

Libérer l'IA en entreprise : opportunités et stratégies

Une étude mondiale menée auprès de
1 100 dirigeants et spécialistes technologiques,
complétée par des entretiens avec 28 DSI



Table des matières

3	À propos de ce rapport
4	Enquête
5	Remerciements
6	Avant-propos d'Andy Kofoid, President of global field operations, Databricks
7	Résumé
10	Chapitre 1 : La libération de l'IA
11	Les dirigeants, et plus encore les opérationnels, estiment que l'IA a prouvé sa valeur
12	La puissance à la portée de tous
14	Prendre plus rapidement la bonne décision
16	Adapter l'approche au niveau de risque
17	Bancs d'essai
19	Chapitre 2 : Récolter les fruits
19	Une démesure raisonnable ?
22	Intelligence des données
23	Mesurer les gains concrets de l'IA
25	La vertu de la patience
27	Chapitre 3 : La refonte des infrastructures
28	Aucune économie d'échelle à l'horizon
30	La bonne méthode pour les bonnes données
33	Le poids de la maintenance
34	Perspectives des professionnels
36	Chapitre 4 : Trouver votre avantage
36	Composer un assortiment
38	Choisir des modèles
42	Chapitre 5 : Garde-fous et gouvernance
44	Les centres d'excellence, catalyseurs d'une coordination à marche forcée
46	Équilibrer l'humain et la machine
50	Conclusion

À propos de ce rapport

Ce rapport d'Economist Impact, réalisé à la demande de Databricks, associe une enquête mondiale menée auprès de 715 cadres techniques et 385 spécialistes des données et de l'intelligence artificielle (IA). Ces personnes travaillent dans les domaines du data engineering, de la data science et de l'architecture d'entreprise.

Il présente également des éclairages issus d'entretiens avec 28 dirigeants d'organisations de premier plan exerçant dans 11 secteurs d'activité : services financiers, santé / sciences du vivant, vente au détail / biens de consommation, secteur public, fabrication, transports, énergie, technologie, services

professionnels, médias / divertissement et télécommunications. L'échantillon inclut des entreprises nativement digitales de divers secteurs. La plupart des personnes interrogées travaillent au sein de sociétés multinationales dont le siège social est implanté dans l'un des dix pays de l'enquête, répartis sur les continents suivants : Amériques, Asie-Pacifique et Europe.

Le rapport procède à une évaluation complète de la mise en œuvre de l'IA du point de vue des dirigeants et des professionnels. Il met ainsi en évidence les catalyseurs et les goulets d'étranglement de la démarche, tout comme les stratégies de réussite les plus prometteuses.

Enquête

L'enquête mondiale au cœur de cette étude, élaborée et menée par Economist Impact à la demande de Databricks, a sondé 1 100 participants aux parcours variés, détaillés ci-dessous. Elle a été menée entre juillet et août 2024.

Les personnes interrogées travaillent dans de grandes entreprises (chiffre d'affaires annuel de 500 millions de dollars ou plus) ou dans d'importantes organisations du secteur public, basées dans 19 pays d'Amérique du Nord, d'Europe et d'Asie-Pacifique (ASEAN).

Fonction des participants

- Directeur des systèmes d'information
- Chief technology officer
- Chief data/analytics officer
- Data scientist en chef
- Architecte de données/d'entreprise en chef
- SVP/VP/Responsable de l'informatique/ des données/de l'ingénierie
- Data scientist (ou poste similaire)
- Data engineer (ou poste similaire)
- Architecte d'entreprise (ou similaire)

Secteur d'activité des participants

- Services financiers, banque et assurance
- Secteur public
- Santé, produits pharmaceutiques et sciences du vivant
- Vente au détail et biens de consommation
- Industrie manufacturière (dont l'automobile)
- Médias et divertissement
- Énergie, pétrole et gaz
- Télécommunications

Localisation des participants

Amériques

- États-Unis

Asie-Pacifique

- Australie
- Inde
- Japon
- Malaisie
- Philippines
- Singapour
- Corée du Sud
- Thaïlande

Europe

- Danemark
- Finlande
- France
- Allemagne
- Italie
- Pays-Bas
- Norvège
- Espagne
- Suède
- Royaume-Uni

Remerciements

Nous tenons à remercier les dirigeants suivants de nous avoir consacré du temps et apporté leurs éclairages :

- **Chalee Asavathiratham**, ancien Chief digital banking officer, **Siam Commercial Bank**
- **Wassym Bensaid**, Directeur logiciel, **Rivian**
- **Sanjay Bhakta**, Directeur des produits et des technologies, **Condé Nast**
- **Ian Botts**, Chief technology officer, **Fanatics Betting & Gaming**
- **Bernd Bucher**, Directeur des systèmes d'information, **Novartis**
- **Roman Bugaev**, Chief technology officer, **Flo Health**
- **Ting Cai**, Directeur de l'IA et des données, **Rakuten Group**
- **Juan Jose Casado**, Chief digital officer, **Repsol**
- **Kushal Chakrabarti**, Chief data officer, **Opendoor**
- **Helen Choi**, Chief digital and information officer, **CJ CheilJedang**
- **Carol Clements**, Chief digital and technology officer, **JetBlue**
- **Ken Finnerty**, President of IT and data analytics, **UPS**
- **Jon Francis**, Chief data and analytics officer, **General Motors**
- **Leonel Garciga**, Directeur des systèmes d'information, **Armée américaine**
- **Scott Hallworth**, Chief data and analytics officer, **HP**
- **Andy Hill**, Chief data officer, **Unilever**
- **Mohit Kapoor**, Group chief technology officer, **Mahindra**
- **Jeff Martin**, Senior vice president and chief data officer, **TD Bank Group**
- **Melissa Pint**, Chief digital information officer, **Frontier**
- **Senthil Ramani**, Global lead, data and AI, **Accenture**
- **Gereurd Roberts**, Directeur général du groupe, Seven Digital, **Seven West Media**
- **Robbert Van Rutten**, Directeur des systèmes d'information, **Shell**
- **Takaaki Sato**, Senior executive vice president, chief technology officer, **NTT Docomo**
- **Amit Sharma**, Chief technology officer, **Dream Sports**
- **Ryan Snyder**, Vice-président senior et directeur des systèmes d'information, **Thermo Fisher Scientific**
- **Sara Vaezy**, Chief strategy and digital officer, **Providence**
- **Greg Ulrich**, Chief AI and data officer, **Mastercard**
- **Darrin Vohs**, Directeur monde des systèmes d'information, **Molson Coors**

Avant-propos d'Andy Kofoid, President of global field operations, Databricks

La trajectoire actuelle de l'intelligence artificielle (IA), en termes d'investissement, d'adoption et de résultats, indique clairement que cette technologie va devenir une partie intégrante de chaque entreprise, quel que soit le secteur d'activité. Une autre réalité s'impose également : pour réussir dans ce nouveau paysage, les dirigeants d'entreprise ne pourront pas se contenter d'adopter le modèle d'IA le plus récent et le plus performant.

Dans chaque secteur, les gagnants seront ceux qui auront adopté une approche globale, rassemblant la gestion des données, la sécurité, la gouvernance, la culture et l'expertise. Tous ces domaines sont abordés dans les conclusions de ce rapport d'Economist Impact, qui a interrogé 1.100 dirigeants et spécialistes technologiques de 19 pays et huit secteurs d'activité, ainsi que 28 hauts dirigeants. Ce rapport déborde d'informations destinées à aider les décideurs technologiques et les professionnels des données et de l'IA à tracer leur propre parcours dans le domaine de l'IA.

Si ce rapport nous enseigne une chose, c'est que les entreprises souhaitent massivement enrichir des modèles d'IA générative avec leurs propres données. Pour cela, elles doivent moderniser leur plateforme de données, résoudre le problème de la fragmentation et dépasser les limites que pose l'utilisation d'un seul grand modèle commercial. L'avenir fait également la part belle aux systèmes et agents d'IA à plusieurs composants et faisant usage de modèles personnalisés, chacun étant optimisé pour des tâches spécifiques. Cette approche délivre une plus grande précision, réduit les coûts, améliore les performances et renforce la sécurité – autant de facteurs cruciaux pour des applications d'entreprise basées sur l'IA.

Ce rapport révèle également l'importance de la démocratisation comme moteur de l'IA. Les organisations souhaitent mettre les données à la disposition d'un plus large éventail de salariés et donner à leurs collaborateurs les moyens de créer et d'affiner des modèles d'IA en phase avec les besoins commerciaux quotidiens. Imaginez que vous puissiez interroger vos données en anglais ou dans votre langue maternelle. Vous pourriez par exemple demander : « Comment se porte notre activité européenne par rapport aux objectifs de fin d'exercice ? » et obtenir instantanément

une réponse détaillée et certifiée exacte, tenant véritablement compte des particularités de votre entreprise.

La sécurité et la gouvernance ont un rôle fondamental à jouer dans le nouveau paysage de l'IA. Les entreprises se précipitent pour mettre en œuvre des solutions d'IA, mais elles se demandent aussi comment protéger les données sensibles et rester en conformité avec des réglementations qui évoluent sans cesse. Elles vont parfois trop vite, mais on sait à quel point la lenteur peut être nuisible à notre époque. Il est impératif de mettre en place des processus robustes de sécurité et de gouvernance pour ouvrir la voie à l'innovation sans se laisser paralyser par un excès de prudence, surtout dans le contexte de l'évolution de l'environnement réglementaire entourant l'IA.

Tout ce qui précède nous amène à une conclusion : l'avenir ne repose pas uniquement sur des modèles d'IA entraînés sur les données du Web. L'avenir réside dans une véritable intelligence des données, autrement dit dans des systèmes qui comprennent et exploitent en profondeur les datasets spécifiques – et souvent sensibles – d'une organisation. Une IA qui comprend le contexte propre à votre entreprise. Grâce à cette approche basée sur des systèmes d'agents exploitant des modèles spécialisés, les entreprises peuvent créer des solutions adaptées à leurs besoins et à leurs domaines, et faire de leurs données propriétaires un puissant avantage compétitif.

La révolution de l'IA ne fait que commencer, et une chose est sûre : les organisations qui sauront exploiter efficacement leurs données et créer des systèmes d'IA en phase avec leurs besoins seront les leaders de demain. Chez Databricks, nous sommes fiers de soutenir les recherches qui mettent en lumière les opportunités et les stratégies qui définiront l'ère de l'IA.

Andy Kofoid,
President of global field operations, Databricks



Résumé

Dans les années à venir, des dépenses d'investissement colossales atteignant 1 000 milliards de dollars sont prévues pour fournir les centres de données, les puces, l'énergie et l'infrastructure nécessaires à la révolution de l'intelligence artificielle (IA).¹ Ce chiffre est en grande partie dû à l'avènement de l'IA générative, mais il faut savoir que les entreprises développent en réalité leurs capacités d'IA depuis des années. Des avancées telles que le deep learning et les réseaux de neurones ont permis aux spécialistes de domaines variés, de la biotechnologie à la finance, d'analyser de vastes ensembles de données pour y découvrir des tendances et des éclairages inédits.

Ce sont la démocratisation et l'échelle des opérations qui distinguent notre époque. Parce que l'IA générative offre une interface intuitive en langage naturel, la puissance de l'IA est devenue accessible à tous les profils professionnels. Pour réussir le passage du projet pilote au déploiement à l'échelle de l'entreprise, il faut une infrastructure robuste capable de répondre aux besoins en données et en calcul de l'IA, une stratégie de main-d'œuvre définissant le juste équilibre entre humains et machines, ainsi qu'une stratégie de retour sur investissement (ROI) appropriée.

Surtout, l'excellence en matière d'IA ne peut être atteinte sans une gouvernance forte et une conception minutieuse de l'interaction homme-machine. Les décisions et les résultats doivent être conformes à la législation émergente qui existe dans de nombreux domaines : protection de la vie privée, sécurité et protection des consommateurs.

Ce rapport d'Economist Impact, réalisé à la demande de Databricks, associe une enquête mondiale menée auprès de 715 cadres techniques et 385 spécialistes des données et de l'IA travaillant dans les domaines du data engineering, de la data science et de l'architecture d'entreprise. Il présente également des éclairages issus d'entretiens avec 28 dirigeants d'organisations de premier plan exerçant dans 11 secteurs d'activité.

Mêlant tendances chiffrées et insights qualitatifs, le rapport procède à une évaluation complète de la mise en œuvre de l'IA du point de vue des dirigeants et des professionnels. Il met ainsi en évidence les catalyseurs et les goulets d'étranglement de la démarche, ainsi que les stratégies de réussite les plus prometteuses.

1 Goldman Sachs, « Gen AI: too much spend, too little benefit? » (IA générative : trop de dépenses, pas assez de bénéfices ?), 2024, <https://www.goldmansachs.com/insights/top-of-mind/gen-ai-too-much-spend-too-little-benefit>

Les dirigeants et les professionnels croient tous au pouvoir de l'IA, mais ils estiment que les investissements sont insuffisants. Qu'il s'agisse d'accélérer le développement de médicaments ou d'ouvrir l'accès au crédit aux personnes financièrement exclues, les entreprises trouvent un large éventail d'applications à l'IA et l'IA générative. Selon notre enquête, 85 % des organisations utilisent activement l'IA générative dans au moins une fonction commerciale. Ce chiffre atteint 97 % dans les entreprises dont le chiffre d'affaires est supérieur à 10 milliards de dollars. Les équipes informatiques sont les plus enthousiastes, tandis que la fonction juridique se montre la plus réticente. L'IA générative est pour l'instant principalement réservée aux projets internes, mais d'ici 2027, 99 % des décideurs s'attendent à ce qu'elle soit aussi bien adoptée dans des cas d'usage externes qu'internes. Sept répondants sur dix considèrent l'IA comme essentielle à leurs objectifs stratégiques à long terme, et ils ne sont que 18 % à penser que son intérêt est surestimé. Malgré cette dynamique, un répondant sur cinq seulement estime que son entreprise investit suffisamment dans les domaines techniques et non techniques actuellement.

L'IA offre des avantages à court terme en matière de productivité et d'efficacité, mais les leaders utiliseront l'IA sur le long terme pour créer de la valeur, des revenus et des modèles commerciaux innovants. Jusqu'à présent, l'IA s'est surtout distinguée par son impact sur la productivité. On observe d'ailleurs que l'IA générative a été massivement expérimentée dans les fonctions à fort potentiel d'automatisation : l'informatique (91 %), le marketing (85 %), les ventes et le service client (83 %) et les opérations (80 %). Les data scientists déclarent même à 82 % utiliser l'IA pour coder. Mais ce n'est pas le seul bienfait cité par les personnes interrogées, qui évoquent également la réduction des coûts, l'amélioration de l'expérience des employés et l'attraction des talents, entre autres. Les dirigeants privilégient toutefois une vision à plus long terme

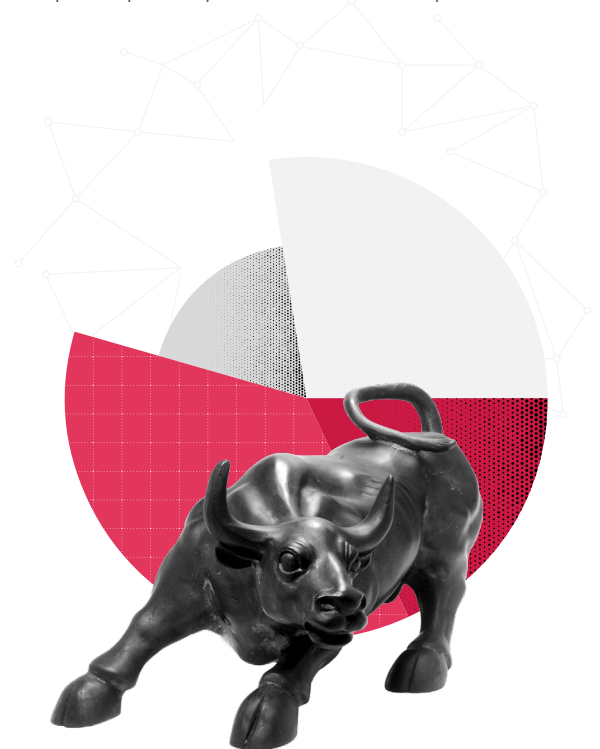
du changement de paradigme que représente l'IA générative. Pour eux, les considérations stratégiques telles que l'innovation dans les modèles économiques, le positionnement sur le marché et les normes environnementales, sociales et de gouvernance sont les critères les plus importants lorsqu'ils évaluent l'intérêt commercial de l'IA. Jusqu'à présent, l'augmentation des revenus reste l'indicateur le moins efficace pour justifier un investissement. Les experts affirment qu'il faudra du temps pour rentabiliser l'IA, étant donné les impératifs d'expérimentation, d'itération et de refonte de l'infrastructure numérique. Pourtant, au fil du temps, c'est bien la capacité à générer des revenus qui permettra aux leaders de se démarquer. Il faudra alors chiffrer la valeur financière avec davantage de précision.

Exploiter l'IA à grande échelle exige une refonte de l'infrastructure : la plupart des entreprises n'ont pas confiance dans leur architecture actuelle. De nombreuses entreprises utilisent aujourd'hui des technologies largement dépassées. Elles sont seulement 22 % à affirmer que leur architecture actuelle peut prendre en charge les charges de travail de l'IA sans modification. Et 48 % des data engineers passent la majeure partie de leur temps à rétablir les connexions aux sources de données. Le problème pourrait s'aggraver avec la démocratisation de l'IA, qui va certainement se traduire par une prolifération de pilotes, elle-même source de congestion, de complexité et d'opacité dans les données et les infrastructures. La récompense sera à la hauteur de l'effort : d'après certaines des entreprises interrogées dans le cadre de cette étude, ce n'est qu'une fois leur socle de données en place qu'elles ont pu faire preuve de créativité dans l'élaboration de cas d'usage et rentabiliser leurs investissements. Les inquiétudes concernant la sécurité des données, souvent enfermées dans des silos ou des systèmes fragmentés, empêchent certaines entreprises de se lancer dans des expérimentations plus ambitieuses.

Deux tiers des organisations trouvent prometteuse l'intégration de leurs propres données à des modèles d'IA générative. Les modèles d'IA sont de plus en plus faciles à acquérir et à exploiter. Les leaders sauront associer les modèles et les outils à leurs meilleurs atouts : leurs données propriétaires et leur savoir-faire. Pour cela, il faut trouver la bonne palette d'outils, en combinant des modèles open source et fermés, par exemple, et acquérir une visibilité sur les datasets internes pour les maîtriser parfaitement. Les organisations explorent différents types de modèles et près de sept sur dix expérimentent ou ont entièrement déployé des modèles open source d'IA générative. 96 % d'entre elles pensent le faire d'ici 2027. Les experts mettent toutefois en garde contre les modèles prêts à l'emploi qui affichent des performances médiocres quand il s'agit de s'adapter à un lexique technique spécifique, et qui peuvent offrir moins de contrôle et de sécurité. Mais ces modèles commerciaux, non connectés aux données d'entreprise, sont encore courants : près de la moitié (45 %) des data scientists utilisent de grands modèles de langage (LLM) sans génération augmentée par récupération (RAG).

Peu d'entreprises utilisent véritablement l'IA générative « en production » pour des questions de coût, de compétences et de gouvernance. Seuls 37 % des dirigeants estiment que leurs applications d'IA générative sont prêtes à passer en production, et ce chiffre tombe à 29 % lorsqu'on interroge les opérationnels. Pour les data scientists, les principales contraintes sont le coût (41 %), les compétences (40 %), la qualité (37 %) et la gouvernance (33 %). Seul un répondant sur six estime que son organisation peut mobiliser suffisamment de talents en IA. La moitié des data engineers déclarent que la gouvernance mobilise la majeure partie de leur temps, et plus de la moitié des architectes d'entreprise (53 %) voient dans les défis liés à la confidentialité des données et aux violations de sécurité les plus grands risques de l'expansion de l'IA. Les dirigeants s'accordent sur une chose : dans les années à venir, les gagnants seront ceux qui passeront de l'expérimentation à la production et à la monétisation à grande échelle.

Inéluctable, l'adoption massive de l'IA impose de calibrer soigneusement l'équilibre entre l'intelligence humaine et celle des machines pour permettre la démocratisation des écosystèmes d'IA. Les dirigeants affirment que l'IA peut soutenir les collaborateurs, mais pas les remplacer. Les personnes interrogées citent les systèmes unifiés et les « centres d'excellence » comme des outils essentiels pour équilibrer gouvernance et habilitation, soulignant le rôle irremplaçable de l'expérience et du jugement dans la supervision des résultats de l'IA. Deux tiers des organisations affirment qu'elles expérimentent encore pour trouver le juste équilibre entre les humains et l'IA. Beaucoup investissent également dans des outils en libre-service et des assistants IA pour permettre à tous leurs collaborateurs de devenir des data scientists. Les architectes d'entreprise considèrent l'IA générative comme une force de démocratisation. Ils sont près de 60 % à prédire que d'ici trois ans, le langage naturel sera le principal moyen – si ce n'est le seul – d'interagir avec des ensembles de données complexes pour le personnel non technique.



Chapitre 1 :

La libération de l'IA

Les organisations font tout leur possible pour exploiter la puissance de l'IA. Elle misent pour cela sur les modèles génératifs qui ouvrent de nouvelles portes dans tous les secteurs d'activité et toutes les fonctions de l'entreprise. Que ce soit pour soutenir la découverte de médicaments ou pour alimenter des assistants intelligents, un large éventail de modèles et de techniques décuplent la productivité des équipes, guident les décisions stratégiques et repoussent les frontières de l'innovation. À l'heure où les entreprises passent de l'expérimentation à la mise en œuvre, l'avenir appartient à celles qui sauront exploiter les solutions d'IA à grande échelle.

Deux ans après le lancement public de ChatGPT, toutes les organisations cherchent à faire des algorithmes un avantage compétitif. Les entreprises couplent l'IA générative à la puissance de l'IA dite « traditionnelle », capable de reconnaître des modèles, de traiter des ensembles de données volumineux et variés, ou de faire des prédictions basées sur des données historiques. Le résultat : des systèmes en mesure de raisonner et de converser en mêlant différents contextes, mais aussi de produire de nouveaux résultats imitant la réflexion humaine, tout cela en langage naturel.²

La grande percée de l'IA générative est d'avoir mis à la portée de tous ou presque une technologie jusque-là réservée à des experts. « Le développement exponentiel de l'IA générative est dû à son accessibilité : elle est accessible à l'utilisateur moyen, » explique Darrin Vohs, Global chief information officer chez Molson Coors, une entreprise canadienne de boissons.

De nombreuses entreprises expérimentaient déjà l'IA bien avant le lancement de ChatGPT, en particulier les grands acteurs de domaines techniquement avancés. Novartis, le géant pharmaceutique suisse, utilise le machine learning (ML) pour découvrir de nouveaux médicaments et l'analyse prédictive pour améliorer les essais cliniques.^{3,4} Flo Health, une application mondiale de santé féminine grand public, utilisée par plus de 380 millions de personnes, repose sur un algorithme utilisant des techniques de ML traditionnelles. « Les avancées des LLM sont vraiment passionnantes pour nous, » affirme Andy Hill, Chief data officer chez Unilever, acteur majeur des biens de consommation. « Mais c'est toujours en les associant à des techniques d'apprentissage supervisé plus traditionnelles que les LLM apportent réellement de la valeur à une entreprise comme la nôtre. »

2 MIT News | Massachusetts Institute of Technology, « Explained: Generative AI » (IA générative : une explication), 2023, <https://news.mit.edu/2023/explained-generative-ai-1109>

3 Novartis, « The art of drug design in a technological age » (L'art de la conception de médicaments à l'ère des technologies), 2021, <https://www.novartis.com/stories/art-drug-design-technological-age>

4 Novartis, « Novartis' commitment to the ethical and responsible use of Artificial Intelligence (AI) Systems » (Novartis s'engage pour une utilisation éthique et responsable des systèmes d'intelligence artificielle), 2020, https://www.novartis.com/sites/novartis_com/files/novartis-responsible-use-of-ai-systems.pdf

Les modèles génératifs enrichissent les capacités de l'IA grâce à leurs nombreuses applications. Même les équipes possédant une grande expérience de l'IA lui ont ajouté une nouvelle couche de fonctionnalités. Thermo Fisher Scientific, une société internationale des sciences de la vie, intègre l'IA dans ses instruments depuis près de dix ans. Selon Ryan Snyder, Directeur des systèmes d'information de l'entreprise, l'IA générative vient enrichir cette solide base IA/ML en permettant une interrogation plus approfondie des données et des expérimentations plus audacieuses.

Les dirigeants, et plus encore les opérationnels, estiment que l'IA a prouvé sa valeur

Si le développement de l'IA a connu ses traversées du désert, les participants à notre enquête se montrent extrêmement optimistes : 73 % d'entre eux affirment en effet que l'IA générative est incontournable pour leurs objectifs stratégiques à long terme. Ils ne sont que 18 % à estimer que l'IA est surfaite, même si les cadres, convaincus à 20 % que le battage médiatique autour de l'IA est excessif, sont un peu plus prudents que les spécialistes des données et de l'IA (seulement 13 %). Dans les années à venir, c'est la capacité de passer de

l'expérimentation à la mise en œuvre à grande échelle qui fera la différence entre gagnants et perdants.

Au milieu de cet optimisme prudent, l'adoption rapide de l'IA générative est impossible à ignorer. Une écrasante majorité (85 %) des organisations interrogées dans le cadre de notre enquête testent ou utilisent l'IA actuellement. D'ailleurs, les taux d'adoption augmentent avec la taille de l'entreprise : chez celles dont le chiffre d'affaires est supérieur à 10 milliards de dollars, il atteint même 97 % (tests et utilisation). Cette tendance suggère que les grandes organisations pourraient consolider leur avantage compétitif grâce à l'IA, tandis que celles qui tardent à investir risquent de passer à côté du prochain virage technologique majeur. « Il y a toujours une part de vérité derrière le battage médiatique. Quand la bulle des dot-coms a éclaté, beaucoup d'entreprises ont fait faillite, mais celles qui proposaient réellement des produits de valeur ont tiré leur épingle du jeu. Je pense que nous allons observer le même phénomène en ce qui concerne l'IA générative, » prédit M. Vohs de Molson Coors. « L'IA générative va devenir omniprésente, à l'instar d'Internet. Sa valeur résidera dans sa capacité à résoudre des problèmes commerciaux. »

« L'IA générative va devenir omniprésente, à l'instar d'Internet. Sa valeur résidera dans sa capacité à résoudre des problèmes commerciaux. »

Darrin Vohs, Directeur monde des systèmes d'information, Molson Coors



La puissance à la portée de tous

À ce jour, ce sont les équipes informatiques qui utilisent le plus l'IA générative : selon notre enquête, 62 % d'entre elles l'ont entièrement déployée. Le marketing et les ventes, service client inclus, s'affichent comme des pionniers de l'adoption, devant les autres fonctions commerciales (Figure 1). Les entreprises interrogées pour cette étude font état de gains de productivité tangibles, notamment au sein des équipes techniques.

Flo Health a ainsi observé une augmentation de 24 % de l'efficacité du data engineering, les assistants IA réduisant les erreurs et accélérant la génération de documents. Repsol, une multinationale espagnole du secteur de l'énergie, a mené une expérience d'IA générative mobilisant 200 développeurs et plus de 200 000 lignes de code dans 20 langages différents. Les améliorations de productivité enregistrées, chiffrées à 7 % en moyenne, ont parfois atteint 30 %, selon la technologie employée. Et la qualité des résultats est élevée. Le leader technologique japonais Rakuten rapporte de son côté que de grandes quantités de code produites par son outil d'IA interne ont pu être acceptées sans modification.

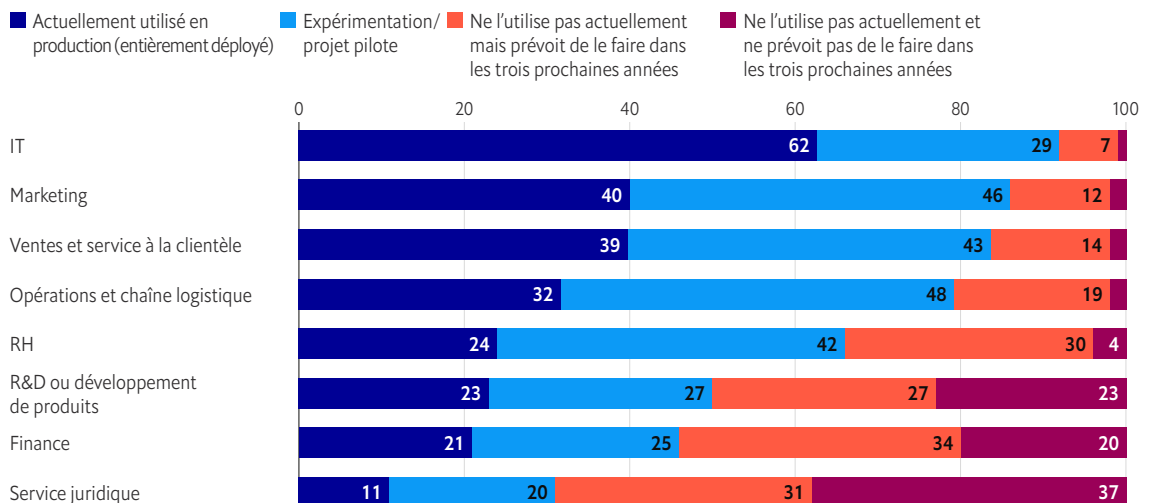
Si l'utilisation de l'IA générative peut sembler naturelle aux personnes occupant des rôles techniques, les développeurs notamment, l'idée d'un assistant au service de tous séduit de plus en plus. C'est le concept d'IA agentique : autrement dit, des agents artificiels dotés d'une interface en langage

naturel et capables de planifier et d'exécuter des séquences d'actions pour le compte d'un utilisateur. Ce concept est amené à occuper une place croissante dans la vie quotidienne et à dépasser les applications de synthèse vocale comme Alexa et Siri en termes de périmètre et d'intelligence. Ces agents pourraient réserver des vacances, apporter des conseils et stimuler la créativité, mais aussi négocier des litiges juridiques et lutter contre les pratiques commerciales douteuses.⁵ Dans notre enquête mondiale, plus de 97 % des architectes d'entreprise prédisent que le personnel non technique utilisera des outils en langage naturel pour interagir avec des ensembles de données complexes au cours des trois prochaines années. Ils sont près de six sur dix (58 %) à affirmer que le langage naturel sera le principal moyen d'y parvenir, sinon le seul.

La majorité des fonctions commerciales représentées dans le cadre de notre enquête ont entièrement déployé l'IA générative ou sont dans une phase d'expérimentation active (Figure 1). Les fonctions internes les plus hésitantes restent le service juridique et les finances : dans ces deux domaines, moins de la moitié des répondants parlent d'un déploiement total de l'IA ou d'expérimentation active. Cette prudence n'est pas nécessairement infondée, compte tenu des contraintes réglementaires et de l'ampleur des enjeux. Pourtant, de nombreux cabinets d'avocats tentent d'adopter des outils d'IA pour la rédaction et la révision de contrats, par exemple.

Figure 1 : Stade d'adoption de l'IA générative par fonction métier

Pourcentage de dirigeants



Source : Economist Impact

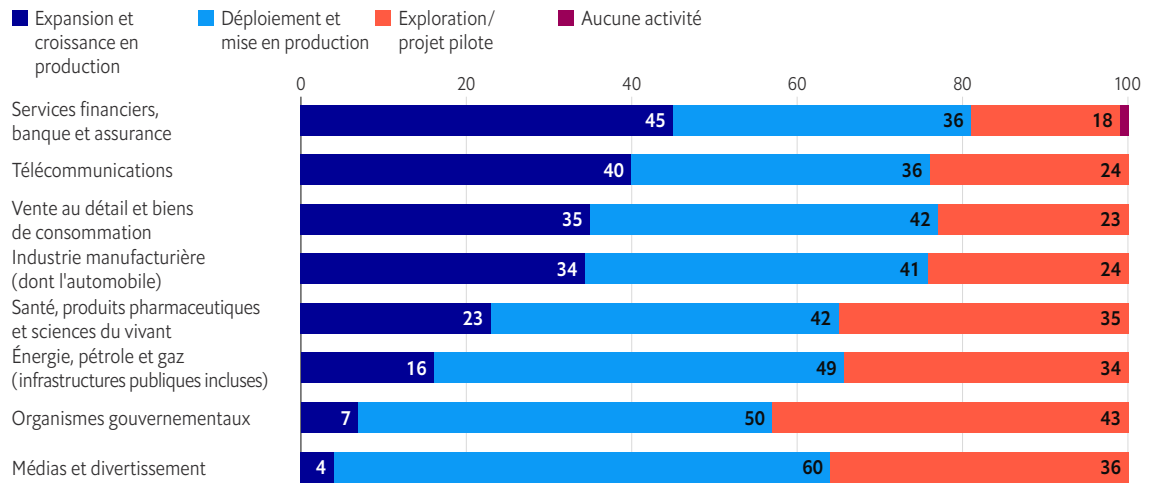
5 Par exemple : <https://donotpay.com/>

Si les équipes « finances » des entreprises affichent un certain retard, le secteur des services financiers devance, lui, tous les autres en ce qui concerne l'implémentation de l'IA générative. Parmi ces institutions, 45 % en sont déjà à étendre et amplifier leurs cas d'usage, et 36 % de plus en déploient activement (Figure 2). Les pionniers sont également nombreux dans les secteurs des télécommunications, de la vente au détail et de la

fabrication, ce qui laisse entrevoir un potentiel de croissance important avec le perfectionnement progressif des approches. En queue de peloton, le secteur public et l'industrie des médias et du divertissement n'ont pas autant développé leurs cas d'usage à grande échelle. Mais la majorité des représentants de ces secteurs commencent à déployer activement l'IA générative en production.

Figure 2 : Stade de mise en œuvre de l'IA générative par secteur

Pourcentage de dirigeants



Source : Economist Impact

Dans la plupart des pays, environ 70 % des organisations ont dépassé la phase pilote de l'IA générative (Figure 3). Les progrès sont toutefois beaucoup plus inégaux lorsqu'on aborde la phase de croissance.

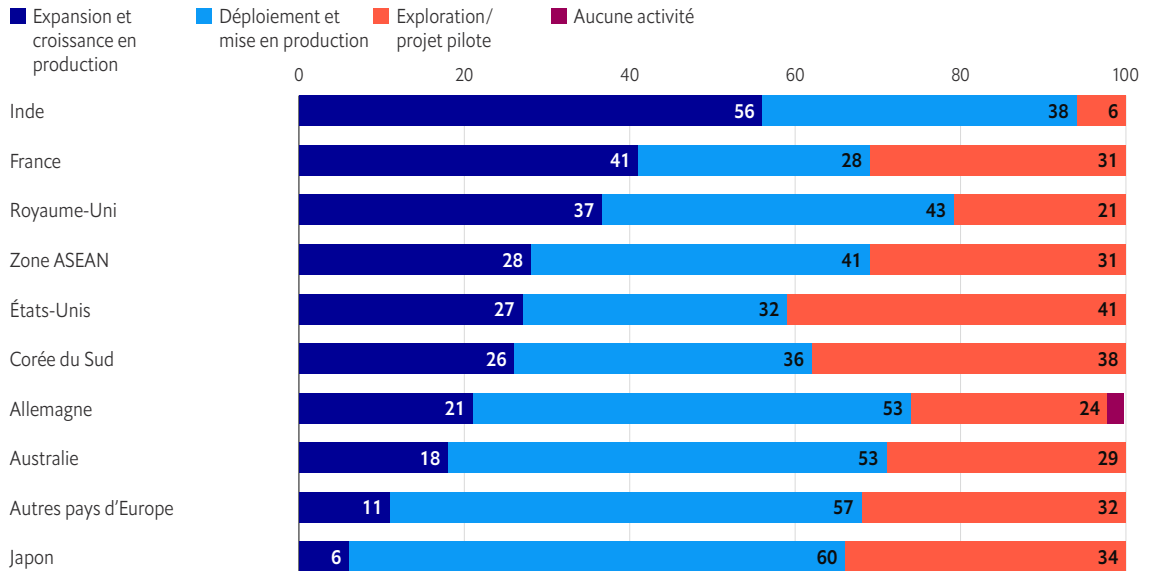
En Inde, par exemple, la majorité des entreprises (56 %) développent l'IA générative en s'appuyant sur des modèles publics et open source, bien plus que les organisations des autres pays.

Au Japon, en revanche, elles ne sont que 6 % à avoir atteint la phase d'expansion, ce qui peut s'expliquer par le fait que le pays privilégie des modèles plus complexes et personnalisés. La moitié des dirigeants japonais ont déclaré qu'ils entraînaient des modèles personnalisés sur des données propriétaires : c'est bien plus que dans tout autre pays.

Mais ce n'est pas forcément un signe de faiblesse pour les États-Unis : nos données suggèrent que bon nombre de ces expérimentateurs pourraient être en réalité des « leaders de l'IA générative », car ils sont aussi ceux qui affichent le plus de confiance quant au degré de préparation de leurs applications. C'est également le cas des entreprises européennes, dans une moindre mesure. Ces résultats mériteraient une enquête plus approfondie, mais il se pourrait que les grandes entreprises américaines soient plus enclines à prendre des risques et à explorer des cas d'utilisation innovants, au lieu d'implémenter des usages « testés et approuvés » le plus rapidement possible.

Figure 3 : Stade de mise en œuvre de l'IA générative par pays/région

Pourcentage de dirigeants



Source : Economist Impact

La région ASEAN comprend la Malaisie, les Philippines, Singapour et la Thaïlande

Les autres pays d'Europe comprennent le Danemark, la Finlande, l'Italie, les Pays-Bas, la Norvège, l'Espagne et la Suède.

Prendre plus rapidement la bonne décision

On comprend aisément que la productivité soit le premier sujet d'attention : on sait la mesurer et son impact est réel. Mais les dirigeants utilisent à la fois l'IA/ML classique et l'IA générative pour guider la prise de décision stratégique dans un environnement opérationnel instable. Notre enquête fait ressortir la gestion des risques comme un domaine stratégique de premier plan. Dans les entreprises de toutes tailles, l'IA dans son ensemble éclaire les décisions, selon 68 % des participants, mais aussi l'entrée sur le marché (66 %) ainsi que les ventes et partenariats stratégiques (66 %).

En revanche, les grandes entreprises utilisent plus souvent l'IA pour les décisions stratégiques entourant le développement de produits (67 %), l'embauche et la planification des effectifs (67 %), ainsi que les investissements et l'allocation de capital (61 %), que ne le font les entreprises de taille moyenne (55 %, 53 % et 48 % respectivement). Pour beaucoup de grandes entreprises, l'IA est incontournable quand il s'agit de mettre au jour des informations cachées. C'est d'ailleurs l'expérience

de Mahindra Group, un conglomérat multinational indien. « Nous possédons de nombreuses sociétés cotées, » explique Mohit Kapoor, Chief technology officer du groupe. « Nous veillons à ce que les données soient toujours identifiables pour pouvoir les isoler les unes des autres. Nous pouvons aussi les anonymiser pour obtenir des informations sur l'ensemble de nos entreprises. »

Pour des entreprises qui exercent dans un monde de plus en plus imprévisible, l'IA apporte un véritable atout défensif en renforçant la résilience et en éclairant la prise de décision. UPS, une société multinationale de transport, utilise l'IA générative pour aider les entreprises dans 200 pays et territoires sur des questions telles que la compréhension des codes douaniers pour les produits.⁶ « En intégrant la science de la décision dans les logiciels, nous éliminons une part de bruit et d'incohérence dans nos processus. C'est essentiel pour une entreprise internationale comme la nôtre, » explique Ken Finnerty, President of IT and data analytics chez UPS. « Nous garantissons la cohérence et nous améliorons la qualité. Nous prenons plus rapidement des décisions, mais surtout ce sont les bonnes décisions. »

6 Mastercard, « Mastercard accelerates card fraud detection with generative AI technology » (Mastercard accélère la détection des fraudes à la carte grâce à la technologie de l'IA générative), 2024, <https://www.mastercard.com/news/press/2024/may/mastercard-accelerates-card-fraud-detection-with-generative-ai-technology/>

« Nous analysons 1 000 milliards de points de données à l'échelle mondiale pour prédire si une transaction est authentique ou non. Nous améliorons constamment ces capacités de prédiction en ajoutant de nouvelles technologies. On ne peut jamais déterminer avec 100 % de certitude si une transaction est frauduleuse, mais notre objectif est de permettre aux banques de prendre une décision éclairée avec un maximum de confiance. »

Greg Ulrich, Chief AI and data officer, Mastercard

JetBlue, une compagnie aérienne américaine, a développé des jumeaux numériques. Ces répliques de systèmes physiques basées sur le ML servent à améliorer ses capacités prédictives dans un secteur très sensible aux perturbations inattendues, des événements météorologiques aux problèmes techniques. « Dans le transport aérien, le nombre d'événements inattendus pouvant survenir au cours d'une journée est presque illimité. Notre jumeau numérique est donc extrêmement précieux pour anticiper les problèmes et préparer des solutions, » déclare Carol Clements, Chief digital and technology officer de JetBlue. Cette pratique a amélioré l'efficacité des opérations de l'entreprise et réduit le coût des perturbations.

Mastercard utilise l'IA générative pour améliorer et accélérer la détection des fraudes par carte. « Nous analysons 1 000 milliards de points de données à l'échelle mondiale pour prédire si une transaction est authentique ou non. Nous améliorons constamment ces capacités de prédiction en ajoutant de nouvelles technologies, » explique Greg Ulrich, Chief AI and data officer de l'entreprise. « On ne peut jamais déterminer avec 100 % de certitude si une transaction est frauduleuse, mais notre objectif est de permettre aux banques de prendre une décision éclairée avec un maximum de confiance. »

General Motors (GM) utilise le ML pour analyser les données de diagnostic afin de prédire les problèmes avant qu'ils ne s'aggravent. Dès la

chaîne de production, son logiciel peut prévoir si une batterie de voiture risque de présenter une défaillance, par exemple. Cela permet de la retirer de la chaîne et de la réparer avant de l'installer. « C'est aussi un atout pour la fidélisation de la clientèle : nous évitons qu'un client vive une mauvaise expérience avec son premier véhicule électrique, un incident qui pourrait éroder sa confiance dans la marque, » explique Jon Francis, Chief data and analytics officer du constructeur automobile.

L'IA peut également aider les entreprises à élargir leur clientèle sans prendre de risques excessifs. Mahindra Group utilise l'IA pour analyser des données publiques et propriétaires afin d'accorder, ou non, des prêts, en s'appuyant sur des scores de crédit susceptibles de favoriser l'inclusion financière. Siam Commercial Bank, une banque basée en Thaïlande, a utilisé le ML pour réduire les coûts de souscription de crédit de 20 %, comme l'explique Chalee Asavathiratham, l'ancien directeur de la banque digitale. « Aujourd'hui, demander un prêt revient presque à obtenir un numéro de crédit précalculé. Le processus d'accord est beaucoup plus simple et efficace. »

Unilever intègre l'intelligence artificielle dans son processus de prise de décision. Son laboratoire Horizon3 se consacre à élaborer et expérimenter de nouveaux concepts et projets d'IA.⁷ Le programme Sky, en particulier, s'appuie sur l'analytique avancée pour optimiser son portefeuille de produits. Cela lui permet de prendre des décisions plus rapides et plus intelligentes sur ce qu'il faut mettre en rayon, comme l'explique M. Hill. « Je crois que nous avons beaucoup progressé dans nos efforts pour exploiter l'intelligence des données, et il nous reste encore beaucoup d'opportunités à saisir. »

« Je crois que nous avons beaucoup progressé dans nos efforts pour exploiter l'intelligence des données, et il nous reste encore beaucoup d'opportunités à saisir. »

Andy Hill, Chief data officer, Unilever

7 Unilever, « Unilever Launches Global AI Lab in Toronto » (Unilever lance un laboratoire mondial d'IA à Toronto), 2023, <https://www.unilever.ca/news/press-releases/2023/unilever-launches-global-ai-lab-in-toronto/>

Adapter l'approche au niveau de risque

Que ce soit pour gagner en productivité ou pour créer des prototypes, les entreprises choisissent soigneusement leurs projets pilotes. Sara Vaezy, Directrice de la stratégie et du numérique de Providence, une organisation de santé basée aux États-Unis, explique son approche consistant à être « très méthodique et réfléchi », au lieu de lancer des produits et prier pour que tout se passe bien. Elle détaille sa démarche : « Nous avons mis en place une structure de gouvernance et un environnement technique permettant de réaliser des expériences de façon très méthodique et réfléchie. Nous les instrumentons correctement pour en comprendre intimement les rouages. »

De nombreuses entreprises préfèrent élaborer des applications internes avant de proposer des produits aux clients. La multinationale financière canadienne TD Bank Group, par exemple, a déployé en interne un chatbot en IA générative pour le personnel de son centre d'appels. Les entreprises font preuve d'une prudence compréhensible lorsqu'il s'agit de mettre des agents de conversation à la disposition de leurs clients. Souvent rapides et efficaces, ils restent imprévisibles et peuvent produire des contenus offensants et des informations erronées, nuisant à la fois à la confiance des consommateurs et à la réputation des entreprises.

« Nous avons mis en place une structure de gouvernance et un environnement technique permettant de réaliser des expériences de façon très méthodique et réfléchie. Nous les instrumentons correctement pour en comprendre intimement les rouages. »

Sara Vaezy, Chief strategy and digital officer, Providence

Les entreprises qui explorent des applications externes de l'IA générative ne peuvent ignorer les perceptions des consommateurs si elles veulent éviter que ces outils pratiques ne deviennent hors de contrôle. Une IA ressentie comme envahissante ou manipulatrice va susciter chez les utilisateurs de la méfiance et des inquiétudes quant au respect de leur vie privée.⁸ C'est également le cas des recommandations erratiques : on a vu des IA suggérer aux utilisateurs de manger des cailloux ou de mettre de la colle sur leur pizza.⁹ Certains robots ont même été incités à faire la promotion de marques concurrentes.¹⁰ « Lorsqu'il s'agit d'informations susceptibles d'induire une action chez l'utilisateur, si celui-ci s'interroge sur un symptôme, par exemple, il faut indéniablement faire preuve d'une grande prudence, » déclare Roman Bugaev, Chief technology officer de Flo Health.

Le ministère américain de la Défense (DoD) établit une distinction entre les applications administratives et de terrain, les enjeux étant bien plus élevés lorsqu'il s'agit de guider des décisions militaires. « C'est une chose d'utiliser un LLM dans les bureaux pour vérifier que des documents sont bien conformes aux règlements du DoD. Mais l'aide à la décision sur le terrain, pour identifier un ennemi à abattre, par exemple, implique des enjeux plus importants, » explique Leonel Garciga, Directeur des systèmes d'information de l'armée américaine. La prise de décision restera du ressort du responsable concerné, et les humains interviendront toujours dans les processus impliquant l'IA.

Les expériences internes offrent un terrain d'expérimentation sécurisé. « Si vous introduisez des produits sans garantie de fiabilité, vous risquez d'éroder la confiance des clients, » affirme Wassym Bensaid, Directeur des logiciels chez Rivian, constructeur de véhicules électriques. « Le lancement d'un produit est le fruit d'une longue maturation du prototype de départ. Nous ne lançons des produits qu'après avoir fait toutes les vérifications nécessaires en termes de qualité et de fiabilité, afin que notre technologie soit un assistant de confiance pour nos clients. »

8 BBC, « Snapchat's My AI chatbot: The privacy concerns » (Robot conversationnel My AI de Snapchat : inquiétudes concernant le respect de la vie privée), 2023, <https://www.bbc.co.uk/news/technology-67027282>

9 BBC, « Glue pizza and eat rocks: Google AI search errors go viral » (Collez votre pizza et mangez des cailloux : les erreurs de l'IA de Google font le tour des réseaux), 2024, <https://www.bbc.co.uk/news/articles/cd11gzejgz4o>

10 Carscoops, « Chevy Dealers' ChatGPT Bots Recommend Teslas, BMWs, Fords, Toyotas And Rivians » (Les robots ChatGPT des concessionnaires de Chevrolet recommandent des modèles de Tesla, BMW, Ford, Toyota et Rivian), 2023, <https://www.carscoops.com/2023/12/chevy-dealers-ai-chatbots-are-recommending-teslas-bmws-fords-toyotas-and-rivians/>

Bancs d'essai

S'inspirant de l'industrie du logiciel, de nombreuses entreprises privilégient des cycles de développement et d'itération courts pour étudier les possibilités et les pièges de l'IA. « Il n'y a pas de bonne réponse, il y en aura toujours plusieurs. Commencez à petite échelle, procédez par itération, apprenez de vos erreurs, demandez des commentaires et continuez d'avancer, » conseille Melissa Pint, Chief digital information officer chez Frontier, une entreprise de télécommunications. Cette stratégie misant sur la prudence permet aux entreprises d'explorer le potentiel de la technologie sans prendre de risque inutile. « Moins de 20 % des projets pilotes d'IA aboutissent en production, » rappelle Gereurd Roberts, Directeur général du groupe Seven Digital, filiale de la société de médias australienne Seven West Media. « Il faut multiplier les expérimentations pour comprendre ce qui va produire de la valeur. »

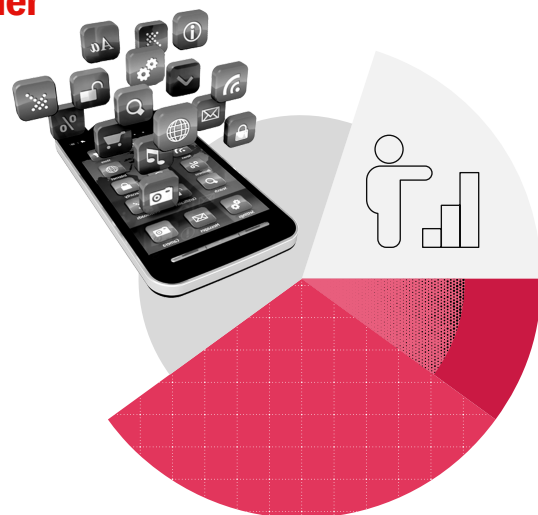
Comme le fait remarquer Kushal Chakrabarti, Chief data officer chez Opendoor, une société immobilière en ligne, « si vous créez rapidement un prototype de faible précision, vous pouvez l'utiliser pour répondre à certaines questions et réfléchir à son évolution, en améliorant sa précision et sa fiabilité au fil du temps. C'est une

version plus performante du développement à rebours. » Les experts sont unanimes sur un point : il ne faut pas déployer cette technologie par pur principe. « Beaucoup d'erreurs sont dues à des problématiques mal définies ou à une mauvaise compréhension de la technologie, en l'absence d'expertise, » explique M. Garcia. « J'ai un conseil fondamental : agissez de façon réfléchie, commencez à petite échelle, limitez le champ d'application du projet et définissez clairement le problème à résoudre. »

Les entreprises envisagent un large éventail d'applications de l'IA avant de ne retenir que les meilleurs cas d'usage. « Nous avons mis en place un processus de validation, une fonction centralisée qui [étudie les idées d'implémentation de l'IA générative dans toute l'entreprise.] Le nombre d'idées augmente sans cesse jour après jour, » explique M. Ulrich, de Mastercard. « Avec ce processus, nous voyons de nouvelles idées apparaître de deux façons. Tout d'abord, quand nous comprenons ce qui peut créer de la valeur, le pourcentage d'idées validées augmente. Ensuite, même si le pourcentage de celles qui sont réellement commercialisées reste faible, il ne faut pas en déduire que nous ne sommes pas actifs : rappelons que le nombre total d'idées au départ augmente sans cesse. »

« Moins de 20 % des projets pilotes d'IA aboutissent en production. Il faut multiplier les expérimentations pour comprendre ce qui va produire de la valeur. »

Gereurd Roberts, Directeur général du groupe, Seven Digital, Seven West Media



Senthil Ramani, Responsable mondial des données et de l'IA chez Accenture, un cabinet de conseil international, explique que son entreprise voit émerger une distinction entre les cas d'usage. Il y a les enjeux de base, qui font référence aux cas d'utilisation les plus simples : optimisation des opérations du centre de contact client, des fonctions marketing ou des workflows de programmation, par exemple. Les paris stratégiques désignent les cas d'usage plus innovants et expérimentaux.

Les environnements de sandbox offrent un mécanisme pratique d'expérimentation de l'IA : ces espaces protégés permettent non seulement de tester les aspects techniques d'un modèle ML pour détecter les hallucinations, les biais et les performances, mais également d'évaluer son impact commercial et sa validité. Scott Hallworth, Chief data and analytics officer chez HP, une entreprise technologique, est un fervent partisan de cette idée. Il a adopté l'approche sandbox dès le début de l'essor de l'IA générative : tous les collaborateurs de HP ont eu accès à un environnement dans lequel ils pouvaient explorer et tester des solutions susceptibles de faciliter la vie des salariés et des clients de l'entreprise. Aujourd'hui, HP exploite 75 sandboxes privées pour des usages divers. Elles visent toutes à explorer et tester les possibilités de l'IA pour l'entreprise au sein d'un environnement sûr, sécurisé et robuste.



Principales conclusions

Adoptez une approche expérimentale proportionnelle aux enjeux. Conscientes des risques et des inconnues, les entreprises se concentrent sur des cas d'usage internes dans des domaines à faible enjeu pour progresser sur la courbe de maturation. Elles ont tout intérêt à privilégier les projets de petite taille, à la portée limitée et ciblant un problème spécifique.

Misez sur les sandboxes et la supervision centralisée des projets pilotes pour coordonner les expérimentations. Laisser chaque idée prendre vie peut nuire à la cohérence et la gouvernance à long terme des projets d'IA. Les sandboxes, les environnements de test et les comités d'évaluation et de validation des expériences peuvent contribuer à mettre de l'ordre dans ce foisonnement.

Dans un monde instable, l'IA peut distinguer le signal du bruit. Les entreprises tributaires des chaînes d'approvisionnement internationales et sensibles aux fluctuations économiques et politiques s'appuient sur l'IA pour éclairer la prise de décision.

Chapitre 2 : Récolter les fruits

À l'heure où les dépenses consacrées à l'IA grimpent en flèche, les entreprises ont la tâche difficile de trouver le juste équilibre entre l'enthousiasme, une certaine discipline dans les stratégies d'investissement et des attentes réalistes en matière de rentabilité. Alors que les améliorations opérationnelles internes ont été l'objectif initial, les entreprises s'intéressent désormais à la création d'applications d'IA externes. Le but est de stimuler la croissance des revenus et de libérer de la valeur en puisant dans des sources de données jusque-là inexploitées. Mais il faut de la patience et de la nuance pour mesurer le succès de l'IA et avoir une vision claire des possibilités de création de valeur à long terme face aux gains immédiats et ce, dans un futur résolument axé sur l'IA.

Une démesure raisonnable ?

Au début de 2024, une note de recherche de Goldman Sachs avertissait que les 1 000 milliards de dollars de dépenses prévues dans l'IA générative risquaient de dépasser les perspectives de revenus.¹¹ Selon le scénario le plus pessimiste, l'IA générative serait incapable de résoudre les problèmes complexes justifiant de tels investissements et de tels coûts. Les dirigeants interrogés dans le cadre de ce rapport voient d'un bon œil cette remise en question de la proposition de valeur de l'IA générative qui fait, selon eux, partie d'un cycle naturel. Selon eux, la désillusion suit toujours le battage médiatique, puis de véritables preuves de valeur apparaissent au fil du temps.

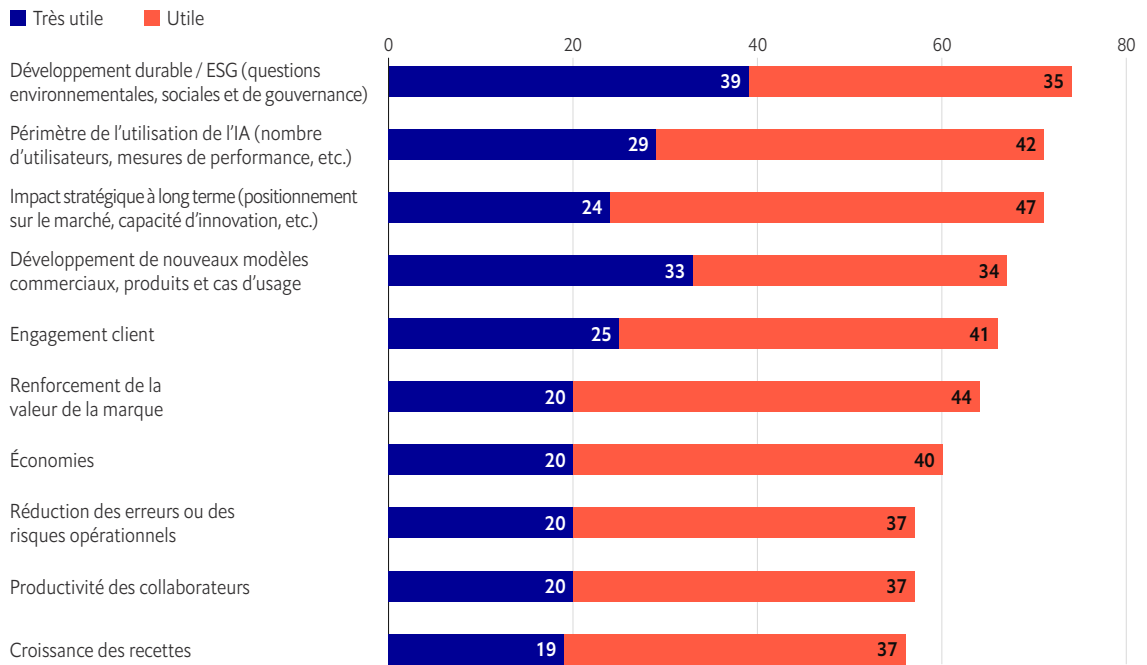
Notre enquête a révélé que la croissance des revenus a été jusqu'à présent un indicateur négligeable pour évaluer l'intérêt des investissements dans l'IA : 19 % des répondants seulement estiment qu'elle a été un facteur important dans la décision d'investissement (Figure 4).



11 Goldman Sachs, « Gen AI: too much spend, too little benefit? » (IA générative : trop de dépenses, pas assez de bénéfices ?), 2024, <https://www.goldmansachs.com/insights/top-of-mind/gen-ai-too-much-spend-too-little-benefit>

Figure 4 : La croissance des revenus n'est pas encore le principal moteur des décisions d'investissement en IA

Dirigeants indiquant que les indicateurs suivants ont aidé leur entreprise à défendre l'intérêt d'un investissement dans l'IA
Pourcentage de dirigeants



Source : Economist Impact

Les dirigeants privilégient traditionnellement la vision à long terme et craignent les conséquences d'un investissement insuffisant. Lors d'une conférence avec des analystes en août, Sundar Pichai, PDG d'Alphabet, a déclaré que les risques liés à un sous-investissement dans l'IA étaient « considérablement plus importants » que ceux d'un surinvestissement. Microsoft, Alphabet, Amazon et Meta devraient engager plus de 200 milliards de dollars de dépenses d'investissement en 2024, dont la majeure partie servira à construire des infrastructures d'IA.¹²

Les investisseurs prévoient toutefois que le temps des preuves succèdera bientôt à celui des promesses. Le fossé qui sépare actuellement les investissements et la génération de revenus sera alors envisagé sous un angle bien plus critique.¹³ Et il faudra sans doute non pas des trimestres, mais des années, pour que ces preuves arrivent dans de nombreuses entreprises pionnières. Mais les fournisseurs et les entreprises qui proposent des

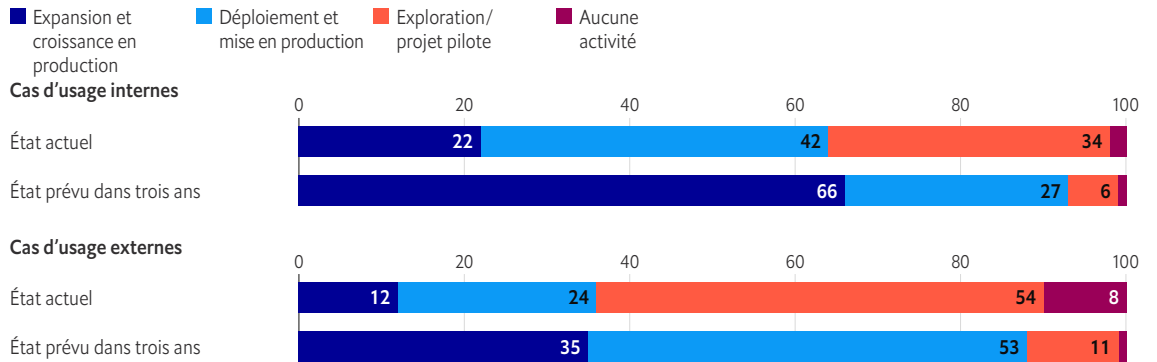
offres d'IA générative concrètes affichent déjà des gains de revenus, tout comme les industries qui fournissent l'infrastructure sous-jacente : cloud computing, puces et datacenters. Le marché du cloud, par exemple, pourrait atteindre un volume de 2 000 milliards de dollars grâce aux workflows de l'IA, selon une prévision.¹⁴ Parallèlement à cela, notre enquête révèle que six architectes d'entreprise sur dix misent de plus en plus sur les services cloud pour l'IA, avec une évolution vers le multi-cloud pour un peu plus d'un tiers d'entre eux.

Notre enquête met également en évidence une évolution des plans de déploiement des entreprises sur les trois prochaines années : après l'amélioration des opérations internes, ils s'orientent désormais vers les cas d'usage externes (Figure 5). Ce virage pourrait accroître le potentiel de création de revenus de l'IA, plus que la simple optimisation des activités quotidiennes.

12 Channel News Asia, « Nvidia fails to impress growth-hungry investors, shares fall » (Nvidia ne parvient pas à impressionner les investisseurs avides de croissance, l'action chute), 2024, <https://www.channelnewsasia.com/business/nvidia-shares-fall-growth-hungry-investors-4573336>
 13 Investors' Business Daily, « AI Stocks: Tech Giants, Cloud Titans Face 'Show Me' Moment. Apple Unveils iPhone 16 » (Cours de l'IA : les géants de la tech et du cloud sont désormais tenus de présenter des preuves. Apple dévoile l'iPhone 16), 2024, <https://www.investors.com/news/technology/artificial-intelligence-stocks/>
 14 Goldman Sachs, « Cloud revenues poised to reach \$2 trillion by 2030 amid AI rollout » (Les revenus du cloud devraient atteindre 2 000 milliards de dollars d'ici 2030 grâce au déploiement de l'IA), 2024, <https://www.goldmansachs.com/insights/articles/cloud-revenues-poised-to-reach-2-trillion-by-2030-amid-ai-rollout>

Figure 5 : Cas d'usage internes et externes de l'IA générative, aujourd'hui et en 2027

Pourcentage de dirigeants



Source : Economist Impact

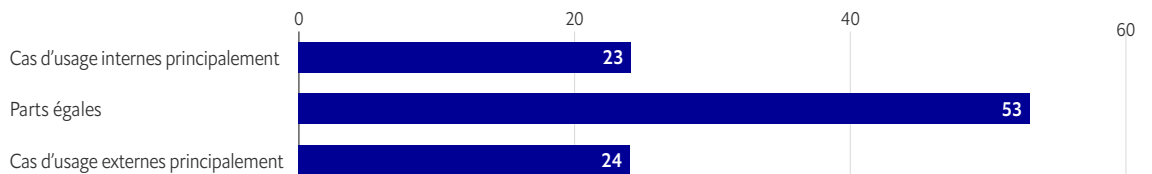
En effet, de nombreux participants à l'enquête (37 %) entrevoient plus de bénéfices dans les cas d'usage externes que les applications internes. C'est d'autant plus remarquable que seulement 24 % des organisations affectent la majorité de leurs investissements à des cas d'utilisation externes (Figure 6). On peut en déduire que certaines organisations, probablement celles qui ont le plus étendu et développé leurs projets d'IA, commencent à rentabiliser massivement leurs applications d'IA générative externes. Lorsque les 88 % restants atteindront ce stade dans les prochaines années, les cas d'usage externes pourraient devenir la principale source de ROI.

M. Ramani d'Accenture souligne que la productivité, facile à mesurer, est un motif raisonnable d'adoption initiale. Mais les organisations appelées à devenir les leaders de l'IA de demain seront celles qui dépasseront ce stade et comprendront que l'IA doit servir à « la croissance des revenus, à la création de nouvelles opportunités, ainsi qu'à l'attraction et la rétention des talents ».

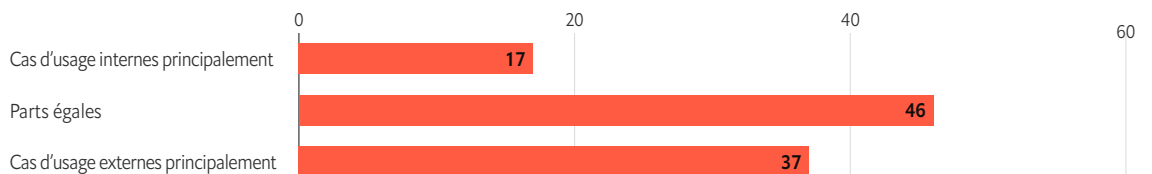
Figure 6 : Investissements et bénéfices

Pourcentage de data scientists

Cas d'usage de l'IA générative recevant le plus d'investissements



Cas d'usage de l'IA générative présentant le plus d'avantages



Source : Economist Impact

Intelligence des données

La croissance et les revenus seront au rendez-vous pour les entreprises qui sauront extraire le maximum de valeur de leurs données. Les données non structurées, notamment, ont toujours été difficiles à traiter, à gérer et à utiliser, mais elles recèlent une valeur considérable. Pensez par exemple, dans le contexte médical, à la possibilité de combiner des images de diagnostics, des évaluations textuelles et des valeurs numériques provenant d'examens et de mesures cliniques. Les compagnies d'assurance peuvent utiliser les données des satellites, du GPS ou des objets connectés pour enrichir la modélisation des risques et donc la tarification de tout type de dommage, des canicules aux accidents cardiaques.

C'est justement la quantité d'informations que les modèles d'IA actuels peuvent récolter et exploiter qui les rend aussi attrayants. Les entreprises évoquent en effet un cercle vertueux : les insights extraits en permanence des données grâce à l'IA leur permettent d'améliorer continuellement leurs produits et services. Le constructeur Rivian intègre sa pile technologique dans des véhicules électriques fabriqués par Volkswagen. Il obtient ainsi l'accès à un volume de données beaucoup plus important et à des informations plus détaillées sur les modèles d'utilisation, comme l'explique M. Bensaid, Chief software officer de l'entreprise. « Nous allons mettre notre technologie au service d'un éventail bien plus large de véhicules et l'adapter à différentes marques et architectures VW, grâce à la modularité et la flexibilité intrinsèques de notre architecture logicielle, » ajoute-t-il. Ce volume de données contribuera à accélérer l'électrification de millions de véhicules.

« L'avenir est à une compréhension complète des clients. Les applications, les produits et les offres commerciales qu'ils souhaitent voir apparaître sont de plus en plus personnalisés. »

Sanjay Bhakta, Directeur des produits et des technologies, Condé Nast

Ces approches data-driven favorisent une tendance à la personnalisation des produits dans tous les grands secteurs d'activité, afin de mieux répondre aux besoins et aux comportements des clients. « L'avenir est à une compréhension complète des clients. Les applications, les produits et les offres commerciales qu'ils souhaitent voir apparaître sont de plus en plus personnalisés, » explique Sanjay Bhakta, Chief product and technology officer de Condé Nast, acteur majeur du secteur des médias.

Le secteur public pourrait être le gagnant inattendu de cette évolution en raison de ses vastes réserves de données inexploitées. Selon les estimations du cabinet de conseil McKinsey, il serait possible de générer jusqu'à 3 000 milliards de dollars de valeur économique à l'échelle mondiale en améliorant l'utilisation des données ouvertes.¹⁵ Lorsque Transport for London a partagé des données sur la fourniture de ses services, plus de 600 applications sont apparues dans des domaines tels que la planification des itinéraires et la logistique. Ces applications ont permis de réaliser jusqu'à 95 millions de livres sterling d'économies.¹⁶ De son côté, le ministère américain de la Défense a su exploiter ses vastes données propriétaires pour traiter plus de 10 millions d'enregistrements de paiement. Cela lui a permis d'identifier 12,7 milliards de dollars de paiements abusifs depuis 2020 et de prendre les mesures appropriées.¹⁷



15 McKinsey, « How government can promote open data and help unleash over \$3 trillion in economic value » (Comment le gouvernement peut encourager les données ouvertes et contribuer à libérer plus de 3 000 milliards de dollars de valeur économique), [https://www.mckinsey.com/~media/mckinsey/industries/public%20and%20social%20sector/our%20insights/how%20government%20can%20promote%20open%20data/how_govt_can_promote_open_data_and_help_unleash_over_\\$3_trillion_in_economic_value.pdf](https://www.mckinsey.com/~media/mckinsey/industries/public%20and%20social%20sector/our%20insights/how%20government%20can%20promote%20open%20data/how_govt_can_promote_open_data_and_help_unleash_over_$3_trillion_in_economic_value.pdf)

16 Deloitte, « Assessing the value of TfL's open data and digital partnerships » (Évaluation de la valeur des données ouvertes et des partenariats numériques de TfL), 2017, <https://content.tfl.gov.uk/deloitte-report-tfl-open-data.pdf>

17 CDAO, « Advana Industry Day » (Journée de l'industrie Advana), <https://storage.tradewindai.com/pdfs/Advana-Industry-Day.pdf>

« Une partie du processus consiste à trouver des occasions de tester ces produits, afin d'obtenir rapidement des informations avant de faire un investissement plus important. Ce cycle d'expérimentation doit occuper une grande partie du parcours. »

Jon Francis, Chief data and analytics officer, General Motors

Mesurer les gains concrets de l'IA

Les organisations doivent développer des cadres pour évaluer la rentabilité de leurs initiatives et justifier les dépenses. Notre enquête a révélé que les plus grandes entreprises s'appuient essentiellement sur des indicateurs clés de performance (KPI) et sur des évaluations post-impact ou post-déploiement. Les petites entreprises misent également sur les KPI, notamment sur le chiffre d'affaires par employé, la fidélisation de la clientèle et les coûts d'exploitation.

M. Francis (General Motors) explique que son entreprise adopte une approche très disciplinée de l'IA et du ML. Ces derniers sont exploités uniquement dans des domaines où ils peuvent générer des revenus supplémentaires ou réduire les coûts. « Il n'est pas possible d'adopter une technologie juste pour l'adopter. Il faut qu'elle ait de la valeur et un impact pour l'organisation. » Le constructeur automobile basé à Détroit collecte des données sur des projets pilotes à petite échelle ainsi que des preuves de concept. « Une partie du processus consiste à trouver des occasions de tester ces produits, afin d'obtenir rapidement des informations avant de faire un investissement plus important. « Ce cycle d'expérimentation doit occuper une grande partie du parcours, » ajoute M. Francis. Mais l'entreprise reste focalisée sur les résultats commerciaux. « Nous devrions être en mesure d'établir le rendement de chaque dollar, chaque centime de calcul. »

Mahindra Group s'intéresse à des défis qui vont au-delà du simple gain de productivité. Comme l'affirme M. Kapoor de Mahindra, l'organisation « se focalise sur les problèmes difficiles à résoudre », en particulier dans les ateliers, pour réduire le temps moyen de réparation, accroître le laps de temps moyen entre deux pannes ou améliorer l'expérience client. Dans ce troisième domaine, l'entreprise mise sur des innovations telles que l'assistance par IA générative pour accélérer l'entretien des véhicules. Condé Nast évalue le retour sur investissement à l'aide d'une liste de résultats incluant la réduction des coûts, l'augmentation de la vitesse et de l'efficacité, l'amélioration de l'expérience du public ou l'augmentation du temps passé sur le site. Unilever consacre ses investissements en IA à la productivité, à la créativité et à la croissance. Seven Digital a choisi une approche similaire. « Nous adoptons une approche pragmatique en utilisant l'IA pour améliorer notre modèle économique – l'efficacité, les résultats et le perfectionnement des collaborateurs – dans une optique de transformation organisationnelle, » explique M. Roberts, Directeur général du groupe.

Chalee Asavathiratham, ancien Chief digital banking officer de la Siam Commercial Bank, affirme qu'il était impératif que les équipes d'IA travaillent avec la finance pour mesurer les résultats et le retour sur investissement de leurs projets sans se laisser séduire par l'effet de nouveauté. Il lance un avertissement : selon lui, les équipes d'IA pourraient « négliger de s'interroger sur les résultats et la rentabilité » d'un projet d'IA. « [Les équipes d'IA] se laissent parfois séduire [par] le côté tendance des projets, » explique-t-il.

Mme Pint, de Frontier, présente un argument similaire. « Il n'y a pas de stratégie technologique. Il ne peut y avoir qu'une stratégie commerciale, soutenue par la technologie. Quoi que vous fassiez, commencez par comprendre parfaitement la stratégie commerciale avant d'y relier votre stratégie technologique. »

« Il n'y a pas de stratégie technologique. Il ne peut y avoir qu'une stratégie commerciale, soutenue par la technologie. Quoi que vous fassiez, commencez par comprendre parfaitement la stratégie commerciale avant d'y relier votre stratégie technologique. »

Melissa Pint, Chief digital information officer, Frontier

Chez Opendoor, les dirigeants en sont venus à la conclusion que la productivité ne devait pas être l'indicateur principal de rentabilité, mais qu'ils devaient se concentrer sur la croissance du chiffre d'affaires et l'amélioration de l'expérience client. Certes, l'utilisation de l'IA génère des gains de productivité dans l'entreprise, mais ils n'occupent pas une place centrale dans son approche. « Notre priorité est de faire un meilleur usage de l'intuition humaine et de gagner en cohérence pour améliorer l'expérience client, » affirme M. Chakrabarti d'Opendoor. « C'est mieux pour les collaborateurs, les clients et l'entreprise dans son ensemble. »

En mettant en place un socle de données solide, les entreprises se donnent les moyens de poursuivre des objectifs d'innovation plus ambitieux en matière de création de revenus et de modèles commerciaux. « Nous avons commencé à utiliser l'IA en créant un socle de données robuste. Puis nous avons exploré une application commerciale axée sur la tarification variable de certains produits complémentaires, » se souvient Mme Clements de JetBlue. « Nous avons d'abord utilisé le modèle sur un ensemble très restreint de marchés. Nous avons été immédiatement frappés par la valeur extraordinaire qu'il a produite dès le départ. Ce cas d'usage a été le catalyseur nous permettant de comprendre la puissance de l'IA et son potentiel pour la transformation de notre activité. Il a stimulé la créativité des équipes de toute l'organisation, qui se sont mises à chercher d'autres moyens d'exploiter la technologie afin de générer encore plus de valeur. »

Autre objectif tout à fait valable, le contrôle des coûts. Dans les entreprises dont les dépenses d'exploitation sont importantes, une optimisation de quelques points de pourcentage peut se traduire par des millions de dollars d'économies, comme l'ont observé certains des dirigeants que nous avons interrogés. Ces économies opérationnelles peuvent être réinvesties dans l'innovation.





« Il est très difficile d'institutionnaliser l'IA et l'analyse des données auprès des collaborateurs. Ce n'est pas quelque chose que l'on peut mesurer à court terme, il faut réfléchir sur une période de trois ans. »

Helen Choi, Chief digital and information officer, CJ CheilJedang

La vertu de la patience

Pour déterminer le retour sur investissement, il faut fixer des objectifs et établir des délais réalistes. Il faut pour cela s'appuyer sur une compréhension précise des capacités de l'IA, des défis de mise en œuvre et du degré de préparation de l'organisation. « Quand il s'agit de transformation numérique et d'IA, de nombreux dirigeants s'attendent à voir des performances et des résultats à court terme, » explique Helen Choi, Chief digital and information officer de CJ CheilJedang, une entreprise sud-coréenne de biens de consommation. « Mais il est très difficile d'institutionnaliser l'IA et l'analyse des données auprès des collaborateurs. Ce n'est pas quelque chose que l'on peut mesurer à court terme, il faut réfléchir sur une période de trois ans. Parce que ce n'est pas seulement une question de données ou de mise en œuvre technologique, » précise Mme Choi. « Pour tirer une réelle valeur commerciale de la technologie numérique, il faut un long parcours d'innovation qui englobe les méthodes de travail et la gestion du changement à l'échelle de l'organisation. »

« L'un des principaux défis réside dans le fait que les modèles de monétisation et d'affaires devant permettre de créer de la valeur avec l'IA générative ne sont pas encore bien établis, » explique Takaaki Sato, senior executive vice president and chief technology officer de NTT Docomo, société japonaise de télécommunications. « Il est actuellement difficile de savoir s'il faut investir dans des LLM généralistes ou dans des LLM spécialisés de plus petite taille. Le retour sur investissement n'est pas encore clair. »

Même si les entreprises peuvent se prémunir de dépenses excessives en accordant une attention extrême aux retours, il peut être intéressant pour elles de faire un pas en arrière avant de faire deux pas en avant. Dream Sports en donne un bon exemple : dans cette société indienne de sports fantasy, certains des algorithmes actuellement utilisés ont d'abord produit des indicateurs commerciaux négatifs. « Naturellement, les équipes des produits et de la finance s'en sont inquiétées. [Mais nous avons] su persévérer, persister et les éduquer, » se souvient Amit Sharma, chief technology officer de l'entreprise. Dans des secteurs

réglementés comme l'industrie pharmaceutique et la biotechnologie, il peut être difficile d'obtenir des résultats rapides. Il faut en effet des années pour qu'une innovation pharmaceutique passe du laboratoire au patient. « L'IA accélère le processus d'identification des molécules synthétisables et le passage aux essais de phase deux ou trois, » observe M. Bucher de Novartis. « Mais il faudra peut-être cinq à dix ans pour déterminer l'efficacité de l'IA, une fois la sécurité du produit confirmée. »

Le rythme soutenu des progrès de l'IA et de l'innovation incite également à faire preuve d'une certaine patience stratégique. Certains experts font une analogie avec le temps qu'il a fallu pour que Google se démarque des nombreux autres moteurs de recherche dans les années 1990 – soit six ans après Archie, le premier produit.¹⁸

Jeff Martin, senior vice president and chief data officer de TD Bank Group, tient à mettre en garde contre « un excès de confiance dans les perspectives d'avenir ». Comme il l'ajoute, les entreprises doivent partir du principe qu'elles seront obligées de s'adapter chaque fois qu'elles prendront une décision, « donc assurez-vous d'être prêts, flexibles et agiles ».

« L'IA accélère le processus d'identification des molécules synthétisables et le passage aux essais de phase deux ou trois. « Mais il faudra peut-être cinq à dix ans pour déterminer l'efficacité de l'IA, une fois la sécurité du produit confirmée. »

Bernd Bucher, Directeur des systèmes d'information, Novartis

Principales conclusions

Les gains de productivité et d'efficacité représentent des victoires rapides et faciles à chiffrer. Mais les vrais leaders poursuivront des résultats plus ambitieux. La génération de revenus, l'attraction des talents et le développement de nouveaux produits et services figurent parmi ces objectifs de plus longue portée. Certaines entreprises créeront également de la valeur en utilisant l'IA pour extraire leurs données propriétaires et en faire de nouveaux usages, notamment pour développer des produits plus personnalisés.

Les organisations doivent parvenir à un équilibre délicat entre le retour sur investissement et le délai nécessaire à l'obtention d'une véritable valeur. Les entreprises devront bientôt démontrer la rentabilité financière de leurs investissements en IA, mais le temps imparti devra être correctement équilibré. L'expérimentation prend du temps et on ne peut empêcher certains projets d'échouer. Les entreprises ont besoin d'un point de départ pour orienter la modernisation de leur infrastructure de données. Les contraintes réglementaires limitent également la vitesse de déploiement dans des secteurs comme celui de la santé. Et il est tout à fait judicieux de prendre le temps d'observer l'évolution du paysage des fournisseurs et des produits d'IA.

18 SEO Mechanic, « The Complete History of Search Engines » (Une histoire complète des moteurs de recherche), 2023, <https://www.seomechanic.com/complete-history-search-engines/>

Chapitre 3 :

La refonte des infrastructures

Dans leurs efforts pour exploiter le potentiel de l'IA, les entreprises sont confrontées à un défi stratégique : leur infrastructure de données est extrêmement mal préparée, comme un système d'éclairage au gaz dans un grand immeuble moderne. Les silos, la latence et les enjeux de sécurité sont autant de freins au déploiement de l'IA. Que ce soit pour accéder à des données précieuses, mais non structurées ou pour trouver leur chemin dans la complexité des solutions cloud, les organisations doivent repenser leur architecture de données.

La puissance des modèles n'est pas le seul grand tournant de l'ère de l'IA : la structure sous-jacente, nécessaire pour les déployer à grande échelle, doit elle aussi se transformer. Dans bon nombre de grandes entreprises dans le monde, l'infrastructure est comparable à de la plomberie de l'époque haussmannienne. Obsolète et incapable de répondre aux exigences de l'IA, elle montre des signes de fuite.

Parmi les organisations interrogées dans le cadre de notre enquête, 22 % seulement pensent leur architecture actuelle pleinement capable de prendre en charge les exigences uniques des charges de travail de l'IA. Et elles ne sont pas plus de 23 % à déclarer que leur architecture actuelle intègre pleinement les applications d'IA aux données commerciales utiles. Même parmi les plus grandes entreprises que nous avons interrogées, ces taux ne dépassent pas 28 % et 27 % respectivement. Ces retards entravent les progrès, et il est probable qu'ils entraînent des problèmes de qualité. Les dirigeants sont ainsi 37 % seulement – et les praticiens encore moins à 29 %

– à estimer que les applications d'IA générative de leur organisation sont prêtes pour la production.

Les exigences des charges de travail d'IA et d'analytique avancée étaient déjà source de problèmes d'infrastructure avant l'avènement de l'IA générative. Cela était dû en grande partie à la diversité des données générées par les entreprises et à la variété des mécanismes de stockage élaborés pour les gérer.

De nombreuses entreprises utilisaient des data warehouses dans le cloud pour conserver des données fortement structurées, bien classées et étiquetées, faciles à interroger et gérées par des équipes informatiques expertes. Mais il faut savoir que les données non structurées – incompatibles avec les data warehouses – peuvent représenter jusqu'à 90 % des données collectées par les organisations.¹⁹ Avec les progrès du ML et de l'IA, les données non structurées ont acquis davantage de valeur, les systèmes ML étant capables de repérer des modèles dans de vastes quantités de données jusque-là négligées.

Parmi les organisations interrogées dans le cadre de notre enquête, 22 % seulement pensent leur architecture actuelle pleinement capable de prendre en charge les exigences uniques des charges de travail de l'IA.

19 MIT Sloan, « Tapping the power of unstructured data » (Exploiter la puissance des données non structurées), 2021, <https://mitsloan.mit.edu/ideas-made-to-matter/tapping-power-unstructured-data>

Les data lakes, capables de stocker de grandes quantités de données structurées et non structurées, ont émergé comme alternative permettant aux organisations d'exploiter des données non structurées. Mais l'adoption de ces nouvelles solutions a fragmenté les données entre data lakes et data warehouses. Les conséquences se sont rapidement fait sentir : formation de silos, prise en charge inégale des cas d'usage et incompatibilité des modèles de sécurité. Pour les entreprises, il est devenu beaucoup plus difficile d'utiliser efficacement leurs données. Pour résoudre ces problèmes, un nouveau modèle, baptisé data lakehouse, a fini par apparaître. Reprenant certains aspects des data lakes et des data warehouses, il repose sur un logiciel open source et utilise des normes ouvertes.

Parmi les contraintes d'architecture à résoudre, les organisations envisagent de traiter la vitesse en priorité. Le traitement des données en temps réel est en effet désigné comme la lacune principale par 47 % des participants, suivi par la mise en place de pipelines de données robustes et la sécurité (Figure 7). Les data engineers confirment le point de vue des architectes : ils désignent eux aussi les problèmes de streaming de données comme le principal défi pour assurer la qualité des données.

Le streaming en temps réel des données de production est crucial pour les applications d'IA

qui doivent pouvoir les interroger immédiatement. Mahindra Group, par exemple, intègre des données en temps réel sur les performances des panneaux solaires. Un algorithme oriente les équipes de maintenance et de nettoyage des installations d'énergie renouvelable du groupe afin d'optimiser l'efficacité de leurs interventions. De la même façon, Dream Sports utilise l'IA pour déterminer le moment idéal pour organiser des concours lucratifs autour d'événements sportifs. Pour cela, l'entreprise doit connaître avec précision le nombre d'utilisateurs qui forment une équipe fantasy pour un événement spécifique à un instant T.

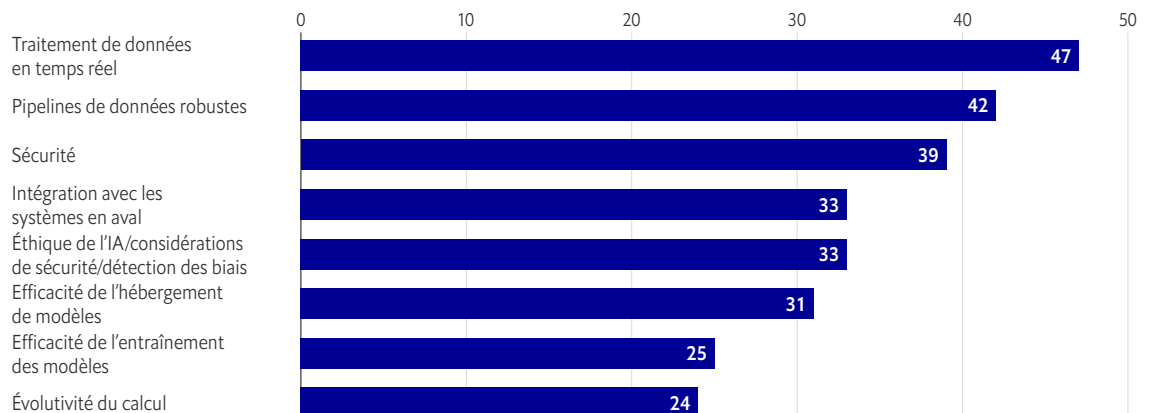
Aucune économie d'échelle à l'horizon

Trouver la bonne infrastructure peut devenir plus difficile au fil de la croissance des organisations. Les entreprises diversifiées, composées de multiples unités commerciales, de même que les conglomérats rassemblant des sociétés entièrement indépendantes, doivent déterminer où vont résider les données. En centralisant leurs données dans des référentiels globaux, elles peuvent obtenir des informations sur l'ensemble du groupe. Mais en choisissant de gérer les données avec des unités individuelles, les équipes d'analytique et de data science auront davantage de flexibilité pour les mettre au service de leurs propres besoins.

Figure 7 : Limites de l'architecture

Principales lacunes identifiées par les participants dans leur architecture actuelle, à combler dans les 3 à 5 prochaines années

Pourcentage d'architectes d'entreprise



Source : Economist Impact

Au sein du Mahindra Group, les fonctions principales telles que le cloud et la cybersécurité sont gérées en commun, mais chaque unité commerciale prend ses propres décisions d'architecture. Le groupe adopte une approche axée sur les problèmes avant les solutions, mettant l'accent sur la collaboration entre les équipes commerciales et technologiques. Son entreprise de construction automobile utilise un data mesh pour stocker les informations utiles aux conducteurs – l'historique des trajets, notamment – séparément des données de fabrication. Elle a tout de même la possibilité de combiner les deux datasets si nécessaire. Cette approche réduit la latence des requêtes des conducteurs en limitant la recherche des informations pertinentes. Elle permet aussi d'identifier les problèmes de la chaîne d'approvisionnement. Par exemple, si plusieurs conducteurs signalent un même problème, l'entreprise peut retracer l'historique de fabrication du composant concerné dans chaque voiture. L'équipe financière de Mahindra, quant à elle, utilise une architecture plus flexible qui lui permet de prendre en compte simultanément les prêts personnels, les dépôts à terme, les assurances, les hypothèques et les prêts aux petites et moyennes entreprises pour évaluer les risques.



On aurait tort de penser que les start-ups et les entreprises digital natives échappent aux problèmes d'infrastructure. Dream Sports, par exemple, interrogeait encore des bases de données de production pour ses besoins de reporting plusieurs années après sa création et n'a mis en place un magasin de données qu'après avoir trouvé une adéquation produit-marché. Ces entreprises se distinguent toutefois par une détermination remarquable à refondre entièrement leur architecture de données s'il le faut, sans doute parce que leur dette technique reste limitée.

Rivian, un constructeur de véhicules électriques, a intégré les informations de l'ensemble de ses systèmes complexes – chaîne d'approvisionnement, opérations commerciales et services financiers – dans un environnement unifié, permettant à tous les niveaux de l'organisation de prendre des décisions basées sur les données. Les bénéficiaires sont sensibles non seulement pour les applications d'IA et d'IA générative, mais aussi dans tout l'éventail des processus commerciaux liés aux données, des tableaux de bord à l'analytique, en passant par le profil et l'historique des clients.

L'IA générative peut exercer une pression sur l'infrastructure pour certains cas, en raison des exigences de nouveaux cas d'usage qui reposent souvent sur un outillage spécifique. C'est notamment le cas de la génération augmentée par récupération, ou RAG. Ce processus consiste, pour le LLM, à vérifier la réponse qu'il a générée en s'appuyant sur une source de connaissances externe, puis à l'affiner. Lorsqu'elles développent leurs propres LLM, les organisations utilisent des données internes telles que les interactions avec le support client, comme source d'informations de référence. Mais pour être exploitables par le ML, ces données internes doivent être nettoyées et préparées, ce qui nécessite des outils et des technologies spécifiques.

Dans leur empressement à adopter l'IA générative, de nombreuses entreprises dupliquent les données ou les systèmes pour réaliser les opérations de nettoyage et de préparation, en ajoutant un pipeline de données supplémentaire à l'infrastructure de l'entreprise. Elles ajoutent ainsi une couche de complexité à une infrastructure de données déjà compliquée.



« Le développement se fait presque en vase clos. Et chaque application développe sa propre infrastructure et ses propres outils pour ses opérations de grand modèle de langage ou de machine learning, » explique Mme Vaezy de Providence.

En effet, dans un monde rempli d'IA générative, les infrastructures de données risquent de se recouper et de devenir de plus en plus difficiles à manipuler et à suivre. Lorsqu'on leur a demandé ce qui améliorerait le plus leur productivité, les data engineers ont évoqué en priorité « la simplification des connexions aux sources de données pour l'ingestion », « l'utilisation d'une solution unifiée à la place d'une palette d'outils » et « une meilleure visibilité sur les pipelines de données pour localiser et résoudre les problèmes ».

Sans un socle de données solide, la mise en œuvre prématurée des applications traditionnelles de l'IA pourrait mettre en péril les opérations des entreprises et les exposer à des risques de sécurité et de réputation. « En mars 2023, quand l'engouement [pour l'IA générative] était à son comble, je savais que nous allions nous retrouver face à un problème : tout le monde allait se lancer avec enthousiasme et potentiellement exposer l'entreprise à des risques indésirables, » explique M. Hallworth de HP.

La bonne méthode pour les bonnes données

De nombreuses organisations misent massivement sur les services cloud pour prendre en charge les charges de travail de l'IA, comme en témoignent 61 % des architectes d'entreprise dans notre enquête. Les solutions multicloud sont privilégiées (37 %) pour des raisons telles que le contrôle, la sécurité, la propriété des données, la conformité et le coût.

Depuis leurs premiers investissements dans l'IA, les entreprises interrogées ont souvent choisi de déplacer une partie des charges de travail hébergées dans des environnements de cloud public vers le cloud privé. Les ressources, serveurs en tête, y sont dédiées à une entreprise spécifique au lieu d'être réparties entre plusieurs clients simultanément. En raison de leurs avantages en termes de contrôle, de sécurité, de propriété des données et de conformité, les infrastructures de cloud privé peuvent devenir une option intéressante pour l'intégration des applications d'IA générative. Les règles de conformité et les données sensibles nécessaires à l'entraînement et à l'exploitation des modèles exigent souvent une infrastructure sur mesure.

D'autre part, l'évolutivité et la flexibilité offertes par les systèmes de cloud public peuvent être de précieux atouts, d'autant que les coûts initiaux sont plus faibles. Parce qu'ils permettent des cycles de développement plus courts, intègrent des frameworks techniques, facilitent les expérimentations à grande échelle et peuvent évoluer rapidement, les systèmes de cloud public ont tout à fait fait leur intérêt. Mais il faut trouver un juste équilibre entre rentabilité, contrôle, sécurité et flexibilité, et cette question est loin d'avoir une réponse définitive.

« Vous pouvez avoir toute l'IA du monde, elle ne vous apportera aucune valeur ajoutée si elle s'appuie sur des données fragiles. »

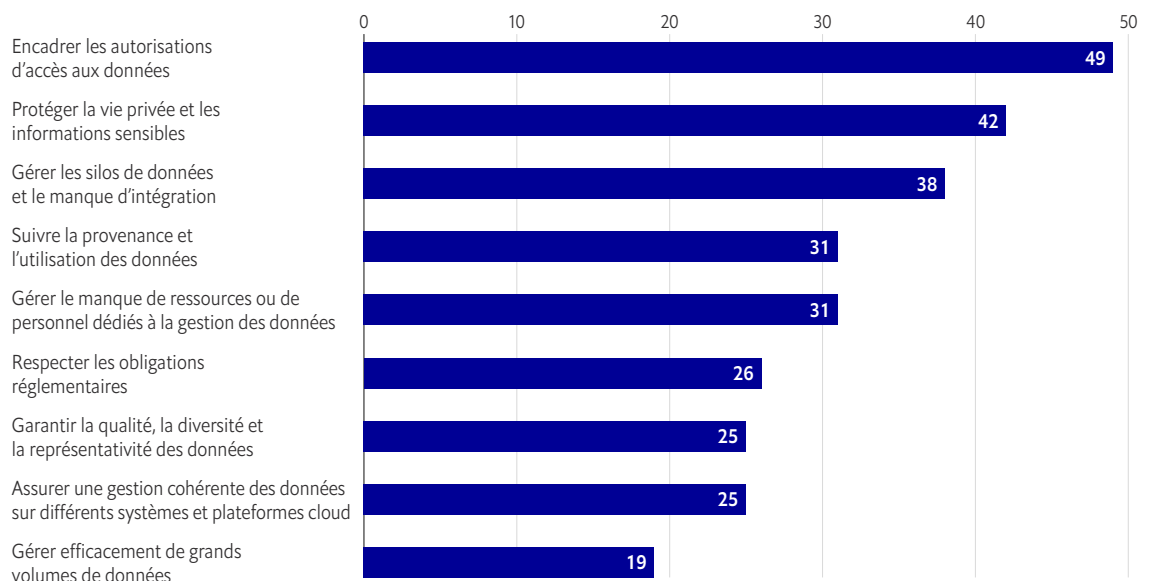
Carol Clements, Chief digital and technology officer, JetBlue

Au-delà de l'entreposage des données se pose l'épineux problème de leur exactitude et de leur sécurité, qui passe notamment par des contrôles d'accès appropriés. N'oublions pas en effet que des données incorrectes ou contradictoires pourraient amener un modèle d'IA générative à fournir des informations inexactes, et que des contrôles laxistes pourraient entraîner des violations de conformité. Comme l'explique Mme Clements de JetBlue : « Vous pouvez avoir toute l'IA du monde, mais elle ne vous apportera aucune valeur ajoutée si elle s'appuie sur des données fragiles. »

« En fin de compte, tout repose sur des données. « Si vos données ne sont pas correctement catégorisées, vous aurez toutes les peines du monde à tirer profit de cette technologie. Vous devez donc vous concentrer sur les fondamentaux de l'ingénierie et de la gestion des données, » affirme de son côté M. Vohs de Molson Coors. En ce qui concerne ces fondamentaux, les architectes d'entreprise identifient la sécurité, la confidentialité des données et les silos comme les trois grands défis de la gestion des données pour les applications d'IA (Figure 8).

Figure 8 : Les défis liés à l'encadrement des données

Principaux défis des architectes d'entreprise en matière de gestion et de contrôle des données pour les applications d'IA
 Pourcentage d'architectes d'entreprise



Source : Economist Impact

Parmi les défis introduits par l'IA générative, la quantité d'informations à envoyer dans le cloud pour analyse est source d'inquiétude. Certains hésitent à transférer davantage de données dans le cloud sans avoir clairement défini le cas d'utilisation. Cette réticence se manifeste particulièrement dans des secteurs comme l'énergie et la banque. Ceux-ci ont déjà accumulé une lourde dette technique et de nombreux silos de données, souvent dus à une multiplicité de systèmes hérités ou à des fusions et acquisitions successives. L'énergéticien espagnol Repsol et la banque canadienne TD mentionnent également les coûts potentiels liés aux importations massives de données. Il s'agit des coûts directs liés au cloud, bien sûr, mais aussi des coûts indirects découlant des ressources humaines nécessaires à la préparation des données ainsi qu'à la création

et à la maintenance des pipelines de transfert.

Auparavant, les entreprises avaient tendance à négliger les données non structurées, alors qu'elles représentaient la grande majorité de leurs actifs. Mais ces données ont acquis une importance croissante grâce aux outils d'IA qui facilitent leur exploitation.²⁰ UPS utilise un data mesh pour classer les données, doublé d'outils d'analyse qui permettent aux data scientists et aux équipes métier de lancer des projets de machine learning (MLOps). « Il est particulièrement puissant, car il peut combiner différents modèles ML dans un court délai. « Si demain nous voulions y intégrer un modèle d'IA générative, nous pourrions le faire sans problème. Nous avons en effet déjà toutes les plateformes et l'orchestration nécessaires, » explique M. Finnerty d'UPS.

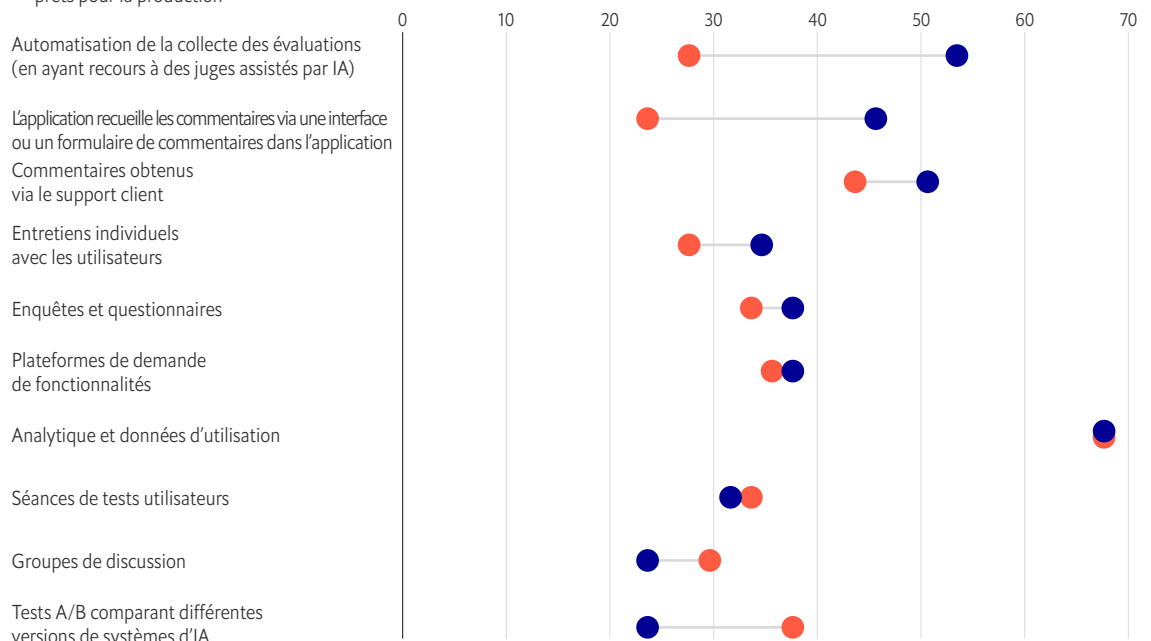
Figure 9 : Qualité de l'IA générative : leaders et retardataires

Stratégies employées pour évaluer la qualité des réponses de l'IA générative

Pourcentage de data scientists

● Leaders = ceux dont les outils d'IA générative sont prêts pour la production

● Retardataires = ceux dont les outils d'IA générative ne le sont pas



Remarque : pour les retardataires, n = 50 ; pour les leaders, n = 37 ; les informations sont fournies à titre purement indicatif.

Source : Economist Impact

20 MIT Sloan, « Tapping the power of unstructured data » (Exploiter la puissance des données non structurées), 2021, <https://mitsloan.mit.edu/ideas-made-to-matter/tapping-power-unstructured-data>

Le poids de la maintenance

Dans l'arsenal technologique, l'IA générative se distingue des autres utilisations des données par la quantité d'entretien et de maintenance qu'elle impose. Les LLM, par exemple, souffrent d'une « dérive » avec le temps : les informations sur lesquelles ils ont été entraînés deviennent obsolètes et leurs performances se détériorent.²¹ M. Hallworth de HP compare d'ailleurs les modèles d'IA aux voitures qui commencent à perdre de la valeur dès qu'elles quittent le parking du concessionnaire.

Cette maintenance prend principalement trois formes. Premièrement, l'exhaustivité, l'actualité et la validité des ensembles de données doivent être constamment vérifiées et confirmées. Ces contrôles peuvent être effectués à l'aide d'outils open source et d'interfaces de programmation d'applications (API), directement intégrés aux ensembles de données et aux clouds. Il est également possible d'appliquer des processus d'assurance qualité à l'ensemble des pipelines de données pour empêcher dès le départ l'introduction de données erronées dans les systèmes de l'organisation.

Selon notre enquête, les très grandes entreprises (dont le chiffre d'affaires dépasse les 10 milliards de dollars) utilisent le ML plus fréquemment que les autres pour détecter les anomalies dans les données. S'agissant de la maintenance des données spécifiquement destinées aux modèles d'IA générative, les « leaders » (les entreprises disposant d'outils d'IA générative prêts pour la production) sont également plus nombreux que les autres à utiliser l'IA pour évaluer les données (Figure 9).

Deuxièmement, dans le cas des systèmes d'IA générative, les sources de données externes ou « encyclopédies » utilisées par les LLM pour affiner leurs réponses doivent être régulièrement actualisées pour préserver la qualité de leurs informations.

Et il faut enfin constamment tester les modèles eux-mêmes pour en vérifier l'efficacité. La figure 9 met en avant l'exemple de la collecte d'évaluation automatisée, qui consiste à utiliser des algorithmes de ML pour quantifier l'exactitude des résultats générés par un LLM par rapport aux réponses attendues.²²

Principales conclusions

Un cinquième seulement des organisations disposent d'une architecture capable de prendre en charge les charges de travail de l'IA. C'est un obstacle majeur à la mise en oeuvre opérationnelle de l'IA. Le traitement des données en temps réel représente une grande source de valeur ajoutée. Très recherché par les entreprises, il n'est toutefois possible qu'au prix d'une refonte de l'infrastructure. Une infrastructure unifiée aide l'IA à montrer tout son potentiel, non seulement en tant qu'outil de productivité, mais aussi pour sa contribution à l'amélioration et l'accélération de la prise de décision.

La taille ou le degré de maîtrise du numérique ne sont pas nécessairement synonymes de réussite dans la transformation de l'infrastructure. Les grandes entreprises ont des décisions plus complexes à prendre en matière de centralisation et de décentralisation. Et les entreprises digital natives, souvent les premières à adopter les logiciels de pointe, peuvent avoir des environnements plus complexes.

L'IA demande beaucoup d'entretien. L'exhaustivité, l'actualité et la validité des ensembles de données doivent être constamment vérifiées et confirmées à l'aide d'outils open source et d'API directement intégrés aux ensembles de données et aux clouds. Il faut également appliquer des processus d'assurance qualité à l'ensemble des pipelines de données pour empêcher dès le départ l'introduction de données erronées dans les systèmes de l'organisation. Les sources de données externes utilisées par les LLM pour affiner leurs réponses doivent être régulièrement actualisées pour préserver la qualité de leurs informations. Et il faut constamment tester les modèles eux-mêmes pour en vérifier l'efficacité.

21 IBM, « What Is Model Drift ? » (Qu'est-ce que la dérive des modèles ?), 2024, <https://www.ibm.com/topics/model-drift>

22 Microsoft, « Evaluation of generative AI applications » (Évaluation des applications d'IA générative), 2024, <https://learn.microsoft.com/en-us/azure/ai-studio/concepts/evaluation-approach-gen-ai>

Perspectives des professionnels

Les architectes d'entreprise sur la corde raide

Le rôle des architectes de données et d'IA est complexe, passionnant et crucial dans les organisations modernes axées sur la technologie. Pour Gregor Hohpe, qui a occupé des postes de direction chez des géants de la technologie comme Google et Amazon Web Services, l'architecte logiciel est « un technicien de haut niveau qui s'épanouit dans la complexité, mais recherche la simplicité ». Les architectes ont pour mission de concevoir, mettre en œuvre et superviser une architecture de données et d'IA à la fois évolutive et sécurisée. Ils doivent pour cela posséder une parfaite connaissance des technologies actuelles et émergentes.

Notre enquête a révélé que seulement 22 % des organisations disposent d'architectures qui prennent pleinement en charge les charges de travail d'IA sans modifications, et que seulement 23 % peuvent connecter les applications d'IA aux données pertinentes de l'entreprise sans changements. Il appartient désormais aux architectes de combler cette lacune et d'aider les organisations à parvenir à une intégration complète de l'IA. Le traitement des données en temps réel, mentionné

par près de la moitié (47 %) des architectes comme un besoin essentiel, est en tête des lacunes identifiées. Les autres préoccupations concernent la confidentialité des données et les failles de sécurité. Pour renforcer la gouvernance, les architectes misent avant tout sur des API personnalisées (58 %), sur des processus d'implication humaine (55 %) et sur l'intégration de modèles ML dans des outils de données familiers (52 %). Cette approche met en évidence la nécessité de conjuguer solutions techniques et supervision humaine dans la mise en œuvre de l'IA, pour assurer l'intégration parfaite des capacités d'IA dans la structure des opérations organisationnelles.

Les technologies qui améliorent l'accessibilité et l'efficacité sur différentes plateformes gagnent du terrain. Le traitement du langage naturel (NLP) est à l'avant-garde de cette tendance : 58 % des architectes estiment en effet qu'il deviendra l'outil principal (voire le seul) de la création de pipelines de données. Ils sont 67 % à affirmer la même chose s'agissant du développement des workflows, et 52 % pour ce qui est de la génération de rapports. C'est le signe d'une évolution vers des outils plus intuitifs et conviviaux dans les opérations de données et d'IA.

Dans le climat économique actuel, qui incite à réduire toujours plus le coût total de possession, ces fonctions multiplateformes revêtent un intérêt stratégique. Mettre l'accent sur l'accessibilité et l'efficacité ne permet pas seulement de réduire les coûts, mais aussi de favoriser la démocratisation des données et de l'IA, et donc de mettre ces puissants outils et ces informations à la disposition d'un public plus large au sein de l'organisation.

Les défis rencontrés par les architectes des données et d'IA sont nombreux, mais les opportunités le sont tout autant. Dans leurs efforts pour élaborer des stratégies pour mettre en œuvre une IA éthique, gérer les écosystèmes de données et démocratiser l'accès aux données, les architectes devront établir un équilibre délicat entre innovation, sécurité et accessibilité.

Les data engineers mettent de l'ordre dans le chaos des données

L'ingénierie des données est soumise à une pression immense. Elle doit répondre aux nouvelles demandes des entreprises modernes tout en relevant deux défis majeurs : maintenir une qualité de données élevée pour les utilisateurs finaux et garantir la sécurité et la gouvernance des données. Ces objectifs s'ajoutent aux responsabilités des professionnels de la discipline, qui doivent déjà gérer la plupart des aspects du cycle de vie de l'ingénierie des données.

L'impératif d'unification, de simplification et de démocratisation n'a jamais atteint de tels sommets dans l'ensemble de la discipline de l'ingénierie des données. Aujourd'hui, près des deux tiers des organisations dépendent entièrement des data engineers pour tous les aspects de la création et de la gestion du pipeline de données. Et près de la moitié des ingénieurs passent la majeure partie de leur temps à configurer et à réparer des connexions aux

sources de données. Le risque de fatigue, voire d'épuisement professionnel est élevé chez les data engineers qui doivent gérer d'innombrables sources de données et utiliser une myriade d'outils pour configurer des pipelines utilisables.

L'émergence du NLP comme outil de démocratisation de la création de pipelines de données et de workflow s'avère prometteuse : plus de 97 % des architectes d'entreprise prédisent en effet que le NLP complétera ou remplacera les méthodes traditionnelles de création de pipelines et d'orchestration des workflows. Ce changement signifie une chose : à l'avenir, les data engineers pourront de plus en plus compter sur l'automatisation intelligente pour gérer les aspects complexes du traitement des données en temps réel et de la gestion des pipelines. Ils seront ainsi libres de se consacrer à des tâches plus stratégiques.

En tête de la liste de souhaits de la plupart des équipes de données et d'IA ? Une simplification des connexions aux sources de données, des outils d'IA générative intégrés pour accélérer l'écriture du code et des plateformes unifiées pour réduire le recours aux systèmes fragmentés. Les architectes attendent beaucoup du traitement des données en temps réel et des pipelines de données robustes. Près de 75 % d'entre eux considèrent l'un ou l'autre comme un objectif essentiel pour les capacités de données et d'IA de leur organisation. Avec la généralisation des architectures de streaming, la question de la latence et de la disponibilité des informations en temps réel va devenir un enjeu majeur. En témoignent les 46 % d'entreprises interrogées ayant converti un grand nombre de pipelines en batch en flux en temps réel au cours l'année écoulée.

En résumé, l'ingénierie des données se trouve à un tournant critique. Le domaine gagne en efficacité grâce à l'automatisation et au traitement des données en temps réel, mais son évolution est encore freinée par les défis fondamentaux de la sécurité, de la gouvernance et de la qualité des données.

Chapitre 4 :

Trouver votre avantage

À l'heure où l'IA devient la norme dans les entreprises, celles-ci doivent désormais trouver leur propre voie en intégrant des données propriétaires et une expertise spécialisée aux modèles de fondation. Leur objectif : trouver le compromis idéal entre les aspects pratiques du prêt à l'emploi et la puissance des solutions sur mesure, tout en préservant leur atout le plus précieux : leurs données.

Les modèles de fondation se généralisent, et leur adoption cesse d'être un catalyseur pour devenir la norme. Un avertissement revient souvent dans les débats actuels : l'IA ne remplacera pas les humains. Mais les humains qui l'utiliseront prendront la place des autres. Il en va de même pour les entreprises. Chacune doit trouver son avantage, en composant sa palette unique de données, d'exigences et de savoir-faire. « Il faut s'appuyer sur ce qu'on possède déjà. Dans notre cas, ce sont nos connaissances médicales, notre expérience utilisateur, notre connexion aux dispositifs corporels et l'IA générative, que nous utilisons pour obtenir des insights ou répondre à des questions, » explique M. Bugaev de Flo Health.



Composer un assortiment

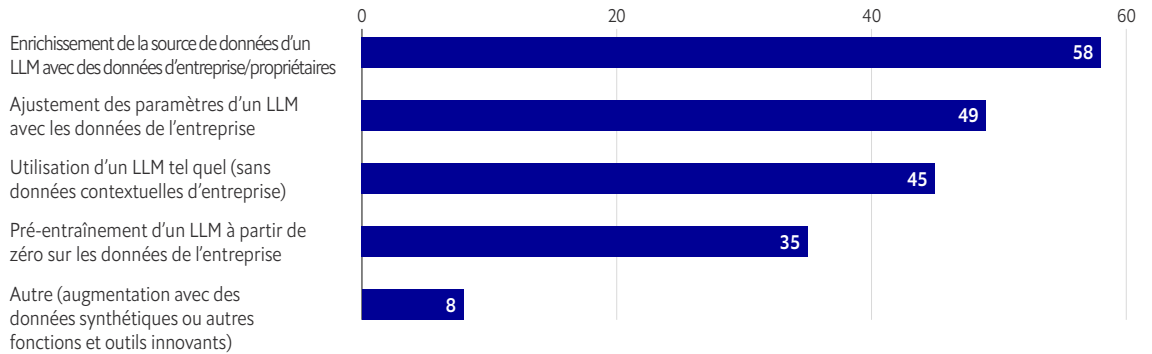
Selon notre étude, deux tiers des organisations trouvent prometteuse l'intégration de leurs propres données à des modèles d'IA générative. Dans la plupart des cas, elles vont composer une palette unique en fonction des workflows et des besoins à prendre en charge. Pour prendre un exemple, 58 % des data scientists enrichissent leurs LLM avec des données propriétaires à l'aide du processus de RAG. Celui-ci consiste, pour un LLM, à utiliser une source externe qualifiée lors de l'inférence ou de la production d'un résultat. En revanche, 45 % déclarent utiliser des LLM prêts à l'emploi sans les connecter à leurs données (Figure 10). Et ils sont 21 % à pratiquer les deux approches.

M. Hallworth de HP explique ce qui les distingue en faisant une analogie avec une pizza. Les modèles d'IA prêts à l'emploi sont pratiques et font gagner du temps, mais un modèle fait maison, qui demande davantage de temps de préparation, aura meilleur goût. Une IA de la catégorie « pizza surgelée » pourrait rédiger le premier jet d'un texte marketing ou une synthèse de réunion. Mais on réservera à un modèle fait maison les tâches plus sophistiquées, comme des interactions avec les clients, la génération d'insights sur les produits ou des prévisions.

Figure 10 : Créer et enrichir les LLM avec les données d'entreprise

Nature des projets d'IA générative actuels parmi les data scientists

Pourcentage de data scientists



Source : Economist Impact

« La plus grande partie de la valeur de l'IA est encore enfouie sous la surface, » explique M. Ramani d'Accenture. « Les avantages immédiats en termes de productivité ne sont que la pointe émergée de l'iceberg. Vous n'avez pas encore atteint le cœur de votre chaîne de valeur principale, où réside toute la promesse. Et c'est là que se trouvent vos données propriétaires. »

« La plus grande partie de la valeur de l'IA est encore enfouie sous la surface. Les avantages immédiats en termes de productivité ne sont que la pointe émergée de l'iceberg. Vous n'avez pas encore atteint le cœur de votre chaîne de valeur principale, où réside toute la promesse. Et c'est là que se trouvent vos données propriétaires. »

Senthil Ramani, Global lead, data and AI, Accenture

Si les modèles de fondation sont un moulin à café, les données de l'entreprise sont les grains à transformer. En utilisant ces données dans le cadre d'un système composite²³ les organisations peuvent obtenir des résultats d'IA qui tiennent compte des spécificités de leur domaine ou de leur situation. Il est déjà courant de combiner plusieurs modèles pour répondre à différents cas d'usage, comme le montre notre enquête (Figure 12), tout comme le fait d'intégrer des données propriétaires à des modèles d'IA générative, pratiqué ou exploré par la plupart des participants.

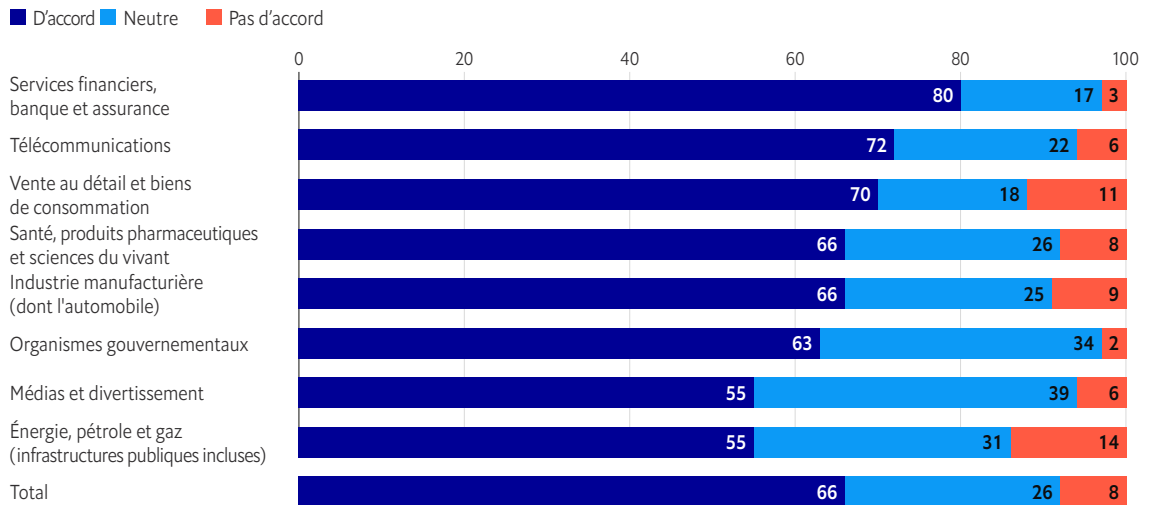
Les répondants issus du secteur des services financiers étaient de loin les plus nombreux (80 %) à voir un potentiel important dans l'intégration des données propriétaires aux modèles d'IA générative (Figure 11). Les points de vue sur ce sujet diffèrent considérablement d'un secteur à l'autre. Les divergences concernent principalement la valeur des données organisationnelles, la complexité de leur traitement, les risques liés à leur utilisation dans les modèles d'IA générative et la volonté culturelle ou opérationnelle de mettre en œuvre ces technologies.

23 Bair, « The Shift from Models to Compound AI Systems » (Des modèles aux systèmes d'IA composites), 2024, <https://bair.berkeley.edu/blog/2024/02/18/compound-ai-systems/>

Figure 11 : L'intérêt des données propriétaires dans les différents secteurs d'activité

Participants qui pensent ou non que leur organisation trouve prometteuse l'intégration de ses propres données dans des modèles d'IA générative.

Pourcentage de l'ensemble des participants



Source : Economist Impact

Choisir des modèles

La sélection des modèles est décisive lorsqu'il s'agit d'adapter l'IA aux actifs d'une organisation. Les modèles d'IA open source, par exemple, autorisent quiconque à « les étudier, les modifier, les utiliser et les partager ». ²⁴ La définition de l'open source dans le domaine de l'IA fait l'objet de débats. La question est notamment de savoir si elle s'étend aux données utilisées pour créer des modèles. Il reste néanmoins que la présence d'outils open source a élargi l'accès à l'IA. ²⁵ Le modèle Llama de Meta, par exemple, a été téléchargé 300 millions de fois et permet aux organisations de créer des modèles d'IA adaptés à leurs besoins. ^{26,27}

Selon notre enquête, l'approche hybride combinant IA open source et IA propriétaire représente une stratégie courante : 75 % des organisations utilisent des modèles open source et fermés en fonction de facteurs propres à leur secteur, et 89 % prévoient de le faire d'ici 2027. Les modèles à source fermée sont généralement déployés pour des applications internes, tandis que les modèles open source sont de plus en plus privilégiés pour les cas d'utilisation externes.

Le choix de l'open source s'explique par plusieurs facteurs selon les organisations interrogées. La transparence du développement arrive en tête de liste avec 52 %, suivie de la disponibilité des ressources communautaires open source, des possibilités de développement de propriété intellectuelle (PI), de l'indépendance vis-à-vis des fournisseurs et d'un meilleur contrôle sur la personnalisation, la confidentialité et la sécurité. Cela confirme les avantages offerts par les modèles open source en termes de flexibilité et de contrôle.

24 Open Source Initiative, « The Open Source AI Definition – draft v. 0.0.8 » (Définition de l'IA open source – version préliminaire 0.0.8), <https://opensource.org/deepdive/drafts/the-open-source-ai-definition-draft-v-0-0-8>

25 Economist Impact, « Open sourcing the AI revolution » (La révolution de l'IA en open source), 2024, <https://impact.economist.com/perspectives/technology-innovation/open-sourcing-ai-revolution>

26 The Economist, « Meta is accused of "bullying" the open-source community » (Meta accusé « d'intimidations » à l'égard de la communauté open source), 2024, <https://www.economist.com/business/2024/08/28/meta-is-accused-of-bullying-the-open-source-community>

27 Microsoft, « Tiny but mighty: The Phi-3 small language models with big potential » (Petits mais costauds : le gros potentiel des petits modèles de langage Phi-3), 2024, <https://news.microsoft.com/source/features/ai/the-phi-3-small-language-models-with-big-potential/>

Mais les modèles propriétaires ont, eux aussi, leurs atouts. Le gain de performance, cité par 59 % des participants, est un facteur décisif dans le choix d'un modèle à source fermée. Mais ce n'est pas le seul avantage : les répondants évoquent également des interfaces et une documentation plus conviviales, des fonctionnalités exclusives et une intégration transparente avec les systèmes propriétaires existants. Certains remarquent également que l'adaptation et le développement de modèles open source exigent davantage d'expertise et de compétences internes que les modèles fermés prêts à l'emploi.²⁸

Les leaders de l'industrie développent des cadres pour éclairer le choix des modèles. Pour aider les organisations, Accenture a mis au point son « switchboard », des outils et un cadre propriétaires permettant d'identifier les modèles à utiliser en fonction du cas d'usage métier et de plusieurs critères comme la pertinence, le coût, la précision et la latence. Ces facteurs aident à déterminer si une organisation a besoin de modèles open source, de pointe (frontier) ou autre.

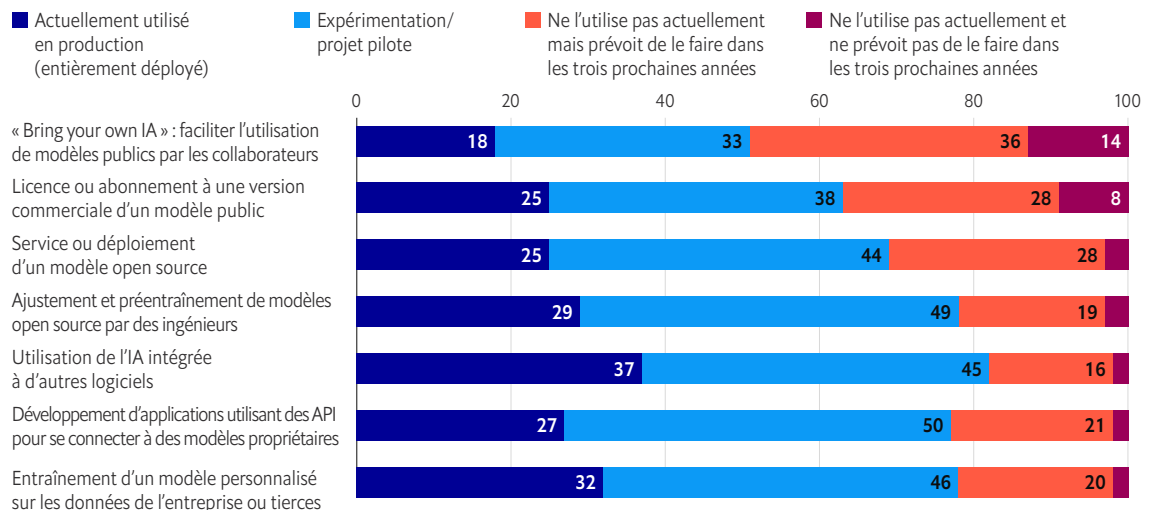
Relativement peu d'entreprises voient l'intérêt de développer leur propre LLM à partir de zéro. « Nous ne pensons pas qu'une entreprise de notre taille puisse rapidement créer un LLM interne véritablement intéressant, » déclare M. Bugaev de Flo Health. De la même façon, les entreprises ne pourront pas tirer un avantage tangible des seuls modèles prêts à l'emploi.

« Nous pensons que les partenariats entre les entreprises technologiques et les secteurs pharmaceutique et de la santé sont indispensables, » déclare M. Bucher de Novartis. « Les entreprises technologiques possèdent des plateformes solides, les capacités de données et les talents en IA indispensables pour créer des modèles utilisables. Les sociétés pharmaceutiques possèdent une expertise approfondie dans les domaines de la sécurité, des essais cliniques et de la production. C'est essentiel pour le développement de médicaments potentiels. L'IA est le catalyseur qui décuple nos efforts combinés. En misant sur nos points forts respectifs, nous pouvons réaliser ensemble des percées majeures. »

Figure 12 : Les organisations combinent différents modèles d'IA générative

Adoption des modèles d'IA générative, par type

Pourcentage de dirigeants



Source : Economist Impact

28 Economist Impact, « Open sourcing the AI revolution » (La révolution de l'IA en open source), 2024, <https://impact.economist.com/perspectives/technology-innovation/open-sourcing-ai-revolution>

Les entreprises et les organisations ont accès au meilleur carburant des modèles d'IA : des données soigneusement choisies et pertinentes, bien supérieures aux données provenant de l'Internet public. Dream Sports, par exemple, a été déçue par l'efficacité des LLM tiers, même en utilisant les applications les plus récentes, en raison de la spécificité de la terminologie de l'entreprise. Les modèles doivent être entraînés sur les données internes.

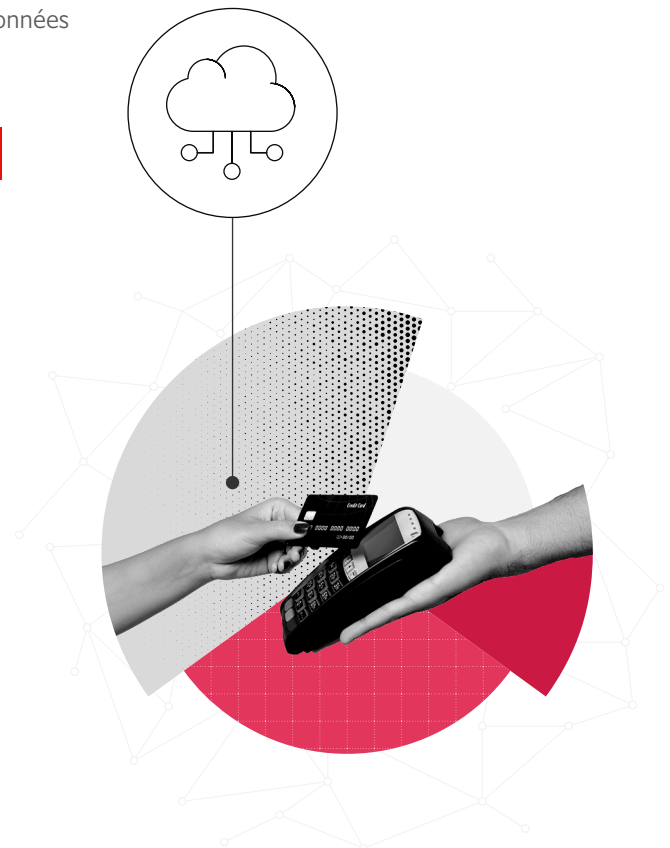
Parce qu'elles sont de meilleure qualité, les données d'entreprise améliorent l'efficacité et les performances des modèles. Rakuten a ainsi considérablement réduit la puissance de calcul nécessaire au développement de son modèle de fondation de deep learning et des intégrations. La société les a entraînés sur la base des données de transaction de son site de commerce électronique, comme l'explique Ting Cai, Directeur de l'IA et des données de l'entreprise. « Nous encodons davantage d'informations dans un nombre réduit de dimensions. Nous disposons en effet de données

de transaction de haute qualité qui traduisent plus précisément les intérêts de nos utilisateurs. Cela nous permet de créer des modèles plus efficaces. »

Shell adapte également des modèles de fondation pour travailler sur des données internes, comme les interactions avec les clients ou les contrats passés. « Je ne m'inquiète pas trop des hallucinations dans les modèles de langage qu'utilise Shell en raison des contrôles que nous avons mis en place avant le déploiement. Ces modèles sont limités ; ils ne peuvent fonctionner qu'avec nos données contrôlées qui ont subi des contrôles rigoureux. Ils sont par ailleurs placés sous une supervision humaine appropriée avant d'être intégrés à nos processus de travail, » explique Robbert Van Rutten, DSI du géant de l'énergie. « Nous travaillons également au développement de nos propres modèles en préentraînant ou en ajustant des LLM pour minimiser davantage les risques. »

« Nous encodons davantage d'informations dans un nombre réduit de dimensions. Nous disposons en effet de données de transaction de haute qualité qui traduisent plus précisément les intérêts de nos utilisateurs. Cela nous permet de créer des modèles plus efficaces. »

Ting Cai, Directeur de l'IA et des données, Rakuten Group



« L'avenir du marché est aux petits LLM privés. Mais l'IA générative est une technologie nouvelle. Nous ne savons pas encore lancer des initiatives et les mettre en production avec agilité. »

Juan Jose Casado, Chief digital officer, Repsol

Les entreprises peuvent naturellement utiliser des modèles tiers pour procéder à des essais initiaux avant d'investir davantage de ressources techniques. Repsol, par exemple, a injecté ses données internes dans des LLM fournis par des sociétés comme OpenAI, Microsoft et Google plutôt que de créer intégralement un pipeline d'IA générative. L'entreprise voulait alors seulement expérimenter des cas d'usage et explorer le potentiel de la technologie. « L'avenir du marché est aux petits LLM privés, » prédit Juan José Casado, Chief digital officer de l'entreprise énergétique espagnole. « Mais l'IA générative est une technologie nouvelle. Nous ne savons pas encore lancer des initiatives et les mettre en production avec agilité. »

L'attrait des modèles développés en interne réside également dans le contrôle qu'il laisse aux entreprises sur leurs données. Beaucoup s'inquiètent des fuites de données dans les LLM à source fermée, craignant notamment que des informations sensibles ou confidentielles ne soient injectées dans des modèles externes. « Nous avons des opinions bien arrêtées sur la propriété et la protection de nos données, d'où notre grande prudence quant à leur utilisation, » explique Mme Clements de JetBlue. « Nous ne les laissons pas quitter JetBlue. »

Principales conclusions

Pour créer de la valeur et acquérir un avantage, les entreprises vont combiner de l'IA prête à l'emploi avec des modèles open source et internes. L'IA fera bientôt partie de l'outillage de base. Les organisations qui excelleront sauront combiner différents modèles en les associant à leurs données et leur savoir-faire uniques. Elles seules auront accès au meilleur carburant des modèles d'IA : des données soigneusement choisies et pertinentes.

L'IA promet l'accès à des données exclusives jusque-là hors de portée ou inexploitées.

Il sera ainsi possible de créer des modèles plus performants grâce à des gains de précision, d'efficacité et de pertinence. L'IA prête à l'emploi sera toujours utile pour expérimenter à moindre coût et obtenir des résultats à faible valeur ajoutée. Mais pour créer véritablement de la valeur en bénéficiant d'un contrôle accru et d'une sécurité renforcée, les entreprises devront combiner des modèles open source et fermés. Les grandes structures pourraient voir un intérêt au développement intégral de leurs propres modèles.

Chapitre 5 :

Garde-fous et gouvernance

Dans un contexte d'omniprésence croissante de l'IA, les organisations font face à un ensemble complexe de risques liés aussi bien aux violations de données qu'à la conformité réglementaire. Les entreprises développent des approches de gouvernance multifacettes. Elles mettent en place des mesures institutionnelles telles que les centres d'excellence en IA, adoptent des principes comme la supervision humaine systématique, et déploient des solutions techniques innovantes – données synthétiques, RAG – pour améliorer la fiabilité de l'IA. Le défi consiste à trouver le juste équilibre pour parvenir à exploiter le pouvoir de transformation de l'IA. Il faut aussi maintenir de solides garanties et calibrer le rapport entre jugement humain et intelligence artificielle, dans un environnement réglementaire qui évolue constamment.

La sécurité des données et la gouvernance sont devenues des priorités pour le conseil d'administration au cours de la dernière décennie, du fait de l'ampleur des attaques de cybersécurité – et de la vigueur des sanctions. Une série de réglementations telles que le Règlement européen sur la résilience opérationnelle numérique (DORA) augmentent également les risques liés aux défauts de gouvernance. Ces priorités sont sensibles dans notre enquête : la confidentialité des données et les violations de sécurité représentent la principale préoccupation de 53 % des architectes d'entreprise. La sécurité et la gouvernance font partie, elles, des trois grands défis du data engineering pour 50 % des ingénieurs.

« De nombreuses problématiques, comme la protection de la vie privée, les mesures de sécurité, les problèmes de propriété intellectuelle, les questions éthiques et le problème des deepfakes, doivent être abordées de front, » déclare M. Sato de NTT Docomo. « Il faut impérativement maximiser les avantages des technologies d'IA en surmontant ces défis grâce à la collaboration de toutes les parties concernées, » ajoute-t-il.

Dans le monde entier, la législation se durcit. Pour l'année 2024, 762 documents législatifs ayant trait à l'IA ont été promulgués sur le seul territoire des États-Unis.²⁹ C'est un véritable défi pour les entreprises qui doivent expérimenter de nouveaux systèmes alors que les règles qui encadrent leur utilisation changent constamment.

Selon Ian Botts, Chief technology officer chez Fanatics Betting & Gaming, la gouvernance de l'IA au sein d'une organisation doit réunir trois composants pour être efficace. « Dans un premier temps, il faut mettre en place des garde-fous qui définissent ce qui est autorisé ou non, afin que des informations

La confidentialité des données et les violations de sécurité représentent la principale préoccupation de 53 % des architectes d'entreprise. La sécurité et la gouvernance font partie, elles, des trois grands défis du data engineering pour les ingénieurs.

29 Regulatory Transparency Project, « A Sensible Approach to State AI Policy » (Une approche raisonnable de la politique de l'État en matière d'IA), https://rtp.fedsoc.org/blog/a-sensible-approach-to-state-ai-policy/?utm_content=311457390&utm_medium=social&utm_source=twitter&hss_channel=tw-574405888

personnelles ne puissent pas être utilisées dans un modèle pour cibler spécifiquement un individu en particulier. Il faut ensuite établir et développer les mécanismes qui vont garantir l'application de ces garde-fous. Et troisièmement, il faut mettre en place un véritable service de la gouvernance des données, sur le modèle d'une fonction de conformité, pour assurer le contrôle des données et la transparence des entreprises. Je pense que la question de la transparence est extrêmement importante car elle permet de prendre de meilleures décisions. »

Les innovations dans le domaine des données peuvent aider les entreprises à faire face aux risques de gouvernance, à commencer par les biais et le manque de fiabilité des modèles. Les organisations n'ont par exemple pas toutes accès aux données nécessaires pour créer les outils dont elles ont besoin, ce qui limite la qualité de leurs systèmes d'IA. Le recours à des données synthétiques, générées à l'aide d'un modèle mathématique ou d'un algorithme spécialement conçu, peut être une solution pour ces tâches de data science.³⁰ Il peut améliorer les performances et la fiabilité des projets en augmentant le volume et la diversité des données d'entraînement, notamment en incluant des classes ou des catégories de données sous-représentées. La réduction des écarts entre les genres et l'élimination des préjugés racistes sont des domaines prometteurs pour les données synthétiques.³¹ Elles peuvent également faire baisser les coûts d'acquisition de données, éliminer les goulets d'étranglement et contribuer à la préservation de la confidentialité. Certes, cette pratique comporte des risques et des inconvénients. Mais en mettant en œuvre des mesures de protection adéquates, les données synthétiques peuvent améliorer la qualité des résultats de l'IA. Elles restent malgré tout une ressource sous-exploitée selon notre enquête : 7,5 % seulement des data scientists

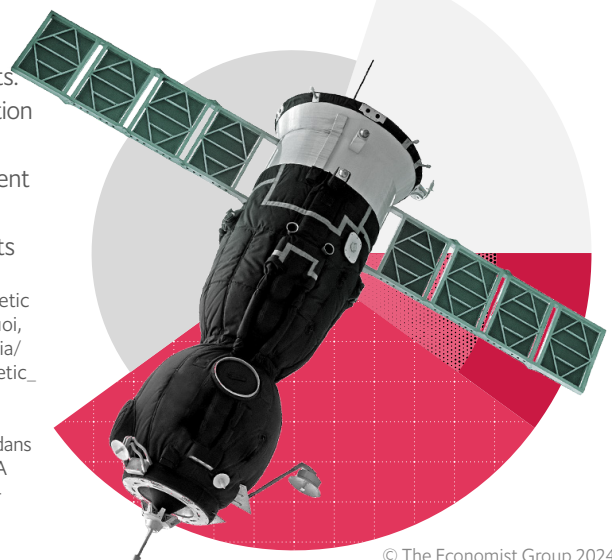
30 The Alan Turing Institute and The Royal Society, « Synthetic data - what, why and how? » (Données synthétiques : quoi, pourquoi et comment ?), https://royalsociety.org/-/media/policy/projects/privacy-enhancing-technologies/Synthetic_Data_Survey-24.pdf

31 UNU Macau, « Bridging the Gender Data Gap: Harnessing Synthetic Data for Inclusive AI » (Comblent le fossé du genre dans les données : miser sur les données synthétiques pour une IA inclusive), 2024, <https://unu.edu/macau/blog-post/bridging-gender-data-gap-harnessing-synthetic-data-inclusive-ai>

les utilisent dans leurs projets d'IA générative.

Autre technique utile pour contrôler la qualité des données et enrichir les connaissances : la génération augmentée par récupération, ou RAG. « Grâce à cette technique, les modèles d'IA générative de Rakuten récupèrent des données d'inventaire et de tarification en temps réel pour intégrer les dernières informations et connaissances de l'entreprise dans des applications comme la recherche, les systèmes de recommandation et les publicités, » explique M. Cai, Chief AI and data officer de l'entreprise.

Ian Botts de Fanatics explique également comment ses modèles d'IA et d'IA générative servent à détecter et corriger les problèmes de gouvernance liés au jeu responsable. « Nous voulons que le système de pari soit un plus pour votre communauté, » explique-t-il. « Si [nos modèles] repèrent avec une très grande précision qu'un utilisateur présente manifestement des dépenses supérieures à ses moyens, ou des modèles de comportements inquiétants, [nous] pourrions l'inciter subtilement à jouer à des jeux gratuits ou moins chers, à éviter de continuer sur cette pente ou à faire une pause. Nous pourrions lui donner des informations. C'est un investissement supplémentaire qui en vaut la peine... si cela permet à quelqu'un de retrouver un comportement sain et viable pour lui. L'IA peut repérer et comprendre les problèmes, et l'IA générative peut permettre d'exercer une influence [de façon individualisée]. »



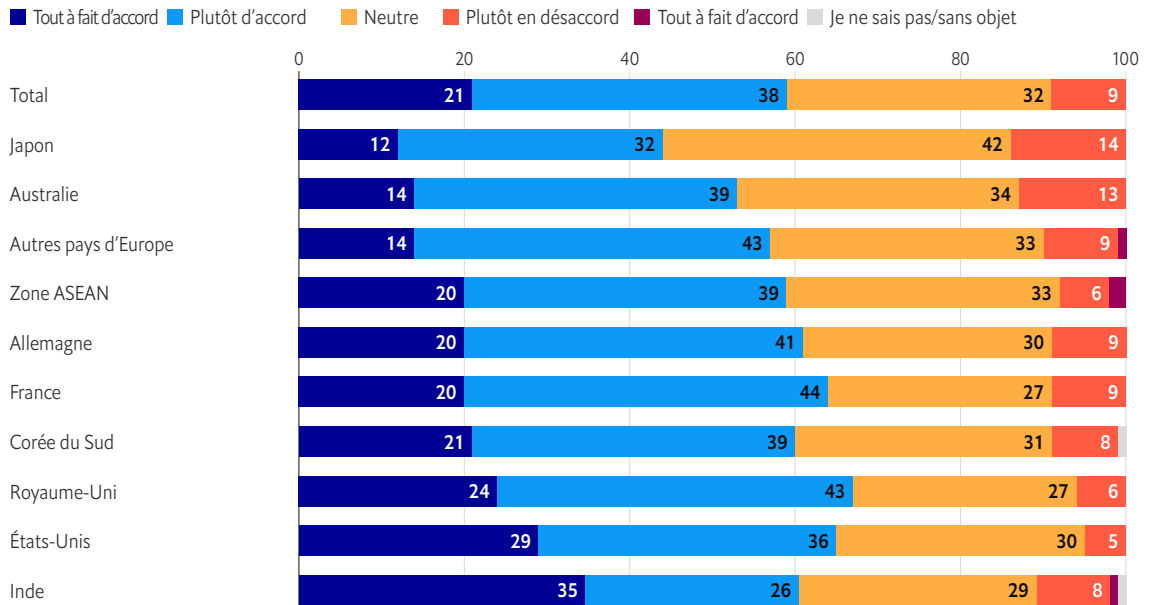
Les centres d'excellence, catalyseurs d'une coordination à marche forcée

Au-delà des innovations techniques, qui font partie de facto de l'arsenal de gouvernance, les entreprises doivent aussi envisager de créer de nouvelles structures et institutions internes pour coordonner les expériences d'IA et intégrer les bonnes pratiques dans leurs activités. Interrogés dans le cadre de cette étude, de nombreux dirigeants de grandes multinationales ont rappelé à quel point il était important de se doter de systèmes unifiés, de processus clairs et d'une gouvernance solide.

Notre enquête a toutefois révélé que les progrès étaient lents dans le domaine : 40 % des participants qualifient d'insuffisants les processus de leur organisation visant à garantir la sécurité et la conformité de l'IA (Figure 13). Des entreprises de toutes tailles affichent le même ressenti à propos de la qualité de leur gouvernance. On observe toutefois des écarts prononcés entre les différentes régions : les participants de l'Inde (35 %) et des États-Unis (29 %) étaient plus nombreux que la moyenne mondiale (21 %) à affirmer que leurs processus étaient suffisants.

Figure 13 : Quatre personnes sur dix estiment que la gouvernance de l'IA est insuffisante au sein de leur organisation

Mon organisation a mis en place des processus suffisants pour garantir la sécurité et la conformité de l'IA. [D'accord ou pas d'accord]
 Pourcentage de l'ensemble des participants



Source : Economist Impact
 La région ASEAN comprend la Malaisie, les Philippines, Singapour et la Thaïlande
 Les autres pays d'Europe comprennent le Danemark, la Finlande, l'Italie, les Pays-Bas, la Norvège, l'Espagne et la Suède.

« L'IA générative nous offre l'opportunité de réinventer tout un processus métier. Elle encourage les décideurs à remettre en question les décisions passées et à accepter qu'il puisse exister une autre approche plus performante. »

Ryan Snyder, Vice-président senior et directeur des systèmes d'information, Thermo Fisher Scientific

Les centres d'excellence (COE) s'affirment comme l'approche de choix pour opérer une surveillance systématique. Les COE sont devenus de puissants catalyseurs pour la génération d'insights et la création d'applications d'IA dans les organisations. Au-delà de leur rôle de centre de commande de l'IA, ils surveillent, mettent en œuvre et diffusent les bonnes pratiques dans l'ensemble des départements et des unités commerciales. Ils favorisent également l'exploration et l'innovation dans le développement et l'opérationnalisation de l'IA.³²

M. Ramani d'Accenture explique que son entreprise a collaboré avec une banque nationale pour créer un centre d'excellence. L'objectif était de définir des directives et des protocoles pour encadrer l'adoption et l'utilisation de l'IA dans les unités commerciales de la banque. Si ces centres peuvent apporter de la cohérence aux pratiques et mettre en place des directives et des garde-fous, leur rôle n'est pas seulement de freiner le développement de l'IA. Au contraire, selon M. Ramani, ils peuvent même devenir de véritables moteurs pour la diffusion de l'IA dans l'ensemble de l'entreprise. En réunissant les différents services, des RH aux équipes techniques en passant par le service juridique, ils libèrent son potentiel. Le COE est une initiative collaborative qui rassemble des équipes multidisciplinaires de différents horizons autour d'un objectif commun. Le centre d'excellence Analytics de Shell, par

exemple, rassemble des équipes aux multiples facettes et appuie une communauté de plusieurs centaines de data scientists et de milliers de « passionnés d'IA », comme les appelle l'entreprise.

Techniquement, les COE développent, mettent en production et supervisent les modèles d'IA en s'intéressant à leur impact et à leur rentabilité. Ils supervisent les opérations de base, notamment en lien avec les données et le ML, et cherchent des moyens d'optimiser l'architecture et l'infrastructure pour gagner en efficacité opérationnelle. Plus important encore, ils unifient la vision et les ambitions en matière d'IA de l'ensemble des unités commerciales et des canaux de l'entreprise, tout en garantissant le respect des règles de sécurité et des normes.

Des initiatives telles que les COE illustrent à quel point les entreprises peuvent profiter de l'avènement de l'IA pour revoir tous les aspects de leur gouvernance et de leur prise de décision. « L'IA générative nous offre l'opportunité de réinventer tout un processus métier, » déclare M. Snyder de Thermo Fisher Scientific. « Elle encourage les décideurs à remettre en question les décisions passées et à accepter qu'il puisse exister une autre approche plus performante. »

32 Deloitte, « Is your AI center of excellence still a center of experimentation » (Votre centre d'excellence en IA est-il toujours un pôle d'expérimentation ?), <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/consulting/us-is-your-ai-center-of-excellence-still-center-of-experimentation.pdf>

Équilibrer l'humain et la machine

La marche en avant des machines suscite depuis longtemps des craintes de pertes d'emplois généralisées, l'automatisation menaçant de rendre les travailleurs humains obsolètes. Pourtant, à l'heure où les technologies d'IA gagnent du terrain dans tous les secteurs d'activité, la réalité se montre plus nuancée. De nombreux dirigeants d'entreprise insistent sur le fait que l'adoption de l'IA n'entraînera pas de licenciements. « L'IA nous rend plus efficaces et nous permet de créer davantage d'articles. Notre intention n'est pas de l'utiliser pour générer du contenu, mais bien pour automatiser des tâches manuelles ou sans intérêt, » précise M. Bhakta de Condé Nast. « Plutôt qu'une dizaine d'articles par semaine, nous pourrions en produire trente. C'est un réel avantage pour notre entreprise. Nous ne cherchons pas à réduire les effectifs, nous cherchons à nous développer. Pouvoir faire plus avec nos équipes actuelles représente le meilleur scénario possible. »

L'adoption massive de l'IA va accroître la demande en rôles spécialisés : data scientists, développeurs et experts MLOps, notamment. Les entreprises versent déjà des primes aux salariés qui possèdent des compétences liées à l'IA. Et ceux qui acquièrent des compétences dans le domaine constatent une augmentation moyenne de 21 % de leur salaire.^{33,34} Selon une prévision pour 2030, 375 millions de travailleurs pourraient avoir besoin de changer de voie professionnelle, en raison des bouleversements liés à la numérisation, l'automatisation et l'IA dans leur secteur d'activité.³⁵

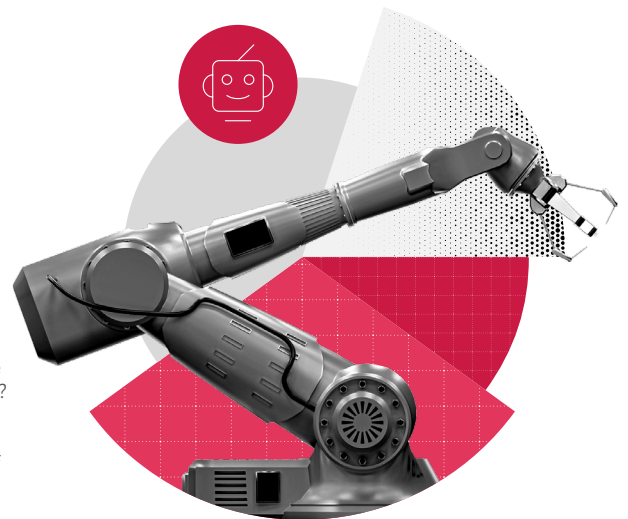
Selon notre enquête, le recrutement reste l'un des plus grands défis pour les entreprises qui déploient l'IA. Seule une entreprise sur six a confiance dans sa capacité à attirer l'expertise nécessaire pour exploiter pleinement les technologies de l'IA. Ceux qui y parviennent déploient l'IA à un rythme beaucoup plus soutenu.

Parallèlement à ces considérations, la prolifération de l'IA générative est également considérée comme une force de démocratisation. Plus de 97 % des architectes d'entreprise prédisent que le personnel non technique utilisera des outils en langage naturel pour interagir avec des ensembles de données complexes au cours des trois prochaines années, et ils sont près de six sur dix (58 %) à affirmer que le langage naturel sera le principal moyen d'y parvenir, sinon le seul. Unilever, par exemple, investit à la fois dans les talents et dans la culture à travers un programme qui a formé près de 20 000 personnes à l'utilisation de l'IA générative. Pour une entreprise qui exerce dans un secteur dynamique marqué par une forte demande en nouvelles compétences, c'est une étape décisive, selon M. Hill.

« Une culture des données et de l'IA aide tous les secteurs de l'entreprise à comprendre l'importance que nous accordons aux décisions basées sur les données. Et c'est ce qui nous aidera à obtenir les insights nécessaires pour améliorer les performances, » explique M. Roberts de Seven West Media. « [Notre] priorité est de former et de perfectionner nos équipes et nos talents à l'utilisation de l'IA générative pour banaliser ce produit et les processus qui l'entourent dans nos pratiques. »

Seule une entreprise sur six a confiance dans sa capacité à attirer l'expertise nécessaire pour exploiter pleinement les technologies de l'IA. Ceux qui y parviennent déploient l'IA à un rythme beaucoup plus soutenu.

- 33 Foote Partners LLC, « 2024 IT Skills & Certifications Pay Index » (Indice de rémunération des compétences et certifications informatiques 2024), <https://footepartners.com/collections/pay-index>
- 34 Fabian Stephany, Ole Teutloff, « What is the price of a skill? The value of complementarity » (Combien coûte une compétence ? La valeur de la complémentarité), Research Policy, Volume 53, édition 1, 2024, <https://doi.org/10.1016/j.respol.2023.104898>.
- 35 McKinsey, « Retraining and reskilling workers in the age of automation » (Formation et reconversion des travailleurs à l'ère de l'automatisation), 2018, <https://www.mckinsey.com/featured-insights/future-of-work/retraining-and-reskilling-workers-in-the-age-of-automation>



Des entreprises comme Condé Nast et Frontier misent avant tout sur l'engagement des collaborateurs. Elles mettent en avant les avantages de l'automatisation afin qu'elle ne soit plus perçue comme une menace. Cette approche montre comment l'IA, loin de supprimer des postes, peut alléger les charges de travail et améliorer les performances professionnelles.

Des entreprises intègrent des outils d'IA dans leurs logiciels de bureau pour faciliter certaines tâches. « Notre objectif est de démocratiser les données. Tous les membres de l'entreprise doivent être en mesure de mobiliser des données pour éclairer leurs décisions... même s'ils ne sont pas ingénieurs ni data scientists, » souligne M. Bensaid chez Rivian. « Traditionnellement, l'IA était réservée à des équipes spécialisées dans des environnements experts, » rappelle M. Bucher de Novartis. « L'IA générative est accessible à de nombreux utilisateurs, via des outils explicites et des cas d'usage intelligibles. »

Une fois qu'une nouvelle technologie a été développée et qu'elle a démontré son efficacité, il reste une étape cruciale : mettre au point des outils en libre-service afin que les employés puissent en bénéficier, comme l'explique M. Casado de Repsol. « Les techniciens n'ont pas de problème avec la technologie ; le véritable défi sera de faire en sorte qu'elle soit utilisable par tous, » confie-t-il.

Les produits d'intelligence artificielle de Repsol donnent les clés de l'IA aux salariés. Ils leur permettent de s'attaquer à des tâches de données complexes en toute autonomie, sans connaissances techniques.³⁶ Les équipes peuvent analyser les performances et optimiser les opérations et les changements organisationnels en utilisant des processeurs en langage naturel. Gages d'autonomie et d'agilité, ces outils permettent à Repsol de rationaliser ses opérations tout en transformant ses collaborateurs en « data scientists citoyens ».

Les plateformes en libre-service donnent aux collaborateurs la possibilité de gérer des tâches jusque-là réservées aux spécialistes, comme le traitement de la paie, la planification de réunions ou des calculs complexes. Pour les employeurs, les avantages sont nets : une efficacité accrue, moins d'intermédiaires et des coûts plus bas. Mais il ne suffit pas de mettre l'IA à la disposition des équipes. La démocratisation de l'IA nécessite une meilleure maîtrise des données. Pour les architectes interrogés dans le cadre de notre enquête, le contrôle de l'accès aux données (49 %) et la protection des données sensibles (42 %) figurent en tête des grands défis en matière de gestion des données.

**« Notre objectif est de démocratiser les données
Tous les membres de l'entreprise doivent
être en mesure de mobiliser des données
pour éclairer leurs décisions... même s'ils ne
sont pas ingénieurs ni data scientists. »**

Wassym Bensaid, Directeur logiciel, Rivian



36 Repsol, « The RAIP Project develops Repsol's own AI solutions with company-wide impact and value » (Avec le projet RAIP, Repsol développe ses propres solutions d'IA pour produire un impact et de la valeur à l'échelle de l'entreprise), <https://www.repsol.com/en/technology-and-digitalization/digital-transformation/digital-program/raip/index.cshhtml>

« L'idée selon laquelle la maîtrise des données est une compétence spécialisée devient rapidement obsolète, » explique M. Garciga de l'armée américaine. Tous les salariés doivent comprendre les règles de protection des données, l'utilisation des plateformes et le stockage centralisé. L'armée américaine a intégré ces principes dans la formation de ses cadres et admet que la gestion des données sensibles est cruciale dans ses utilisations de l'IA. L'armée l'utilise notamment pour examiner les documents à déclassifier en vertu de la loi sur la liberté d'information. Selon cette loi, les documents importants liés à des événements comme la guerre en Afghanistan doivent être rendus publics. « L'IA peut améliorer la précision de ce processus et nous aider à traiter davantage de documents, dans un souci de transparence maximale, » explique M. Garcia.

« L'idée selon laquelle la maîtrise des données est une compétence spécialisée devient rapidement obsolète. Tous les salariés doivent comprendre les règles de protection des données, l'utilisation des plateformes et le stockage centralisé. L'armée américaine a intégré ces principes dans la formation de ses cadres et admet que la gestion des données sensibles est cruciale dans ses utilisations de l'IA. »

Leonel Garciga, Directeur des systèmes d'information, Armée américaine

Le jugement humain sera toujours indispensable pour surveiller les résultats de l'IA, et les entreprises doivent trouver le bon équilibre entre l'humain et la machine. Deux tiers des organisations interrogées dans notre enquête affirment qu'elles expérimentent encore pour trouver le juste équilibre entre les humains et l'IA. Elles sont 80 % à penser que ces efforts mériteraient d'être plus intenses. M. Sato de NTT Docomo souligne que « les personnes qui utilisent l'IA générative doivent être en mesure d'évaluer ses résultats de manière indépendante et de les gérer de manière responsable ».

Selon notre enquête, c'est dans le secteur des services financiers que les progrès sont les plus avancés, car l'IA doit respecter des directives réglementaires strictes. « Les organisations commencent à utiliser l'IA pour éclairer leurs décisions clés, mais elles doivent toujours rester dans le cadre de modèles de risque convenus. Par exemple, toute approche adoptée par TD pour exploiter l'IA doit respecter notre tolérance au risque et nos obligations réglementaires, » explique M. Martin de TD. C'est également le cas pour les interactions avec les clients. « Dans le monde d'avant, la confiance se nouait en agence, en apprenant à connaître le directeur et en parlant du produit, » explique M. Kapoor de Mahindra Group. « Aujourd'hui, la confiance s'obtient par la cybersécurité et la confidentialité des données ; le client doit savoir que ses données ne seront pas utilisées à mauvais escient. »

Dans les secteurs sensibles, caractérisés par des profils à haut risque et une faible appétence au risque, une erreur de la machine provoque généralement des perceptions extrêmement négatives. Comme le remarque M. Bugaev de Flo Health, il arrive également aux professionnels de santé de commettre des erreurs ou d'avoir différentes réponses à la même question, comme les modèles d'IA. Mais l'on se montre beaucoup moins tolérant vis-à-vis des erreurs et des variations de l'IA que de celles des soignants – et c'est tout à fait compréhensible.

« Les personnes qui utilisent l'IA générative doivent être en mesure d'évaluer ses résultats de manière indépendante et de les gérer de manière responsable ».

Takaaki Sato, Senior executive vice president, chief technology officer, NTT Docomo



Principales conclusions

Les innovations en matière de données améliorent la gouvernance de l'IA. Des techniques telles que les données synthétiques et la RAG aident les organisations à compenser les limites des données, à réduire les biais et à améliorer la qualité des systèmes d'IA. Elles ont le pouvoir de rehausser les performances, de préserver la confidentialité des données et d'enrichir les modèles d'IA avec des informations en temps réel.

Les centres d'excellence (COE) assurent une supervision systématique de l'IA. Les COE jouent le rôle de centre de commande pour la mise en œuvre de l'IA. Ils surveillent le respect des bonnes pratiques et favorisent l'innovation dans tous les départements. Ils unifient la vision qu'une organisation a de l'IA, garantissent le respect des normes de sécurité et optimisent les opérations fondamentales telles que les MLOps. Les COE rassemblent des équipes multidisciplinaires, encouragent la collaboration et réinventent les processus métier.

Il faut trouver le juste équilibre entre l'expertise humaine et les capacités de l'IA. Les organisations misent sur la formation des collaborateurs et la création d'outils d'IA en libre-service pour démocratiser l'utilisation des données. Si l'adoption de l'IA permet de gagner en efficacité, le jugement humain reste indispensable pour surveiller les résultats qu'elle produit et gérer les données sensibles. Les entreprises doivent parvenir au bon mélange entre l'intelligence humaine et celle de la machine, en particulier dans les secteurs hautement réglementés où la confiance et la gestion des risques sont primordiales.

Conclusion

Les cas d'usage internes et les projets pilotes restent un terrain d'essai incontournable.

Si l'on comprend qu'elles soient désireuses de déployer l'IA à grande échelle, les entreprises restent conscientes des risques et des inconnues. Les experts préconisent de privilégier les cas d'usage internes dans un premier temps, via des sandboxes et des itérations. Cette phase d'apprentissage permet également de déterminer quelle combinaison de modèles et d'outils fonctionne le mieux selon les objectifs à atteindre. La sélection des projets pilotes doit être déterminée par les seuils de risque, afin de résoudre des problèmes clairs dans un périmètre bien défini.

Moderniser l'infrastructure. Pour bâtir un socle de données et des fondations techniques robustes, les entreprises doivent investir dans la refonte de leur infrastructure pour accueillir de grands volumes de données variées. Les systèmes de stockage de données traditionnels, particulièrement rigides, sont être mis à rude épreuve : ils sont incapables de gérer la diversité des données, leur utilisation en temps réel et leur exploitation à grande échelle à l'ère de la démocratisation. Les grandes entreprises et les grands groupes doivent notamment trouver le juste équilibre entre centralisation et décentralisation pour faire le meilleur usage des plateformes cloud.

L'IA a besoin de KPI. Pour éviter les dépenses inconsidérées et les désillusions qui en découleraient inévitablement, les entreprises doivent développer des indicateurs de rentabilité intelligents et une approche disciplinée des dépenses. La productivité, le contrôle des coûts, le revenu par travailleur, l'expérience utilisateur et la réduction de l'épuisement professionnel du personnel sont autant de métriques pertinentes. Les leaders sauront utiliser l'IA pour générer de nouvelles sources de revenus et inventer des modèles commerciaux innovants.

Misez sur des calendriers flexibles pour évaluer les rendements. Les KPI et les métriques ont leur importance, mais les entreprises ne doivent pas nécessairement rechercher des retours rapides. Les essais et les erreurs sont inévitables dans les expérimentations. Les entreprises des secteurs fortement réglementés comme celui de la santé doivent exercer leur activité au sein de systèmes structurés de surveillance réglementaire. Il faut du temps pour développer les ensembles de données de haute qualité et la gouvernance nécessaires pour alimenter une IA efficace. Et il est tout à fait judicieux de prendre le temps d'observer l'évolution de l'écosystème des produits et des services d'IA.

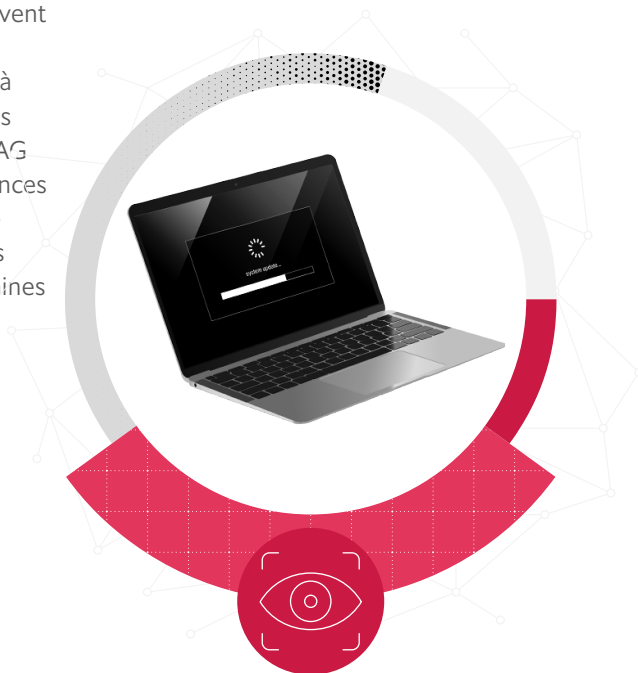
Combinez les modèles pour affûter votre meilleur atout : vos données internes et votre savoir-faire organisationnel.

Les modèles de fondation se généralisent, et leur adoption cesse d'être un catalyseur pour devenir la norme. Toutes les organisations doivent trouver leur avantage en composant une palette unique de données et de savoir-faire. Dans les secteurs hautement spécialisés, les modèles prêts à l'emploi ne seront pas à la hauteur des exigences de qualité et de conformité. Les entreprises peuvent combiner des IA open source et en source fermée, de même que les clouds privés et publics, en fonction de leurs critères de performance, de gouvernance, de coût et de capacité.

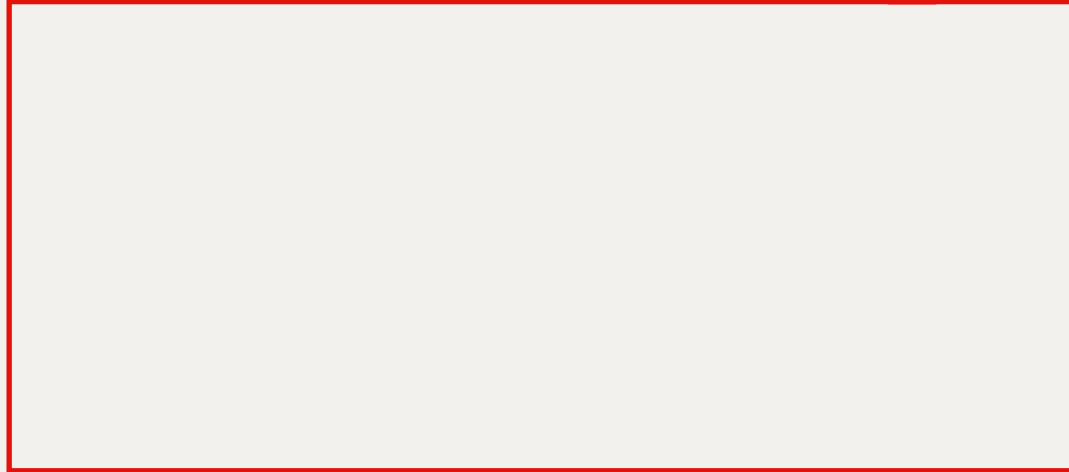
Mettez en place des cadres de gouvernance solides.

Les organisations ont besoin de directives, de protocoles et de mécanismes de surveillance clairs, appuyés par des innovations institutionnelles comme les centres d'excellence (COE) pour gérer le déploiement de l'IA et atténuer les risques. Les COE peuvent également avoir un rôle de catalyseur en améliorant la coordination et la cohérence à l'échelle globale. Les innovations techniques telles que les données synthétiques et la RAG peuvent également renforcer les performances des modèles. Il est essentiel de reconnaître l'importance de la supervision humaine des systèmes d'IA, en particulier dans les domaines où les enjeux sont particulièrement élevés.

Équilibrez l'humain et la machine. Il est essentiel d'impliquer les salariés dans la conception et le déploiement des systèmes d'IA pour s'assurer de leur adhésion et les encourager à s'en servir. Il faut donc communiquer sur le fait que l'IA va les aider dans leur travail, et non les remplacer. Identifier des ambassadeurs de l'IA peut également contribuer à créer une dynamique. En misant sur des programmes et des exercices aux dimensions modestes pour développer de nouvelles compétences et aptitudes, les entreprises pourront combler le fossé qui sépare les spécialistes de l'IA du reste de leurs effectifs. L'IA doit également être intégrée aux programmes de leadership. Enfin et surtout, les organisations doivent encourager l'esprit critique pour éviter les biais d'automatisation et maintenir le jugement humain au premier plan.



Même si nous mettons tout en œuvre pour vérifier leur exactitude et leur fiabilité, Economist Impact ne saurait être tenu responsable des informations, opinions ou conclusions exposées dans le présent rapport. Les conclusions et opinions exprimées dans le rapport ne reflètent pas nécessairement celles du sponsor.



LONDRES

The Adelphi
1-11 John Adam Street
London WC2N 6HT
Royaume-Uni
Tél : (44) 20 7830 7000
E-mail : london@economist.com

GENÈVE

Rue de l'Athénée 32
1206 Geneva
Suisse
Tél : (41) 22 566 2470
Fax : (41) 22 346 93 47
E-mail : geneva@economist.com

SINGAPOUR

8 Cross Street
#23-01 Manulife Tower
Singapour
048424
Tél : (65) 6534 5177
Fax : (65) 6534 5077
E-mail : asia@economist.com

NEW YORK

900 Third Avenue
16th Floor
New York, NY 10022
États-Unis
Tél : (1.212) 554 0600
Fax : (1.212) 586 1181/2
E-mail : americas@economist.com

DUBAÏ

Office 1301a
Aurora Tower
Dubai Media City
Dubai
Tél : (971) 4 433 4202
Fax : (971) 4 438 0224
E-mail : dubai@economist.com

WASHINGTON DC

1920 L street NW Suite 500
Washington DC
20002
E-mail : americas@economist.com

HONG KONG

1301
12 Taikoo Wan Road
Taikoo Shing
Hong Kong
Tél : (852) 2585 3888
Fax : (852) 2802 7638
Email: asia@economist.com