La plupart des entreprises prévoient d’adopter une stratégie d’IA multi-modèles, alors que près de six sur dix ont subi des coûts liés à l’IA significativement supérieurs aux prévisions

Les coûts de l’IA constituent désormais une préoccupation pour 53 % des dirigeants, selon une nouvelle étude « flash » sur l’économie de l’IA publiée par le Capgemini Research Institute et réalisée en septembre 2026 auprès de 500 hauts dirigeants de grandes entreprises. Au-delà des coûts directs tels que les *tokens*, les dirigeants manquent de visibilité sur le coût total d’un déploiement sécurisé de l’IA, notamment la supervision humaine, le besoin de retravailler le contenu produit et la gouvernance. Pour y répondre, les entreprises renforcent leurs contrôles des coûts tout en explorant des stratégies multi-modèles. Alors qu’elles déploient l’IA à plus grande échelle et se préparent à l’essor des systèmes agentiques, les entreprises doivent accorder plus d’attention à la valeur métier générée par la transformation de l’IA.

Un manque de visibilité sur les coûts de l’IA pour l’entreprise peut entraîner des conséquences financières majeures

Au cours de l'année écoulée, 59 % des dirigeants indiquent que les dépenses de leur entreprise en matière d'IA ont été significativement plus élevées que prévu ou budgété. Dans les entreprises concernées, les dépenses ont dépassé les prévisions ou le budget de 24 % en moyenne, et près d'une entreprise sur cinq (21 %) ont dépassé leur budget IA de plus de 30 %. Les cadres dirigeants sont davantage préoccupés par les coûts de l'IA qu'il y a un an, et seules 56 % d'entre eux ont confiance dans leur capacité à pleinement anticiper l'évolution de leurs coûts à mesure que l'adoption de l'IA se généralise.

Les cadres dirigeants manquent de visibilité sur les coûts à l'échelle de l'entreprise. La moitié d'entre eux déclarent que les dépenses liées à l'IA sont validées sans visibilité suffisante sur le coût total. Ils manquent aussi de confiance dans leur capacité à estimer précisément le coût total de leurs projets actuels d'IA, en particulier les coûts indirects, tels que la supervision humaine, le retravail et la gouvernance.

Cela peut engendrer des conséquences fâcheuses pour les entreprises concernées. Deux tiers des entreprises ayant dépassé leur budget IA ont subi des répercussions financières négatives, telles qu'une baisse du flux de trésorerie disponible ou une réduction des dépenses discrétionnaires. Dans ce contexte, les directeurs financiers jouent désormais un rôle plus important dans la validation des principaux investissements liés à l'IA.

« *À mesure que l'IA s'intègre davantage aux processus métiers, les entreprises doivent cesser de considérer son coût comme une simple question technologique ou d'achats et commencer à gérer l'IA comme un ensemble de résultats créateurs de valeur*, déclare Volker Darius, Directeur mondial de Capgemini Invent. *La maîtrise des coûts à elle seule ne permettra pas de rentabiliser l'IA, et la création de valeur sans contrôle des coûts ne sera pas tenable non plus. Les entreprises doivent mesurer simultanément les coûts, la qualité et la valeur métier, utiliser le modèle le plus adapté à chaque besoin et prendre en compte la supervision humaine nécessaire pour garantir des résultats fiables et sécurisés. En mesurant continuellement l'*AI Yield*, elles peuvent prendre des décisions d'allocation du capital plus éclairées et déployer l'IA à grande échelle avec davantage de confiance* . »

Les stratégies multi-modèles s'imposent dans les stratégies d'IA

Face à l'augmentation inattendue des dépenses liées à l'IA, plus de la moitié des entreprises (53 %) ont renforcé le contrôle de leurs dépenses, soit en resserrant les budgets et en renforçant le contrôle des initiatives concernées, soit en instaurant des contrôles des dépenses IA à l'échelle de l'entreprise, soit encore en réduisant le périmètre ou l'ambition du projet. Seule une entreprise sur cinq a repensé ses outils ou ses processus, notamment en réévaluant quelles activités devraient être pilotées par l'IA plutôt que par l'humain, ou en changeant de solution IA, de modèles utilisés, ainsi que leur configuration ou leur routage.

Au-delà de ces réponses immédiates, plus de trois cadres dirigeants sur cinq prévoient que leur entreprise adoptera une stratégie multi-modèles d'ici un an, ce qui permettra de sélectionner dynamiquement les modèles en fonction du coût, des capacités, de la latence, du niveau de risque et d'autres exigences métiers. Ils prévoient également d'utiliser davantage de modèles propriétaires moins coûteux dès lors qu'ils affichent une performance suffisante, ainsi que des modèles plus petits ou spécialisés pour certaines tâches spécifiques. Par ailleurs, près de trois entreprises sur cinq prévoient d'étendre leur utilisation des modèles *open-weight*.[1]

L'*AI Yield* replace le retour sur investissement au premier plan

Si deux tiers des dirigeants déclarent avoir confiance dans leur capacité à vérifier la valeur ajoutée liée à l'utilisation de l'IA, la plupart des entreprises ne sont pas en mesure de relier directement leurs dépenses d'IA aux résultats métiers obtenus. Cela s'explique soit parce qu'elles ne les mesurent pas formellement, soit parce qu'elles suivent uniquement les coûts ou l'utilisation de l'IA sans mesurer les résultats associés, soit parce qu'elles suivent ces indicateurs séparément. Parmi les défis majeurs pour obtenir la validation et le financement d'investissements en IA, 57% des cadres dirigeants indiquent qu'il est difficile de démontrer la valeur ajoutée de l'IA pour l'entreprise.

A peine une entreprise sur cinq calcule l'*AI Yield*, un indicateur qui vérifie la valeur ajoutée de l'utilisation de l'IA en tenant compte de l'ensemble des coûts nécessaires à son déploiement sécurisé. Les bénéfices métiers les plus fréquemment cités sont la productivité des collaborateurs ou la capacité dégagée, suivis de la réduction des coûts opérationnels, et l'accélération des processus ou des cycles d'activité. Moins d'un tiers des répondants (32 %) cite la croissance du chiffre d'affaires, notamment grâce à de nouveaux produits, services ou modèles économiques.

Les indicateurs utilisés pour piloter l'économie de l'IA devraient donc évoluer. Aujourd'hui, 60 % des entreprises comparent principalement les dépenses totales d'IA au budget, mais seules 44 % prévoient de continuer à prioriser ce mode de calcul au cours des 6 à 12 prochains mois. Le taux de réussite dès la première exécution[2] devrait devenir l'indicateur prioritaire, suivi par le coût par tâche ou workflow réalisé. Ces indicateurs placent les résultats obtenus, plutôt que la seule consommation d'IA, au cœur de l'économie de l'IA.

L'*AI Compass* de Capgemini

Cette étude « instantanée » fait suite à la première édition de ***The AI Compass***, la nouvelle publication de Capgemini consacrée à l'économie de l'IA. Intitulée *The Era of AI Yield: Mastering Tokenomics to Maximize Enterprise Value* (L'ère de l'*AI Yield* : maîtriser l'économie des *tokens* pour maximiser la valeur créée pour l'entreprise), elle examine comment les entreprises peuvent dépasser les seuls indicateurs de consommation et de coûts de l'IA afin de mesurer l'*AI Yield* : le rendement économique de l'utilisation de l'IA qui tient compte de l'ensemble des coûts nécessaires

à son déploiement sécurisé. Cette publication examine également les implications pour les dirigeants et les responsables technologiques en matière d'allocation du capital, de gouvernance, de routage des modèles, d'ingénierie des workflows et de supervision humaine.

Pour accéder à l'étude « instantanée » en intégralité, cliquez ici : https://www.capgemini.com/insights/research-library/ai-economics/.

[1] Un modèle open-weight est un modèle d'intelligence artificielle dont les paramètres appris (les poids) sont rendus publics et téléchargeables, mais dont le code source d'entraînement et les données d'apprentissage restent généralement privés.

[2] Le taux de réussite dès la première exécution correspond au pourcentage de tâches, workflows ou résultats obtenus grâce à l'IA réalisés avec succès dès la première tentative, sans retravail, escalade, correction ni intervention humaine supplémentaire.

Méthodologie du rapport

En septembre 2026, le Capgemini Research Institute a interrogé 500 hauts dirigeants issus d'entreprises réalisant plus d'un milliard de dollars de chiffre d'affaires annuel, dans 10 pays (Allemagne, Chine, Espagne, États-Unis, France, Inde, Italie, Japon, Pays-Bas et Royaume-Uni) et couvrant 13 secteurs d'activité. Les répondants représentaient principalement les fonctions de stratégie et de gouvernance IA, de plateformes et d'ingénierie, de création de valeur, de finance technologique et de gestion des coûts liés à l'IA. L'enquête a été complétée par deux entretiens approfondis avec des experts du secteur.