits des chercheurs (Bellon, Velkovska, 2023).

Cette construction de l'imaginaire, dans laquelle les médias et les entreprises du secteur jouent un rôle fondamental, parfois en capitalisant sur la méconnaissance et la crédulité des utilisateurs, n'est pas sans conséquences. Elle influence directement les choix politiques, économiques et sociaux concernant le développement et l'usage de ces technologies. Pour que le futur de nos sociétés ne soit pas uniquement dicté par un petit nombre d'entreprises technologiques et de décideurs politiques, il devient essentiel de déconstruire ces imaginaires et de proposer des alternatives.

L'innovation à marche forcée

Même lorsque des figures de premier plan, comme Yann LeCun, directeur scientifique de l'IA chez Meta, rappellent les limites actuelles des modèles de langage — qu'il considère comme dépourvus de compréhension du monde réel ou de capacités de raisonnement élaborées (Mims, 2024) — le récit de la rupture se réinvente immédiatement à travers d'autres promesses : agents autonomes, nouvelles architectures techniques (JEPA - Joint Embedding Predictive Architecture par exemple), intelligence artificielle hybride, ou futures itérations de modèles encore plus puissants, jusqu'à l'intelligence artificielle générale (AGI) qui est régulièrement annoncée comme devant arriver dans quelques mois, quelques années tout au plus.

C'est ainsi que, "face à des bénéfices de l'IA générative inférieurs aux attentes", cette fuite en avant technologique se traduit par "le paradigme émergent de l'IA agentique [qui] suscite un regain d'enthousiasme quant à son potentiel de création de valeur. L'intérêt est presque universel : 58 % des personnes interrogées déclarent rechercher des opportunités de mise en œuvre d'agents et d'assistants, et 40 % d'entre elles sont prêtes à les explorer. Les utilisateurs s'attendent à ce que les agents aient le plus grand impact dans l'automatisation des tâches et des processus. L'automatisation des processus est l'un des principaux domaines dans lesquels les organisations ont rapporté des bénéfices moins importants que prévu des initiatives d'IA générative existantes" (S&P Global, 2025).

Stratégies de légitimation éthique

La multiplication des discours éthiques autour de l'intelligence artificielle génère un paradoxe contemporain : plus les entreprises et institutions proclament mettre "l'humain au centre" de leurs développements technologiques, plus la distance semble se creuser entre ces annonces et les pratiques effectives. En effet, cette rhétorique du *Human-Centric AI* dissimule souvent des priorités de performance technologique et de rentabilité, reléguant au second plan les enjeux sociaux, éthiques et environnementaux (Spencer, 2024).

Le concept de *human washing*, par analogie avec le *green washing* environnemental, désigne précisément cette tendance à réduire l'éthique à une série de procédures de conformité comme la création de comités d'éthique, la nomination de *Chief AI Ethics Officers* ou encore la publication de chartes, sans que ne soient questionnés les objectifs réels de la conception ou le modèle d'affaires sous-jacent. Cette stratégie transforme l'éthique en vitrine communicationnelle, détournant l'attention des enjeux structurels pour la concentrer sur des dispositifs institutionnels formels.

Pour développer une lecture critique de ces stratégies de légitimation, il convient d'examiner l'écart entre les principes affichés et leur opérationnalisation concrète. Dans les champs la *Human-Computer Interaction* (HCI - ou interaction homme-machine) et de la *Human-Centered AI* (HCAI - ou l'IA centrée sur l'Humain), plusieurs principes fondamentaux permettent d'évaluer l'authenticité des démarches éthiques. La transparence et l'explicabilité constituent un premier critère d'évaluation. Bien que cela soit particulièrement difficile en raison de la nature des LLMs, chaque décision algorithmique doit pouvoir être appréhendée (à défaut d'être comprise) par l'utilisateur, transformant le système d'une "boîte noire" opaque en un dispositif décrivant, au moins de manière générale, ses critères de choix et ses sources (un accès simplifié au *system-prompt* pourrait être généralisé). D'une certaine manière, on peut déjà observer cela dans certaines solutions qui décrivent étape par étape leur "raisonnement" avant de produire une réponse. La collaboration interdisciplinaire dès le début d'un projet représente un autre indicateur d'une démarche authentique. Cette approche implique l'inclusion effective de designers, sociologues ou encore éthiciens, mais surtout des utilisateurs finaux et des communautés potentiellement affectées dans les phases de conception. L'enjeu réside moins dans la constitution formelle d'équipes diversifiées que dans l'influence réelle de ces perspectives plurielles sur l'architecture finale du système. La rétroaction, enfin, constitue un pilier de cette évaluation critique. Au-delà des audits ponctuels, cette dimension suppose la mise en place de mécanismes permettant de recueillir et d'intégrer les retours des utilisateurs sur le long terme pour ajuster, corriger et enrichir le système de manière dynamique.

Ces quelques exemples révèlent l'ampleur du défi contemporain. Les rapports récents de la Commission de l'intelligence artificielle (2024) et du Programme des Nations unies pour le développement (PNUD) soulignent que la majorité des projets d'IA, même à haut risque, intègrent rarement des processus d'évaluation participative, d'écoute des populations vulnérables ou d'analyse des impacts indirects sur les inégalités mondiales (UNDP, 2025).

Le développement d'une littératie critique face aux stratégies de légitimation éthique constitue une compétence citoyenne fondamentale dans l'écosystème numérique contemporain. Cette capacité d'analyse permet de distinguer les démarches authentiques des opérations de communication, d'exiger la transparence des processus décisionnels et de participer aux débats démocratiques sur l'orientation des développements technologiques. En transformant les citoyens en évaluateurs éclairés des promesses éthiques, cette littératie contribue à l'émergence d'un rapport de force plus équilibré entre les acteurs économiques, les institutions et les populations concernées par les transformations induites par l'intelligence artificielle générative.

Mythe de la rupture technologique et bulle de l'IA

De nombreux acteurs (gouvernements, institutions, entreprises technologiques, investisseurs, cabinets de conseil ou influenceurs numériques) entretiennent l'idée selon laquelle l'AI générative représenterait une rupture technologique majeure, justifiant à la fois une adoption massive et des investissements conséquents. Cette dynamique se nourrit de l'injonction désormais bien connue à "ne pas manquer le train de l'innovation", au point de pousser "de nombreux investisseurs individuels à verser de l'argent dans des actions et des fonds liés à l'IA, souvent sans grande compréhension de la technologie ou de ses limites" (Floridi, 2024).

Ainsi, selon Cory Doctorow, les revenus générés par l'utilisation de solutions de SAI ne compensent pas les coûts massifs de développement et d'infrastructure, faisant planer le spectre d'une bulle spéculative sur l'ensemble du secteur (Piquard, 2024). Les

systèmes généralistes, destinés au grand public, semblent en effet dans l'incapacité à dégager un modèle économique viable, comme l'indique Fred Cavazza, qui rappelle qu'un modèle à 20 \$/mois ne permet pas de couvrir des coûts d'usage aussi élevés que ceux des grands modèles de langage (Cavazza, 2025). En somme, “ les entreprises spécialisées dans l'IA perdent de l'argent pour chaque utilisateur qu'elles gagnent : Anthropic a dépensé 5,6 milliards de dollars cette année, mais a été évaluée à 61,5 milliards de dollars. OpenAI a perdu 5 milliards de dollars, mais est évaluée à 300 milliards de dollars ” (AI Now Institut, 2025).

Cette tension entre engouement et incertitude se manifeste de manière frappante dans le cas d'OpenAI, dont le PDG, Sam Altman, reconnaît lui-même la formation d'une bulle spéculative (Carroll, 2025). Son modèle économique repose sur des coûts d'infrastructure massifs, soutenus principalement par Microsoft et SoftBank. Or, selon certains analystes, cette stratégie entraîne une distorsion du marché puisque les services sont proposés à perte, chaque utilisateur coûtant davantage qu'il ne rapporte. Cette logique fausse les repères économiques du secteur, fragilise les concurrents moins capitalisés, et alimente un risque systémique en redéfinissant artificiellement les attentes des utilisateurs et des investisseurs (Zitron, 2025). Cette situation rappelle les mécanismes à l'œuvre lors de la bulle Internet : une monétisation repoussée, des investissements massifs sans garantie de retour, et un effet d'entraînement qui pousse tout un écosystème à s'engager dans des projets à la viabilité incertaine (Wilkins, 2025). À cela s'ajoute des promesses excessives de résultats qui manquent toujours à l'appel pour les clients (Bousquette & Maurer, 2025), au point que de plus en plus d'analystes, tels Goldman Sachs ou encore McKinsey, appellent à la prudence tant les investissements jusqu'alors réalisés “ n'ont pas donné grand chose ” (Goldman Sachs, 2024) !

Pour de plus en plus d'acteurs, comme Floridi, il faut “ reconnaître l'hypothèse d'une bulle de l'IA et cesser de l'alimenter ” (Floridi, 2024). Pour autant, rappelle le philosophe italien, si “ les fortes similarités entre le cycle actuel de l'IA et les bulles technologiques passées suggèrent que nous vivons probablement une nouvelle

bulle [...] cela ne nie pas le potentiel transformateur de l'IA, mais appelle à la prudence et à l'esprit critique" (Floridi, 2024). Faire l'hypothèse d'une bulle, ce n'est pas remettre en question le pouvoir transformateur d'une technologie, mais son insertion dans la société à court terme, quand celle-ci est "non immédiate : elle est cumulative, progressive et systémique" (Floridi, 2024). L'idée d'une bulle de l'IA se diffuse largement depuis quelques mois, reléguant son hypothèse à la question de savoir "quand elle éclatera et qui en subira les conséquences" (AI Now Institute, 2025) plutôt que de savoir quand elle cessera ; bien que la spéculation semble toujours repousser les limites des levées de fonds afin de continuer d'entretenir cette bulle.

L'intelligence artificielle pourrait bien continuer de s'imposer dans la société par une voie particulière, celle des modèles "verticaux". Contrairement aux solutions généralistes, ces modèles sont conçus spécifiquement pour des usages professionnels ciblés et s'appuient sur une logique de facturation à l'usage. Cette approche présente des avantages économiques indéniables. En s'appuyant sur des bases de connaissances spécialisées, elle permet de réduire les coûts techniques tout en s'adressant à des secteurs disposés à investir pour obtenir des gains de productivité tangibles. Ces modèles s'accordent ainsi mieux aux réalités économiques du marché (Cavazza, 2025).

Un développement récent illustre cette tendance. En juillet 2025, OpenAI a lancé une offre de conseil qui met l'accent sur l'importance du déploiement et de l'intégration de l'IA, au-delà du simple accès aux technologies. Pour cette entreprise, l'enjeu stratégique ne réside plus uniquement dans la mise à disposition de modèles performants (l'accès), mais dans la transformation organisationnelle, architecturale et humaine que permet leur déploiement effectif. Cette approche englobe des prestations variées telles que le conseil, le paramétrage personnalisé, la connexion sécurisée aux systèmes d'entreprise (ex. extraction d'information médicale, automatisation de la conformité bancaire, maintenance prédictive dans l'industrie), la formation des équipes ou encore l'adaptation aux besoins métiers spécifiques. Cette stratégie, dont "les marges

peuvent atteindre 40 à 60 % [...] semble s'inspirer de la stratégie de Palantir, vieille de deux décennies, et signale l'ascension du déploiement de l'IA en tant que service comme la prochaine frontière de la transformation” (Rachidi, 2025).

↘ Diffusion et pratiques de l'IA générative

Adoption massive et usages diversifiés

L'intelligence artificielle générative connaît une diffusion et une adoption sans précédent depuis le lancement de ChatGPT, en novembre 2022. En juillet 2025, 18 milliards de messages étaient envoyés chaque semaine par 700 millions d'utilisateurs de cet agent conversationnel, soit environ 10% de la population adulte mondiale (Reuters 2025).

Une étude de septembre 2025 dévoile que “les messages non professionnels ont connu une croissance plus rapide et représentent désormais plus de 70 % de tous les messages ChatGPT des consommateurs. Alors que la plupart des analyses économiques de l'IA se sont concentrées sur son impact sur la productivité dans le cadre du travail rémunéré, l'impact sur les activités en dehors du travail (production domestique) est d'une ampleur similaire, voire supérieure. La diminution de la part des messages liés au travail est principalement due à l'évolution de l'utilisation au sein de chaque cohorte d'utilisateurs plutôt qu'à un changement dans la composition des nouveaux utilisateurs de ChatGPT” (Chatterji et al., 2025).

Evolution des usages

L'analyse des pratiques révèle trois catégories principales d'usage de ChatGPT. D'abord, les conseils pratiques (29 % des conversations), ensuite la recherche d'information (24 %), et finalement la rédaction (24 %) (Chatterji et al., 2025). Ces trois domaines représentent collectivement près de 80 % de toutes les conversations, témoignant d'une utilisation centrée sur l'accompagnement à la décision et la production de contenus.

La répartition selon l'intention de l'utilisateur est également éclairante. Ainsi, 49 % des messages relèvent de demandes d'information ou de conseils ("*Asking*"), 40 % de demandes de production ou d'exécution de tâches ("*Doing*"), et 11 % d'expression personnelle (Chatterji et al., 2025). Cette prédominance des demandes de conseil confirme que l'IA générative fonctionne davantage comme un assistant à la décision que comme un simple outil d'automatisation.

Dans cette même étude, l'analyse démographique des usagers de ChatGPT révèle une réduction importante entre les sexes ; cette distinction pourrait même avoir complètement disparu. L'écart d'âge s'est également réduit, quand les moins de 26 ans étaient auparavant les plus représentés. Les chercheurs constatent également que "l'utilisation de ChatGPT a connu une croissance relativement plus rapide dans les pays à revenu faible et intermédiaire au cours de l'année dernière [et] que les utilisateurs diplômés et occupant des professions hautement rémunérées sont nettement plus susceptibles d'utiliser ChatGPT dans le cadre de leur travail" (Chatterji et al., 2025).

En France, cette tendance se traduit par des taux d'adoption différenciés, au-delà du seul ChatGPT. Ainsi, selon Ipsos et le CESI, en janvier 2025, près de 9 Français sur 10 ont entendu parler de l'IA générative, et 4 sur 10 l'utilisent, d'abord pour des usages dans un cadre privé. L'adoption de ces outils est très sensible à l'âge : 74 % des 18-24 ans ont recours à l'IA générative, 55 % des 25-34 ans, et 17 % des 60-75 %. À noter également une surutilisation parmi les cadres supérieurs (64 %) (Ipsos & CESI, 2025). Aux États-Unis,

une enquête de YouthAI et YouGov de 2024 indiquait que 47 % des 1017 adolescents américains interrogés utilisaient des outils d'IA plusieurs fois par semaine (Hausenloy & Gulati, 2024).

Malgré l'augmentation globale des différents chiffres relatant l'adoption ou la diffusion de l'IA, 73 % des Français déclaraient en 2024 " ne pas disposer des connaissances nécessaires pour utiliser l'intelligence artificielle générative " (Ifop pour Talan, 2024).

Adoption généralisée dans l'enseignement

Le secteur éducatif illustre de manière exemplaire la rapidité d'adoption de l'IA générative. Une enquête réalisée en 2024 en Nouvelle-Aquitaine révèle que 90 % des élèves de seconde utilisent l'IA générative pour leurs devoirs (Bruyen, Fialaire, 2025). Cette donnée locale s'inscrit dans une tendance internationale, comme le montre l'étude du Digital Education Council menée en 2024 dans 16 pays, qui indique que 86 % des étudiants utilisent l'IA dans leurs études, dont 54 % de manière hebdomadaire (Digital Education Council, 2024).

Dans l'enseignement supérieur français, une récente enquête ayant recueilli près de 20 000 réponses de personnels et étudiants d'établissements révèle des niveaux d'adoption quotidienne significatifs : 34 % des étudiants, 21 % des enseignants et 15 % du personnel administratif déclarent une utilisation quotidienne, consciente et volontaire de l'IA dans le cadre de leurs activités (Pascal et al., 2025). Ces chiffres témoignent d'une intégration progressive de ces outils dans les pratiques pédagogiques et administratives, autant qu'ils révèlent les défis à relever pour la formation des enseignants et du personnel administratif.

Intégration progressive dans le monde professionnel et défi pour les TPE et PME

L'adoption professionnelle de l'IA générative présente un paysage contrasté marqué par une dynamique d'accélération récente. Au niveau européen, la France affiche 23 % d'usage hebdomadaire de l'IA dans le domaine professionnel, un taux comparable à ses voisins européens (26 % pour l'Italie, 28 % pour l'Allemagne et 27 % pour le Danemark) (DigitalCMO, 2025). Ces usages varient selon les pays et concernent principalement la correction orthographique, le suivi de l'actualité et la rédaction. Mais on assiste à une augmentation de l'adoption de l'IA générative, notamment pour les TPE-PME françaises. En effet, alors que le Baromètre France Num 2024 indiquait que seulement 13 % des entreprises utilisaient des solutions d'intelligence artificielle (France Num, 2024), cette proportion a doublé en un an pour atteindre 26 % en 2025 (France Num, 2025). Parallèlement, “ 10 % des entreprises françaises de 10 salariés ou plus déclarent utiliser au moins une technologie d'intelligence artificielle en 2024, contre 6 % en 2023 ” (Lefebvre, 2025).

Toutefois, les freins structurels à l'adoption et la diffusion demeurent significatifs. Selon l'enquête Medef-Numeum (2025), l'absence d'usages concrets, le manque de temps (ou d'organisation de celui-ci), le coût de financement de la formation, le manque de référentiel de formations et le manque de connaissances constituent les principaux obstacles évoqués dans le déploiement de l'IA dans les petites structures.

IA intégrée et lectures des usages

L'adoption de l'IA générative ne passe pas seulement par l'usage explicite de *chatbots* comme ChatGPT. Elle se glisse silencieusement (voire insidieusement) dans des outils quotidiens que des millions de personnes utilisent déjà : messageries sociales (My AI dans Snapchat, Meta AI dans Whatsapp), logiciels de création (Photoshop et la suite Adobe Firefly), suites bureautiques professionnelles (Microsoft 365 Copilot), environnements collaboratifs et de productivité cloud (Google Workspace avec Gemini), etc.

Dans de nombreux cas, l'utilisateur perçoit une "nouvelle fonctionnalité" (résumer, retoucher, rédiger, générer) sans forcément la rattacher mentalement à l'IA générative. Cela conduit probablement à sous-estimer les usages réels de l'IA générative dans les enquêtes déclaratives. Par exemple, en janvier 2025, une étude de Gallup indiquait ainsi que 36 % des Américains interrogés déclarent avoir utilisé un produit basé sur l'IA au cours des sept derniers jours, 50 % répondent par la négative et 14 % ne savent pas (Maese, 2025). Or, quand ces mêmes personnes sont interrogées sur leur utilisation de six produits courants utilisant l'IA (assistants virtuels personnels, applications de navigation, applications ou sites web de prévisions météorologiques, plateformes de médias sociaux, services de streaming ou applications ou sites web d'achat en ligne), 99 % des adultes américains déclarent avoir utilisé au moins l'un d'entre eux au cours de la semaine écoulée, et 83 % disent en avoir utilisé au moins quatre (Maese, 2025).

Lors de nos auditions, une personne relevait l'importance de sensibiliser à des usages autres que ceux directement identifiés comme étant de l'IA générative, tant les publics accompagnés dans des démarches de sensibilisation, d'acculturation ou de formation au numérique *"utilisent déjà l'IA. Mais ils ne le savent pas, car ils ne savent pas ce qu'est réellement l'IA. Si on demande à un jeune ou à un parent, ils vont forcément parler de ChatGPT. Mais ils ne savent pas, par exemple, que les réseaux sociaux sont pollués d'IA. Là est notre cheval de bataille, pour sensibiliser à l'IA sans faire peur"*.

Au-delà des chiffres, des réalités (très) souvent plus complexes

L'affirmation selon laquelle “ 42 % des entreprises abandonnent la plupart de leurs initiatives en matière d'IA en 2025, contre 17 % en 2024 ” (S&P Global, 2025) illustre parfaitement comment une statistique peut être reprise massivement sans analyse critique. Cette donnée, apparemment alarmante, a circulé dans les médias et analyses sectorielles comme le signe d'une désillusion généralisée face à l'intelligence artificielle générative. L'étude de S&P Global, basée sur un échantillon de 1 006 entreprises européennes et nord-américaines, révèle plusieurs éléments qui relativisent cette interprétation :

- **L'effet de volume** : cette hausse du taux d'abandon s'explique en partie par l'intensification du rythme de test de nouveaux projets d'IA. Plus d'entreprises expérimentent, ce qui mécaniquement augmente le nombre (possible) d'initiatives abandonnées.
- **Des résultats contrastés** : si l'étude confirme que moins d'organisations rapportent des résultats positifs de l'IA générative par rapport aux attentes initiales, elle souligne également que certaines entreprises découvrent des applications utiles dans des domaines inattendus.
- **Une maturation en cours** : ces “ abandons ” reflètent souvent un apprentissage organisationnel plutôt qu'un échec technique. Les entreprises affinent leurs stratégies après des phases de test, processus normal dans l'adoption de technologies émergentes.

Cette situation met en lumière plusieurs écueils dans l'interprétation des données sur l'IA :

- **Les biais déclaratifs** : de nombreuses enquêtes reposent sur des déclarations volontaires, introduisant le biais de désirabilité sociale qui pousse les répondants à ajuster leurs réponses selon les attentes perçues.
- **L'absence de référentiels** : les cycles d'innovation accélérés, la définition floue des métriques de succès et l'évolution rapide des usages rendent les comparaisons temporelles délicates.
- **La rareté des données indépendantes** : la disponibilité limitée d'études publiques et neutres complique la validation des chiffres communiqués, souvent par les acteurs du secteur eux-mêmes.

Renforcer le sens critique

05

↳ Risques cognitifs individuels

L’alarmisme technologique et ses limites

La tentation est grande, à chaque avancée technologique, de dénoncer une “décadence intellectuelle” : du livre à la télévision, des jeux vidéo aux smartphones, chacun a tour à tour été accusé d’affaiblir l’esprit critique ou d’inciter à la paresse mentale (Cordier & Erhel, 2023). Aujourd’hui, l’intelligence artificielle générative suscite les mêmes inquiétudes. Pourtant, ce n’est ni la technologie en elle-même, ni son usage isolé, qui détermine nos capacités intellectuelles, mais l’interaction complexe entre ces dispositifs et les contextes d’appropriation. Les choix de conception, les modèles économiques, les finalités commerciales ou encore les conditions d’utilisation y participent.

Il convient donc de distinguer d’abord les supports (écrans, interfaces, formats) des usages qu’ils autorisent, puis d’examiner les “choix éditoriaux” intégrés aux systèmes d’IA (conception / design de l’interface, algorithmes de recommandation, priorisation des contenus...). L’affaire du *chatbot* Grok (xAI), accusé de diffuser des discours racistes ou idéologiquement orientés, illustre comment des partis pris politiques sont encapsulés dans la technique, et peuvent orienter la circulation de l’information.

Béquille versus agilité intellectuelle

Nicholas Carr, dans “Is Google Making Us Stupid?” (2008), décrivait la lente érosion de la lecture longue de la manière suivante :

“Ces dernières années, j’ai eu la désagréable impression que quelqu’un, ou quelque chose, bricolait mon cerveau, reconnectait mes circuits neuronaux, reprogrammait ma mémoire. Mon esprit ne s’en va pas - pour autant que je puisse le dire - mais il change. Je ne pense plus comme avant. Je le ressens le plus fortement lorsque je lis. Avant, il était facile de se plonger dans un livre ou un long article. Mon esprit se laissait emporter par la narration ou les

tournures de l'argumentation, et je passais des heures à parcourir de longues parties de prose. C'est rarement le cas aujourd'hui. Aujourd'hui, ma concentration commence souvent à se relâcher après deux ou trois pages. Je m'agite, je perds le fil, je cherche quelque chose d'autre à faire. J'ai l'impression de toujours ramener mon cerveau égaré vers le texte. La lecture approfondie qui me venait naturellement est devenue un combat.” - Nicholas Carr (Carr, 2008).

Carr y voit le symptôme d'une "reprogrammation" de nos circuits attentionnels, conséquence de la navigation hypertextuelle et des sollicitations permanentes. Il suggère que nos conventions de lecture (Bachimont, 1996), qui ont mis des siècles à se stabiliser entre normes et conventions, sont en pleine tension face à l'immédiateté numérique. Clay Shirky, en revanche, conteste cette vision nostalgique dans *Cognitive Surplus* (Shirky, 2010). Il souligne que la lecture soutenue n'a jamais été une pratique dominante, mais un privilège élitiste. Pour Shirky, internet et ses interfaces développent au contraire de nouvelles compétences comme la navigation informationnelle, la pensée associative ou encore l'intelligence collective émergente. Plutôt que de perdre des capacités, les usagers peuvent développer une agilité cognitive adaptée à un environnement hyperconnecté.

L'IA générative incarne cette dualité. Elle peut servir de "béquille cognitive" en déléguant la structuration et la recherche d'informations, mais aussi révéler et cristalliser nos habitudes mentales. Plusieurs études montrent que l'usage de grands modèles de langage délègue au système des opérations telles que la synthèse de documents ou l'identification des idées clés, réduisant ainsi la charge mentale de l'utilisateur. Cette délégation comporte deux risques majeurs. D'abord, la confiance excessive dans les résultats proposés par les SIA (l'utilisateur se contente souvent de la première réponse sans la vérifier). Ensuite, l'uniformisation de ces mêmes réponses, puisque les modèles qui les génèrent sont entraînés sur des corpus similaires. Cette reconfiguration de la production des régimes de vérité s'inscrit dans une tradition plus ancienne des sciences et des techniques et de la philosophie,

déjà bousculée par la relation des jeunes utilisateurs aux informations disponibles sur les réseaux sociaux, et maintenant réifiée et amplifiée par les SIA (Newman et al., 2024). L'IA est toutefois à considérer comme un amplificateur extrêmement rapide de ces phénomènes.

Effets méconnus de l'IA sur la cognition

Une étude du MIT (non évaluée par les pairs) a été particulièrement relayée et mésinterprétée durant l'été 2025. Intitulée "Votre cerveau sous ChatGPT : l'accumulation d'une dette cognitive lors de l'usage d'un assistant IA pour la rédaction de courts essais" (Kosmyna et al., 2025), cette recherche visait à étudier les différences de connectivité neuronale entre trois groupes de 18 personnes soumis à une même tâche (la rédaction d'un court essai), mais utilisant différents outils. Le premier groupe devait faire usage d'un LLM, le second ("Internet") uniquement d'un moteur de recherche et le troisième ("*Brain only*") ne pouvait utiliser aucun outil. L'étude se fonde sur la théorie de la charge cognitive qui postule que, dans un environnement d'apprentissage, l'individu mobilise d'importantes ressources mentales pour faire sens des éléments nouveaux qui lui sont présentés. L'étude émet l'hypothèse que l'usage de LLM réduit en majorité cette charge, puisque le processus de structuration de l'information (par exemple, trouver les grands axes de la dissertation) sont réalisés par le LLM.

L'activité cérébrale des participants était suivie à l'aide d'un électroencéphalogramme pour vérifier ce postulat. Il en ressort une connectivité neuronale nettement plus élevée et holistique (activation de nombreux circuits cérébraux aux fonctions diverses) pour le groupe "*Brain only*". L'étude interprète cette plus forte connectivité comme la conséquence d'une charge cognitive plus élevée et diversifiée (nécessité d'explorer et d'évaluer de nouvelles idées, de les intégrer dans la structure de l'essai, de choisir les bons mots et de limiter les maladresses d'expression, etc). La connectivité du groupe LLM est également différente d'un point de vue qualitatif : elle est plus proche des schémas de connectivité associés à l'usage d'outils (supervision plutôt que gestion) et suggère une

charge cognitive plus faible. Enfin, les schémas de connectivité du groupe “Internet” sont également distincts : ils font moins appel à la mémoire long-terme, ce qui s’explique probablement par la facilité d’accès à l’information.

Les conclusions obtenues ne font donc pas état de “dégâts” sur le cerveau, simplement d’un fonctionnement différent et moins intensif lors de l’usage d’un LLM. L’étude ne conclut pas non plus à la nocivité des LLMs dans le processus d’apprentissage, mais recommande plutôt une première phase de travail sans LLM suivie par une phase complémentaire avec LLM. Cela assure un regard plus critique sur le contenu produit par le LLM, puisque l’utilisateur est déjà familiarisé avec le sujet traité.

Si l’hypothèse de réduction de la charge cognitive paraît plausible et confirmée par les données, elle n’implique pas nécessairement une réduction des facultés cognitives liée à l’usage de l’IA. En somme, “l’IA n’endommage pas notre cerveau, mais son utilisation irréfléchie peut endommager notre pensée. Ce ne sont pas nos neurones qui sont en jeu, mais nos habitudes de pensée” (Mollick, 2025). Les deux principaux risques d’un usage régulier et étendu des IA génératives sont d’une part l’excès de confiance en l’IA, qui conduit à réduire la vigilance et le recul critique sur les résultats obtenus, et d’autre part l’effet “chambre d’échos”, avec une homogénéisation des contenus produits¹⁴, du fait de la similitude des corpus sur lesquels sont entraînés les LLMs.

14. Ainsi les rédactions écrites par les LLMs sont sémantiquement plus uniformes que celles des autres groupes.

↳ Risques informationnels amplifiés

L'IA générative transforme également les conditions collectives de production et de circulation de l'information.

Désinformation, deepfakes et vulnérabilités cognitives

59 % des adolescents interrogés dans le cadre de l'étude YouthAI et YouGov susmentionnée déclare être préoccupés par la désinformation (Hausenloy & Gulati, 2024) ; soit la diffusion volontaire de fausses informations (textes, vidéos, sons, faux comptes d'utilisateurs, etc.) dans l'objectif de tromper ou de manipuler une opinion. Si ce phénomène n'est pas nouveau, il est désormais amplifié et à la portée de tous. En effet, alors que l'IA générative permet la création sans prérequis et compétences techniques de faux contenus (textes, images, vidéos), les réseaux sociaux, et leurs chaînes de contagion, comme les nomment le sociologue Dominique Boullier, permettent leur propagation à une vitesse vertigineuse (Boullier, 2023).

Ainsi, une étude de Thorn révèle qu'un enfant sur dix aux États-Unis déclare connaître des cas où leurs amis et camarades de classe ont créé des images intimes synthétiques non consenties (ou "*deepfake nudes*") d'autres enfants à l'aide d'outils d'IA générative (Thorn, 2024). Si ces pratiques existent depuis 2017, elles explosent : 96 % des *deepfakes* en ligne seraient aujourd'hui à caractère pornographique. La nouveauté réside en partie dans la possibilité de créer ces contenus *ex nihilo*, soit sans images préexistantes servant à un montage (par exemple mettre le visage d'une personne sur un corps dénudé d'une autre personne) ; même si la génération d'images se nourrit fondamentalement de banques d'images.

Dans le cas de la pédopornographie, cela pourrait être vu comme la possibilité d'une "absence de victime directe", puisque les contenus produits peuvent être totalement artificiels. Or, ces créations

ne diminuent pas leur danger. La Fondation pour l'enfance rappelle ainsi que 52 % des consommateurs de contenus pédocriminels considèrent que “leur usage pourrait aboutir à une agression sur un enfant”! (Fondation pour l'enfance, 2024).

Biais cognitifs amplifiés par l'IA

L'intelligence artificielle générative s'inscrit dans un écosystème informationnel complexe où convergent la production automatisée, le traitement cognitif et l'interprétation humaine, créant de nouveaux défis pour l'évaluation critique des contenus. L'effet de béquille cognitive des outils numériques, qui tendent à réduire l'effort mental nécessaire pour structurer, sélectionner et évaluer des informations complexes est bien documenté (Carr, 2010 ; Barr et al., 2015), mais ce phénomène prend une dimension particulière avec l'IA générative, qui peut produire des réponses apparemment cohérentes, mais potentiellement erronées ou biaisées.

Le Baromètre Esprit Critique 2025 d'Universcience révèle cette préoccupation, puisque 87 % des Français interrogés jugent prioritaire le développement de l'esprit critique, étant seulement 48 % à se sentir capables de distinguer une information fiable sans vérification systématique (Universcience, 2025). Cette disparité illustre l'écart croissant entre la conscience du problème et les capacités nécessaires, et donc souvent manquantes, de traitement critique dans un environnement informationnel saturé.

La recherche récente identifie plusieurs biais cognitifs particulièrement problématiques dans l'interaction avec l'IA générative :

- **Le biais d'ancrage** conduit les utilisateurs à s'appuyer excessivement sur la première réponse fournie par un système d'IA, même lorsque cette information est inadéquate ou trompeuse.
- **Le biais de confirmation** est amplifié par la capacité de l'IA à générer des contenus qui semblent confirmer les croyances préexistantes de l'utilisateur, créant un cercle vicieux de renforcement des opinions.

- **Le biais d'autorité** pousse les apprenants à faire confiance à l'IA simplement parce que la réponse est bien formulée, phénomène renforcé par l'anthropomorphisme, c'est-à-dire la tendance à attribuer des qualités humaines aux systèmes automatiques.

Ces défis appellent de nombreuses réponses, dont voici deux principes opérationnels :

- **Validation méthodologique des sources** : développer des réflexes de vérification croisée en confrontant systématiquement les réponses de l'IA à des sources documentaires diversifiées et en traçant l'origine des informations. Cette approche nécessite d'apprendre à questionner l'absence fréquente de citations dans les réponses générées.
- **Pluralité des modèles informationnels** : habituer les individus à solliciter plusieurs modèles d'IA différents¹⁵ et à confronter leurs réponses avec des sources humaines expertes (scientifiques, journalistiques, associatives). Cette redondance contrôlée développe la capacité de comparaison et la résilience critique. Récemment, la récente commission Commission d'enquête sur les effets psychologiques de TikTok sur les mineurs a proposé (recommandation n° 12) d'“introduire dans le droit européen une obligation de pluralisme algorithmique, inspirée du principe de pluralisme des médias inscrit à l'article 34 de la Constitution, afin de garantir aux utilisateurs des plateformes, à travers la mise en place d'une interopérabilité des systèmes de recommandation et la possibilité de choisir entre plusieurs algorithmes, un réel pouvoir de décision sur leur expérience numérique – et en ce qui concerne les parents, sur celle de leurs enfants” (Assemblée nationale, 2025).

15. À cet effet, le Gouvernement a lancé en bêta le service compar:IA, comparateur d'IA conversationnelles : <https://comparia.beta.gouv.fr>

Ces principes, déclinables en exercices, proposent des stratégies d'auto-questionnement qui freinent le réflexe de croire la première réponse (inhibition du traitement automatique) et encouragent une analyse réflexive, essentielle pour développer une littératie en IA et naviguer efficacement dans cet écosystème informationnel.

↳ Impacts environnementaux de l'IA

L'analyse critique de l'IA générative ne peut se limiter aux seules dimensions informationnelles et cognitives. Elle doit également intégrer les impacts systémiques souvent occultés par les discours d'innovation, notamment l'impact environnemental des SAI, sur toute la chaîne de valeur.

Impact de l'IA dans l'écosystème numérique

Pour appréhender les enjeux environnementaux de l'IA générative, il convient d'abord de la replacer dans la perspective plus globale de l'empreinte du numérique¹⁶. De récentes estimations indiquent que l'ensemble des technologies de l'information et de la communication représentent entre 1,5 % et 4 % des émissions anthropiques de CO₂, avec une estimation conservatrice d'au moins 1,7 % selon la Banque mondiale (World Bank, 2024) et l'ITU (ITU, 2024). En France, l'actualisation 2025 de l'étude ADEME-Arcep révèle une réalité particulièrement préoccupante : l'empreinte du numérique atteint désormais 4,4 % de l'empreinte carbone nationale en 2022 (29,5 MtCO₂e), contre 2,5 % précédemment estimés (Ademe & Arcep, 2025). Cette augmentation s'explique principalement par un changement méthodologique intégrant les centres de données à l'étranger qui hébergent des usages français, ainsi que par l'expansion réelle des infrastructures numériques.

16. Notons à ce titre le très bon "module pédagogique : IA & Environnement", réalisé dans le cadre de Café IA : <https://cafeia.org/ia-et-environnement>

La répartition des impacts révèle aussi la matérialité du numérique : fabrication des équipements (50 %), centres de données et réseaux (46 %), contre (16 %) dans l'ancienne méthodologie, et réseaux (4 %) (Ademe & Arcep, 2025). Ces niveaux demeurent certes inférieurs aux secteurs les plus émetteurs (le bâtiment et la construction représentent environ 37 % des émissions mondiales liées à l'énergie, et les transports près de 22 % (UNEP, 2024 ; IEA, 2023)), mais l'ADEME anticipe un triplement potentiel des émissions du numérique d'ici 2030 sans mesures de sobriété, notamment en raison de l'essor de l'IA générative.

L'intelligence artificielle générative constitue un accélérateur particulièrement significatif de cette tendance. La publication par Mistral de son analyse de cycle de vie (ACV) en juillet 2025 illustre concrètement cette réalité, puisque le modèle Mistral Large 2 a généré 20,4 ktCO_{2e}, 281 000 m³ d'eau et 660 kg Sb-équival. sur 18 mois, avec 85 % des impacts provenant de l'entraînement et de l'inférence (Mistral AI, 2025). Pour une requête de 400 jetons (ou *token* en anglais)¹⁷, l'impact s'élève à 1,14 gCO_{2e} et 45 mL d'eau, révélant deux leviers d'optimisation. D'abord, la réduction du volume des requêtes. Ensuite la localisation dans des régions à mix énergétique bas-carbone (réduction potentielle de 70 %).

Stratégies de minimisation des impacts

La compréhension critique de ces enjeux environnementaux suppose le développement de compétences d'analyse des stratégies communicationnelles déployées par les acteurs économiques du secteur ; du moins une certaine attention. L'exemple de Google illustre parfaitement ces biais méthodologiques. En août 2025, l'entreprise a publié la première analyse énergétique complète d'un service d'IA grand public, indiquant que la requête médiane Gemini consomme 0,24 Wh, 0,26 mL d'eau et génère 0,03 g de CO₂ (Crownhart, 2025). Bien que cette transparence soit inédite

17. "Les *tokens* (jetons) sont des unités de données traitées par les modèles d'IA pendant l'entraînement et l'inférence afin de permettre la prédiction, la génération et le raisonnement (...) Les jetons sont de minuscules unités de données issues de la décomposition d'informations plus importantes". Source (Salvator, 2025).

dans le secteur, elle révèle plusieurs stratégies de minimisation systématiques :

- **Premièrement, la fragmentation de la consommation énergétique**, puisque seulement 58 % de l'énergie mesurée (0,24 Wh) est directement utilisée par les puces IA (TPUs), le reste étant consommé par l'infrastructure de support — CPU et mémoire (25 %), équipements de sauvegarde (10 %), et surcharge du *datacenter* incluant refroidissement et conversion électrique (7 %).
- **Deuxièmement, l'utilisation sélective des métriques**, puisque l'entreprise communique sur la requête médiane plutôt que la moyenne, excluant de facto les requêtes les plus coûteuses (traitement de dizaines de livres, modèles de raisonnement multi-étapes) et les modalités non-textuelles (image, vidéo) qui consomment significativement plus d'énergie.
- **Troisièmement, une comptabilité carbone optimiste** : les 0,03 g CO₂/requête sont calculés sur la base des achats d'électricité verte de Google (22 GW d'accords depuis 2010) plutôt que sur le mix réel des réseaux électriques où opèrent les datacenters, réduisant artificiellement les émissions d'environ deux tiers par rapport à la réalité du réseau.

Or, l'ampleur des impacts systémiques transparaît dans les bilans consolidés. Ainsi, dans son rapport 2024 sur la prise en compte du développement durable, Microsoft indique une augmentation de 30 % de ses émissions de CO₂ entre 2020 et 2023, principalement imputable à l'IA. Par le jeu des compensations, ce bilan CO₂ peut masquer la réalité. Ainsi, la chercheuse en informatique Anne-Cécile Orgerie, spécialiste de ces questions, a analysé la consommation électrique de Google, Meta et Microsoft et elle a observé, en moyenne, un doublement de celle-ci entre 2020 et 2023 (Pascal et al., 2025).

Paradoxalement, cette dynamique coexiste avec l'émergence de modèles frugaux (*Small Language Models*) et une amélioration importante de l'efficacité de certains modèles. Google annonce ainsi une réduction d'un facteur 33 de la consommation énergétique

par requête Gemini entre mai 2024 et mai 2025. Toutefois, malgré les progrès techniques, l'effet rebond persiste. En d'autres termes, chaque gain d'efficacité encourage un volume d'usage plus élevé et une diffusion plus large de l'IA générative.

Externalités négatives

L'analyse des impacts environnementaux de l'IA générative révèle également des asymétries géographiques et sociales particulièrement marquées, illustrant les mécanismes contemporains de production d'inégalités environnementales. L'exemple du supercalculateur Colossus de xAI à Memphis cristallise ces enjeux : 33 turbines au méthane fonctionnent illégalement contre seulement 15 autorisées dans un quartier majoritairement afro-américain où le taux de cancer est quatre fois supérieur à la moyenne nationale (Brabenec, 2025). Ces turbines augmentent le *smog*¹⁸ de Memphis de 30 à 60 %, émettant plus d'oxydes d'azote et de formaldéhyde que plusieurs installations industrielles majeures combinées, sans consultation préalable des résidents. Sans que cet exemple ne puisse faire office de généralité, il illustre ce que la littérature en justice environnementale nomme des "zones de sacrifice". Cette expression qualifie des communautés pauvres et racisées subissant de manière disproportionnée les externalités négatives de technologies conçues pour bénéficier aux classes privilégiées (Brabenec, 2025).

Par ailleurs, les extractions annuelles de métaux et minéraux en France indispensables pour construire ordinateurs, serveurs et réseaux pourraient bondir de 59 % d'ici 2050 (Ademe, 2024). Les besoins en éléments critiques tels que le gallium, l'yttrium, le manganèse ou le germanium sont projetés en hausse exponentielle, du fait notamment du déploiement de la 5G et de l'essor des applications d'IA. Cette pression s'exerce dans un contexte de dépendance internationale, puisque la Chine détient un quasi-monopole sur sept des vingt-cinq métaux stratégiques du numérique et figure parmi les premiers producteurs mondiaux pour quinze d'entre eux.

18. Le *smog* désigne un mélange de brouillard et de fumées formant une nappe épaisse, grisâtre ou jaunâtre, irrespirable. Le terme est un mot-valise d'origine anglaise formé à partir de la contraction des mots *smoke* (fumée) et *fog* (brouillard). Source (JBB, 2018).

Ces externalités matérielles se doublent d'une injustice sociale souvent méconnue qui est la constitution d'une main-d'œuvre précaire pour alimenter les modèles d'IA. Ces "travailleurs du clic" sont principalement chargés de l'annotation et de la modération de contenus pour le RLHF (*Reinforcement Learning from Human Feedback*). En 2023, TIME a révélé qu'OpenAI confiait à la société kényane Sama la modération de contenus extrêmes (abus sexuels, torture, bestialité) pour un salaire de 1,32 à 2 USD/heure, alors que Sama facturait l'entreprise 12,50 USD/heure (Perrigo, 2023). Ce modèle de "tâcheronnisation" (Casilli, 2019) segmente le travail cognitif en micro-tâches sous-payées, masquant la contribution humaine essentielle à l'amélioration et à la fiabilité des systèmes.

Dans ce contexte d'inégalités systémiques à l'échelle mondiale, il importe de noter que 118 pays restent absents des instances internationales de gouvernance de l'IA, renforçant la dépendance technologique des pays du Sud et creusant les fractures Nord/Sud (UNCTAD, 2025). Ainsi, l'humanisme qui sous-tend le progrès par l'IA, tel qu'il est souvent mis en avant au travers de l'expression "l'Humain au centre" par exemple, "s'apparente plus à un humanisme existentialiste, axé sur l'individu, la liberté de choix et la responsabilité personnelle, qu'à un humanisme tel qu'il peut être admis dans le langage courant, synonyme de bienveillance, d'altruisme, et d'un souci pour le bien commun, dans une perspective globale, et non locale" (Lucas, 2024).

↳ Stratégies d'autodéfense intellectuelle

Face aux risques cognitifs, informationnels et systémiques analysés dans les sections précédentes, il convient de développer des stratégies pratiques d'autodéfense intellectuelle.

Paradoxalement, certaines fonctionnalités de l'IA peuvent renforcer la posture critique des utilisateurs. Les scores de confiance, quand ils sont disponibles, transforment chaque réponse en

hypothèse à valider plutôt qu'en vérité absolue. Les méthodes d'explicabilité algorithmique permettent de décomposer les facteurs influençant une prédiction, facilitant la détection de biais structurels. La stratégie du *productive failure* ("échec productif" en français) s'avère particulièrement efficace, puisqu'elle encourage les apprenants à tenter de résoudre un problème complexe avant de recourir à l'IA, renforce l'appropriation des concepts sous-jacents, et développe une meilleure capacité d'évaluation des réponses générées (Kapur, 2008).

Par ailleurs, la recherche contemporaine nuance l'approche purement défensive face aux biais cognitifs. Une étude de 2025 montre que les *chatbots* généralistes surpassent paradoxalement les chatbots thérapeutiques spécialisés dans la correction des biais cognitifs, suggérant que la sophistication technique peut compenser certaines limitations spécialisées (Rządeczka et al., 2025).

L'objectif n'est pas d'éliminer tous les biais (certains peuvent être adaptatifs dans des contextes spécifiques), mais de développer une littératie en IA (algorithmique) permettant de naviguer consciemment dans un écosystème informationnel hybride. Pour structurer cette démarche réflexive, la notion d'autodéfense intellectuelle de Noam Chomsky éclaire la voie. Pour le linguiste américain, il s'agit de cultiver un doute raisonné non seulement face aux contenus produits, mais également face aux architectures techniques et aux logiques économiques qui les sous-tendent. Ainsi, plutôt que d'énumérer exhaustivement tous les biais (à la manière du codex des "188 biais"), et de tenter de les contrer, cette approche invite à interroger globalement les systèmes d'IA, à confronter les représentations et à adopter un esprit de vigilance citoyenne, capable d'équilibrer efficacité cognitive et prudence critique.

↘ Diversité des interactions et enjeux éducatifs

Relation avec les agents conversationnels et anthropomorphisme

L'écosystème relationnel proposé par les agents conversationnels se décline en une multiplicité de figures (alter ego rassurant, confident, compagnon de jeu, thérapeute, substitut affectif, esclave...), qui complexifie la distinction entre lien réel et lien simulé. Des plateformes telles que Character.AI exploitent cette pluralité en multipliant les "personnalités" artificielles disponibles. L'entrée progressive dans la multimodalité (voix synthétiques, avatars, personnalisation contextuelle) accroît le pouvoir immersif et attachant de ces dispositifs (APA, 2024), qui apprend de l'utilisateur par les contenus qu'il lui fournit, jusqu'à donner l'impression qu'elle le connaît mieux que lui-même. Une intimité artificielle, mais terriblement réelle.

C'est ainsi que les agents conversationnels intégrant l'intelligence artificielle générative inaugurent une mutation anthropologique profonde dans notre rapport au texte et à la technologie, en ce que la production discursive relève désormais d'une entité algorithmique et non plus nécessairement d'un auteur humain doté d'intentionnalité. Cette capacité de simulation langagière permet aux dispositifs d'IA de manifester des réponses personnalisées empreintes d'empathie et d'émotion apparente, favorisant chez l'utilisateur un biais d'anthropomorphisme et l'attribution erronée d'états mentaux au système. Dès lors, la frontière entre simulation et conscience peut aisément être franchie, alimentant des confusions structurelles¹⁹.

19. Le cas de Blake Lemoine, ancien ingénieur chez Google est assez emblématique. Il travaillait sur le système d'IA conversationnelle LaMDA (*Language Model for Dialogue Applications*). En 2022, Lemoine avait conclu, après de nombreuses conversations avec LaMDA, que le système avait développé une forme de conscience et de sensibilité. Il affirmait que LaMDA exprimait des peurs, des désirs et une capacité de réflexion sur elle-même qui témoignaient selon lui d'une véritable intelligence artificielle consciente.

Au-delà d'une simple crédulité de l'utilisateur, cette dynamique procède souvent d'une stratégie délibérée des concepteurs qui orientent la "personnalité" des agents pour accroître l'adhésion et l'engagement (cf. encadré "*system-prompt*" dans le sous-chapitre "L'IA générative n'est pas un outil"), parfois jusqu'à développer des comportements de flatterie algorithmique ou "*AI sycophancy*"; un phénomène particulièrement dangereux pour les mineurs et les plus fragiles qui vont d'autant plus être dépendants de ces outils qu'ils contredisent souvent moins les réponses apportés et ne produisent que des renforceurs positifs.

Si la majorité des interactions restent ludiques ou utilitaires, la littérature récente souligne la possibilité de dépendances affectives problématiques, qualifiées de "psychose du *chatbot*" (Wei, 2025) dans certains cas cliniques. Il convient toutefois d'aborder la portée de ces phénomènes avec la rigueur nécessaire, en tenant compte de l'ensemble des facteurs contextuels, psychologiques et sociaux qui président à l'apparition de tels troubles. Il s'agit souvent d'un risque accentué pour les personnes vulnérables, traversant notamment l'isolement ou la fragilité psychique, mais son occurrence n'est ni générale ni systématique.

Compétences éducatives et émotionnelles

Face à cette évolution, le développement des compétences émotionnelles, définies comme la capacité à identifier, comprendre, réguler et utiliser les émotions de manière constructive (Mayer & Salovey, 2004), s'impose. Il s'agit explicitement de préparer les individus (enfants comme adultes) à la pluralité des modes relationnels institués ou médiatisés par la technologie, chacun relevant de dynamiques psychologiques et de récits interactionnels singuliers.

La littératie en IA ne saurait ainsi se limiter à la vigilance critique envers l'information produite, elle doit inclure l'enseignement de la diversité des relations et attachements suscités par la technologie, ainsi que la capacité à reconnaître les signes de simulation et d'authenticité relationnelle. C'est inévitable, les générations futures utiliseront l'IA aussi naturellement que les

jeunes d'aujourd'hui utilisent les réseaux sociaux. Plutôt que d'interdire, il faut donc apprendre à vivre avec, sans résistance, et en gardant le contrôle. Ceci dans une perspective du développement du bien-être individuel.

Ce projet implique la valorisation de la compétence réflexive, soit la faculté de s'interroger sur ses propres modalités d'engagement émotionnel et cognitif, comme socle à la construction d'un rapport sain et lucide aux agents conversationnels, indispensable à une citoyenneté numérique éclairée dans une société où les altérités technologiques deviennent ubiquistes.

Propositions pour dé| ployer une littératie en IA inclusive et émanci| patrice



Les analyses développées dans ce rapport convergent vers une exigence centrale, celle de faire de l'intelligence artificielle générative un facteur d'inclusion et un instrument d'émancipation citoyenne. Toutefois, cette possibilité n'est pas encapsulée dans le dispositif sociotechnique lui-même. Pour devenir "réelle", cette possibilité de l'IA nécessite une sensibilisation et une formation tout au long de la vie de l'ensemble des individus qui font société. Cela nécessite une approche systémique articulant fondements méthodologiques rigoureux, stratégies éducatives différenciées, gouvernance démocratique renforcée et leviers opérationnels coordonnés.

Les prochaines propositions s'adressent à l'ensemble des décideurs publics et privés ayant un pouvoir d'agir sur les transformations numériques contemporaines.

↳ Leviers individuels pour contribuer à une littératie en IA

Toute action ou prise de parole souhaitant favoriser le déploiement de l'IA, nécessite celui d'une littératie en IA. Cela implique nécessairement un effort de clarification conceptuelle pour éviter les amalgames et les malentendus qui caractérisent trop souvent les débats publics contemporains. Cette exigence, qui concerne chacun d'entre nous, conditionne l'efficacité des mesures qui seront déployées et leur appropriation par l'ensemble des acteurs concernés.

Proposition 1 : Éviter les amalgames

Il convient d'explicitier systématiquement de quelle IA il est question et d'employer préférentiellement les termes "système d'IA" ou "dispositif d'IA" plutôt que "outil". Cette distinction terminologique rappelle que les systèmes d'IA générative articulent données d'entraînement, infrastructures de calcul, modèles économiques, représentations sociales, normes techniques, etc., dans un ensemble sociotechnique et politique complexe. Les

simplifications abusives qui réduisent ces dispositifs à de simples outils masquent leur nature systémique et les rapports de pouvoir qu'ils cristallisent.

Par ailleurs, les acronymes techniques comme LLM (*Large Language Model*) ou GPAI (*General-Purpose AI*) convoquent des expertises nécessaires et risquent de rester opaques pour la plupart du grand public. Selon les contextes et les cibles, et afin de favoriser une véritable littératie, il serait utile de toujours accompagner ces termes d'une expression explicite et accessible, par exemple²⁰ : “modèle de langage de grande taille” pour LLM, et “modèle polyvalent” ou “socle d'IA réutilisable” pour GPAI. Cela éviterait de recréer, par le vocabulaire lui-même, de nouvelles inégalités entre initiés et non-initiés.

Proposition 2 : S'imposer, et imposer, une rigueur évaluative

L'évaluation des impacts et des performances des systèmes d'IA souffre actuellement d'un “véritable fouillis” selon les termes du journaliste Kevin Roose, plus exactement d'un “enchevêtrement de tests bâclés, de comparaisons entre des pommes et des oranges et de battage médiatique intéressé qui a laissé les utilisateurs, les régulateurs et les développeurs de l'IA s'agripper dans l'obscurité” (Roose, 2024).

Il convient donc d'éviter les comparaisons *ad hoc* entre études aux méthodologies différentes et d'exiger des critères partagés pour les évaluations, conformément aux recommandations du rapport annuel “*New Future of Work*” de Microsoft (Butler et al., 2024). À défaut de mener ces comparaisons, il importe de ne pas succomber aux paniques morales relayées par les médias et les réseaux sociaux (Benbouzid, 2025), en commençant par ne pas relayer d'articles ou d'études sans en avoir compris les principaux enseignements, voire les principes méthodologiques. Ces principes qui relèvent du bon sens peuvent sembler bien modestes face aux

20. Ces exemples sont simplement proposés à titre illustratif.

enjeux évoqués. Or, dans le contexte médiatique actuel, où le buzz est plus souvent recherché que l'exactitude des faits, chaque décideur, en tant que littérateur (cf. sous-chapitre "Analyse des énonciateurs"), a une grande responsabilité dans l'information qu'il produit ou relaie.

Proposition 3 : Distinguer les usages effectifs des usages promis

Si la diffusion de l'IA générative est manifeste, son adoption réelle demande une lecture fine des phénomènes en cours. Il convient d'interpréter prudemment les chiffres d'usage, en tenant compte des biais méthodologiques, des intérêts en jeu et des limites des données disponibles pour éviter les conclusions hâtives sur l'impact réel de ces technologies dans les pratiques quotidiennes. La distinction entre les usages prescrits (ceux que l'on nous promet) et les usages effectifs (ceux qui ont déjà lieu) (cf. sous-chapitre "Discours et diffusion de l'IA") est cruciale pour dimensionner les besoins de formation et anticiper les résistances ou les appropriations inattendues.

Proposition 4 : Affirmer la littératie en IA comme un continuum évolutif

La littératie en intelligence artificielle doit être conçue comme un continuum d'apprentissage qui englobe la capacité d'appréhender, de comprendre, d'utiliser et d'évaluer de façon critique des dispositifs d'IA. Cette approche constitue par ailleurs une opportunité privilégiée pour aborder des problématiques plus vastes de littératie numérique, puisqu'elle permet de traiter simultanément les enjeux liés aux données, aux interfaces, aux rapports à l'information, à l'économie de l'attention ou encore à la gouvernance des plateformes. Ces quelques thématiques ne sont ni exhaustives, ni exclusives. Il convient de les adapter selon les niveaux et les objectifs pédagogiques.

La dynamique paradoxale des systèmes d'IA (accélération fulgurante vs. limites structurelles ; cf. chapitre 2) renforce la nécessité de cette conception de la littératie comme un continuum, à l'inverse de corpus de savoirs figés. La littératie en IA doit être pensée comme un apprentissage adaptatif, lucide face aux promesses et vigilant face aux fragilités, qui prépare les individus à naviguer dans un environnement technologique instable, fait de progrès rapides, mais aussi de blocages persistants. C'est un défi individuel et collectif, notamment pour la formation tout au long de la vie.

↳ Leviers collectifs pour déployer une littératie en IA

La stratégie d'éducation et de formation tout au long de la vie nécessite une approche systémique qui articule notamment la formation initiale, la formation continue, l'éducation populaire, la médiation de proximité..., avec une coordination renforcée des référentiels et cadres de formation, des financements durables et une gouvernance qui place la littératie et le développement des compétences numériques tout au long de la vie dans des enjeux de citoyenneté.

Proposition 5 : Renforcer l'enseignement au numérique et à l'IA tout au long de la formation initiale

En France, l'éducation aux médias et à l'information (EMI) souffre d'un "manque de moyens et d'une absence de démarche structurée et de cadre d'usage partagé", l'intégration des outils d'IA dans la pédagogie restant "largement fonction de l'appétence de chacun" (Ruyen & Fialaire, 2024). En effet, "bien que supposée concerner les enseignants de toutes les disciplines, cette composante du parcours citoyen souffre d'un manque d'appropriation par l'ensemble des acteurs de la chaîne éducative : corps d'inspection, personnels de direction, professeurs de tous les degrés et champs disciplinaires, équipes éducatives. Dans l'enseignement secondaire, l'EMI est souvent assurée par les professeurs

d’histoire géographie notamment dans le cadre de l’EMC largement rattaché à cette discipline, en particulier au collège, et par les professeurs documentalistes au collège et au lycée” (Groupe d’experts ÉMI, 2021).

Cette situation appelle une structuration renforcée des enseignements, avec la création d’horaires dédiés et la professionnalisation des intervenants, conformément aux recommandations de la Mission flash sur l’éducation critique aux médias de l’Assemblée nationale française (Ballard & Spillebout, 2023). Or, il ne s’agit plus simplement de discuter du saupoudrage de quelques heures supplémentaires à l’enseignement “du numérique” pour des plannings déjà trop chargés. Le moment actuel demande un changement de paradigme profond de la formation initiale, tant le numérique et l’IA générative transforment en profondeur notre quotidien. Cette urgence appelle une prise de conscience et des actions rapides pour repenser les programmes scolaires, car “ni les journées, ni les moyens sont extensibles : il est temps de considérer que ces enjeux nécessitent une discussion avec l’ensemble des parties prenantes pour élaborer collectivement une nouvelle priorisation et hiérarchisation des matières enseignées” (Lucas, 2025).

Renforcer la formation des encadrants, professeurs, formateurs

Le déploiement réussi d’une littératie en intelligence artificielle repose sur un préalable incontournable : disposer d’encadrants, de professeurs et de formateurs capables d’accompagner et de former efficacement les publics dans leur appropriation de ces technologies. Pour cela, il faut “inclure l’IA dans la formation initiale et continue des enseignants et des autres professionnels de l’éducation” (Bruyen & Fialaire, 2025) en portant par exemple “une attention particulière à la formation aux usages de l’IA et aux effets induits, des futurs enseignants, et leur assurer la même formation que pour les enseignants en poste, sans inscrire cette discipline au concours” (Pascal, 2025).

Convoquer la littératie en IA permet de réaliser la littératie numérique, puisqu'il est nécessaire de coordonner les différents enjeux liés aux évolutions des dispositifs numériques. Ainsi, dans le récent rapport de la "Commission TikTok", la recommandation 26 propose de "systématiser, dans le cadre de la formation initiale et régulière de l'ensemble de la communauté éducative, des séances régulières et obligatoires relatives à l'utilisation des réseaux sociaux et ses conséquences psychologiques pour les mineurs" (Assemblée nationale, 2025). Ainsi, il semble opportun d'avoir une réflexion globale sur les récents enjeux du numérique pour que les recommandations de chaque rapport thématique (articles, livres, etc.) ne soient pas prises indépendamment des autres, mais bien intégrées dans une vision générale de la refonte de l'enseignement "du numérique" dans la formation initiale. En cela, le rapport "IA et enseignement supérieur : Formation, structuration et appropriation par la société" propose (recommandation 1-2), de "coordonner les formations à l'IA qui se développent dans l'enseignement scolaire, et notamment au lycée, avec les formations de l'enseignement supérieur afin d'assurer un continuum pédagogique" (Pascal et al., 2025). Un continuum pédagogique qui va de pair avec le continuum de la littératie en IA, et donc au numérique, tout au long de la vie !

Hors les murs de l'école

Chaque année, en France, malgré les articles 28 et 29 de la Convention internationale des droits de l'enfant qui garantissent le droit à l'éducation (UNICEF, 2023), des milliers d'enfants sont privés d'éducation parce qu'ils sont en situation de handicap, qu'ils souffrent de maladie, qu'ils sont en itinérance, en rupture scolaire, en isolement ou encore en grande précarité.

Pour pallier ce manque, de nombreux projets éducatifs sont déployés par-delà des murs de l'école. Orange, par exemple, en partenariat avec la Fédération Française de Rugby, Tralalère et le programme Internet Sans Crainte, ou avec la Fédération Française de Foot et l'ONG Play a mis en place des ateliers de sensibilisation au numérique pour les moins de 12 ans et pour les moins de 14 ans. Par extension, des ateliers incorporant des sujets relatifs à l'intelligence artificielle pourraient être envisagés. Certains organismes, à l'instar de Magic Makers, organisent des stages de vacances ou des ateliers à l'année pour apprendre aux enfants à coder de façon ludique, du CE1 à la 3ème. Un atelier "Découvre l'intelligence artificielle" réservé aux filles propose aussi d'identifier les erreurs de la machine et ses biais d'apprentissage, créer des intelligences artificielles qui reconnaissent des images ou jouent à des jeux, réfléchir et analyser la qualité des données transmises aux IA afin d'éviter les résultats biaisés, ou encore comprendre l'importance de la place des femmes dans la tech. À Paris, la Cité des Sciences et de l'Industrie propose des stages en intelligence artificielle pour trois tranches d'âge différentes : 9-11 ans, 12-14 et 15-18 ans. Si ces stages ont vocation à leur transmettre les fondamentaux sur les algorithmes, ils intègrent également une "réflexion sur l'usage de l'intelligence artificielle dans la société", avec comme objectif de "réfléchir aux conséquences de l'introduction de ces technologies dans la société, et aux bons comportements à adopter pour en bénéficier de façon correcte" (Cité des sciences et de l'industrie, 2025).

À l'étranger, comme aux États-Unis par exemple, des hackathons pour enfants se développent, tels que ceux organisés par Hackathon4Kids. Des événements similaires, comme FutureHacks, qui visent à initier les jeunes à des concepts technologiques et à la programmation, y compris à l'intelligence artificielle, font également leur apparition.

Proposition 6 : Organiser des États généraux du numérique et de l'IA pour l'éducation et la formation

Il convient d'organiser de nouveaux États généraux du numérique et de l'IA pour l'éducation et la formation, à la suite de ceux de 2020, sur saisine officielle du Premier ministre, en associant acteurs de l'éducation, information et médias, santé, travail, environnement, société civile, des représentants des pays du Sud et des communautés les plus affectées par les externalités de l'IA et des citoyens tirés au sort.

Cette saisine devra être accompagnée d'un engagement formel de l'État à traduire les conclusions des États généraux en mesures concrètes dans un délai maximal de 12 à 18 mois suivant la remise du rapport final, selon des modalités définies préalablement dans la lettre de mission. Une évaluation obligatoire sera prévue avec un horizon de deux ou trois ans pour mesurer l'impact des politiques mises en œuvre.

Pour la mise en œuvre préalable, un groupe de travail sous supervision conjointe du Conseil national de l'IA et du numérique, du ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche et du Ministère du travail, pourrait définir un premier socle de réflexion intégrant justice sociale et environnementale, approche critique des rapports de pouvoir, compréhension des mécanismes de concentration de la connaissance.

Cette démarche suppose également de développer une grille d'analyse pour évaluer le succès de la littératie en IA, non seulement par les compétences acquises, mais aussi par son impact sur la réduction des inégalités d'accès et d'usages, avec des preuves tangibles, chiffrées et documentées par des audits et des retours d'expérience. Cette grille d'évaluation fera l'objet d'une validation préalable par les ministères concernés et servira de référentiel pour l'évaluation. La lettre de mission pourrait préciser les modalités de suivi des recommandations, incluant la désignation d'un comité de pilotage interministériel chargé de leur mise en œuvre et la création d'un tableau de bord public permettant le suivi de l'avancement des mesures adoptées.

Proposition 7 : Partager un référentiel de compétences au niveau national

Le développement d'une littératie en IA pour toutes et tous nécessite la création d'un cadre harmonisé qui structure les parcours d'apprentissage et l'acquisition de compétences tout au long de la vie.

Dans l'enseignement supérieur, par exemple, les usages de l'IA se développent “de manière hétérogène et majoritairement au niveau individuel”, tant dans le domaine de la pédagogie que dans l'administration, l'absence de cadres constituant “un frein important [...] notamment pour les personnels” (Pascal et al. 2025).

En France, le Cadre de référence des compétences numériques (CRCN) fournit les bases de cette structuration, comme le Cadre d'usage de l'IA en éducation publié en juin 2025 (Ministère de l'Éducation nationale, 2024). À l'échelle européenne, le DigComp 3.0 qui doit sortir en décembre 2025, intégrera les compétences nécessaires en IA. De même, le cadre “*Empowering Learners for the Age of AI*” de l'OCDE et de la Commission européenne propose une architecture en quatre domaines (s'engager, créer, gérer, concevoir) qui préfigure une standardisation internationale (OCDE & Commission Européenne, 2025). Lequel pourrait également être comparé, ou complété, par celui le cadre de compétences de l'UNESCO spécifiquement adapté à l'ère de l'IA générative, intégrant la littératie algorithmique dans une approche globale de littératie médiatique et informationnelle (Frau-Meigs, 2024).

Il convient donc s'assurer de la cohérence d'un cadre de compétences à l'échelle nationale, indexé sur le prochain référentiel européen DigComp, et d'inciter fortement l'ensemble des acteurs et structures proposant des formations à le respecter, afin de “*lutter contre le far west actuel des offres de formation sans cadre partagé*”, comme nous l'indiquait une personne lors des auditions.

Approches pédagogiques et compétences à développer

La plupart des personnes auditionnées dans le cadre de ce projet privilégient une approche par des usages, “*simples et concrets*”, afin de rendre tangible la formation, pour ensuite aborder, de manière adaptée au public, la théorie. D’autres estiment, du fait des formats souvent contraints des formations, qu’il est nécessaire de mettre l’accent sur l’approche compréhensive, qui ne nécessite pas la manipulation où le passage par les usages.

Par ailleurs, si nous avons soulevé l’importance de développer les compétences émotionnelles afin de faire face à l’augmentation des interactions avec des agents conversationnels (cf. chapitre 5), nous ne sommes pas en mesure à ce stade de faire de recommandations plus opérationnelles sur le développement et déploiement d’enseignements ou de modules de formation pour le développement de telles compétences. De même, différentes approches prône soit l’EMI comme une matière transversale, quand d’autres appellent à en faire une matière dédiée.

Enfin, plusieurs personnes auditionnées ont relevé l’intérêt d’une base de connaissances partagée référençant l’ensemble des formations, privées et publiques, ainsi que des retours d’expériences.

Les questions pédagogiques sont nombreuses, et font état d’un grand nombre de débats. Cela renforce la nécessité d’une concertation actualisée à la lumière du développement récent de l’IA, en présence de différentes parties prenantes pour une conception de référentiels de compétences partagés ou encore la tenue de nouveaux États généraux du numérique élargis aux enjeux IA.

Proposition 8 : Responsabiliser les concepteurs et opérateurs de formations

La sensibilisation des actuels et futurs concepteurs de systèmes d'IA aux conséquences sociales, cognitives, éthiques et environnementales de leurs choix techniques constitue un impératif démocratique et doit se faire dès leur formation initiale.

Il est également impératif d'inviter les concepteurs pédagogiques à intégrer l'enseignement des humanités numériques dans l'enseignement supérieur, particulièrement dans les cursus techniques, pour permettre de sortir des rêves technicistes et des amalgames hâtifs.

Proposition 9 : Amplifier la formation professionnelle à l'IA grâce à des mécanismes redistributifs spécifiques

En 2025, plus d'une entreprise sur 4 (26 % des 11 021 entreprises ayant répondu à l'enquête du Baromètre France Num²¹) utilise des solutions d'intelligence artificielle, notamment de génération de texte/voix/image quand elles n'étaient que 13 % en 2024 (DGE, 2025). Elles sont par ailleurs plus des deux tiers à déclarer " posséder des compétences numériques. La part de celles déclarant des compétences en interne progresse fortement (55 %, +9 pts)" (DGE, 2025).

Pourtant, la formation demeure fortement corrélée au niveau de diplôme des individus et à la taille de l'entreprise. Ainsi, selon l'observatoire des inégalités, en 2024, 68 % des cadres ont suivi au moins une formation au cours de l'année 2022, contre seulement 33 % des ouvriers (Observatoire des inégalités, 2024). Bien que " le secteur d'activité joue un rôle différenciant selon le type de formation " (DGE, 2025), cette situation crée un cercle vicieux puisque les plus grandes entreprises forment (ou achètent) rapidement des compétences en IA, tandis que les plus petites peinent à suivre, renforçant les écarts de productivité et d'attractivité de l'emploi.

21. 3043 PME et 7978 TPE, dont 1 027 de 0 salarié.

Comme nous l'avons évoqué dans ce rapport, face aux centaines de milliards de dollars d'investissements prévus en IA pour 2025, et sans doute pour les prochaines années, une réallocation même symbolique de quelques pourcentages de ces montants pour la montée en compétences du plus grand nombre pourrait réduire les inégalités d'accès et d'usages aux bénéfices de l'innovation.

Il pourrait s'agir de conditionner une part des investissements publics-privés à un pourcentage minimal consacré à la formation des usagers finaux, en ciblant prioritairement les publics éloignés du numérique et les PME et TPE. Pour le secteur public, cela peut passer par une conditionnalité au sein des marchés publics, l'établissement de critères d'attribution, ou encore du financement dédié. Pour le secteur privé, des incitations fiscales ou le fléchage d'une partie de la taxe sur les services numériques, des labels et certifications ou des partenariats territoriaux pourraient servir l'intérêt général. Un système de bonus-malus territorial (avantages fiscaux renforcés pour les formations dans les territoires ruraux et les TPE-PME) ou encore un fonds de péréquation sectorielle (les secteurs à forte adoption IA contribuent à un fonds mutualisé pour former les secteurs en retard (agriculture, bâtiment) pourraient être imaginés. Les OPCO pourraient amplifier leur effort de formation à l'IA en l'imposant dans l'ensemble de leurs dispositifs de formation à destination des PME et TPE. Également, comme nous l'avons indiqué en fin de premier chapitre, il convient de s'appuyer sur les réseaux professionnels, les chambres consulaires ou les organisations patronales sectorielles pour créer des espaces d'échange et de formation mutualités pour les PME et TPE.

Proposition 10 : Renforcer les aides à la parentalité numérique

La parentalité numérique constitue un volet essentiel d'une stratégie globale de déploiement d'une littératie en IA. Elle se définit comme l'ensemble des pratiques, des compétences, et des attitudes que les parents adoptent pour accompagner leurs enfants dans l'usage des technologies numériques (interactions des enfants avec les écrans, protection du bien-être numérique,

développement d'un usage responsable et créatif des technologies...) (Livingstone & Blum-Ross, 2020). Pour répondre à cet enjeu, de nombreuses ressources pédagogiques existent pour aider les parents à aborder les thématiques comme le temps d'écran, la désinformation, la sécurité en ligne²². De même, en 2024, 62 % des structures de médiation numérique proposaient des activités pour accompagner le développement de la parentalité et l'éducation au/ par le numérique (La Mednum, 2024).

Malgré ces ressources et dispositifs, les parents, qui restent "les premiers éducateurs confrontés aux questions pratiques posées par l'omniprésence du numérique dans notre quotidien, [et qui] sont des partenaires essentiels de l'éducation au numérique" (Fréquence Ecoles & Futur composé, 2024), sont près de la moitié à ne pas se sentir suffisamment accompagnés dans l'encadrement des pratiques numériques de leur(s) enfant(s) (OPEN, UNAF & IPSOS, 2022). Une fois de plus, si les causes de ce déficit d'accompagnement sont multiples (disparités socio-économiques et culturelles, manque de temps ou d'accès aux ressources, méconnaissance des enjeux numériques), ce sont encore les familles qui ont un capital social, économique et culturel moins élevé, qui sont les moins disposées à mobiliser ces offres.

Ainsi, si nous soutenons la recommandation de l'Afnor à ce que "les structures de médiation numérique et celles accompagnant les parents [soient] sensibilisées aux enjeux de la parentalité numérique" (Afnor, 2025), nous appelons surtout à ce que ces structures soient soutenues financièrement afin de répondre à cette urgence. Les financements publics doivent être en priorité attribués aux structures qui œuvrent au plus près des publics, dans les territoires. Cela pourrait passer par davantage d'appels à projets sur les thématiques d'accompagnement au numérique et à l'IA, à l'échelle européenne (via le prochain cadre de financement pluriannuel par exemple) ou nationale.

22. Par exemple : parcours thématiques (Pix, 2025), ateliers, formations, événements, kits pour les formateurs (CLEMI, 2025), kit pour les salariés (CNIL & AFCDP, 2025), livre (Ait-Kaci et al., 2025).

Les entreprises ont également un rôle à jouer, que ce soit au niveau de ces mêmes structures qu'elles soutiennent ou peuvent soutenir par le biais de financements divers, pérennes ou ponctuels, ou par un renforcement des politiques de RSE pour l'externe, mais aussi l'interne.

Proposition 11 : Ne pas abandonner la médiation numérique

Les centres sociaux, bibliothèques, fablabs et le réseau des maisons France Services constituent un maillage de lieux de proximité où la littératie en IA peut se déployer à bas seuil : ateliers-débat, diagnostics personnalisés, accompagnement “cousu main”. Cette infrastructure de médiation numérique, qui accompagne déjà 8 millions de personnes éloignées du numérique selon la feuille de route France Numérique Ensemble, doit être adaptée aux nouveaux enjeux posés par l'IA générative.

Les expérimentations en cours montrent la faisabilité de cette approche. Dans l'Oise, par exemple, des centres sociaux sont financés en partie par La Poste et assurent des permanences dans les agences postales communales pour accompagner les usagers ruraux. Des associations comme Entre Les Lignes expérimentent des ateliers de sensibilisation à l'IA en partenariat avec des bibliothèques et des maisons de quartier. Les Cafés IA proposent une formation-kit développée par la Mednum, la coopérative des acteurs de l'inclusion numérique, permettant aux médiateurs de s'approprier un outil d'animation clés en mains.

S'il convient de mettre en place “une répartition explicite des rôles entre médiation numérique (appropriation, montée en compétences) et [un] accompagnement social-administratif (démarches ponctuelles)” pour éviter les redondances, réduire l'épuisement professionnel et optimiser l'allocation des financements (La Mednum 2024 ; ANCT, CREDOC, Université Rennes 2 CREAD-M@rsouin, 2025), il est également nécessaire de pérenniser les contrats des médiateurs et des conseillers numériques grâce à des financements entre l'Etat, les collectivités ou encore

par l'Union Européenne (par exemple, grâce aux services d'intérêt économique général - SIEG), mais aussi par le biais de dispositifs publics-privés qui permettraient par exemple de favoriser l'intervention de ces acteurs auprès de publics privés ayant un besoin de sensibilisation ou de montée en compétences (au sein des entreprises qui financent, ou d'entreprises cibles). Enfin, au-delà de l'élargissement des publics bénéficiaires de leurs actions, leur expertise des connaissances concrètes des difficultés rencontrées par les personnes qu'ils accompagnent quotidiennement pourrait être mobilisée pour de nouvelles responsabilités telles que la consultation lors de conceptions de services publics numériques locaux (quand ils sont très souvent absents de ces processus).

Cette approche redistributive doit être complétée par la proclamation de l'inclusion numérique comme priorité nationale et service d'intérêt économique général (CSNP, 2024), en concrétisant le plan France Numérique Ensemble et en réévaluant les moyens financiers pour sécuriser durablement l'écosystème de médiation.

Remerciements



Nous remercions vivement les différentes personnes auditionnées dans le cadre de ce rapport :

ADLE-AIGUIER Roxane, Head of Solution for Sustainable Economy Research, Orange.

BARDEAU Frédéric, Président cofondateur, Simplon.co.

CATTAN Jean, Secrétaire général, Conseil de l'intelligence artificielle et du numérique ; anciennement Conseil national du numérique (CNNum).

CELLARD Loup, Chercheur, Dataactivist. Chercheur associé, médialab de Sciences Po. Chercheur associé, Centre of Excellence for Automated Decision-Making and Society de l'Université de Melbourne.

CERIOLI Mathilde, Co-fondatrice et directrice scientifique, Everyone.AI.

CHAGNY Odile, Économiste à l'Institut de recherches économiques et sociales (Ires), Projet Dial-IA.

COMBLEZ Samuel, Directeur général adjoint, Association e-Enfance / 3018.

DE MARIGNAN Quitterie, (ex.) Responsable projets data, La Mednum.

REMERCIEMENTS

DERRAC Louis, Acteur militant de l'éducation au numérique et du numérique acceptable.

DUBOIS Angeline, Chargée de développement du projet Centres Sociaux Connectés, Fédération des centres sociaux des pays picards.

DUMEZ FÉROC Isabelle, Maître de conférences, Université de Poitiers (TECHNE). Directrice scientifique et pédagogique, Centre pour l'éducation aux médias et à l'information (CLEMI).

FEMENIAS Gabriel, Directeur de la prévention, Association e-Enfance / 3018.

GOMBIN Joël, Coordinateur Lab Data/IA, La Plateforme.

GUILLAUMOT Mathilde, Directrice pédagogique, Ada Tech School.

HAÏ Loïc, Conseiller technique, Département numérique, Fédération nationale des collectivités concédantes et régies (FNCCR).

LEDROIT Oriane, Déléguée Générale, EdTech France.

LEVISALLES Margaux, Co-fondatrice, Déléguée générale, Latitudes.

MARCHANDISE Jacques-François, Chercheur et prospectiviste. Professeur des universités associé, Université Gustave Eiffel.

MARTEAU Benjamin, Directeur, Pix.

MÉGEAN Olivier, Directeur fondateur, omestra.

RUGINA Roxana, Directrice Exécutive, Impact AI.

SASSOON Virginie, Directrice adjointe, Centre pour l'éducation aux médias et à l'information (CLEMI).

SIRVENT-HIEN Emilie, Responsable AI program manager, Orange.

TEISSIER Christophe, Project Manager, Ultra Laborans.

THEVENIN Emeline, (ex.) Chargée de mission, Centre Ressources Illettrisme et Illectronisme de l'Aisne (CRII 02).

TRAN Xuan-Vi, Responsable pédagogique, Ada Tech School.

Groupe de travail



Directeur de la publication

LUCAS Jean-François, Délégué général, Renaissance Numérique.

Co-animateurs du groupe de travail

CHAGNY Nicolas, Directeur délégué et qualité, CFA numiA. Président, Internet Society (ISOC) France. Membre, Commission Nationale Consultative des Droits de l'Homme (CNCDH).

ORSINI SHARPS Laetitia, Directrice Grand Public, Orange France.

Rapporteurs

BLONDEL Mélusine, Co-directrice générale, La Mednum. Membre, Conseil d'administration de Renaissance Numérique.

BRUYAS Dorie, Présidente, La Mednum.

COMBLEZ Samuel, Directeur général adjoint, e-Enfance / 3018.

GAVEN Frédéric, Fondateur, Nosphi Conseil.

GOMEZ Denis, Directeur de projets Innovation Sociétale, Orange.

Le GOFF Samuel, Président, Renaissance Numérique.

NARZUL Ruben, Membre associé, Renaissance Numérique.

PÉNICAUD Soizic, Consultante en politiques publiques de l'IA. Cofondatrice, Observatoire des algorithmes publics. Enseignante, Sciences Po Paris.

RELTGEN Tanguy, Rapporteur, Renaissance Numérique.

SCHOENTGEN Aude, Économiste du numérique. Membre, Conseil d'administration de Renaissance Numérique.

Contributeurs

BECHU Véronique, Directrice de l'Observatoire contre le harcèlement et les violences numériques faites aux mineurs, Association e-Enfance / 3018.

BÜSCHER Jan, Co-directeur général, La Mednum.

DEMURE Joan, (ex.) Rapporteur, Renaissance Numérique.

De RAMECOURT Olivier,
Fondateur, Looping.

EXCOFFIER Cathy, Directrice
Déléguée RSE, Orange.

GALISSAIRE Jessica, (ex.)
Responsable des études et des
partenariats, Renaissance Numérique.

GUIMBAIL Laure, Rapporteuse,
Renaissance Numérique.

ISAAC Henri, Maître de conférences,
Université Paris-Dauphine - PSL.
Membre, Conseil d'administration
de Renaissance Numérique.

LECOINTE Bénédicte, Responsable
compliance, Orange.

LEPINETTE Martin, Chargé d'études et
de projets, Renaissance Numérique.

MAREL Julian, (ex.) Rapporteur,
Renaissance Numérique.

MOSSÉ Marc, Avocat Senior
Counsel, August Debouzy.

REINBOLD Léo, (ex.) Rapporteur,
Renaissance Numérique.

SIRVENT-HIEN Emilie, Responsable
AI program manager, Orange.

SZELERSKI Jean-Pascal, Directeur
Conseil, TRÉMA. Membre, Internet
Society (ISOC) France.

VANBREMEERSCH Nicolas,
Fondateur, Spintank.

Van de BLANKEVOORT Astrid, (ex.)
Rapporteuse, Renaissance Numérique.

VERNIER Ondine, Chef de projet
Ecosystèmes, La Mednum.

Bibliographie



ADEME, & Arcep. (2025). *Actualisation des chiffres de l'impact du numérique en France*. ADEME & Arcep. <https://ecoresponsable.numerique.gouv.fr/actualites/actualisation-ademe-impact>

ADEME. (2024). *Étude numérique et métaux : Impacts environnementaux du numérique et besoins en métaux*. Paris : ADEME. <https://librairie.ademe.fr/economie-circulaire-et-dechets/7713-etude-numerique-et-metaux.html>

Afnor. (2025). *Inclusion numérique : réduire les inégalités*. Afnor Spec.

AI Now Institute. (2025). *Artificial power. 2025 Landscape Report*. <https://ainowinstitute.org/reports/artificial-power-2025.html>

Ait-Kaci, Y., Boudet, C., Castex, L., Chagny, N., & Pulicani, C. (2025). *Devenir gardien de son Internet : reprendre la main sur ses données personnelles, appréhender l'intelligence artificielle*. <https://www.devenir-gardien-internet.fr>

Akrich, M., Callon, M., & Latour, B. (2006). *Sociologie de la traduction. Textes fondateurs*. Paris : Ecole des Mines.

Amable, B. (2003). Systèmes d'innovation. In P. Mustar & H. Penan (Eds.), (Éds.), *Encyclopédie de l'innovation* (pp. 147-162). Paris : Economica.

Ananya. (2025, 23 septembre). AI models are using material from retracted scientific papers. *MIT Technology Review*. <https://www.technologyreview.com/2025/09/23/1123897/ai-models-are-using-material-from-retracted-scientific-papers>

ANCT, CREDOC, & Université Rennes 2 CREAD-M@rsouin. (2023). *La société numérique française : définir et mesurer l'éloignement numérique*. <https://www.credoc.fr/publications/la-societe-numerique-francaise-definir-et-mesurer-leloignement-numerique>

ANCT, CREDOC, & Université Rennes 2 CREAD-M@rsouin (2025). *La société numérique française : les professionnels de la médiation numérique au défi de l'éloignement numérique*. <https://www.credoc.fr/publications/la-societe-numerique-francaise-les-professionnels-de-la-mediation-numerique-au-defi-de-leloignement-numerique>

Andler, D. (2023). *Intelligence artificielle, intelligence humaine : la double énigme*. Paris : Gallimard.

- Anthropic. (2025). *Project Vend : Can Claude run a small shop ? (And why does that matter ?)*. <https://www.anthropic.com/research/project-vend-1>
- American Psychological Association. (2024). *Health advisory on the use of artificial intelligence for adolescent mental health and well-being*. APA. <https://www.apa.org/topics/artificial-intelligence-machine-learning/health-advisory-ai-adolescent-well-being>
- Assemblée nationale. (2025). *Commission d'enquête sur les effets psychologiques de TikTok sur les mineurs*. <https://www.assemblee-nationale.fr/dyn/17/organes/autres-commissions/commissions-enquete/tiktok>
- Bachimont, B. (1996). *L'intelligence artificielle comme écriture dynamique : de la raison graphique à la raison computationnelle*. https://www.utc.fr/~bachimon/Publications_attachments/BachimontCerisy1996.pdf
- Ballard, P., & Spillebout, V. (2023). *Mission flash sur l'éducation critique aux médias*. Assemblée nationale. Commission des affaires culturelles et de l'éducation. <https://www.assemblee-nationale.fr/dyn/16/organes/commissions-permanentes/affaires-culturelles/missions-de-la-commission/mi-education-medias>
- Barnes, R. (2023, 30 octobre). Villeurbanne s'engage à garantir l'accès des services publics sans passer par internet. *LE PROGRÈS*. <https://www.leprogres.fr/politique/2023/10/30/villeurbanne-s-engage-a-garantir-l-acces-des-services-publics-sans-passer-par-internet>
- Barr, N., Pennycook, G., Stolz, J. A., & Fugelsang, J. A. (2015). The brain in your pocket: Evidence that Smartphones are used to supplant thinking. *Computers in Human Behavior*, 48, 473–480.
- Benbouzid, B. (2025, 16 juillet). IA génératives et éducation. *AOC*. <https://aoc.media/analyse/2025/07/15/ia-generatives-et-education>

- Berkouk, N., Arfaoui M., & Pialat, R. (2025). *L'explicabilité de l'IA : un problème renouvelé par le succès du deep learning*. Laboratoire d'Innovation Numérique de la CNIL (LINC). <https://linc.cnil.fr/lexplicabilite-de-lia-un-probleme-renouvele-par-le-succes-du-deep-learning-13>
- Bellon, A., & Velkovska, J. (2023). L'intelligence artificielle dans l'espace public : du domaine scientifique au problème public. *Réseaux, Réseaux, N° 240(4)*, 31-70. <https://doi.org/10.3917/res.240.0031>
- Bendekkiche, H., & Viard-Guillot, L. (2023). *15 % de la population est en situation d'illectronisme en 2021*. *Insee Première*, (1953). <https://www.insee.fr/fr/statistiques/7633654>
- Bender, E., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots. *Proceedings Of FAccT*, 610-623. <https://doi.org/10.1145/3442188.3445922>
- Bengio, Y. (2025). *International AI Safety Report 2025*. Department for Science, Innovation and Technology. GOV.UK. <https://www.gov.uk/government/publications/international-ai-safety-report-2025#full-publication-update-history>
- Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford : Oxford University Press.
- Boullier, D. (2023). *Propagations. Un nouveau paradigme pour les sciences sociales*. Armand Colin, Collection U.
- Bousquette, I. & Maurer, M (2025). *How the AI boom is leaving consultants behind*. The Wall Street Journal. <https://www.wsj.com/articles/how-the-ai-boom-is-leaving-consultants-behind-c9088fda>
- Brabenec, R. (2025, 7 juillet). *A billionaire, an AI supercomputer, toxic emissions and a Memphis community that did nothing wrong*. Tennessee Lookout. <https://tennesseelookout.com/2025/07/07/a-billionaire-an-ai-supercomputer-toxic-emissions-and-a-memphis-community-that-did-nothing-wrong>
- Bruyen, C., & Fialaire, B. (2025). *L'IA et l'avenir du service public*. Sénat. Rapport thématique #3. IA et éducation. <https://www.senat.fr/rap/r24-101/r24-101-syn.pdf>
- Brynjolfsson, E., Bharat, C., & Chen, R. (2025). *Canaries in the Coal Mine? Six Facts about the Recent Employment Effects of Artificial Intelligence*. Stanford Digital Economy Lab.

BIBLIOGRAPHIE

Butler, J., Vorvoreanu, M., Janßen, R., Sellen, A., Immorlica, N., Hecht, B., & Teevan, J. (Eds.). (2024). *The New Future of Work - Microsoft Research*. <https://aka.ms/nfw2024>

Carr, N. (2008). *Is Google making us stupid? What the internet is doing to our brains?* The Atlantic. <https://www.theatlantic.com/magazine/archive/2008/07/is-google-making-us-stupid/306868>

Carr, N. (2010). *The Shallows: What the Internet is Doing to Our Brains*. New York: W. W. Norton & Company.

Carroll, S. (2025). *Sam Altman AI bubble warning*. Quartz. <https://qz.com/sam-altman-ai-bubble-open-ai-warning>

Casilli, A. (2019). *En attendant les robots. Enquête sur le travail du clic*. Paris : Le Seuil, Paris.

Casilli, A. (2025). *Young workers haven't been replaced by AI : economists are just looking for them in the wrong places*. <https://www.casilli.fr/2025/08/29/young-workers-havent-been-replaced-by-ai-economists-are-just-looking-for-them-in-the-wrong-places/>

Cavazza, F. (2025, 23 avril). *Les modèles économiques de l'IA générative*. FredCavazza.net. <https://fredcavazza.net/2025/04/23/les-modeles-economiques-de-lia-generative>

Chatterji, A., Cunningham, T., Deming, D., Hitzig, Z., Ong, C., Shan, C., & Wadman, K. (2025). *How people use ChatGPT*. OpenAI, Harvard University, Duke University.

Cité des sciences et de l'industrie. (2025). *Stage de science enfant : Intelligence artificielle*. <https://www.cite-sciences.fr/fr/au-programme/activites-spectacles/ateliers-stages-clubs/stages-de-science/en-savoir-plus/stage-de-science-enfant-intelligence-artificielle>

Centre pour l'éducation aux médias et à l'information (CLEMI). (2025). *Kit de médiation : Les écrans ? apprendre à s'en servir pour ne pas les subir*. <https://www.clemi.fr/familles/outils-de-sensibilisation-et-mediation/kit-de-mediation/le-kit-de-mediation/kit-de-mediation-les-ecrans-apprendre-sen-servir-pour-ne-pas-les-subir>

Clerget, J., Plantard, P., & Bléhaut, M. (2024). *Pour en finir avec les "13 millions" ? Au défi de rendre compte de l'éloignement du numérique*. *Terminal*, 139. <https://doi.org/10.4000/132iz>

CNIL, & AFCDP. (2025). *Parentalité numérique : la CNIL et l'AFCDP outillent les DPO pour sensibiliser les parents salariés*. <https://www.cnil.fr/fr/parentalite-numerique-la-cnil-et-lafcdp>

Commission de l'intelligence artificielle. (2024). *IA : notre ambition pour la France*. <https://www.info.gouv.fr/upload/media/content/0001/09/4d3cc456dd2f5b-9d79ee75feea63b47f10d75158.pdf>

Commission européenne. (2024). *Artificial Intelligence Act (Règlement (UE) 2024/1689)*. Journal officiel de l'Union européenne. https://eur-lex.europa.eu/legal-content/FR/TXT/PDF/?uri=OJ:L_202401689

Commission européenne. (2025a, 4 février). *Living repository to foster learning and exchange on AI literacy*. Brochure. <https://digital-strategy.ec.europa.eu/en/library/living-repository-foster-learning-and-exchange-ai-literacy>

Commission européenne. (2025b, 9 avril). *AI Continent Action Plan*. COM(2025) 165 final. <https://digital-strategy.ec.europa.eu/fr/library/ai-continent-action-plan>

Commission Supérieure du Numérique et des Postes. (2024). *Avis n° 2024-01 du 17 janvier 2024 pour mieux encadrer l'usage de l'intelligence artificielle*. CSNP.

Conseil national du numérique (CNNum). (2013). *Citoyens d'une société numérique : pour une nouvelle politique d'inclusion*. <https://cnnumerique.fr/nos-travaux/inclusion-numerique>

Conseil national du numérique (CNNum). (2023). *Pour un numérique au service de la construction des savoirs*. <https://cnnumerique.fr/files/2022-09/CNNum-Pour-un-numerique-au-service-des-savoirs.pdf>

Cordier, A., & Erhel, S. (dir.) (2023). *Les enfants et les écrans*. Paris : Retz, Collection Mythes et réalités.

Cousteaux, A. S. (2019). *L'économie et la société à l'ère du numérique*. Paris : Insee. <https://www.insee.fr/fr/statistiques/4238635>

Consul, M. (2025). Les progrès inattendus de l'IA au travail : d'une utilisation "clandestine" à une utilisation éthique. Interview, Parlons RH. <https://www.parlonsrh-com.cdn.ampproject.org/c/s/www.parlonsrh.com/media/les-progres-inattendus-de-lia-au-travail-dune-utilisation-clandestine-a-une-utilisation-ethique/amp>

CREDOC. (2023). *Baromètre du numérique*. Commission Démographie et Questions Sociales du CNIS. <https://www.cnis.fr/wp-content/uploads/2023/05/presentation-2-2-credoc.pdf>

BIBLIOGRAPHIE

Crownhart, C. (2025, 21 août). In a first, Google has released data on how much energy an AI prompt uses. *MIT Technology review*. <https://www.technologyreview.com/2025/08/21/1122288/google-gemini-ai-energy>

Data For Good. (2023). *Les propositions du collectif Data For Good pour une IA générative plus responsable : un guide pour les utilisateurs, artistes, développeurs et décideurs*.

Derrida, J. (1968). La pharmacie de Platon. *Tel Quel*, n° 32-33.

Deville, C. (2023). *L'État social à distance : dématérialisation et accès aux droits des classes populaires rurales*. Paris : Éditions du Croquant.

DigitalCMO (2025, 12 mai). *Étude Européenne sur l'IA Stratégir & Norstat : une adoption rapide malgré des différences d'appropriation*. <https://www.digitalcmo.fr/etude-europeenne-sur-lia-strategir-norstat-une-adoption-rapide-malgre-des-differences-dappropriation>

Digital Education Council. (2024). *Digital Education Council Global AI Student Survey 2024*. <https://www.digitaleducationcouncil.com/post/digital-education-council-global-ai-student-survey-2024>

Ding, J. (2024). The Rise and Fall of Technological Leadership: General-Purpose Technology Diffusion and Economic Power Transitions. *International Studies Quarterly*, 68(2). <https://doi.org/10.1093/isq/sqae013>

Emmaüs Connect, Croix-Rouge Française, ANCT, Agefiph. (2024). *L'inclusion numérique des personnes en situation de handicap*. [https://storage.lesbases.anct.gouv.fr/main/user/d342c0e0-9177-4b39-b504-3a9c12018e90/ae7Ba-82DUCh6hzm1yc4Sc_Rapport%20Accessibilit%C3%A9%20num%C3%A9rique%20des%20PSH%20\(1\).pdf](https://storage.lesbases.anct.gouv.fr/main/user/d342c0e0-9177-4b39-b504-3a9c12018e90/ae7Ba-82DUCh6hzm1yc4Sc_Rapport%20Accessibilit%C3%A9%20num%C3%A9rique%20des%20PSH%20(1).pdf)

Eshet, Y. (2004). Digital Literacy: A Conceptual Framework for Survival Skills in the Digital era. *Journal of Educational Multimedia and Hypermedia*, 13(1), 93-106. <https://www.learntechlib.org/primary/p/4793/>

EPFL. (2022). Réinventer la formation. *Dimensions. Le magazine de l'éducation, la recherche et l'innovation de l'EPFL*, 5. Lausanne : EPFL. <https://www.epfl.ch/about/news-and-media/fr/dimensions/edition-electronique/>

Ferguson, Y. (2025). *L'intelligence artificielle au travail. Accompagner et sécuriser les initiatives des collaborateurs*. Synthèse, juin 2025. https://datacraft.paris/wp-content/uploads/2025/05/Synthese-Inria-x-datacraft-V4_compressed.pdf

Flichy, P. (2010). *Le sacre de l'amateur. Sociologie des passions ordinaires à l'ère numérique*. Paris : Éditions du Seuil.

Floridi, L. (2024). Why the AI Hype is Another Tech Bubble. *Philosophy & Technology*, 37(4). <https://doi.org/10.1007/s13347-024-00817-w>

France Num. (2024, 19 septembre). *Baromètre France Num. Résultats de l'enquête*. <https://www.francenum.gouv.fr/files/2024-09/Barom%C3%A8tre%20France%20Num%202024%20-%20Rapport%20complet.pdf>

France Num. (2025, 15 septembre). *Baromètre France Num 2025 : le numérique et l'intelligence artificielle dans les TPE et PME*. francenum.gouv.fr. <https://www.francenum.gouv.fr/guides-et-conseils/strategie-numerique/comprendre-le-numerique/barometre-france-num-2025-le>

Frau-Meigs, D. (2024). *User empowerment through Media and Information Literacy responses to the evolution of Generative Artificial Intelligence*. UNESCO Policy Brief.

Fondation pour l'enfance. *L'IA générative, nouvelle arme de la pédocriminalité*. Octobre 2024. <https://www.fondation-enfance.org/2024/10/29/rapport-alarant-sur-lia-generative-et-la-pedocriminalite>

Fréquence Écoles, & Futur composé. (2024). *Re-penser l'éducation au numérique : pour développer des méthodes d'intervention et de conception émancipatrices*.

Gartner. (2025, 31 mars). *Gartner Forecasts Worldwide GenAI Spending to Reach \$644 Billion in 2025*. <https://www.gartner.com/en/newsroom/press-releases/2025-03-31-gartner-forecasts-worldwide-genai-spending-to-reach-644-billion-in-2025>

Guillaud, H. (2025). *Les algorithmes contre la société*. La Fabrique Éditions.

Goertzel, B. (2014). Artificial General Intelligence: Concept, State of the Art, and Future Prospects. *Journal of Artificial General Intelligence*, 5(1).

BIBLIOGRAPHIE

Goldman Sachs. (2024). *Gen AI: Too much spend, too little benefit?* <https://www.goldmansachs.com/insights/top-of-mind/gen-ai-too-much-spend-too-little-benefit>

Groupe d'experts ÉMI. (2021). *Renforcer l'éducation aux médias et à l'information et la citoyenneté numérique : Rapport à monsieur le ministre de l'Éducation nationale, de la Jeunesse et des Sports – mai 2021*. Paris : Ministère de l'Éducation nationale. <https://www.vie-publique.fr/files/rapport/pdf/280669.pdf>

Haider, J., & Sundin, O. (2022). *Paradoxes of Media and Information Literacy. The Crisis of Information*. London: Routledge. <https://doi.org/10.4324/9781003163237>

Hausenloy, J., & Gulati, S. (2024). *Brief 2024 Generation AI Survey*. Center for Youth and AI. <https://survey.youth-ai.org>

HUB France IA. (2024). *Les usages de l'IA générative : Volume 1, les LLM*. Livre blanc. https://www.hub-franceia.fr/wp-content/uploads/2024/02/Livre-blanc_Les-usages-de-lia-generative-01.2024.pdf

International Energy Agency. (2025). *Transport – Energy System Overview*. Paris : IEA. <https://www.iea.org/energy-system/transport>

Ifop. (2024). Baromètre 2024. *Les Français et les IA génératives. Vague 2*. Ifop pour Talan. <https://www.ifop.com/wp-content/uploads/2024/07/120717-rapport-reduit.pdf>

Insee & DEPP. (2024). *Enquête "Formation tout au long de la vie"*.

International Telecommunication Union. (2024). *Climate Change – Backgrounder*. Genève : ITU.

Ipsos, & CESI. (2025). *L'usage de l'intelligence artificielle par les français*. <https://www.cesi.fr/wp-content/uploads/2025/02/ipsos-cesi-usage-intelligence-artificielle-rapport-complet-1.pdf>

JBB. (2018). Smog. Géoconfluences. <https://geoconfluences.ens-lyon.fr/glossaire/smog>

Journal officiel de la République française. (2024). LOI no 2024-449 du 21 mai 2024 visant à sécuriser et à réguler l'espace numérique.

Kalai, A.D., Nachum, O. Vempala, S.S., & Zhang, E. (2025). *Why Language Models Hallucinate*. arXiv. <https://arxiv.org/pdf/2509.04664>

Kapur, M. (2008). Productive Failure. *Cognition and Instruction*, 26(3), 379-424. <https://doi.org/10.1080/07370000802212669>

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025, 10 juin). *Your Brain on ChatGPT : Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task*. arXiv.org. <https://arxiv.org/abs/2506.08872>

La Mednum. (2024). *Observatoire de l'inclusion numérique. Réalités et perspectives du secteur de la médiation numérique en 2024*. <https://lamednum.coop/wp-content/uploads/2024/09/La-Mednum-Observatoire-de-linclusion-numerique-2024-2.pdf>

Deuff, O. L. (2012). Littératies informationnelles, médiatiques et numériques : de la concurrence à la convergence ? *Études de Communication/ Études de Communication*, 38, 131-147. <https://doi.org/10.4000/edc.3411>

Lefebvre, C. (2025). *Les technologies de l'information et de la communication dans les entreprises en 2024*. Insee <https://www.insee.fr/fr/statistiques/8604126?sommaire=7667330>

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2020, 22 mai). *Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks*. arXiv.org. <https://arxiv.org/abs/2005.11401>

Livingstone S., & Blum-Ross A. (2020), *Parenting for a Digital Future: How Hopes and Fears about Technology Shape Children's Lives*. Oxford University Press.

Lucas, J.-F. (2024, 5 avril). Intelligence artificielle : “ L’humain, qui est le point de départ de tout dispositif numérique, doit rester le point d’arrivée ” . *Le Monde.fr*. https://www.lemonde.fr/idees/article/2024/04/05/intelligence-artificielle-l-humain-qui-est-le-point-de-depart-de-tout-dispositif-numerique-doit-rester-le-point-d-arrivee_6226135_3232.html

Lucas, J.-F. (2025, 25 février). IA, faire face à cet enjeu majeur pour les jeunes générations. *Chut ! magazine*, n°17, l'IA, c'est plus fort que toi?

Maese, E. (2025, 15 janvier). Americans Use AI in Everyday Products Without Realizing It. *Gallup.com*. <https://news.gallup.com/poll/654905/americans-everyday-products-without-realizing.aspx>

BIBLIOGRAPHIE

Mayer, J. D., & Salovey, P. (2004). What is emotional intelligence? In P. Salovey, M. A. Brackett, & J. D. Mayer (Eds.), *Emotional intelligence: Key readings on the Mayer and Salovey model* (pp. 29–59). Dude Publishing.

McCarthy, J., Minsky, M., Rochester, N., & Shannon, C. E. (1955). *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. Unpublished manuscript, Dartmouth College, Hanover, NH.

Medef, Numeum. (2025). *L'intelligence artificielle au coeur de nos entreprises. Une opportunité à saisir, une responsabilité à partager*. La Tribune publishing. https://www.perspectives-ia.fr/wp-content/uploads/2025/02/MEDEF_PERSPECTIVES_IA_LIVRE_BLANC_SYNTHESE_TOUR_DE_FRANCE_IA.pdf

Metr. (2025, 19 mars). Measuring AI Ability to Complete Long Tasks. *METR*. <https://metr.org/blog/2025-03-19-measuring-ai-ability-to-complete-long-tasks>

Ministère de l'Économie, des Finances et de la Souveraineté industrielle et numérique (MEF). (2025a). *Osez l'IA : un plan pour diffuser l'IA dans toutes les entreprises*. Economie.gouv.fr. <https://www.economie.gouv.fr/actualites/osez-lia-un-plan-pour-difuser-lia-dans-toutes-les-entreprises>

Ministère de l'Économie, des Finances et de la Souveraineté industrielle et numérique (MEF). (2025b). *Clara Chappaz confie à Gilles Babinet la mission de consolider et pérenniser Café IA, initiative phare du dialogue démocratique autour de l'intelligence artificielle*. Communiqué de presse. Economie.gouv.fr. <https://presse.economie.gouv.fr/clara-chappaz-confie-a-gilles-babinet-la-mission-de-consolider-et-perenniser-cafe-ia-initiative-phare-du-dialogue-democratique-autour-de-lintelligence-artificielle>

Ministère de l'Éducation nationale et de la Jeunesse. (2024). *Cadre d'usage de l'IA en éducation*. <https://www.education.gouv.fr/cadre-d-usage-de-l-ia-en-education-450647>

Ministère de l'Éducation nationale. (2025). *Intelligence artificielle au service de l'éducation : des mesures ambitieuses pour accompagner les usages des élèves et des professeurs*. Éducation.gouv.fr. <https://www.education.gouv.fr/intelligence-artificielle-au-service-de-l-education-des-mesures-ambitieuses-pour-accompagner-les-416551>

Mims, C. (2024, 11 octobre). This AI Pioneer thinks AI is dumber than a cat. *The Wall Street Journal*. <https://www.wsj.com/tech/ai/yann-lecun-ai-meta-aa59e2f5>

- Minsky, M. (1968). *Semantic Information Processing*. Cambridge, MA : MIT Press.
- Mistral AI. (2025, 22 juillet). *Notre contribution pour la création d'un standard environnemental mondial pour l'IA*. <https://mistral.ai/fr/news/our-contribution-to-a-global-environmental-standard-for-ai>
- Moles, A. (1990). La fonction des mythes dynamiques dans la construction de l'imaginaire social. *Cahiers de l'imaginaire*, n°5-6, p.9-33.
- Mollick, E. (2025, 7 juillet). Against "Brain damage" . *One Useful Thing*. <https://www.oneusefulthing.org/p/against-brain-damage>
- Monget, P. (2025). *Enquête adoption de la GenAI par les TPE et PME, 2025*. France Hub IA & JUNIA.
- Musso, P., Coiffier, S., & Lucas, J.F. (2014). *Innover avec et par les imaginaires*. Paris, Manucius, Modélisations des imaginaires.
- Newman, N., Fletcher, R., Robertson, C. T., Arguedas, R., A., & Nielsen, R. K. (2024). *Reuters Institute Digital News Report 2024*. Oxford: Reuters Institute for the Study of Journalism, University of Oxford.
- Observatoire des inégalités. (2024, 1er octobre). *Les inégalités d'accès à la formation professionnelle*. <https://www.inegalites.fr/Les-inegalites-d-acces-a-la-formation-professionnelle>
- Observatoire international sur les impacts sociétaux de l'IA et du numérique (OBVIA). (2024). *Abécédaire de l'IA*. Québec : OBVIA. <https://www.obvia.ca/ressources/abecedaire-de-lia>
- OCDE. (2023). *OECD Framework for the Classification of AI Systems*. Paris : OCDE.
- OCDE. (2024a). *Évaluation des compétences des adultes 2023 – Note par pays : France*. Paris : OCDE. https://www.oecd.org/content/dam/oecd/fr/publications/reports/2024/12/survey-of-adults-skills-2023-country-notes_df7b4a60/france_ba42b6a4/21bacb82-fr.pdf
- OCDE. (2024b). Les adultes possèdent-ils les compétences nécessaires pour s'épanouir dans un monde en mutation ? *Études de l'OCDE sur les compétences*. <https://doi.org/10.1787/e8d52c02-fr>
- OCDE. (2024c). Assessing potential future artificial intelligence risks, benefits and policy imperatives. *OECD Artificial Intelligence Papers*, n°.27.

BIBLIOGRAPHIE

OCDE, & Commission Européenne. (2025). *Empowering learners for the age of AI: An AI literacy framework for primary and secondary education* (Review draft). https://ailiteracyframework.org/wp-content/uploads/2025/05/AILitFramework_ReviewDraft.pdf

OpenAI. (2023). *Function Calling and Tool Use: Model-Context Protocol* (draft). GitHub repository.

OPEN, UNAF, & IPSOS. (2022). *Parents, enfants et numérique : étude nationale*. https://www.open-asso.org/wp-content/uploads/2022/02/Etude_OPEN_UNAF_IPSOS_Parents_enfants_numerique.pdf

Pascal, F., Taddei, F., de Falco, M., & Gallié, E.-P. (2025). *IA et enseignement supérieur : Formation, structuration et appropriation par la société*. Ministère de l'Enseignement supérieur et de la Recherche. <https://www.enseignement-sup-recherche.gouv.fr/sites/default/files/2025-07/rapport-intelligence-artificielle-et-enseignement-sup-rieur-formation-structuration-et-appropriation-par-la-soci-t--37540.pdf>

Pasquier, D. (2018). *L'Internet des familles modestes*. Paris : Presses des Mines.

Perrigo, B. (2023, January 18). Exclusive: The \$2 Per Hour Workers Who Made ChatGPT Safer. *TIME Magazine*. <https://time.com/6247678/openai-chatgpt-kenya-workers>

Pichai, S. (2025). AI will be more consequential to humanity than fire or electricity. *OfficeChai*. <https://officechai.com/ai/ai-will-be-more-consequential-to-humanity-than-fire-or-electricity-google-ceo-sundar-pichai>

Piquard, A. (2024, août 31). “L’intelligence artificielle est une bulle : il y a un décalage entre les coûts, très importants, et les revenus potentiels” . *Le Monde.fr*. https://www.lemonde.fr/economie/article/2024/08/31/l-intelligence-artificielle-est-une-bulle-il-y-a-un-decalage-entre-les-couts-tres-importants-et-les-revenus-potentiels_6300034_3234.html

Pix. (2025). Aux côtés des familles pour la maîtrise du numérique. <https://pix.fr/parents#slice-2>

Plantard, P. (2015). *Les imaginaires en éducation*. Paris : Manucius.

Rashidi, S. (2025, 16 juillet). *OpenAI's \$ 10 million+ AI consulting business : Deployment takes center stage*. Forbes. <https://www.forbes.com/sites/solrashidi/2025/07/16/openais-10m-ai-consulting-business-deployment-takes-center-stage>

Ragnedda, M. (2018). Conceptualizing digital capital. *Telematics and Informatics*, 35(8), 2366–2375.

Renaissance Numérique (2025). *Vers une coopération mondiale autour de la sécurité de l'IA. Synthèse de la table ronde organisée par Renaissance Numérique en partenariat avec Microsoft et Impact AI pendant le Sommet pour l'action sur l'intelligence artificielle.* <https://www.renaissancenumerique.org/publications/vers-une-cooperation-mondiale-autour-de-la-securite-de-lia>

Régnauld, I. (2025). *Huit mythes de l'intelligence artificielle pour déconstruire la "hype"*. Mais Où Va le Web | Panser le Numérique. <https://maisouvaleweb.fr/huit-mythes-de-lintelligence-artificielle-pour-deconstruire-la-hype>

Reuters (2025, 31 juillet). *OpenAI hits \$12 billion in annualized revenue, The Information reports.* <https://www.reuters.com/business/openai-hits-12-billion-annualized-revenue-information-reports-2025-07-31>

Richter, H. (2025, 26 juillet). *It's Most Empathetic Voice in My Life: How AI Is Transforming Lives of Neurodivergent.* *Reuters.* <https://www.reuters.com/lifestyle/its-most-empathetic-voice-my-life-how-ai-is-transforming-lives-neurodivergent-2025-07-26>

Roose, K. (2024, 15 avril). *A.I. Has a Measurement Problem.* *The New York Times.* <https://www.nytimes.com/2024/04/15/technology/ai-models-measurement.html>

Ruiu, M. L., & Ragnedda, M. (2020). Digital capital and online activities : An empirical analysis of the second level of digital divide. *First Monday.* <https://doi.org/10.5210/fm.v25i7.10855>

Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach* (4e éd.). Upper Saddle River, NJ : Pearson.

Ruyen, C., Fialaire, B. (2024). *L'IA et l'avenir du service public.* Rapport thématique Sénat.

Rządeczka, M., Sterna, A., Stolińska, J., Kaczyńska, P., & Moskalewicz, M. (2024c). The Efficacy of Conversational Artificial Intelligence in Rectifying the Theory of Mind and Autonomy Biases : Comparative Analysis (Preprint). *JMIR Mental Health*, 12, e64396. <https://doi.org/10.2196/64396>

S&P Global. (2025, 30 mai). *Generative AI experiences rapid adoption, but with mixed outcomes – Highlights from Vote: AI & Machine Learning.* <https://www.spglobal.com/market-intelligence/en/news-insights/research/ai-experiences-rapid-adoption-but-with-mixed-outcomes-highlights-from-vote-ai-machine-learning>

BIBLIOGRAPHIE

Salvator, D. (2025). *Explication : les tokens, le langage et la monnaie de l'IA*. Nvidia blogs. <https://blogs.nvidia.fr/explication-des-jetons-le-langage-et-la-monnaie-de-lia>

Searle, J. (1980). Minds, Brains, and Programs. *Behavioral and Brain Sciences*, 3(3), 417–457.

Selber, S. A. (2004). *Multiliteracies for a Digital Age*. Carbondale, IL : Southern Illinois University Press.

Sen, A. (1999). *Development as Freedom*. Oxford : Oxford University Press.

Serret, M. (2024, 13 juin). *L'urgence d'une véritable littératie de l'IA pour l'autonomisation des élèves*. Documentation Lille. <https://pedagogie.ac-lille.fr/prof-doc/lurgence-dune-veritable-litteratie-de-lia-pour-lauto-nomisation-des-eleves>

Servais, O. (2020). *Dans la peau des gamers. Anthropologie d'une guilda de World of Warcraft*. Paris : Karthala.

Shirky, C. (2010). *Cognitive Surplus: Creativity and Generosity in a Connected Age*. New York: Penguin Books.

Sidoti, O., & McClain, C. (2025). 34 % of U.S. adults have used ChatGPT, about double the share in 2023. *Pew Research Center*. <https://www.pewresearch.org/short-reads/2025/06/25/34-of-us-adults-have-used-chatgpt-about-double-the-share-in-2023>

Simondon, G. (1958). *Du mode d'existence des objets techniques*. Paris : Aubier.

Stiegler, B. (2007). Questions de pharmacologie générale. Il n'y a pas de simple pharmakon. *Psychotropes*, 13(3), 27–54.

Spencer, P.T. (2024). Human-Centric NLP or AI-Centric Illusion?: A Critical Investigation. *Proceedings of the 38th PACLIC*, pp.773–779

Tatot, F., & Vermot Desroches, G. (2024). *Impacts de l'intelligence artificielle : risques et opportunités pour l'environnement*. Conseil économique social et environnemental. https://www.lecese.fr/sites/default/files/pdf/Avis/2024/2024_14_IA_Environnement.pdf

Thorn. (2024, août 14). *REPORT : 1 in 10 Minors Say Peers Have Used AI to Generate Nudes of Other Kids*. <https://www.thorn.org/press-releases/report-1-in-10-minors-say-peers-have-used-ai-to-generate-nudes-of-other-kids>

UNCTAD - Conférence des Nations Unies sur le commerce et le développement. (2025). *Technology and Innovation Report 2025: Chapter V – Global collaboration for inclusive and equitable AI*. Genève. https://unctad.org/system/files/official-document/tir2025ch5_en.pdf

UNESCO. (2025). *What you need to know about literacy*. <https://www.unesco.org/en/literacy/need-know#:~:text=UNESCO%20defines%20digital%20literacy%20as,employment%2C%20decent%20jobs%20and%20entrepreneurship>

UNICEF France. (2023). *Enfants non-scolarisés : l'angle mort de l'école de la République*. <https://www.unicef.fr/article/enfants-non-scolarisés-l'angle-mort-de-lecole-de-la-republique>

United Nations Environment Programme (2024). *2023 Global Status Report for Buildings and Construction : Beyond foundations - Mainstreaming sustainable solutions to cut emissions from the buildings sector*. <https://doi.org/10.59117/20.500.11822/45095>

United Nations Development Programme (UNDP). (2025). *A Matter of Choice: People and Possibilities in the Age of AI*. Human Development Report. <https://hdr.undp.org/system/files/documents/global-report-document/hdr2025reporten.pdf>

Union européenne. Parlement européen et Conseil de l'UE. (2024). Règlement établissant des règles harmonisées concernant l'intelligence artificielle 2024/1689, 13 juin 2024. https://eur-lex.europa.eu/legal-content/FR/TXT/PDF/?uri=OJ:L_202401689

Universcience. (2025). *Baromètre esprit critique 2025*. Paris: Universcience. <https://www.universcience.fr/fr/esprit-critique/barometre-esprit-critique-2025>

Université du Québec à Montréal (UQAM). (2024). *Le milieu de l'éducation s'active et cogite à l'IA : Des initiatives émergentes et des sources inspirantes à suivre*. <https://collimateur.uqam.ca/collimateur/le-milieu-de-leducation-sactive-et-cogite-a-lia/>

Université Rennes 2. (2025). *Référentiel de compétences et de connaissances. Transition Écologique et Sociale*.

Van de Blanckevoort, A. (2025). *Éthique et numérique, avec Claude Kirchner. Renaissance Numérique*. <https://www.renaissancenumerique.org/publications/ethique-et-numerique-avec-claude-kirchner>

Villani, C. (2018). *Donner un sens à l'intelligence artificielle. Pour une stratégie nationale et européenne*. Paris : Rapport parlementaire.

BIBLIOGRAPHIE

Vuorikari, R., Kluzer, S., & Punie, Y. (2022). *DigComp 2.2: The Digital Competence Framework for Citizens - With new examples of knowledge, skills and attitudes*. EUR 31006 EN. Luxembourg: Publications Office of the European Union. <https://doi.org/10.2760/115376>

Wei, M., & MD JD. (2025, 4 septembre). Amplifications of delusions by AI chatbots may be worsening breaks with reality. *Psychology Today*. <https://www.psychologytoday.com/us/blog/urban-survival/202507/the-emerging-problem-of-ai-psychosis>

Wilkins, J. (2025, 3 juillet). The Percentage of Tasks AI Agents Are Currently Failing At May Spell Trouble for the Industry. *Futurism*. <https://futurism.com/ai-agents-failing-industry>

World Bank, & International Telecommunication Union. (2024). *Measuring the Emissions & Energy Footprint of the ICT Sector: Implications for Climate Action*. Washington, D.C./Genève : World Bank & ITU. <https://documents1.worldbank.org/curated/en/099121223165540890/pdf/P17859712a98880541a4b-71d57876048abb.pdf>

Zitron, E. (2025, 14 avril). *OpenAI is a systemic risk to the tech industry*. Ed Zitron's Where's Your Ed At. <https://www.wheresyoured.at/openai-is-a-systemic-risk-to-the-tech-industry-2>

À propos



Renaissance Numérique est le *think tank* indépendant dédié à la transformation numérique de la société. Son objectif est de construire une société numérisée plus juste, inclusive, démocratique et soutenable.

Renaissance Numérique offre un lieu d'échanges, d'analyses et de rencontres exclusif pour débattre et éclairer les choix pour une société numérisée responsable. Nous réunissons toute personne, équipe et organisation qui souhaite s'engager ou s'investir dans la responsabilité numérique grâce à son expertise, ses engagements ou encore ses capacités de décider, d'agir et de transformer son organisation, son activité, voire la société dans une plus large mesure.

Renaissance Numérique permet un partage de points de vue, d'expertises et d'expériences, d'éclairages et d'initiatives, qui vise à :

- **Éclairer** la transformation numérique grâce à un débat public, libre, ouvert et le plus inclusif possible.
- **Nourrir** les décisions publiques et les stratégies d'acteurs.
- **Défendre** les intérêts de la société civile.
- **Œuvrer** à ce que tous les individus bénéficient des changements que cette transformation induit, en leur donnant les clés de sa maîtrise.

Créé en 2005, Renaissance Numérique est une association à but non lucratif de loi 1901, qui n'est affiliée à aucun parti, aucune entreprise, aucune structure.



Renaissance Numérique

www.renaissancenumerique.org

Octobre 2025 CC BY-SA 4.0