

ÉTUDE DE DOSSIER

Concours Assistant 2026

Le déploiement de l'intelligence artificielle par les entreprises est source d'opportunités mais aussi de risques. La Banque de France est activement engagée pour anticiper les défis qui seront à relever en la matière.

À partir des documents qui vous sont proposés, vous répondrez aux questions suivantes :

1. Comment l'intelligence artificielle est-elle appréhendée par les entreprises françaises aujourd'hui ?
2. Quels sont les risques potentiels associés à un déploiement plus généralisé de l'intelligence artificielle ?
3. Quelles actions la Banque de France met-elle en œuvre en matière d'intelligence artificielle ?

Les questions sont indépendantes. Nous vous recommandons toutefois de les traiter dans l'ordre. Il n'est pas nécessaire de recopier l'intitulé des questions. Les réponses doivent être rédigées en langue française. Une attention particulière sera portée à la qualité de la rédaction.

LISTE DES DOCUMENTS JOINTS

1. **Comment les entreprises françaises travaillent déjà avec l'IA**
www.strategies.fr – 08/04/2024 – 2 pages
2. **L'intelligence artificielle, une « révolution » industrielle en suspens**
www.lemonde.fr – 12/09/2023 – 2 pages
3. **Intelligence artificielle : évolution et résultats. Confiance numérique, données et cloud**
www.banque-france.fr – 02/07/2024 – 3 pages
4. **Intelligence artificielle : des aides et des formations pour l'intégrer dans votre entreprise**
www.economie.gouv.fr – 24/01/2025 – 3 pages
5. **Les entreprises françaises et l'IA : « L'adoption à grande échelle reste faible »**
www.entreprendre.fr – 27/07/2023 – 2 pages
6. **Les enjeux de l'intelligence artificielle pour le système financier**
www.banque-france.fr – Rapport sur la stabilité financière – 06/2024 – 10 pages

Comment les entreprises françaises travaillent déjà avec l'IA

Productivité, détection des fraudes, conseils personnalisés... L'intelligence artificielle s'est rapidement imposée dans le quotidien des plus grandes entreprises en France, d'Engie à BNP Paribas.

Depuis le lancement de ChatGPT par la société américaine OpenAI, les innovations se multiplient. France Travail, TotalEnergies, BNP Paribas, Axa, LVMH, L'Oréal ou encore Nestlé : mi-mars, des dizaines de responsables, invités par Microsoft, se sont succédé à la tribune de « *l'IA Tour* » pour partager leur utilisation de l'IA au quotidien. Cet événement organisé à Paris au Carrousel du Louvre affichait complet, signe de l'engouement pour cette technologie qui pourrait bien révolutionner le quotidien de millions d'employés.

Lundi 8 avril, le *Raise Summit* réunit à son tour à Paris plusieurs décideurs, dont Arthur Mensch, co-fondateur de MistralAI, l'entreprise européenne considérée comme la plus avancée dans l'IA générative : un programme capable de formuler des phrases cohérentes et souvent justes à partir de simples questions ou demandes formulées par un utilisateur.

Intégré par Microsoft dans ses logiciels de bureautique via un service payant baptisé *Copilot*, ChatGPT résume des mails ou pages web, facilite les recherches et la rédaction de présentations, permet de préparer des réunions prévues en fonction des participants. Le groupe mise sur ce *Copilot* pour démocratiser l'IA dans les entreprises.

« *Déployer des outils de productivité* » pour les employés, par exemple une « *interface pour rechercher de l'information plus rapidement dans toutes les bases de données de l'entreprise* », est souvent une des premières mises en pratique de l'IA pour les employés, explique à l'AFP Nicolas de Bellefonds, directeur monde de BCG X, l'activité IA du cabinet de conseil BCG. Viennent ensuite les applications ciblées, comme les centres d'appel, le développement informatique ou le marketing.

Développée depuis plus longtemps que l'IA générative, l'IA par apprentissage (machine learning) permet notamment d'analyser des images. Elle contribue ainsi par exemple à la traque de fuites de méthane -un gaz à effet de serre puissant issu des énergies fossiles- ou au diagnostic de certains cancers. Chez Engie, l'IA par apprentissage permet de prédire le rendement d'éoliennes. Dans des centaines de rapports d'accidents du travail, « *nous utilisons l'IA pour détecter des tendances que nous n'avons pas reconnues avant* », a détaillé lors de l'IA Tour de Microsoft Biljana Kaitovic, directrice générale adjointe chargée des systèmes d'information et du digital.

Validation humaine

TotalEnergies a annoncé fin février l'achat de licences Copilot pour ses collaborateurs. Et chez le géant des moteurs Cummins, une IA dessine les plans de plusieurs versions d'un composant en fonction des caractéristiques recherchées (poids, solidité...). Les assureurs ont également recours à l'IA pour déceler la fraude. Et des banques comme BNP Paribas s'appuient sur une analyse automatisée de dossiers clients dans le cadre des obligations réglementaires comme la

lutte contre le blanchiment. L'IA détecte mieux qu'un humain des montages financiers complexes et illégaux, a détaillé à l'AFP Hugues Even, responsable data du groupe.

Quelque 20 000 employés du géant du luxe LVMH utilisent aujourd'hui une IA interne basée sur ChatGPT, baptisée « MaIA », qui permet de synthétiser des documents ou encore de générer des mails. Mais la validation par des humains reste nécessaire. Il faudra un œil expert « *probablement pour très longtemps, pour contrôler ce qui est produit, retoucher et le valider* », note Nicolas de Bellefonds. Ce qui entraîne « *une vraie problématique RH* » : comment faire émerger ces experts si l'IA se charge des tâches ayant, auparavant, permis de former les nouveaux arrivés ?

Le problème des hallucinations de l'IA

Autres problèmes à régler, l'IA peut « *halluciner* », c'est-à-dire inventer des références, ou ses réponses peuvent carrément être hors sujet. Car l'IA ne réfléchit pas comme un humain ; elle ajoute en général derrière chaque mot le plus probable suivant, en fonction d'un corpus de textes que ses créateurs lui ont fait lire. Et l'énorme besoin en énergie pour entraîner et faire fonctionner une IA générative est souvent critiquée.

Mais ces technologies sont « *déjà matures et fonctionnent sur certains sujets bien définis* », ajoute Nicolas de Bellefonds. « *Sur un certain nombre de tâches, on a des gains de vitesse très significatifs* », qui peuvent atteindre 70% dans les relations client, et aussi « *des gains de qualité* », selon BCG. « *Ce qui est sûr, c'est que ça va profondément changer nos manières de travailler* », note-t-il, lui-même utilisateur au quotidien d'une IA générative.

L'intelligence artificielle, une « révolution » industrielle en suspens

Pierre-Yves Gomez, professeur émérite à l'EM Lyon Business School, explore dans sa chronique les questions économiques, politiques et anthropologiques posées par le développement exponentiel de l'IA. Il anticipe notamment les conséquences de sa généralisation.

L'intelligence artificielle (IA) fait de nouveau le buzz. Avec l'irruption de ChatGPT ou la grève des artistes d'Hollywood, inquiets à l'idée d'être remplacés, elle aura alimenté toute l'année l'inépuisable débat entre les technophiles et les technophobes, ceux qui attendent de l'IA une révolution sans précédent et ceux qui ironisent sur le énième emballement médiatico-financier à son sujet, ceux qui espèrent qu'elle produira un dépassement de l'humain et ceux qui s'en effraient.

De fait, une telle approche anthropologique brouille le débat, précisément parce qu'elle considère l'IA de manière anthropomorphique, prêtant aux machines des qualités, des comportements et des dispositions humaines, tout en réduisant, symétriquement, l'être humain à la qualité prêtée à la machine.

Bien évidemment, celui-ci paraît alors toujours surpassé par la machine, puisqu'il l'a justement créée pour dépasser ses propres incapacités. C'est ce que Günther Anders (1902-1992) avait appelé la honte prométhéenne, « *qui s'empare de l'homme devant l'humiliante qualité des choses qu'il a lui-même fabriquées* » (*L'Obsolescence de l'homme*, 1956). Force est de constater que l'homme occidental se délecte dans cette honte.

Un oligopole mondial d'entreprises

Or, l'IA pose plus essentiellement d'importantes questions économiques et politiques. Économiques parce qu'elle préfigure une nouvelle phase de l'industrialisation du monde, qui concerne désormais le traitement à haute fréquence de milliards d'informations numérisées. Elle permet d'industrialiser des réponses crédibles à de multiples questions, de créer des artefacts narratifs ou visuels, de déceler des probabilités de liens entre des occurrences ou de proposer des décisions statistiquement logiques face à des choix.

L'usage d'un tel outil industriel modifie l'appareil productif dans des proportions qui restent à mesurer, sans céder à l'exaltation du nouveau, mais en examinant le rapport entre le coût énorme de l'investissement et l'utilité finale des services produits. L'IA pose aussi et principalement trois questions politiques.

D'une part, une telle industrialisation du traitement d'informations contribue à renforcer les idées reçues, les représentations communes et à enfermer les usagers dans des systèmes de pensée autoréférés qui accentueront la fragmentation de la société.

D'autre part, les investissements nécessaires à la maîtrise de l'IA étant considérables, on verra inévitablement se constituer un oligopole mondial d'entreprises contrôlant cette industrie très sensible.

Nos lieux communs

Enfin, l'usage de l'IA est exponentiel, puisque le traitement massif d'informations s'autoalimente, avec un coût énergétique et écologique qui pourrait vite devenir insupportable.

S'il demeure une dimension anthropologique et culturelle, elle porterait plutôt sur les possibilités d'autoévaluation de notre propre intelligence que suggère l'IA. Si on prend l'exemple de ChatGPT, il apparaît que les réponses qu'apporte l'algorithme à nos requêtes nous semblent d'autant plus « intelligentes » qu'elles nous paraissent familières : rien d'étrange à cela car, par construction, l'algorithme cherche dans les milliards d'informations disponibles les enchaînements logiques les plus courants, qui constituent, au sens propre, nos lieux communs.

Sitôt dépassée la stupéfaction de constater que l'IA « pense comme nous », l'habitude nous conduira sans doute à considérer qu'il n'est pas glorieux de penser comme elle. Raisonner comme une intelligence artificielle pourrait ainsi, dans l'avenir, qualifier de manière péjorative la banalité d'une réflexion.

Pierre-Yves Gomez

Professeur émérite à l'EM Lyon Business School

Intelligence artificielle : évolution et résultats. Confiance numérique, données et cloud

Point Zero Forum - Zurich

2 juillet 2024

Remarques introductives

Denis Beau, Premier sous-gouverneur

Mesdames, Messieurs,

C'est avec grand plaisir que j'ouvre cette table-ronde consacrée à l'intelligence artificielle (IA). L'IA constitue en effet une innovation de rupture, susceptible de conduire à de véritables bouleversements économiques. C'est particulièrement vrai dans le secteur financier, dont l'IA constitue, depuis quelques années déjà, le principal moteur de transformation. L'arrivée de l'IA générative devrait d'ailleurs encore accélérer cette tendance, non seulement en renforçant l'adoption des outils d'IA par les utilisateurs, mais également en accélérant structurellement le rythme de l'innovation (pensons par exemple à la capacité nouvelle de générer du code informatique à partir de requêtes en langage naturel).

Ces évolutions, marquantes, ne vont toutefois pas sans poser de questions, et notamment pour le banquier central et le superviseur financier que je suis. Je souhaite partager avec vous quelques-unes de ces interrogations, avant de vous faire part d'un certain nombre de convictions quant à la manière dont nous devons nous emparer du sujet.

*

1/ D'abord, malgré les progrès récents, la technologie sous-jacente ne paraît pas encore entièrement mûre, notamment s'agissant de l'IA générative. Sur ce sujet, un certain nombre de questions restent aujourd'hui sans réponse. J'en aborderai deux, qui pourront d'ailleurs ultérieurement stimuler les discussions du prestigieux panel réuni ici.

Premièrement, la question des modèles à usage général : quelle sera leur performance dans toute une série de tâches intéressant le secteur financier ? Cette question se pose en réalité à deux niveaux. Est-ce que les modèles généralistes vont devenir la norme pour tous les usages, au détriment des modèles spécialisés ? Les petits modèles bien entraînés -c'est-à-dire plus spécialisés- auront-ils la capacité à tirer leur épingle du jeu face aux gros modèles plus généralistes ? Ces questions de performance ont de nombreuses conséquences potentielles, notamment en matière de concurrence : si les gros modèles généralistes s'imposent dans tous les domaines, nous risquons fort de nous retrouver dans une situation de monopole ou d'oligopole naturel, qui viendrait d'ajouter au caractère déjà largement oligopolistique du marché du cloud.

Deuxièmement, la question des vulnérabilités des systèmes d'IA : si l'on commence à avoir une vision plus précise de la situation, les recherches sur le sujet sont loin d'être terminées. C'est notamment le cas dans le domaine de la cyber-sécurité des modèles d'IA générative, avec la découverte récente des dangers liés à l'« *indirect prompt injection* ». Si cette course de vitesse entre le glaive (le développement de nouvelles techniques d'attaque) et le bouclier

(l'élaboration de parades efficaces) est classique dans le domaine de la sécurité, notre capacité à sécuriser convenablement les systèmes d'IA aura une influence majeure sur la faculté des différents acteurs à utiliser largement cette technologie.

*

2/ Quand bien même les technologies d'IA ne sont pas encore entièrement mûres, il me semble que les banques centrales et les superviseurs financiers doivent s'en emparer sans attendre, et ce pour au moins trois raisons.

D'abord, pour continuer à assurer nos missions efficacement, en faisant plus et mieux. L'IA peut évidemment nous permettre de gagner en efficacité, en automatisant davantage. Mais nous souhaitons aussi offrir de nouvelles capacités aux agents. C'est ainsi que notre outil LUCIA, un système reposant sur l'IA qui offre la capacité d'analyser de larges volumes d'opérations bancaires, nous permet, au cours de nos contrôles sur place, d'évaluer la performance et la pertinence des modèles de LCB-FT développés par les banques.

Ensuite, pour développer une expertise critique sur l'IA. Utiliser l'IA pour nos propres besoins nous permet d'acquérir progressivement une bonne maîtrise de la technologie, et constitue une manière très efficace d'en appréhender correctement les bénéfices et les risques. Les vertus de l'apprentissage par la pratique expliquent d'ailleurs que les utilisations internes de l'IA soient très complémentaires des missions de surveillance des systèmes d'IA déployés par le secteur financier. C'est ainsi que, très récemment, l'ACPR a organisé, avec l'aide du centre d'innovation de la Banque de France, Le Lab, un « Tech Sprint Suptech », un hackathon destiné à explorer ce que l'IA générative peut apporter aux différents métiers du superviseur. En trois jours, cet événement a permis de révéler le potentiel des grands modèles de langage pour la supervision.

Enfin, pour entraîner l'écosystème financier, en envoyant au marché le signal qu'il peut -ou doit- à son tour se lancer. Par exemple, les travaux de pointe menés à la Banque de France sur la cryptographie post-quantique nous permettent de sensibiliser les acteurs privés à la nécessité de se préoccuper de cette menace.

S'il est donc clair à mes yeux que les banques centrales et les superviseurs doivent se saisir des opportunités offertes par l'IA, la question est : comment procéder ?

*

3/ Il me semble que nous devons d'abord poser un principe fondamental de gouvernance : l'IA doit être au service de l'humanité et de la société, et pas l'inverse. De ce point de vue, même s'il ne résout pas tous les problèmes, l'adoption en cours du règlement européen sur l'IA, premier texte contraignant dans le monde fixant les principes de « l'IA de confiance », constitue une avancée à saluer. Ce texte va notamment permettre de renforcer la confiance des consommateurs, tout en donnant de la sécurité juridique aux opérateurs économiques.

Ce principe de gouvernance peut être complété par trois principes opérationnels.

Premièrement, utiliser l'IA de manière proportionnelle et progressive. Avec une règle simple : plus le cas d'usage est critique pour nos métiers, plus nous devons nous atteler à faire nous-mêmes. Pour des institutions comme les nôtres, cela renvoie notamment à la question fondamentale des données : une partie des données des banques centrales et des superviseurs

financiers sont en effet trop confidentielles pour être stockées sur l'infrastructure cloud d'un tiers prestataire.

Deuxièmement, expérimenter sans attendre, même via des cas d'usage simple, pour trouver la bonne manière d'intégrer l'IA dans nos métiers, afin d'aboutir à un « agent augmenté » plutôt qu'à un « agent substitué ». Il faut en effet s'attendre à ce que l'IA redessine largement le schéma des interactions homme-machine. Trouver les bonnes combinaisons permet de favoriser l'adoption des nouveaux outils, en emportant l'adhésion des utilisateurs, ce qui constitue un enjeu crucial.

Troisièmement, coopérer avec d'autres, pour partager les bonnes pratiques opérationnelles, mais aussi pour construire un cadre cohérent de supervision de l'IA. Je pense bien sûr d'abord à la coopération au niveau international, car les enjeux liés à l'IA sont par nature mondiaux. En la matière, si des nuances peuvent exister quant à la manière de procéder, je relève surtout que de nombreuses juridictions expriment des préoccupations proches, ce qui devrait permettre à la coopération internationale d'avancer. Mais il nous faut aussi coopérer avec les autorités d'autres secteurs, en particulier celles de la concurrence, de la cyber-sécurité, des droits fondamentaux, voire de la transition écologique, car les préoccupations liées à l'IA sont largement interconnectées. Ces différentes coopérations constituent à mes yeux une condition essentielle pour faire émerger les modèles d'IA les plus pertinents et les plus résilients, c'est-à-dire pour influencer le développement de la technologie dans le sens de l'intérêt général.

Je vous remercie pour votre attention.

Intelligence artificielle : des aides et des formations pour l'intégrer dans votre entreprise

L'intelligence artificielle se développe de plus en plus au sein des entreprises, tous les secteurs d'activité confondus. Comment fonctionne-t-elle ? Quel levier peut-elle représenter pour la croissance de votre entreprise ? Comment vous y former ? Quels soutiens publics sont disponibles pour vous accompagner ? On vous éclaire.

Intelligence artificielle : quels sont les avantages pour votre entreprise ?

Qu'est-ce que l'intelligence artificielle ?

Le Parlement européen définit l'intelligence artificielle (IA) comme tout outil utilisé par une machine capable de « reproduire des comportements liés aux humains, tels que le raisonnement, la planification et la créativité ».

Grâce à elle, les entreprises peuvent :

- simplifier des tâches complexes et répétitives pour trouver des gains d'efficacité,
- améliorer l'efficacité des procédés et diminuer leurs coûts,
- proposer des outils pour de nouveaux services,
- analyser et exploiter les données issues du *Big Data* (mégadonnées),
- optimiser les campagnes marketing et l'affichage ciblé,
- améliorer le service client : *chatbots* (dialogueurs), assistances virtuelles, etc.

Aujourd'hui, il existe une multitude de solutions IA sur le marché, gratuites ou payantes, que ce soit pour générer du texte, des images, traiter automatiquement des documents, etc.

De nombreuses formations pour mieux appréhender l'IA

Pour vous aider à comprendre comment fonctionne l'IA, ce qu'elle peut vous apporter et comment la mettre en œuvre, il est impératif de vous former, d'autant que les technologies et solutions proposées évoluent très rapidement. France Num recense dans [sa page dédiée](#) des nombreuses formations en ligne, gratuites et accessibles à tous, pour tout comprendre à l'IA.

À savoir

[France Num](#) est une initiative gouvernementale pour la transformation numérique des TPE/PME. Elle vise à apporter une solution concrète à chaque TPE/PME en mettant à disposition toutes les ressources existantes : accompagnement sur tout le territoire, outils, etc. Ce site propose par ailleurs un moteur de recherche de financement dédié aux projets numériques.

Des formations théoriques

Trois MOOCs* gratuits sont proposés, apportant une introduction théorique à l'intelligence artificielle. Elles sont ouvertes à tous y compris des publics qui démarrent sur le sujet :

- [Objectif IA](#), réalisée par OpenClassrooms en partenariat avec l'Institut Montaigne et la Fondation Abeona, en 6 heures.
- [Class'Code IAI](#), coproduit par l'association Class'Code et l'Inria, en 10 heures.
- Enfin, [un MOOC de 30 heures](#) permettant d'approfondir les notions abordées dans les deux autres formations et proposé par l'Université d'Helsinki.

**Le MOOC (Massive Open Online Course, en anglais), est un format pédagogique consistant à mettre en libre accès des formations en ligne pour les apprenants.*

Des formations pratiques

Trois formations en ligne et proposées gratuitement par Bpifrance peuvent également vous aider à évaluer la pertinence et la mise en œuvre de l'IA pour votre entreprise.

Les deux premières sont deux formations de 15 minutes : « [Dans quels cas l'IA est-elle une bonne solution ?](#) » et « [Votre première solution intégrant de l'IA](#) ».

Une [troisième propose un cursus de 10 heures](#) dans le cadre du programme IA Booster.

Des formations ciblées sur l'IA générative

- [ChatGPT expliqué en 5 minutes](#) proposée par l'INRIA
- [Les IA génératives](#) : formation proposée par Bpifrance
- [L'intelligence artificielle générative et moi](#) proposé par le Conservatoire national des arts et métiers (Cnam)

Les formations France Num

France Num propose également des formations pour aider les TPE et PME à débiter avec les outils numériques ou à mieux les utiliser pour maintenir ou développer leur activité. Centrées sur des problématiques concrètes, ces formations, en ligne ou en présentiel, permettent aux petites entreprises d'expérimenter des solutions numériques en réponse à un besoin réel.

Pour accompagner votre entreprise : le programme IA Booster France 2030

Proposé par [Bpifrance](#), le [programme IA Booster France 2030](#) s'adresse aux entreprises françaises (PME et ETI) de tout secteur d'activité, avec une priorité donnée aux entreprises ayant un effectif compris entre 10 et 2 000 collaborateurs et réalisant plus de 250 000 euros de chiffre d'affaires HT.

Le Programme IA Booster France 2030 se décline en quatre prestations :

- Sensibilisation et acculturation aux solutions d'IA
- Diagnostic Data IA

- Mission choix de l'approche IA
- Mission accompagnement à la mise en œuvre de la solution IA

Des aides financières régionales pour développer l'utilisation de l'IA

Plusieurs régions proposent des financements et dispositifs d'accompagnements pour aider les entreprises à se saisir des potentialités de l'intelligence artificielle :

- Pour la région Île-de-France, [le Pack IA](#)
- Pour la Région Sud Provence-Alpes-Côte d'Azur, le [dispositif "Mon assistant IA"](#)
- Pour la région Centre-Val de Loire : le [dispositif CAP'TN - Transformation Numérique](#)
- Pour la région Hauts-de-France, le [Pass Cyber Investissement](#)
- Pour la région Grand Est, le [Diagnostic Intelligence Artificielle](#) ainsi que des aides aux entreprises primo utilisatrices d'intelligence artificielle avec un [volet recrutement](#) et un [volet prestation](#)

Des ressources pour répondre à vos interrogations

L'IA et ses usages soulèvent de nombreuses questions pour les entreprises : sécurité, délais, coûts, retour sur investissement, compétences requises, etc.

Un guide et quatre fiches pédagogiques sur l'IA générative ont été récemment publiés afin d'aider les TPE et PME à y répondre et sont [disponibles en ligne gratuitement sur le site de France Num](#).

Elles abordent les sujets suivants :

- Une intelligence artificielle générative, à quoi ça sert ? Comment ça marche ?
- Cas d'usage de l'IA générative dans les TPE et PME
- Choisir parmi les solutions d'IA générative et comment les utiliser ?
- Prendre les précautions indispensables à un usage responsable et sécurisé.

Vous pouvez également [consulter le guide consacré aux solutions de recherche augmentée grâce à l'IA générative](#) sur le site de la [Direction générale des entreprises](#).

D'autres ressources peuvent également vous intéresser, avec notamment :

- un article sur le site de France Num : [Exploiter l'intelligence artificielle pour améliorer le fonctionnement de sa TPE PME : mode d'emploi](#)
- un webinaire de 45 minutes (en ligne, gratuit) sur le sujet [Intelligence artificielle : opportunité ou menace pour les TPE-PME ?](#)
- un article sur le site de Bpifrance : [Comment intégrer l'IA dans votre entreprise ? 8 étapes clés pour réussir](#)

Les entreprises françaises et l'IA : « L'adoption à grande échelle reste faible »

L'adoption de l'intelligence artificielle est encore timide au sein des entreprises françaises. Comment expliquer une telle réticence ? Éléments de réponse avec Charlotte Douette, data scientist chez SAS France.

Solutions d'intelligence artificielle inefficaces, incohérence des informations, expérience client déconnectée, manque de confiance dans la prise de décision, insuffisance de ressources... Comment expliquer les lacunes de l'IA ?

L'IA fascine la société depuis l'invention du premier ordinateur au milieu du XX^e siècle. C'est une réalité qui s'imisce dans nos vies de tous les jours. L'intelligence artificielle a beaucoup d'avantages. Elle permet par exemple, d'analyser les données, de détecter des tendances et d'effectuer des prédictions. Malgré ces bénéfices, l'IA présente aussi des lacunes.

Pourquoi ? L'intelligence artificielle est un outil qui se base sur une multitude de données. Parfois, les algorithmes d'IA peuvent être confrontés à des données biaisées, fausses voire insuffisantes. Contrairement à l'Homme, la machine est incapable de différencier la bonne de la mauvaise information ou de s'assurer de la licéité de cette dernière. Si les données sont biaisées, l'IA reproduira ainsi ces biais, ce qui pourrait engendrer des conséquences néfastes. En somme, l'IA est une technologie en constante évolution qui ne peut remplacer l'intelligence humaine, mais qui reste tout de même un outil d'aide à la prise de décision.

Qu'apportent les récentes innovations en matière d'IA à l'expérience client ?

Depuis la crise sanitaire, les besoins des consommateurs changent drastiquement et leurs demandes sont devenues de plus en plus exigeantes. Dans un contexte de forte concurrence, les entreprises doivent relever un défi : répondre d'une manière rapide et efficace aux besoins des consommateurs. Personnaliser les échanges clients est un différenciateur essentiel, cela implique une utilisation optimisée des données clients. Les organisations doivent se doter d'outils technologiques qui les aident à définir une stratégie répondant d'une manière personnalisée aux besoins des consommateurs.

Quel est le degré de maturité des entreprises françaises dans l'adoption de l'IA ?

Il est intéressant de mentionner ici trois études faites par des grands cabinets de conseil. Elles donnent un aperçu de la maturité des entreprises dans l'adoption de l'IA. La première est une étude Accenture publiée en 2022. Elle révèle qu'à l'échelle mondiale, seules 12% des entreprises capitalisent sur l'IA pour booster leur croissance et leurs résultats. La deuxième est le dernier rapport annuel du cabinet McKinsey, "The State of AI", publié le 6 décembre 2022. Il affirme que l'adoption de l'IA a plus que doublé depuis 2017 (de 20 à 50 %), date de la première étude de McKinsey sur le sujet. Enfin, la troisième étude KPMG fait en partenariat avec l'Université de Queensland publiée le 22 février 2023, révèle qu'en France, un peu moins d'un tiers des personnes interrogées (31%) a confiance dans l'IA.

L'enseignement tiré de ces études est que les entreprises, y compris françaises, restent méfiantes quant à l'adoption des nouvelles technologies. Ce retard s'explique par deux facteurs : le coût élevé des outils et le manque de personnel qualifié. Ces rapports confirment la récente étude mondiale faite par SAS sur la résilience des entreprises. Cette étude montre que plus d'un dirigeant français sur deux (54%) estime que son entreprise n'est pas assez résiliente pour faire face aux bouleversements économiques et géopolitiques. En revanche, 89% des cadres qui évoluent dans les entreprises les plus résilientes, priorisent l'intelligence artificielle et l'analytique pour relever les défis dans les périodes de crise.

En résumé, bien que de nombreuses entreprises françaises se soient intéressées à l'IA, l'adoption à grande échelle reste tout de même relativement faible et nécessite une meilleure compréhension des avantages et des défis liés à cette technologie et un accompagnement vers cette transformation.

Anaïs Lamarge

3. Les enjeux de l'intelligence artificielle pour le système financier

Par Claire Brousse, Olivier Fliche, Jules Joyez et Julien Uri

Les technologies basées sur l'intelligence artificielle offrent de nombreuses opportunités pour le secteur financier, avec des applications notamment pour le service clients, l'optimisation des processus, la gestion des risques ou les activités de marché. Néanmoins, par sa nature, l'intelligence artificielle présente des enjeux de gestion des données, de modélisation ou de gouvernance. Elle pourrait accentuer des risques pour la stabilité financière, en particulier les risques de procyclicité et de volatilité des marchés, le risque d'une concentration des marchés et le risque de cyberattaques. En outre, des risques de mauvais usages, tels que la manipulation de l'information, pourraient déstabiliser le système financier. Les impacts potentiels de l'intelligence artificielle restent difficiles à estimer, car ils dépendront de l'ampleur et des modalités de son adoption. Néanmoins, une gouvernance et un cadre de gestion des risques adaptés aux enjeux de l'intelligence artificielle sont nécessaires pour en maîtriser les risques. La législation européenne sur l'intelligence artificielle, adoptée en mai 2024, est conçue spécifiquement pour favoriser une intelligence artificielle digne de confiance.

3.1. Le développement de l'intelligence artificielle présente des opportunités pour le système financier et l'économie

(...)

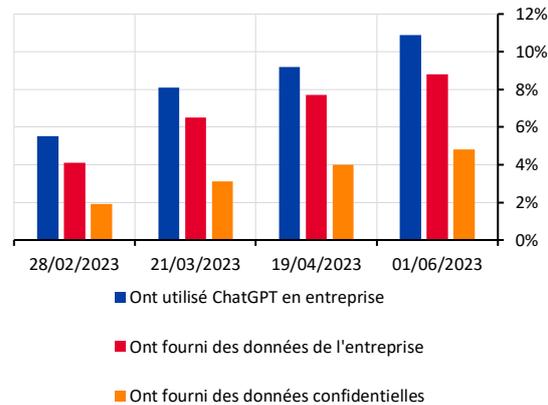
3.2 Des risques inhérents à la nature de l'intelligence artificielle

Si les potentielles opportunités de l'intelligence artificielle sont nombreuses, son déploiement comporte des risques inhérents à sa nature dont certains pourraient être renforcés par l'intelligence artificielle générative. En premier lieu, l'intelligence artificielle soulève des enjeux liés à la qualité et à la confidentialité des données utilisées pour son apprentissage. Le fonctionnement autonome et opaque des modèles d'intelligence implique également le risque d'une compréhension difficile des résultats fournis et de leur fiabilité. Enfin, une concentration excessive des acteurs du secteur exacerberait ces risques pour les utilisateurs et serait préjudiciable au système dans son ensemble.

L'intelligence artificielle, par le fonctionnement de ses algorithmes et son apprentissage, engendre des risques liés à la gestion des données notamment leur confidentialité. Ce risque est particulièrement marqué dans les systèmes d'intelligence artificielle générative qui apprennent, non seulement, à partir des données disponibles sur Internet, mais aussi à partir des données fournies par les utilisateurs (en vue d'améliorer leurs réponses) : ces outils peuvent ainsi transmettre les informations apprises d'un utilisateur en générant des contenus pour d'autres utilisateurs. D'après un rapport de l'entreprise spécialisée Cyberhaven publié en juin 2023, 10,8 % des employés ont utilisé ChatGPT au travail et 4,7 % d'entre eux ont déjà fourni des informations confidentielles de leur entreprise à l'agent conversationnel, le pourcentage ayant même tendance à croître au cours du temps. En juin 2023, les données confidentielles représentaient 11 % du total des données copiées et collées directement dans l'outil. Ces comportements représentent un danger pour les entreprises, dont les informations sensibles peuvent être diffusées largement. En outre, les données transmises à des outils conversationnels peuvent aussi être volées lors de cyberattaques. Pour pallier ce risque et préserver la confidentialité des données financières, plusieurs groupes bancaires dont JPMorgan, Bank of America, Goldman Sachs et Deutsche Bank ont interdit l'utilisation du service en ligne ChatGPT à leurs employés. Néanmoins, ces risques pourraient être moindres dans le cas d'outils réservés à l'usage interne dans les entreprises, avec des règles strictes de gestion des données.

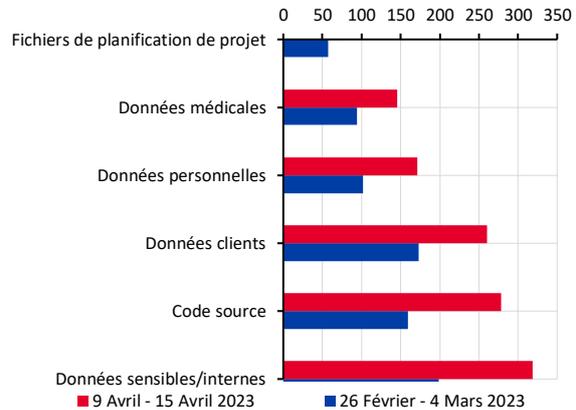
Graphique 3.5 : Adoption de ChatGPT par les salariés

x : axe temporel / y : pourcentage cumulatif



Graphique 3.6 : Nombre d'incidents de transmission de données sensibles à ChatGPT

x : nombre d'incidents par 100 000 employés / y : type d'incident



Source : Cyberhaven

Plus généralement, la qualité des résultats obtenus par des technologies d'intelligence artificielle dépend directement de la qualité des données. Le recours à des données inadéquates lors de l'apprentissage de l'outil, ou lors de son utilisation, peut entraîner des résultats erronés. Certaines technologies sont limitées par leur historique de données ou bien par les sources de données disponibles. À titre d'exemple, GPT 3.5 avait été entraîné sur l'historique des données d'Internet jusqu'à janvier 2022, un utilisateur peut donc obtenir des informations erronées ou obsolètes en interrogeant l'outil (pas de prise en compte de la guerre en Ukraine, par exemple). À ce stade, l'information générée par les outils d'intelligence artificielle générative provient, pour l'essentiel, de contenus rédigés par des humains. Or, une forte adoption de ces outils pourrait mener à un phénomène de boucle : le contenu qu'ils génèrent pourrait devenir l'une des principales sources d'information des modèles d'intelligence artificielle générative. Ce phénomène de récursivité pose un risque de perte de diversité de l'information originale au profit d'une standardisation des réponses fournies potentiellement biaisées, erronées ou obsolètes.

L'intelligence artificielle peut perpétuer les biais historiques provenant des données ou en créer de nouveaux en raison de son fonctionnement. Ceci peut créer des points aveugles, voire entraîner un manque d'équité. Un modèle entraîné sur des données biaisées perpétuera – voire renforcera – ces biais historiques et potentiellement des discriminations. Et ce, d'autant plus qu'un modèle biaisé génère des résultats biaisés qui seront ensuite utilisés dans son processus d'apprentissage et renforceront les biais (par récursivité). À titre d'exemple, l'algorithme utilisé par Apple et Goldman Sachs, pour les décisions d'allocation de lignes de crédit, aurait offert des prêts de montants inférieurs aux femmes qu'aux hommes, toutes choses égales par ailleurs. En outre, les décisions biaisées prises par les modèles d'apprentissage automatique peuvent se produire même avec des données de bonne qualité, en raison d'approximations, de défauts dans la construction de l'algorithme, ou du fait que les corrélations entre variables peuvent être spécifiques aux bases de données utilisées (reflétant, par exemple, les caractéristiques d'un échantillon historique).

En outre, l'intelligence artificielle comporte des risques liés aux modèles qui peuvent remettre en cause la fiabilité des résultats. Par exemple, les « hallucinations » sont des réponses fausses présentées comme des faits par des outils utilisant l'intelligence artificielle générative avec des grands modèles de langage. Ces problèmes d'hallucinations de modèles d'intelligence artificielle générative sont inhérents à la nature des grands modèles de langage, qui estiment une réponse probable à partir des mots fournis par l'utilisateur et peuvent donner des réponses anormales, voire fausses. À titre d'exemple, une entreprise a demandé à ChatGPT de produire un article à propos des résultats de l'entreprise Tesma, ChatGPT a créé un article cohérent mais avec des chiffres financiers inventés et complètement faux. Ce phénomène peut aussi se produire lorsque le modèle est entraîné sur des données fausses. Par ailleurs, une même question posée à deux moments différents peut générer des réponses différentes, mettant en évidence un problème de cohérence temporelle. Un article de la Banque des règlements internationaux (BRI), publié en janvier 2024, illustre également les limites des modèles d'intelligence artificielle générative, qui inventent des réponses fausses lorsqu'ils font face à des problèmes nouveaux.

De fait, les technologies utilisant l'intelligence artificielle produisent des résultats complexes et leur fonctionnement opaque implique des difficultés de compréhension des résultats. Les résultats des systèmes basés sur l'intelligence artificielle sont intrinsèquement complexes à interpréter pour un esprit humain, en raison de la nature de certaines modélisations utilisées (réseaux de neurones, etc.), qui reposent essentiellement sur des relations non linéaires. Les acteurs du secteur de l'intelligence artificielle peuvent parfois eux-mêmes renforcer l'opacité du fonctionnement des outils, afin de protéger leur propriété intellectuelle. Certains outils d'intelligence artificielle constituent ainsi des boîtes noires, dont les paramètres et le fonctionnement ne sont pas visibles pour les entités les intégrant dans leurs services et dont les sources ne sont pas mentionnées.

Une gouvernance claire est nécessaire pour réduire les risques associés à de potentiels biais, au manque de transparence et à la fiabilité de l'outil. Une gouvernance claire nécessite de définir des lignes de responsabilité pour le développement et la supervision des systèmes basés sur l'intelligence artificielle tout au long de leur cycle de vie, du développement au déploiement, et la désignation explicite de la responsabilité des résultats produits par le modèle.

Le manque de concurrence entre les fournisseurs d'outils basés sur l'intelligence artificielle peut exacerber les risques présentés. Les coûts nécessaires au développement de modèles basés sur l'intelligence artificielle peuvent impliquer un risque de dépendance à des fournisseurs oligopolistiques. À ce stade, OpenAI représente 39 % du marché de l'intelligence artificielle générative contre 30 % pour Microsoft, 8 % pour Amazon et 7 % pour Google. La concentration de ces technologies chez un nombre limité de fournisseurs (s'appuyant sur leurs avantages en termes de données ou de puissance de calcul) pourrait amplifier le risque de récursivité d'utilisation de données fournies par des intelligences artificielles pour l'apprentissage des modèles. Les modèles seraient entraînés par des données provenant d'un nombre limité d'outils basés sur l'intelligence artificielle et perdraient en diversité.

L'intelligence artificielle générative peut renforcer les risques de cyberattaques. Même si la plupart des agents conversationnels sont créés avec des mesures de sécurité internes, comme le refus de répondre à une demande directe de création de cyberattaque, ils ne sont généralement pas programmés pour refuser la création de code informatique ou de texte pouvant indirectement servir à des cyberattaques.

3.3 L'intelligence artificielle présente des risques pour la stabilité financière

Ces risques inhérents à l'intelligence artificielle constituent des facteurs potentiels d'amplification de risques existants pour la stabilité financière. En premier lieu, l'utilisation des modèles similaires d'intelligence artificielle par une large partie des participants de marché pourrait renforcer les risques de procyclicité et de volatilité des marchés. Par ailleurs, le déploiement de l'intelligence artificielle pourrait entraîner une concentration des marchés et favoriser l'émergence de nouveaux acteurs d'importance systémique. Enfin, elle renforce le risque de cyberattaques auquel le système financier est particulièrement exposé.

L'intelligence artificielle pourrait accentuer la procyclicité et la volatilité des marchés. Le développement du recours à l'intelligence artificielle pour des stratégies de négociation ou des décisions d'investissement peut créer un risque d'homogénéité des recommandations, ainsi que des comportements moutonniers pouvant mener à un risque accru de procyclicité pour les marchés. L'utilisation des mêmes outils d'intelligence artificielle par de nombreux acteurs de marché peut engendrer un cumul de transactions unidirectionnelles, et donc amplifier les mouvements de marché en cas de mouvement adverse. Si un algorithme recommande à plusieurs acteurs significatifs de vendre un actif, alors la baisse de prix consécutive peut engendrer des recommandations de vente pour d'autres acteurs, et ainsi alimenter la chute du cours de l'actif avec des impacts sur les opérations de couverture associées, et ce, de manière extrêmement rapide (*flash crash*). Ces épisodes de tensions peuvent entraîner d'importants besoins de liquidité pour les acteurs les plus fragiles, qui peuvent notamment se traduire par des ventes forcées d'actifs, avec pour conséquence potentielle une extension des tensions sur d'autres classes d'actifs. Compte tenu des fortes interconnexions directes ou indirectes des institutions financières par le biais des marchés financiers, de tels mouvements peuvent donc avoir des effets systémiques. De plus, en cas de manque de compréhension des modèles, les utilisateurs pourraient avoir des difficultés à prédire leur réaction en fonction des conditions de marché et à ajuster leur stratégie en période de faible performance. En outre, l'intelligence artificielle pourrait renforcer le risque de manipulation de marché. Par exemple, la diffusion de fausses informations facilitée par les technologies d'intelligence artificielle générative permettant de générer des images et du son à partir de contenu existant (appelés *deepfakes*) pourrait engendrer des mouvements de marché massifs et une forte volatilité des marchés.

Plus généralement, une mauvaise estimation des risques pourrait engendrer d'importantes pertes financières pour les institutions financières. Une mauvaise estimation pourrait se traduire par l'octroi excessif de prêts, ou par des erreurs massives de tarification, conduisant à d'importantes pertes pour les institutions financières. De même que pour les modèles statistiques classiques, l'entraînement des modèles d'intelligence artificielle sur des données par définition non exhaustives peut conduire à une amplification des pertes en période de crise. En effet, les modèles étant généralement entraînés sur des données historiques, celles-ci ne sont pas nécessairement représentatives du comportement d'un actif donné lors d'événements extrêmes. De même, l'utilisation d'un modèle avec des données non exhaustives peut engendrer une mauvaise tarification des prêts par les banques ou des pertes systématiques sur des produits d'assurance. De telles erreurs pourraient potentiellement se traduire par une perte de confiance des clients, avec des mouvements de type panique bancaire, et pourraient se présenter à une échelle systémique en cas d'utilisation de modèles similaires par les institutions financières.

Le déploiement de l'intelligence artificielle pourrait se traduire par une concentration des fournisseurs de services, avec des risques de dépendance, et par une concentration du système financier. À l'instar d'une utilisation massive du même logiciel par les acteurs de marché, l'utilisation généralisée d'un petit nombre d'outils d'intelligence artificielle peut engendrer un risque opérationnel en cas de dysfonctionnement ou d'indisponibilité de la technologie les empêchant de mener leurs activités. De plus, la prépondérance d'un fournisseur pourrait amplifier les interconnexions entre les acteurs du système financier et la diffusion de chocs. La domination croissante d'une poignée de systèmes de gestion des risques basés sur l'intelligence artificielle, comme *Aladdin* de BlackRock et *RiskMetrics* de MSCI, peut renforcer cette tendance par une augmentation de la concentration des acteurs spécialisés. En outre, le développement d'outils basés sur l'intelligence artificielle nécessite des investissements initiaux importants et les grandes institutions financières sont plus à même de développer ces technologies que les petits acteurs. Ainsi, ces différences technologiques pourraient conduire à une plus grande concentration au sein du système financier.

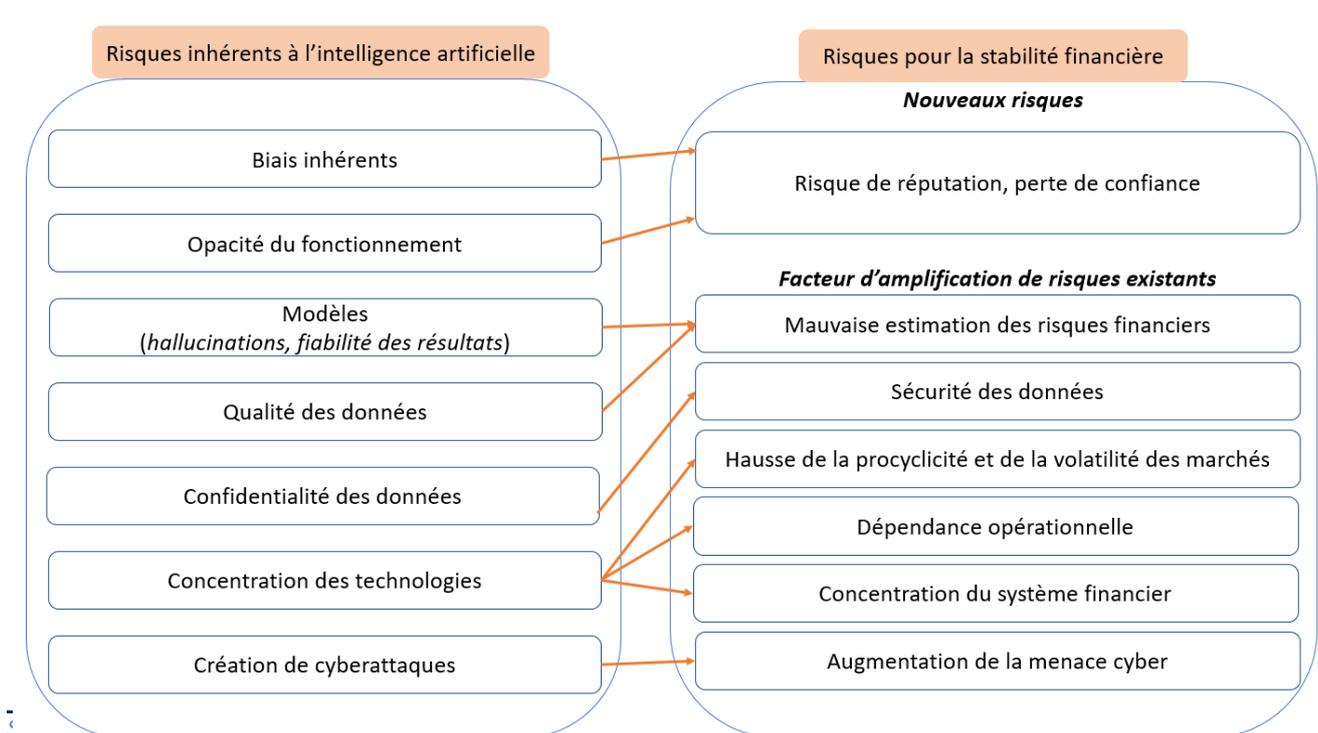
L'intelligence artificielle générative pourrait renforcer le risque cyber, auquel le système financier est très exposé. En effet, l'intelligence artificielle générative permet aux cyberattaquants de créer des attaques plus sophistiquées et plus difficiles à détecter. Cette technologie permet, par exemple, la rédaction automatisée de mails d'hameçonnage (*phishing*) personnalisés et à grande échelle dans plusieurs langues et sans faute d'orthographe. De même, les systèmes d'intelligence artificielle générative permettent d'imiter la voix (voire l'image) d'une personne en vue d'usurper son identité. L'intelligence artificielle générative facilite aussi les attaques reposant sur du code informatique malveillant. Les agents conversationnels sont en effet capables de produire ou de corriger du code informatique, ce qui permet de réaliser des attaques plus sophistiquées. L'intelligence artificielle permet aussi l'industrialisation des cyberattaques en automatisant leur conception et leur exécution car celle-ci est capable d'adapter sa stratégie en fonction des vulnérabilités détectées. En outre, l'intelligence artificielle générative permet une nouvelle catégorie de cyberattaque, nommée *indirect prompt injection*, qui consiste à manipuler les requêtes pour contourner la sécurité intégrée de l'outil afin que l'intelligence artificielle générative récupère elle-même des informations sensibles. Ces attaques sont possibles dès lors qu'une intelligence artificielle a la capacité de se connecter sur des serveurs par une requête, dont elle peut se servir pour transmettre des informations issues de sa conversation avec l'utilisateur. À la différence des méthodes précédentes, l'outil d'intelligence artificielle n'aide pas à produire des outils malveillants mais elle se comporte directement en acteur malveillant.

Néanmoins, des solutions d'intelligence artificielle générative sont en cours de développement pour assister dans la gestion du risque cyber, pour la détection des menaces et le renforcement dynamique des mesures de cybersécurité. La préparation face aux cyberattaques est un élément essentiel de la résilience cyber, et l'intelligence artificielle permet notamment d'analyser les vulnérabilités en direct, de détecter des cyberattaques, de simuler des scénarios d'attaques. L'intelligence artificielle offre aussi la possibilité d'une gestion automatisée des attaques, avec des protocoles de résolution adaptés à la menace.

Les risques inhérents à l'intelligence artificielle, notamment les risques liés aux biais potentiels, aux erreurs ou à l'opacité des modèles, pourraient également avoir des impacts réputationnels sur les institutions financières. Des erreurs ou des biais pourraient devenir systématiques à l'échelle d'un organisme ou d'un marché avec l'utilisation de modèles basés sur l'intelligence artificielle et engendrer une potentielle perte de confiance de la part des clients. Au-delà des risques pour les institutions financières, les biais potentiels et les difficultés de compréhension des résultats fournis peuvent engendrer un risque pour la protection des clients et des investisseurs. Une étude de l'ACPR met en évidence que les explications fournies sous forme de conversation par des *robo-advisors* tendent à augmenter, à tort, la confiance des clients (davantage prêts à suivre les recommandations incorrectes de l'outil) par rapport à des explications sous d'autres formes. Ceux-ci peuvent donc être mal informés des risques liés à un investissement et prendre des décisions inadaptées à leur situation.

La capacité de compréhension des résultats est un facteur essentiel pour une bonne gestion des risques par les établissements financiers et pour les superviseurs. La compréhension du fonctionnement de l'algorithme est nécessaire pour vérifier non seulement les calculs effectués, mais aussi pour étudier les variables prises en compte et le poids assigné à chacune d'entre elles pour l'obtention du résultat. En outre, la compréhension du fonctionnement est une condition nécessaire pour la vérification par les superviseurs que les calculs correspondent aux exigences réglementaires. Dans le cas des modèles internes d'évaluation de risque de crédit des banques, les algorithmes utilisés doivent être facilement interprétables et compréhensibles à la fois par la direction de la banque, afin qu'elle puisse prendre des décisions stratégiques en matière d'allocation de portefeuille de façon éclairée, et par le superviseur en charge de l'application des réglementations prudentielles.

Schéma 3.1 : Typologie des risques de l'intelligence artificielle pour la stabilité financière



3.4 Des risques structurels de manipulation de l'information pourraient avoir un impact sur le système financier

Les outils d'intelligence artificielle générative facilitent la création de fausses informations pouvant être utilisées à des fins de fraude et de déstabilisation géopolitique, avec des effets potentiellement déstabilisateurs pour le système financier.

Les technologies d'intelligence artificielle générative permettent de générer des images et du son à partir de contenu existant pour créer de fausses vidéos et enregistrements à des fins de manipulation et d'escroquerie. Ces contenus sont basés sur un processus appelé *deepfake* en anglais, qui consiste à superposer des images et du son pour générer un contenu falsifié. Début février 2024, l'image et la voix du gouverneur de la banque centrale de Roumanie ont été usurpées pour réaliser une vidéo visant à promouvoir des investissements frauduleux. En septembre 2023, le président de la *Securities and Exchange Commission* (SEC), Gary Gensler, a alerté sur le risque d'importants mouvements de marché unidirectionnels causés par les *deepfakes*, en cas de création de fausses informations menant à une perte de confiance dans un actif.

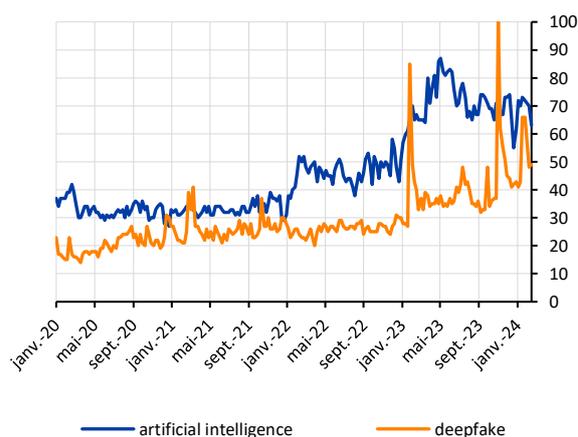
Ces fausses informations créées par les outils d'intelligence artificielle générative peuvent conduire à une perte de confiance vis-à-vis des institutions financières, et entraîner de potentielles paniques bancaires. La crise des banques régionales aux États-Unis, en mars 2023, a mis en exergue l'accélération de la panique des déposants par le biais de messages échangés sur les réseaux sociaux. Dès lors que l'intelligence artificielle générative permet de rendre crédibles de fausses informations concernant des institutions bancaires, leur diffusion peut mener à une panique et à des retraits massifs des dépôts.

Plus généralement, les campagnes de désinformation peuvent entraîner une déstabilisation géopolitique pouvant engendrer d'importantes conséquences politiques, sociales et économiques. L'édition 2024 du rapport

sur les risques mondiaux du Forum économique mondial (FEM) classe le risque de désinformation en tête des menaces les plus sérieuses à court terme (deux ans). Pour limiter le risque de désinformation du grand public avec des *deepfakes* générés par intelligence artificielle générative, de plus en plus d'entreprises du secteur ont décidé de mettre en place des mesures pour faciliter la détection des *deepfakes* ou limiter leur diffusion. C'est le cas d'OpenAI qui a annoncé l'ajout de filigranes sur les images générées par ses systèmes, ou de MidJourney qui a interdit la génération d'images utilisant le visage de personnalités politiques. Pour améliorer la détection de *deepfakes* et empêcher leur diffusion, la sensibilisation du public est nécessaire, ainsi que la collaboration des différents acteurs (entreprises spécialisées, réseaux sociaux, opérateurs téléphoniques, médias).

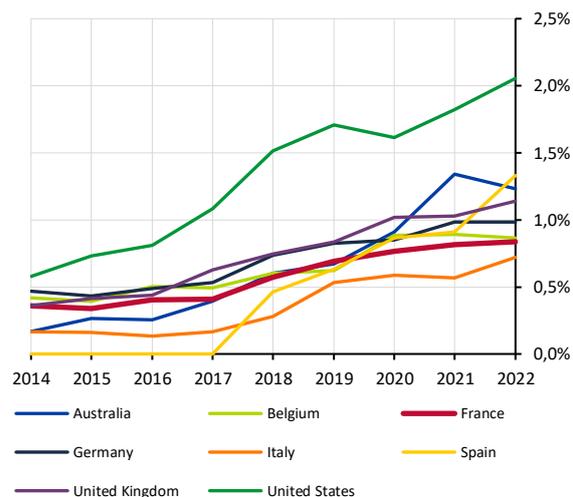
Graphique 3.7 : Volumes de recherches Google des termes intelligence artificielle et deepfakes

x : axe temporel / y : indice



Graphique 3.8 : Part des emplois en intelligence artificielle dans l'ensemble des offres d'emploi

x : axe temporel / y : pourcentage



Source : [Gauche] : Google Trends ; [Droite] : OurWorldInData

Note : Les tendances Google présentées ici sont mesurées sous la forme d'un indice qui prend la valeur 100 pour le point d'intérêt de recherche le plus élevé pour le terme depuis début 2020 dans le monde.

L'automatisation de certaines tâches, permise par l'intelligence artificielle, pourrait conduire à des changements structurels pour l'économie, même si ces effets restent difficiles à estimer avec certitude. Tous les pays et toutes les fonctions économiques ne sont pas également exposés aux changements du marché du travail par l'intelligence artificielle. Les économies développées disposent, d'une part, d'une infrastructure informatique permettant le déploiement de l'intelligence artificielle et, d'autre part, de nombreuses fonctions économiques automatisées ou automatisables (robots dans le secteur industriel, importance des fonctions juridiques, comptables, financières et informatiques) ; ce sont donc les plus susceptibles d'être affectées par l'essor de l'intelligence artificielle. Néanmoins, à court terme, peu de fonctions économiques seront entièrement automatisées par l'intelligence artificielle, et le plus probable est que nombre de métiers intègrent progressivement des composants de ces technologies.

L'intelligence artificielle est ainsi susceptible de conduire à des restructurations sectorielles majeures, qui pourraient alors déstabiliser le secteur financier si elles étaient mal anticipées ou plus rapides que prévu. Le déploiement de l'intelligence artificielle pourrait entraîner d'importants transferts d'emplois de certains secteurs économiques vers d'autres et, par conséquent, mener à un nombre élevé de faillites dans certains secteurs, ou au sein des entreprises elles-mêmes, avec des effets potentiels sur la solvabilité des entreprises ou des ménages concernés. Enfin, sur les marchés, l'adoption croissante et l'engouement autour de l'intelligence artificielle ont fait grimper les valorisations des entreprises liées au secteur en raison de l'optimisme des investisseurs quant aux bénéfices futurs de ces nouvelles technologies (cf. encadré 1.2 chapitre 1).

Le développement de l'intelligence artificielle dans le futur devra prendre en compte les limites imposées par les contraintes environnementales. L'entraînement d'un modèle d'intelligence artificielle nécessite une puissance de calcul importante, qui requiert à son tour une grande quantité d'énergie et implique des émissions de gaz à effet de serre. D'après les recherches d'OpenAI, depuis 2012, la puissance de calcul nécessaire pour l'apprentissage des modèles d'intelligence artificielle a doublé tous les trois mois. Une recherche en ligne utilisant l'intelligence artificielle générative utilise environ 4 à 5 fois l'énergie nécessaire à une recherche en ligne classique sur un moteur de recherche. En outre, les systèmes d'intelligence artificielle générative nécessitent d'importantes quantités d'eau afin de refroidir les processeurs. Le développement des modèles d'intelligence artificielle générative par Google et Microsoft a engendré des augmentations annuelles de consommation d'eau de 20 % et 34 %, respectivement, d'après les rapports environnementaux des entreprises. Les conséquences environnementales de l'intelligence artificielle seront donc des limites potentielles à son déploiement.

Encadré 3.1 : Utilisation de l'intelligence artificielle par la Banque de France

Par Philippe Grad

L'essor de l'intelligence artificielle présente également des opportunités pour l'exercice des missions de la Banque de France. Afin de tirer le meilleur parti de ces nouvelles technologies, tout en maîtrisant les risques, la Banque de France a adopté une démarche d'expérimentation. Celle-ci consiste à tester l'apport potentiel d'une solution technologique donnée sur un cas d'usage spécifique, afin de bien appréhender les mérites et les inconvénients, et de lever au maximum les incertitudes. À l'issue de la phase d'expérimentation, si celle-ci se révèle concluante, le projet passe en production.

Dans ce contexte, une dizaine d'applications intègrent aujourd'hui des fonctionnalités d'intelligence artificielle, parmi lesquelles :

- **N-ACSEL** : modèle de partitionnement (*clustering*) basé sur des données économiques qui permet de « profiler » les territoires. Cette vision, restituée aux responsables locaux, leur permet de mieux connaître leurs forces et leurs faiblesses et de se comparer aux autres territoires ;
- **ScoreIA** : modèle de cote de crédit, incorporé dans l'application d'aide à la cotation des entreprises afin d'être utilisé comme un outil d'aide à la décision par les analystes ;
- **PAI** : outil de mesure des perceptions et des anticipations d'inflation, établi par analyse d'un large panel d'articles de presse ;
- **BASTID** : système de détection de fraudes bancaires ;
- **LUCIA** : un outil pour la lutte contre le blanchiment (LCB-FT), permettant notamment l'analyse de grands volumes d'opérations bancaires et utilisé dans le cadre des missions de contrôle sur place des établissements ;
- **PLUME** : système de retranscription d'enregistrements vocaux vers du texte, utilisé notamment dans le contrôle des pratiques commerciales du secteur financier.

D'autres travaux, plus exploratoires, utilisant de l'intelligence artificielle et d'autres sources originales comme les données satellites, par exemple, sont aussi menés sur l'utilisation des données de pollution pour suivre l'activité industrielle y compris avec des partenaires extérieurs, permettant de couvrir de nombreux pays avec des méthodes homogènes.

La Banque de France bénéficie en effet d'un contexte favorable pour déployer les solutions à base d'intelligence artificielle : des données variées et nombreuses, une plateforme pour centraliser ces données et organiser leur gouvernance, des infrastructures pour l'innovation mais aussi une large expertise technique, notamment dans les équipes de *data scientists* et *data engineers* rattachées aux différents métiers, qui travaillent en coordination avec le LAB d'innovation de la Banque de France.

En parallèle de ces réalisations, le volet méthodologique pour la conception, le développement et l'utilisation des systèmes d'intelligence artificielle a également été investi. Ces travaux ont donné lieu à l'établissement d'un cadre de travail, qui s'applique à la fois aux équipes de développement et aux utilisateurs des systèmes d'intelligence artificielle. Ce cadre, dit « IA de confiance », comprend :

- un code de conduite définissant les principes et la conduite à adopter pour les développeurs et utilisateurs de l'IA ;
- une méthode d'analyse des risques propres à la technologie, à sa mise en œuvre et à son utilisation ;
- une boîte à outils pour faire face aux risques.

La méthode se base sur sept enjeux : i) diversité, non-discrimination et équité ; ii) sécurité, robustesse, fiabilité ; iii) transparence et explicabilité ; iv) respect de la vie privée et protection des données personnelles ; v) environnement de travail, responsabilité sociétale et environnementale ; vi) volet humain et contrôle humain ; vii) responsabilités et responsabilisation.

Ces enjeux, passés en revue dans le cadre des chantiers d'intelligence artificielle, permettent de faire émerger une hiérarchie de risques. Face à ces risques, une boîte à outils de bonnes pratiques est appliquée tout au long du cycle de vie et s'intègre dans notre organisation de contrôle interne.

Ces bonnes pratiques sont organisées en trois blocs correspondant aux étapes de développement : i) l'approvisionnement en données ; ii) le développement et la validation de la solution ; iii) l'exploitation du modèle.

3.5. Pistes de réglementations de l'IA

L'évolution rapide de l'intelligence artificielle et son adoption croissante nécessitent un cadre réglementaire pour limiter les différents risques mentionnés plus haut.

Au sein de l'Union européenne : le choix d'un règlement européen (AI Act). Le Parlement européen et le Conseil de l'Union européenne ont adopté (respectivement le 13 mars et le 21 mai 2024) un règlement (*l'Artificial Intelligence Act* ou *AI Act*) visant à encourager l'innovation et à promouvoir une intelligence artificielle digne de confiance, respectant les droits fondamentaux, la sécurité et les principes éthiques¹⁷², tout en limitant les risques que peuvent présenter les modèles d'intelligence artificielle les plus puissants.

Ce règlement, d'application directe, distingue les systèmes d'intelligence artificielle en fonction du niveau de risque qu'ils présentent :

- Les pratiques d'intelligence artificielle présentant un niveau de *risque inacceptable* pour la sécurité et les droits fondamentaux des citoyens, tels que la manipulation subliminale, le *scoring* social ou la police prédictive, seront interdites par l'*AI Act* (Chapitre II).

¹⁷² Le Règlement général sur la protection des données (RGPD) précise déjà que toute personne a « le droit de ne pas faire l'objet d'une décision fondée exclusivement sur un traitement automatisé, y compris le profilage, produisant des effets juridiques la concernant ou l'affectant de manière significative de façon similaire » (article 22).

- Les systèmes d'intelligence artificielle présentant un niveau de *risque élevé* seront encadrés. Dans le secteur financier, les deux cas d'usage considérés sont, pour les banques, l'évaluation de la solvabilité des personnes physiques et, pour les assurances, l'évaluation de la tarification pour les personnes physiques en matière d'assurance-vie et d'assurance-maladie. Avant d'être mis sur le marché ou déployés, ces systèmes devront notamment faire l'objet d'un enregistrement et d'une déclaration de conformité (portant sur les *items* suivants : processus de gestion des risques ; haute qualité des données ; documentation assurant la traçabilité et l'auditabilité ; degré approprié de transparence et d'interprétabilité ; mesures permettant la surveillance humaine ; robustesse, exactitude, cybersécurité). En France, l'ACPR sera l'autorité compétente (autorité de surveillance du marché) chargée du contrôle de l'utilisation de ces systèmes par les banques et les assurances.

D'autres systèmes d'intelligence artificielle, parfois désignés sous le terme de système à *risque limité*, formellement rassemblés sous l'appellation « certains systèmes d'IA » (Chapitre IV, article 50), devront se conformer à des exigences de transparence, notamment en précisant que leur contenu a été produit par une intelligence artificielle. Ces exigences s'appliqueront notamment aux systèmes qui interagissent directement avec des utilisateurs humains ou qui génèrent des contenus (images, vidéos, textes). **Les législateurs européens ont également pris en compte le cas spécifique des modèles d'intelligence artificielle à usage général.**

Ces modèles, définis comme ceux capables d'exécuter un large éventail de tâches distinctes (article 3 § 63), et qui peuvent être intégrés à d'autres systèmes, sont soumis à des exigences particulières (Chapitre V). Tous les modèles généralistes doivent ainsi respecter des obligations de documentation et de partage d'information avec les déployeurs. Pour les modèles présentant un *risque systémique*¹⁷³, les fournisseurs doivent se conformer à des exigences plus poussées d'évaluation et de transparence, proches de celles incombant aux systèmes à risque élevé (cf. ci-dessus).

L'AI Act sera généralement applicable deux ans après son entrée en vigueur. Toutefois, les dispositions concernant les risques inacceptables et celles concernant les modèles à usage général entreront en application respectivement six et douze mois après l'entrée en vigueur du règlement. L'articulation entre les règles horizontales énoncées par l'AI Act et la réglementation verticale spécifique au secteur financier n'est pas explicitement prévue par le règlement (notamment pour donner de la flexibilité aux superviseurs dans la manière d'intégrer ces nouvelles missions). Cet enjeu d'articulation sera crucial pour les superviseurs financiers, et pourrait notamment être précisé dans des orientations de l'Autorité bancaire européenne (ABE) et de l'Autorité européenne des assurances et des pensions professionnelles (AEAPP).

Les régulateurs internationaux ont entamé des travaux relatifs à l'impact de l'intelligence artificielle sur la stabilité financière. Le Conseil de stabilité financière (CSF) et le Comité de Bâle ont inscrit l'intelligence artificielle parmi les priorités de leur programme de travail pour 2024. D'une part, le CSF prévoit de formuler des recommandations sur la prise en compte des implications potentielles de l'intelligence artificielle sur la stabilité financière, après un premier rapport publié en novembre 2017. Dans un rapport publié en mai 2024, le Comité de Bâle présente les implications de la transformation numérique de la finance, et notamment de l'usage de l'intelligence artificielle pour les banques et leurs superviseurs. Alors que le déploiement de l'intelligence artificielle pourrait se traduire par une plus grande automatisation de certaines tâches, ce rapport insiste notamment sur la nécessité de maintenir un jugement humain dans la gouvernance et la gestion des risques des institutions financières.

Pour être appliquée, une réglementation nécessite toutefois des organes de contrôle efficaces. Les superviseurs financiers doivent ainsi se préparer à auditer les systèmes d'intelligence artificielle déployés dans le secteur financier, et d'abord ceux qui sont les plus risqués au regard des grands objectifs de la réglementation financière (stabilité financière, protection des consommateurs, LCB-FT). Cela constitue un défi à plusieurs égards, et nécessite en particulier d'acquérir une bonne maîtrise de la technologie. Outre des enjeux de ressources humaines, les superviseurs devront ainsi développer une méthodologie *ad hoc* de l'audit des systèmes d'intelligence artificielle. De ce point de vue, une démarche d'expérimentation, éventuellement en collaboration avec le secteur privé, peut aider à parcourir plus rapidement la courbe d'apprentissage. L'atténuation des risques liés à l'intelligence artificielle concerne tous les acteurs du système financier, et nécessite une mobilisation collective.

¹⁷³ Défini notamment par un seuil quantitatif de puissance informatique, fixé actuellement à 10^{25} opérations en virgule flottante par seconde (en anglais (*floating-point operations per second*)).