



La chaîne d'approvisionnement agentique

La plus grande tendance liée aux opérations de la chaîne d'approvisionnement est maintenant à portée de la main

Sarthak Pany | Elizabeth Baker | Mahendra Dedasaniya | Nihar Dalmia
Associés, Deloitte Canada

Avril 2026

Sommaire de gestion

Les obstacles qui empêchaient l'IA appliquée à la chaîne d'approvisionnement d'être déployée dans un contexte de production ont maintenant été surmontés. Pas partiellement. Trois obstacles sur quatre, de manière décisive.

30 %

Hausse de productivité chez
C.H. Robinson, 2023–2024

L'IA prédictive a amélioré la qualité des décisions prises par les humains, l'IA générative a accéléré la façon dont ceux-ci s'y préparaient et l'IA agentique retire désormais certaines décisions du flux de traitement humain. Gartner prévoit que d'ici 2028, un tiers des applications logicielles d'entreprise intégreront l'IA agentique, ce qui fera en sorte que 15 % des décisions opérationnelles quotidiennes seront prises de manière autonome²⁰. Il ne s'agit pas d'une amélioration progressive des capacités, mais plutôt d'un changement structurel dans le fonctionnement des opérations de la chaîne d'approvisionnement. Trois obstacles qui empêchaient les organisations d'atteindre cette réalité ont jusqu'ici été surmontés :

1. Construire n'est plus le défi principal. Les plateformes d'agents à programmation schématisée, notamment Microsoft Copilot Studio et Google Vertex AI Agent Builder, permettent désormais aux professionnels de la chaîne d'approvisionnement de décrire un flux de travail en langage clair et de déployer un agent fonctionnel en quelques jours sans nécessairement détenir d'expertise en science des données. Le bassin de concepteurs s'est élargi pour inclure non seulement des spécialistes, mais aussi des personnes qui comprennent les opérations. Les plateformes d'entreprise que les organisations possèdent déjà – systèmes de planification avancée, PGI et solutions spécialisées de chaîne d'approvisionnement – intègrent désormais l'IA agentique comme capacité standard. La technologie n'est donc plus le facteur limitant principal.

2. Il existe désormais des preuves de l'efficacité à grande échelle. Le parc de plus de 30 agents d'IA de C.H. Robinson a traité plus de 3 millions de tâches d'expédition de fret en 2025, produisant une hausse de productivité documentée de 30 %. Walmart exploite des agents autonomes dans la gestion des stocks, la négociation avec les fournisseurs et la logistique. Les résultats financiers correspondent à une précision de gestion des stocks alimentée par l'IA : une croissance des ventes de 5 % pour une hausse des stocks de seulement 2,6 %¹⁷. Ce ne sont pas des vitrines d'innovation. Ce sont des déploiements en contexte de production dont les résultats sont vérifiables par des directeurs financiers.

3. La discipline de réalisation de valeur a été éprouvée. La plupart des investissements en IA ont échoué non pas parce que la technologie était insuffisante, mais parce que personne n'avait défini les résultats financiers attendus avant d'amorcer la conception. Les organisations qui ont déployé l'IA à grande échelle ont appliqué une approche que les équipes de direction connaissent déjà : un indicateur précis, un point de référence défini et un calendrier d'évaluation. À la section 7 du présent document, cette méthodologie est présentée sous le nom « Clarté, intention et transparence ». Ce n'est pas une théorie nouvelle. C'est la rigueur de l'investissement en capital appliquée au déploiement de l'IA.

Le quatrième obstacle n'a toutefois toujours pas été levé. Plus de 70 % des organisations ont déployé l'IA sans repenser les postes, les flux de travail, ni les pouvoirs décisionnels que l'IA était censée transformer⁶. Les gains d'efficacité individuels s'évaporent avant même de pouvoir se refléter dans les résultats financiers. Les organisations qui ont implanté l'IA ont fait un choix délibéré : elles ont repensé le travail en fonction des capacités des agents et ont confié la responsabilité des résultats à un dirigeant opérationnel plutôt qu'à une équipe de projet de TI. Ce choix ne s'achète pas auprès d'un fournisseur. Il doit être fait par un chef de la chaîne d'approvisionnement.

Le présent document définit cinq décisions séquentielles que les chefs de la chaîne d'approvisionnement doivent prendre pour combler l'écart entre le projet pilote et la mise en production. Les conditions qui bloquaient autrefois ces décisions ont changé. Ce qui reste, c'est la volonté de les prendre.

1. L'écart pour la mise en production

- *Quatre obstacles structurels expliquent pourquoi l'IA dans la chaîne d'approvisionnement a relevé davantage de l'expérimentation que d'une valeur pour l'entreprise.*
- *Trois obstacles ont été franchis. Le quatrième – la refonte du travail – demeure du ressort de la direction, et non de la technologie.*
- *Comprendre ce qui a changé est le préalable à toute décision sur la voie à suivre.*

Obstacle n° 1 : La technologie et le parcours n'étaient pas prêts

Lorsque la plupart des organisations de chaîne d'approvisionnement ont expérimenté l'IA pour la première fois, la technologie exigeait des équipes spécialisées en science des données, des cycles de développement de plusieurs mois et une infrastructure sur mesure. Le parcours de la validation de concept à la production était mal compris, et peu d'organisations disposaient du talent, de la maturité en matière de données ou du niveau de préparation sur le plan architectural nécessaires. Cet obstacle a bel et bien été franchi.

Le contexte auquel faisaient face les premiers adoptants était réellement inhospitalier. L'IA prédictive exigeait une ingénierie de caractéristiques artisanale, des pipelines de modèles personnalisés et un réentraînement continu, des tâches devant être accomplies par des ressources en science des données qui, en plus de se faire rares, travaillaient dans des environnements spécialisés déconnectés des opérations quotidiennes de la chaîne d'approvisionnement. Des cycles de déploiement mesurés en mois, voire en années, étaient courants, et chaque modèle servait une fonction étroite : une prévision de la demande par ici, un signal de maintenance prédictive par là. L'écart entre un modèle fonctionnel et un flux de travail opérationnel intégré était considérable.

Les défis liés aux données aggravaient le problème. Les données de la chaîne d'approvisionnement étaient fragmentées entre les progiciels de gestion intégrés, les systèmes de gestion d'entrepôts, les systèmes de gestion de la formation et les systèmes des fournisseurs, souvent dans des formats incohérents, avec des latences variables et une exactitude discutable. Créer une IA sur cette base exigeait un travail d'ingénierie des données considérable avant même de commencer le développement de modèles. L'IA générative a commencé à alléger ce fardeau : les grands modèles de langage peuvent désormais interpréter des données non structurées, par exemple des courriels, des fichiers PDF, des spécifications et des documents d'expédition, qui nécessitaient auparavant une extraction et une normalisation manuelles avant de pouvoir être intégrées à des modèles analytiques.

La contrainte de talents était tout aussi réelle. L'IA traditionnelle exigeait une combinaison d'expertise statistique approfondie que peu de professionnels de la chaîne d'approvisionnement possédaient, et d'une connaissance du domaine que peu de scientifiques de données souhaitaient acquérir. Cela créait un goulot d'étranglement persistant : les équipes opérationnelles comprenaient le problème, mais ne pouvaient pas créer les solutions, et les équipes techniques pouvaient créer les solutions, mais ne comprenaient pas le contexte opérationnel.

Ce qui a changé : l'obstacle technologique s'est largement résorbé.

Trois évolutions ont significativement atténué cet obstacle :

1. Les plateformes d'entreprise intègrent désormais l'IA agentique. SAP Joule, Oracle Fusion, Kinaxis Maestro et Blue Yonder offrent des agents d'IA en tant que capacités standards, bien que l'étendue et la maturité de ces capacités varient selon la plateforme et le cas d'usage. L'adoption ne requiert plus de pile technologique parallèle.

2. Les plateformes à programmation schématisée ont comprimé les cycles de développement. Microsoft Copilot Studio, Google Vertex AI Agent Builder et Amazon Bedrock Agents permettent le développement d'agents en quelques semaines plutôt qu'en quelques mois, élargissant le bassin de concepteurs pour inclure non seulement des scientifiques de données, mais aussi des développeurs-utilisateurs. Par exemple, une compagnie ferroviaire de fret a utilisé Copilot Studio pour créer un assistant d'IA qui a traité plus de 4 000 conversations clients dans ses 45 premiers jours¹.

3. L'IA générative a allégé le fardeau de la préparation des données. Les grands modèles de langage peuvent travailler avec des données imparfaites et semi-structurées qui auraient été inutilisables dans les pipelines traditionnels d'apprentissage automatique.

Cependant, les chefs de la chaîne d'approvisionnement doivent valider avec soin les affirmations des fournisseurs : Gartner estime que seule une fraction des produits que les fournisseurs décrivent comme relevant de l'IA agentique offrent véritablement des capacités autonomes et prévoit que plus de 40 % des projets d'IA agentique seront abandonnés d'ici 2027 en raison des coûts qui excéderont la valeur produite⁴⁶.

Obstacle n° 2 : Les projets pilotes étaient évalués en fonction des activités, et non des résultats financiers

Selon notre expérience, la majorité des projets pilotes d'IA dans la chaîne d'approvisionnement étaient évalués selon des indicateurs techniques, soit la précision du modèle, la vitesse de traitement et le taux d'adoption par les utilisateurs, plutôt que selon leur apport aux résultats financiers. Une déconnexion structurelle s'est ainsi créée : des projets pouvaient démontrer une performance technique impressionnante tout en générant une valeur négligeable pour l'entreprise.

La cause profonde était l'absence de cadres rigoureux reliant les capacités de l'IA aux résultats financiers. Les organisations traitaient l'investissement en IA comme une dépense technologique aux rendements incertains plutôt que comme un dossier de décision exigeant la même discipline qu'un investissement en capital ou une refonte de processus. Les équipes des projets pilotes optimisaient ce qu'elles pouvaient mesurer – latence d'inférence, pourcentages de précision des prévisions et taux d'automatisation – parce que personne n'avait défini les résultats financiers que les projets devaient produire.

Ce qui a changé : les organisations qui ont adopté l'IA traitent la valeur comme une contrainte de conception, et non comme un espoir à concrétiser après le déploiement.

40 %

des projets d'IA agentique seront abandonnés d'ici 2027 (Gartner)

Les organisations présentées dans ce document, soit C.H. Robinson et Walmart, partagent une approche commune : elles ont défini les résultats financiers attendus avant de créer la solution. C.H. Robinson n'a pas déployé des agents de soumission pour démontrer les capacités de l'IA générative; l'entreprise les a déployés pour réduire le délai de soumission de plusieurs heures à une moyenne de 2 minutes et 13 secondes, parce que les retards de soumission augmentent le prix au comptant de 23 % à 25 %². Walmart n'a pas mis en œuvre le système autonome de rééquilibrage des stocks comme vitrine d'innovation; l'entreprise l'a déployé pour éliminer le gaspillage causé par les surplus stockés dans les mauvais magasins³. Il existe désormais des cadres qui permettent cette rigueur, y compris la méthodologie « Clarté, intention et transparence » détaillée à la section 7, et qui ont été éprouvés en contexte de production.

Obstacle n° 3 : La gouvernance et la confiance ne suivaient pas le rythme

Comme les outils d'IA se multipliaient plus vite que les politiques organisationnelles, les fonctions de chaîne d'approvisionnement faisaient face à un double défi : les cadres de gouvernance conçus pour l'analytique traditionnelle ne pouvaient pas gérer les risques de la prise de décisions autonome, et aucun processus visant à déterminer le niveau de confiance nécessaire pour déléguer de telles décisions à l'IA n'avait été établi.

L'IA fantôme, c'est-à-dire l'adoption d'outils d'IA par des individus ou des équipes en dehors des structures de gouvernance approuvées, est devenue un risque tangible. Le rapport d'IBM de 2025 intitulé *Cost of a Data Breach* a révélé que 63 % des organisations ne disposaient pas de politiques de gouvernance pour gérer l'IA et éviter l'IA fantôme⁴. Dans le domaine de la chaîne d'approvisionnement, les conséquences sont opérationnelles : un agent non autorisé qui modifie les paramètres de réapprovisionnement, renégocie les conditions des fournisseurs ou reclassifie le fret sans piste de vérification crée une brèche que la gouvernance des TI traditionnelle n'a jamais été conçue pour détecter.

La confiance est une dimension moins visible, mais tout aussi importante. Même en présence de structures de gouvernance, les organisations peinent à définir quand et comment accorder leur confiance aux systèmes autonomes. On ne peut pas nécessairement faire confiance à un agent d'approvisionnement qui effectue parfaitement des transactions de dépenses marginales pour négocier des contrats stratégiques avec des fournisseurs. Sans cadre structuré pour accorder graduellement la confiance, le choix par défaut est soit la supervision humaine perpétuelle (ce qui entrave complètement l'efficacité), soit l'autonomie complète prématurée (ce qui crée un risque non géré).

Ce qui a changé : la gouvernance dispose désormais de leviers réglementaires et de cadres pratiques.

Des cadres réglementaires émergent, imposés par la Loi européenne sur l'intelligence artificielle (détaillée à la section 6), et une réglementation nord-américaine est attendue. Plus important encore, des outils de gouvernance pratiques existent désormais. L'échelle d'autonomie présentée à la section 6 offre un mécanisme structuré pour gagner et étendre la confiance de manière progressive, remplaçant le choix binaire entre la supervision humaine complète et l'autonomie prématurée.

Il convient de noter que la gouvernance de l'IA demeure un chantier. Des enquêtes récentes indiquent que l'utilisation non autorisée d'outils d'IA reste répandue, même dans les organisations dotées de politiques officielles, bien que l'ampleur varie considérablement selon les secteurs et la maturité de la mise en application. Les sondages suggèrent que jusqu'à 70 % des entreprises reconnaissent que leur gouvernance de l'IA n'est pas pleinement optimisée³¹. Le changement ne réside pas dans le fait que le problème de gouvernance soit résolu ou non, mais dans le fait que les outils, les cadres et les structures réglementaires pour y remédier existent désormais. Les organisations proactives bâtissent leur capacité de gouvernance en amont de la mise en application, reconnaissant que le coût de la conformité réactive dépasse largement celui d'un investissement hâtif.

Obstacle n° 4 : Les gains individuels ne se sont pas traduits en gains organisationnels

C'est l'obstacle qui n'a toujours pas été surmonté. Tous secteurs confondus, une grande majorité d'organisations, soit plus de 70 % selon les estimations récentes, ont déployé l'IA sans repenser les postes, les flux de travail, ni les pouvoirs décisionnels que l'IA était censée transformer⁶. Il en résulte une tendance que les chercheurs décrivent comme un paradoxe de productivité : les employés déclarent des gains de temps significatifs, mais ces gains ne se traduisent pas en gains de productivité organisationnelle parce que le travail lui-même n'a jamais été restructuré pour absorber la capacité libérée⁷.

+ de 70 %

des organisations ont déployé l'IA sans repenser le travail

La cause profonde n'est ni les données, ni les modèles, ni les talents, bien que la direction jette souvent le blâme sur ces aspects. Les projets pilotes piétinent parce que personne n'est imputable pour définir comment et quand l'IA devrait prendre des décisions. Tant qu'une organisation ne délèguera pas explicitement des pouvoirs décisionnels à l'IA et ne modifiera pas le processus au niveau de la tâche, la mise à l'échelle stagnera indépendamment de la capacité de la technologie.

Le mécanisme est désormais bien documenté, et il est propre aux organisations qui n'ont pas repensé le travail. Un rapport de Workday de janvier 2026 a constaté que si 85 % des employés économisent une à sept heures par semaine avec l'IA, près de 40 % de ces économies de temps sont ensuite perdues à corriger, vérifier et réécrire des résultats d'IA de piètre qualité, un phénomène que les

40 %

du temps épargné grâce à l'IA est perdu à corriger, à vérifier et à réécrire les résultats de l'IA

chercheurs ont qualifié de « taxe de l'IA sur la productivité »³². Une étude distincte portant sur des développeurs logiciels expérimentés a constaté que les outils d'IA les rendaient 19 % plus lents sur les tâches de programmation, alors que les développeurs croyaient être 20 % plus rapides, un écart de 39 points de pourcentage entre perception et réalité³³. Lorsque les équipes assistées par l'IA accomplissent davantage de tâches, le temps consacré à l'approbation et à la révision humaines de ces tâches peut augmenter jusqu'à 91 %, créant de nouveaux goulots d'étranglement ailleurs dans le flux de travail³⁴.

Considérons la différence entre superposer l'IA à un flux de travail et repenser le flux de travail autour de l'IA. Dans l'approvisionnement traditionnel relatif aux dépenses marginales, un achat courant suit un parcours prévisible : demande d'achat, processus de sollicitation à trois soumissions, évaluation

des réponses par l'acheteur, approbation du gestionnaire, émission du bon de commande. Cinq transferts sont effectués pour chaque transaction, quel qu'en soit le montant. Lorsque les organisations superposent un assistant d'IA à ce flux de travail, cet assistant peut rédiger l'appel d'offres plus rapidement ou résumer les soumissions plus efficacement, mais la structure à cinq transferts reste intacte, et la compression du délai total est marginale.

Considérons maintenant le flux de travail repensé : un agent IA gère l'approvisionnement relatif aux dépenses marginales et indirectes de manière autonome selon des balises de gouvernance établies, une liste de fournisseurs approuvés, des limites de dépenses par catégorie et des conditions contractuelles standards. Les acheteurs sont redéployés vers l'approvisionnement stratégique et les négociations complexes où le jugement, la gestion des relations et la créativité commerciale créent une valeur que l'IA ne peut pas reproduire. Selon les estimations du secteur, de 60 % à 80 % des bons de commande courants et conformes aux politiques pourraient faire l'objet d'une chaîne d'approbation selon laquelle la gouvernance est intégrée dès la conception plutôt que seulement à l'exécution. C'est la différence entre accélérer le travail existant et créer une nouvelle méthode de travail.

Cette conclusion est cohérente avec les recherches menées à l'échelle mondiale par Deloitte. Une analyse distincte de l'IA agentique dans les chaînes d'approvisionnement industrielles est parvenue à la même conclusion : les organisations devraient « réimaginer et repenser les processus de la chaîne d'approvisionnement autour des forces uniques des humains et des agents, plutôt que de simplement insérer des agents d'IA dans les modèles opérationnels existants »⁴⁸.

Ce qui n'a pas changé : la plupart des organisations n'ont pas pris la décision au niveau de la direction de repenser le travail autour de l'IA.

Il faut toutefois mentionner qu'on observe un élan encourageant du côté de l'état de préparation de la main-d'œuvre. Plus de 60 % des vice-présidents et directeurs ont suivi une formation en IA en 2025, contre 43 % l'année précédente, et 73 % des dirigeants encouragent désormais activement leurs équipes à utiliser l'IA³⁵. De nombreuses organisations ont créé des équipes consacrées à l'IA et de nouveaux rôles. Cependant, préparation à l'IA et refonte du travail ne sont pas synonymes. Former les employés à utiliser les outils d'IA dans des flux de travail inchangés est une étape nécessaire, mais insuffisante. Une refonte fondamentale signifie réingénierie des processus, pouvoirs décisionnels, structures de rôles et indicateurs de performance autour des capacités de l'IA. Selon ce critère, l'obstacle demeure presque intact.

Cet obstacle ne peut être surmonté par une meilleure technologie, davantage de données ou des cadres de gouvernance améliorés. Il exige une décision délibérée de la direction pour cartographier les flux de travail existants, identifier où le jugement humain est irremplaçable, définir à quoi devraient ressembler les rôles et responsabilités dans un modèle opérationnel augmenté par l'IA, et gérer le changement organisationnel requis. La section 9 de ce document aborde directement cette question.

2. Du projet pilote à la mise en production : des preuves tangibles

- Deux organisations démontrent à quoi ressemble aujourd'hui l'IA dans la chaîne d'approvisionnement en contexte de production.
- Toutes deux ont dépassé le stade du projet pilote en traitant l'IA comme un investissement opérationnel, et non comme une expérience technologique.
- Les éléments probants proviennent de renseignements accessibles au public fournis par les entreprises, de communiqués de presse vérifiés et de déclarations de dirigeants nommément cités. L'analyse est celle de Deloitte.

Avant d'examiner ces preuves en détail, la figure 1 cartographie l'éventail des cas d'utilisation de l'IA dans les quatre domaines de la gestion de la chaîne d'approvisionnement : planifier, approvisionner, fabriquer et livrer. Ces cas d'utilisation sont classés selon leur niveau d'autonomie : assistance, exécution supervisée et autonomie complète. Les exemples présentés illustrent les capacités observées sur le marché. Il s'agit de l'orientation adoptée par le marché, et non d'une recommandation de fournisseurs précis. Chaque cas d'utilisation renvoie à sa source principale dans les notes de fin.

Figure 1 : Spectre d'autonomie des cas d'utilisation de l'IA

De l'assistance à l'autonomie complète dans l'ensemble de la chaîne de valeur.²⁹

[AC] Autonomie complète (agit dans le cadre d'une autorité déléguée) • [ES] Exécution supervisée (agit, l'humain supervise) • [A] Assistance (recommande, l'humain décide)

PLANIFIER	APPROVISIONNER	FABRIQUER	LIVRER
<p>[AC] Équilibrage autonome de la demande et des stocks Ingère des données en temps réel pour réajuster les prévisions par UGS et lancer des transferts/bons de commande.</p>	<p>[A] Analyse et saisie des appels d'offres Extrait les détails clés des devis et appels d'offres pour remplir automatiquement les champs d'approvisionnement.</p>	<p>[AC] Replanification intelligente de la production Réajuste automatiquement les calendriers de production lors de perturbations pour minimiser les temps d'arrêt.</p>	<p>[ES] Optimisation dynamique des itinéraires Planifie et ajuste les itinéraires de transport selon les conditions en temps réel.</p>
<p>[A] Planification de la demande assistée par l'IA Utilise l'historique des ventes et des facteurs externes/locaux pour générer des prévisions quotidiennes d'ingrédients.</p>	<p>[A] Résolution des anomalies de bons de commande Signale les articles manquants, les quantités incorrectes ou les expéditions en retard et suggère des correctifs.</p>	<p>[ES] Maintenance prédictive automatisée Analyse les données IdO pour prévoir les défaillances d'équipement et planifier automatiquement la maintenance.</p>	<p>[AC] Gestion automatisée du fret Suit les expéditions pour entreprendre automatiquement la gestion des exceptions et assurer les communications courantes.</p>

<p>[AC] Rééquilibrage des stocks Surveille les prévisions et les stocks en temps réel pour ajuster dynamiquement la production et la distribution.</p>	<p>[AC] Négociation des achats marginaux Agit comme acheteur numérique pour les achats à faible valeur et négocie selon des règles et référentiels prédéfinis.</p>	<p>[ES] Automatisation des processus externes Automatise les transferts de produits semi-finis envoyés à des fournisseurs externes et gère la documentation.</p>	<p>[AC] Planification des opérations d'entrepôt Prévoit la charge de travail et alloue dynamiquement les ressources (quarts de travail et robots de prélèvement).</p>
<p>[AC] Atténuation des retards d'approvisionnement Détection des retards pour demander automatiquement aux fournisseurs d'accélérer, de replanifier la production ou d'ajuster les plans.</p>	<p>[ES] Recherche de fournisseurs Effectue des recherches mondiales assistées par l'IA et évalue des fournisseurs selon des critères d'approvisionnement complexes.</p>	<p>[ES] Planification de production en temps réel Assure la replanification en atelier en tenant compte de la disponibilité des pièces, des quarts et des contraintes de séquence.</p>	<p>[ES] Copilote de livraison du dernier kilomètre Réordonne dynamiquement les séquences de livraison, consolide les itinéraires et informe les clients.</p>
<p>[AC] Réponse aux perturbations d'approvisionnement Surveille l'approvisionnement en amont, réaffecte automatiquement les matériaux et réserve le transport accéléré.</p>	<p>[A] Analyse des risques contractuels Examine les contrats fournisseurs pour signaler les risques, assurer la conformité et suggérer des clauses améliorées.</p>	<p>[ES] Collaboration humain-robot Orchestre les flux de travail entre les robots mobiles autonomes et les travailleurs humains.</p>	<p>[ES] Contrôle qualité autonome Utilise l'IA générative et la vision par ordinateur pour inspecter automatiquement les articles avant l'emballage.</p>
<p>[ES] Autoapprentissage de la planification des ventes et des opérations Exécute des scénarios en continu sur des données actualisées pour simuler les perturbations et recommander des plans de contingence.</p>	<p>[AC] Traitement des factures sans intervention Automatise l'approbation et le rapprochement des factures avec les bons de commande et les reçus pour signaler les divergences.</p>		<p>[AC] Vérification et paiement automatisés du fret Vérifie automatiquement les factures de fret, repère les écarts et traite les paiements.</p>

Source : Compilation à partir d'annonces d'entreprises, de couvertures d'analystes et d'études de cas vérifiées (2024-2026). Les noms d'entreprises sont des exemples de catégories de cas d'utilisation. Voir les notes de fin pour les citations individuelles.

C.H. Robinson : automatisation du cycle de vie du fret

C.H. Robinson, le plus grand courtier en fret au monde selon son volume, a déployé plus de 30 agents exclusifs d'IA générative qui automatisent le cycle complet d'une expédition de marchandises, de la soumission de prix et de l'attribution des chargements à la planification des rendez-vous, aux vérifications du statut en transit et à la classification du fret à chargement partiel⁸. L'entreprise rapporte que ces agents ont effectué plus de trois millions de tâches d'expédition en 2025 seulement⁹.

Le déploiement a débuté par les flux de travail à haut volume dont les données étaient les plus matures. L'agent de soumission fournit des cotations au comptant personnalisées en 2 minutes et 13 secondes, en moyenne, un processus qui nécessitait auparavant des heures de travail manuel. L'entreprise a déclaré publiquement que les retards de soumission augmentent les coûts au comptant de 23 % à 25 %, fournissant un argument financier direct pour l'amélioration de la rapidité¹⁰. Plus d'un million de soumissions ont été traitées par cet agent.

+ de 3 M

de tâches d'expédition effectuées par les agents d'IA chez C.H. Robinson en 2025

L'agent de commandes convertit les appels d'offres non structurés reçus par courriel en commandes complètes et prêtes pour le système en environ 90 secondes, traitant 5 500 commandes de chargements complets par jour et économisant environ 600 heures de travail manuel quotidiennement¹¹. Dans les opérations de chargement partiel, des agents en tandem, l'un contactant les transporteurs pour suivre le fret manqué, le second analysant et exécutant les étapes suivantes, exécutent 100 interactions simultanées avec les transporteurs, exécutant ainsi des centaines d'expéditions quotidiennes pour plus de 11 000 clients. L'entreprise rapporte que 95 % des vérifications manuelles de suivi ont été automatisées, avec une réduction de 42 % des déplacements de retour inutiles des transporteurs¹².

Le chef de la stratégie et de l'innovation de C.H. Robinson a publiquement rapporté une hausse de productivité de 30 % sur la période 2023–2024, avec un objectif supplémentaire de 15 % pour 2025¹³.

Walmart : chaîne d'approvisionnement agentique de bout en bout

L'approche de Walmart en matière d'IA de la chaîne d'approvisionnement est cohérente avec sa stratégie plus large qui consiste à développer la technologie comme un actif concurrentiel plutôt que comme un centre de coûts. L'entreprise est passée d'une IA centrée sur les modèles, c'est-à-dire des algorithmes individuels répondant à des problèmes isolés, à une architecture centrée sur les systèmes dans laquelle des agents spécialisés opèrent dans différents domaines : gestion des stocks, prévision de la demande, optimisation logistique, négociation avec les fournisseurs et opérations en magasin¹⁴.

Le système autonome de rééquilibrage des stocks identifie automatiquement les surplus dans un emplacement et les redirige vers les magasins où la demande est plus forte. Il s'agit d'un flux de travail pleinement autonome : l'agent surveille en continu l'état des stocks, repère les déséquilibres et amorce les transferts sans intervention humaine pour chaque transaction. La supervision humaine est réservée aux exceptions de politique et aux décisions stratégiques de stock. La stratégie d'automatisation plus large de Walmart, qui englobe la robotique, les logiciels pilotés par l'IA et la logistique autonome, vise une amélioration de 20 % des coûts unitaires moyens, environ 60 % des magasins américains recevant désormais des marchandises en provenance de centres de distribution automatisés¹⁵.

Dans le domaine de l'approvisionnement, le déploiement par Walmart de **Pactum AI** pour la négociation autonome avec les fournisseurs illustre ce que représente l'approvisionnement agentique à l'échelle de l'entreprise. Les agents mènent simultanément des négociations asynchrones pilotées par l'IA avec des milliers de fournisseurs de taille moyenne, obtenant un taux d'accord de 68 % et une réduction moyenne des coûts de 1,5 % par contrat négocié, avec des délais de paiement prolongés¹⁶. Le programme s'étend désormais aux négociations de tarifs de transport. Ce sont des résultats que les processus d'approvisionnement manuels ne pourraient pas atteindre à

une échelle comparable; la composante autonome n'est pas un moteur de recommandations, mais plutôt un agent négociateur qui fait des propositions et des contre-offres pour en arriver à des accords selon des paramètres commerciaux définis.

En 2025, Walmart a lancé « Wally », un outil d'IA agentique pour les marchands, conçu pour identifier de manière autonome les causes profondes des problèmes de stocks, tels que les ruptures de stock, avec une rapidité supérieure à l'analyse manuelle³⁰. L'entreprise s'est également associée à Gatik pour déployer des camions autonomes sans conducteur pour le transport intermédiaire à l'échelle commerciale, une première en Amérique du Nord. Des dizaines de milliers de livraisons autonomes depuis la mi-2025 ont été effectuées³⁶.

Les résultats financiers de Walmart fournissent des éléments corroborants à l'échelle de l'entreprise. Au cours de sa plus récente période de rapport, l'entreprise a réalisé une croissance des ventes de 5 % avec une augmentation des stocks de seulement 2,6 %, un ratio qui correspond à une précision de gestion des stocks pilotée par l'IA, permettant la croissance des revenus sans expansion proportionnelle du fonds de roulement¹⁷. L'optimisation des itinéraires par IA a permis de retrancher près de 50 millions de kilomètres des distances totales parcourues, une capacité que Walmart a depuis commercialisée auprès d'autres entreprises sous forme de logiciel-service¹⁸.

Les deux organisations partagent un schéma commun : elles ont commencé par les flux de travail à haut volume dont les données étaient les plus matures, elles ont prouvé la valeur avec des résultats financiers vérifiables, puis elles ont étendu les activités de manière systématique.

3. Les couches de l'IA, en termes simples

- L'IA dans la chaîne d'approvisionnement n'est pas une technologie unique. Ce sont trois couches distinctes reposant sur un ensemble de données.
- Chaque couche sert un objectif différent. Les flux de travail ne nécessitent pas tous ces trois couches.
- Comprendre ces couches évite les deux erreurs les plus courantes : traiter toutes les IA comme équivalentes ou supposer que l'IA agentique remplace ce qui existe déjà.

Avant d'examiner où l'IA devrait résider sur le plan architectural ou comment elle devrait être gouvernée, les dirigeants de la chaîne d'approvisionnement ont besoin d'une taxonomie claire de ce que fait réellement la technologie. Le marché est bruyant, les fournisseurs qualifient tout d'« IA », et les messages commerciaux brouillent souvent les distinctions entre l'analytique prédictive, l'IA générative et l'IA agentique. La présente section lève le voile sur ces ambiguïtés.

Figure 2 : Les couches de l'IA dans la chaîne d'approvisionnement : trois voies vers la création de valeur

Chaque couche est un point de déploiement autonome. Les organisations n'ont pas toutes besoin de ces trois couches.

<p>1</p> <h4>IA prédictive</h4> <p>Modèles d'apprentissage automatique qui prévoient les résultats, détectent les tendances et optimisent les décisions à partir de données historiques structurées.</p> <p>EXEMPLE Modèles de détection de la demande qui ajustent les prévisions par UGS en temps réel selon les signaux des points de vente, la météo et les calendriers promotionnels.</p>	<p>2</p> <h4>IA générative</h4> <p>Grands modèles de langage et systèmes multimodaux qui synthétisent l'information non structurée, génèrent du contenu et augmentent la prise de décisions humaine.</p> <p>EXEMPLE Un copilote de planification qui résume les rapports d'exception des fournisseurs, rédige des actions recommandées et explique les écarts de prévision en langage naturel.</p>	<p>3</p> <h4>IA agentique</h4> <p>Agents logiciels autonomes qui perçoivent, raisonnent, planifient et exécutent des flux de travail multiétapes selon des balises définies, avec une supervision humaine selon des niveaux d'autonomie gagnée.</p> <p>EXEMPLE Un agent d'opérations de fret qui surveille les exceptions d'expédition, réachemine les chargements, renégocie les tarifs des transporteurs et met à jour les délais de livraison aux clients sans intervention humaine.</p>
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Une base fondamentale : des données de qualité décisionnelle

Toutes les voies exigent des données suffisamment fiables pour le flux de travail concerné, et non des données parfaites à l'échelle de l'entreprise.

Note : Toutes les voies exigent des données de qualité décisionnelle. Les flux de travail ne nécessitent pas tous les trois couches, mais les organisations de grande envergure déploieront généralement les trois, car différents flux de travail exigent différentes capacités. Source : Analyse de Deloitte, 2026.

Couche 1 : IA prédictive

L'IA prédictive est le point de départ de la plupart des organisations de la chaîne d'approvisionnement : prévision de la demande, estimation des délais, maintenance prédictive, estimation des heures d'arrivée et évaluation des risques. Ce sont des modèles d'apprentissage automatique entraînés sur des données historiques pour identifier des tendances et prédire des états futurs. L'IA prédictive est puissante, mais passive. Elle indique au planificateur ce qui est susceptible de se produire; elle ne décide pas de la réponse. Pour de nombreux flux de travail, cela demeure le niveau de capacité IA approprié.

Couche 2 : IA générative

L'IA générative fonctionne sur un principe fondamentalement différent. Plutôt que de prédire des résultats à partir de données structurées, elle génère du nouveau contenu à partir d'entrées non structurées : langage naturel, images, documents, code. Dans le domaine de la chaîne d'approvisionnement, ses applications comprennent l'interprétation des communications avec les fournisseurs, la rédaction de réponses aux appels d'offres, la synthèse de rapports d'exception et la conversion d'offres reçues par courriel en commandes structurées. L'IA générative joue également un rôle de pont : elle traite les données non structurées et désordonnées (courriels, documents PDF, contrats, documents d'expédition) qui constituent une part importante des flux d'information de la chaîne d'approvisionnement, rendant le parcours de préparation des données pour l'IA prédictive et agentique considérablement moins pénible. Concrètement, l'IA générative peut interpréter des dossiers fournisseurs qui ne concordent pas entre les systèmes, normaliser les formats de données sans correspondance manuelle et extraire des champs structurés à partir de documents non structurés tels que des manifestes, des bons de commande et des certificats de qualité. Cette capacité transforme l'équation de la préparation des données : les organisations qui stagnaient en raison de données fragmentées peuvent désormais commencer à déployer l'IA tout en améliorant la qualité des données en parallèle.

Couche 3 : IA agentique

L'IA agentique représente un changement qualitatif, non seulement de l'analyse vers l'exécution, mais des outils vers les agents. Trois propriétés distinguent les systèmes agentiques de l'automatisation traditionnelle des flux de travail : des objectifs persistants (les agents maintiennent leurs objectifs dans toutes les sessions et adaptent leur approche), des pouvoirs décisionnels délégués (les agents agissent dans un cadre d'autorité défini sans attendre d'instruction humaine) et des boucles d'apprentissage continu (les agents s'améliorent à chaque interaction et résultat). Gartner définit les systèmes d'IA agentique comme ceux qui planifient et prennent des actions de manière autonome pour atteindre des objectifs définis par l'utilisateur¹⁹. Contrairement à l'IA prédictive (qui prévoit) ou à l'IA générative (qui crée du contenu), l'IA agentique agit : elle exécute des flux de travail multiétapes, interagit avec les systèmes d'entreprise, prend des décisions selon des paramètres définis et transmet des cas à la supervision humaine lorsqu'elle fait face à des conditions hors de son autorité. Dans le domaine de la chaîne d'approvisionnement, les applications comprennent la négociation autonome avec les fournisseurs, la gestion du cycle de vie du fret, le système autonome de rééquilibrage des stocks et la résolution automatisée des exceptions.

Gartner prévoit que d'ici 2028, un tiers des applications logicielles d'entreprise intégreront l'IA agentique, ce qui fera en sorte que 15 % des décisions opérationnelles quotidiennes seront prises de manière autonome²⁰. Une analyse plus récente de Gartner prévoit que, d'ici 2030, la moitié des solutions de gestion de la chaîne d'approvisionnement auront recours à des agents intelligents pour exécuter de manière autonome les décisions au sein de l'écosystème⁴⁷. Les fournisseurs de la chaîne d'approvisionnement intègrent déjà ces agents : SAP Joule couvre l'ensemble du portefeuille infonuagique de l'entreprise avec plus de 400 cas d'utilisation pilotés par l'IA, les agents Kinaxis Maestro fonctionnent nativement au sein de la plateforme de planification, et les agents opérationnels Blue Yonder surveillent et ajustent les opérations en continu²¹.

Ce changement a une implication plus large : l'IA agentique introduit le concept de main-d'œuvre numérique, où les agents sont traités comme une capacité plutôt que comme des outils. Un agent qui traite 5 500 commandes de chargements complets par jour n'est pas une fonctionnalité logicielle; c'est une ressource de main-d'œuvre avec un débit, un taux d'erreur et une utilisation mesurables. Les organisations qui adoptent ce cadrage – gérer la capacité des agents aux côtés de la capacité humaine – trouvent plus facile de justifier l'investissement, de mesurer le rendement et de planifier la mise à l'échelle.

Les trois couches exigent des données de qualité décisionnelle. Les flux de travail ne nécessitent pas tous les trois couches, mais les chaînes d'approvisionnement de plusieurs milliards de dollars déploieront les trois, car différentes décisions exigent différentes capacités.

4. Où doit résider l'IA : les décisions d'architecture

- L'architecture n'est pas une décision technologique. C'est une décision de portefeuille sur les domaines où investir et ceux où accepter ce que le marché offre.
- Cinq voies. Deux niveaux. Un cadre de gouvernance.
- L'erreur la plus coûteuse est de construire ce qui devrait être acheté.

Déployer l'IA agentique dans la chaîne d'approvisionnement implique un choix stratégique qui façonne la structure de coûts, la rapidité de création de valeur, la différenciation concurrentielle et la dette technique à long terme. La présente section présente un modèle architectural à cinq voies organisé en deux niveaux, intégration et conception, chapeauté par un cadre de gouvernance qui s'applique indépendamment de la voie où réside l'IA.

Figure 3 : Le cadre d'architecture mixte : où l'IA devrait-elle résider?

Miser sur l'intégration pour une mise à l'échelle et sur la programmation schématisée pour une accélération. Réserver la conception sur mesure aux véritables bijoux de la couronne.



p. ex. Microsoft Copilot Studio, Google Vertex AI Agent Builder, Amazon Bedrock Agents
Déploiement rapide par les équipes métiers et les développeurs-utilisateurs. Interfaces en langage naturel (programmation intuitive) rendant la création d'agents conversationnelle. Déploiement en jours, pas en mois.

À ENVISAGER LORSQUE :

La rapidité compte, les utilisateurs métiers peuvent gérer la solution et la logique est simple.

p. ex. Google ADK, LangChain, Azure AI Foundry, CrewAI
Personnalisation et différenciation concurrentielle maximales. Réserver aux avantages stratégiques défendables. Requiert des talents spécialisés.

À ENVISAGER LORSQUE :

Le flux de travail constitue un différenciateur concurrentiel et aucune solution du marché ne comble l'écart.

Les noms de fournisseurs sont fournis à titre illustratif et ne constituent pas une recommandation⁴³. La plupart des organisations opèrent sur 2 à 3 voies.
Source : Analyse de Deloitte, 2026.

Niveau 1 : intégration (tirer parti de ce que les plateformes offrent déjà)

IA intégrée au PGI

Pour les organisations utilisant SAP ou Oracle, la première question est de savoir si le cas d'utilisation peut être traité par les capacités que le fournisseur de la plateforme offre déjà. Les agents de SAP Joule couvrent désormais la planification de la production, la maintenance et l'exploitation ainsi que l'intégration des fournisseurs dans les aspects clés de son portefeuille infonuagique²¹. Oracle Fusion SCM offre des capacités préconçues d'IA pour la gestion des stocks et la détection de la demande.

Cette voie est à envisager lorsque le PGI est le système de référence et que le cas d'utilisation est transactionnel à grande échelle. Les capacités d'IA intégrée gagnent rapidement en maturité, mais de manière inégale entre les fournisseurs. Il importe de valider les affirmations en fonction de la préparation à la production, et non des promesses de la feuille de route.

Plateformes spécialisées de chaîne d'approvisionnement

Les plateformes telles que Kinaxis Maestro, Blue Yonder, Manhattan Active et Coupa offrent une intelligence propre au domaine plus poussée que l'IA intégrée au PGI pour les décisions complexes de chaîne d'approvisionnement : planification concurrente sur des réseaux multiniveaux, gestion agentique des exceptions, optimisation de réseau et flux d'approvisionnement autonomes. Les agents Kinaxis Maestro gèrent la détection d'anomalies, les actions prescriptives et la gestion des perturbations d'approvisionnement nativement au sein de la plateforme de planification²¹.

Pour les organisations utilisant déjà ces plateformes, la première étape consiste à évaluer les capacités agentiques offertes par le fournisseur depuis le dernier cycle de mise à niveau. Bon nombre de ces fonctionnalités sont déjà sous licence et disponibles.

Plateformes spécialisées d'IA native

De plus en plus de plateformes intègrent l'IA dès la conception. **Pactum AI** offre la négociation autonome¹⁶; **FourKites**, la visibilité prédictive⁴⁴; et **Celonis**, l'intelligence des processus⁴⁵. Les trois offrent des capacités précises, mais approfondies pour des problèmes spécifiques à haute valeur, avec une rapidité de création de valeur que les plateformes plus génériques ne peuvent pas toujours égaler.

Cette voie est à envisager lorsqu'un problème précis et à haute valeur doit disposer d'une solution spécialement conçue et qu'aucune solution standard du PGI ou des plateformes spécialisées ne comble l'écart.

Niveau 2 : conception (lorsque les capacités standards sont insuffisantes)

Programmation schématisée/agents métiers

Les plateformes de développement à programmation schématisée (Microsoft Copilot Studio, Google Vertex AI Agent Builder, Amazon Bedrock Agents) permettent aux équipes de la chaîne d'approvisionnement de créer des agents ciblés en jours ou en semaines plutôt qu'en mois. Ces plateformes sont conçues pour le développement métier : les professionnels de la chaîne d'approvisionnement qui comprennent le flux de travail peuvent concevoir, tester et itérer des agents sans nécessairement posséder une expertise approfondie en génie logiciel.

La conception de ces agents devient de plus en plus conversationnelle. Des outils comme Copilot Studio permettent aux professionnels de la chaîne d'approvisionnement de décrire leurs besoins en langage clair et d'itérer en temps réel, une évolution que l'on appelle parfois « programmation intuitive », et qui rend le développement métier accessible à une population plus large à chaque version de la plateforme¹. Pour les cas d'utilisation simples, l'obstacle à l'entrée est presque nul.

Cette voie est à envisager pour les flux de travail à haut volume et forte intensité humaine où la rapidité d'adoption compte, où les utilisateurs métiers peuvent gérer la solution et où la logique est simple : triage du service client, requêtes de visibilité des expéditions, acheminement des exceptions, assistance aux flux de travail.

Agents programmés sur mesure

Le développement d'agents sur mesure à l'aide de trousseaux de développement, comme LangChain, Google Gemini ADK et Microsoft Azure AI Foundry, offre une flexibilité et un contrôle maximaux. Le parc de plus de 30 agents de C.H. Robinson a été conçu selon cette approche²².

Cette voie est à envisager pour les quelque 5 % de flux de travail qui différencient véritablement l'entreprise. L'investissement doit être justifié par la valeur stratégique de la capacité : le développement sur mesure exige un coût initial plus élevé, un délai plus long avant la création de valeur et des talents spécialisés à l'interne. Pour les cas d'utilisation réalisables par des solutions intégrées ou à programmation schématisée, opter par défaut pour le développement sur mesure monopolise des ressources qui devraient être affectées ailleurs.

Le cadre de gouvernance

La gouvernance n'est pas une voie, c'est le cadre qui chapeaute les cinq voies. Les droits décisionnels, les pistes de vérification, la sécurité et les exigences d'interopérabilité s'appliquent indépendamment de la voie où réside l'IA.

L'architecture doit faire l'objet d'une gouvernance. En effet, une politique devrait établir un point de contrôle de gouvernance et exiger une justification explicite avant que le développement sur mesure soit approuvé pour tout cas d'utilisation réalisable par des solutions intégrées ou à programmation schématisée. Sans ce point de contrôle, la tendance naturelle des équipes de développement est de construire sur mesure, parce que les conceptions sur mesure sont plus intéressantes, mieux contrôlables et plus faciles à défendre devant les comités de gestion de budget que les décisions de configuration. Le point de contrôle empêche ce réflexe par défaut qui monopolise des ressources qui devraient être affectées à des flux de travail prioritaires.

La décision en matière d'architecture ne consiste pas à choisir entre concevoir et acheter. Il faut plutôt se demander où l'investissement sur mesure crée un avantage défendable et où il crée un coût et une complexité inutiles.

5. Ensemble de données : des données utilisables valent mieux que des données parfaites

- *La qualité des données est l'obstacle le plus cité au succès de l'IA dans le domaine de la chaîne d'approvisionnement, et l'aspect le moins bien compris.*
- *La norme n'est pas de recourir à des données parfaites. Il faut plutôt s'appuyer sur des données de qualité décisionnelle qui sont adaptées au flux de travail automatisé spécifique.*
- *Les organisations qui visent la perfection des données à l'échelle de l'entreprise attendent indéfiniment.*

La question n'est pas de savoir si la qualité des données compte, mais plutôt de connaître la norme pour amorcer le processus : des données de qualité décisionnelle, suffisamment fiables pour soutenir en toute sécurité une décision déléguée précise, suffisent.

Le coût de l'inaction en matière de préparation des données est quantifiable. L'Institute for Business Value d'IBM a constaté que les problèmes de données non résolus peuvent réduire le rendement des projets d'IA de 18 % à 29 %³⁷. Cependant, le coût d'attendre des données parfaites est tout aussi réel : un report indéfini.

Données de qualité décisionnelle : une norme pratique

L'alternative pratique à l'approche « corriger toutes les données d'abord » est le concept de données de qualité décisionnelle, soit des données suffisamment fiables pour soutenir en toute sécurité une décision déléguée précise. Les données de qualité décisionnelle ne sont pas un compromis, mais une norme de précision : des données suffisamment exactes, opportunes, complètes et contextualisées pour le flux de travail autonome qu'elles servent. Une perfection à l'échelle de l'entreprise n'est pas requise.

Un agent de réapprovisionnement doit s'appuyer sur des informations précises : positions des stocks disponibles, délais de livraison fiables et signaux de la demande actuelle. Il n'a pas besoin de données financières parfaites sur les fournisseurs, de fiches de produits complètes ni des coûts à l'échelle de l'entreprise d'une exactitude comptable. Définir le périmètre de données par flux de travail, plutôt qu'à l'échelle de l'entreprise, est le changement qui permet aux organisations de démarrer le processus.

L'outil opérationnel pour concrétiser cette norme est le jeu de données minimal viable : la plus petite collection de données pertinentes de haute qualité nécessaire pour automatiser la première boucle décisionnelle. Pour le placement des stocks, cela pourrait comprendre les attributs des articles, les dimensions, la durée de conservation, le poids, la position des stocks par emplacement et les signaux de demande à court terme. Si ces champs sont fiables, le processus peut être entamé.

Une mise en garde importante s'impose : démarrer avec un jeu de données minimal viable n'est pas une permission pour négliger l'ensemble de données plus large. Un gain rapide doit être suivi d'une expansion systématique de la qualité et de l'intégration des données, sinon, les organisations risquent de se retrouver avec un ensemble disparate de solutions étroites et fragiles qui fonctionnent de façon isolée, mais ne peuvent pas être mises à l'échelle. L'approche consiste à démarrer étroitement et à prouver la valeur, tout en planifiant simultanément l'expansion de l'ensemble de données afin que le projet pilote d'aujourd'hui devienne la plateforme d'entreprise de demain, et non un déploiement isolé.

Le cas d'utilisation d'abord, les données ensuite

L'erreur de séquençage que commettent la plupart des organisations est de traiter la préparation des données comme un préalable à la sélection du cas d'utilisation, alors que la causalité va dans l'autre sens. Le cas d'utilisation définit les données qui comptent. Définir le cas d'utilisation en premier n'est pas un raccourci; c'est le seul moyen de circonscrire le travail de remédiation avec suffisamment de précision pour qu'il soit exécutable. Les organisations qui commencent par un programme de qualité de données dans l'abstrait le terminent rarement. Les organisations qui commencent avec un flux de travail spécifique pour procéder à l'automatisation savent exactement comment atteindre leur objectif.

Les organisations qui ont déployé l'IA rapportent systématiquement que la définition du jeu de données minimal viable permet de tenir les conversations appropriées entre la chaîne d'approvisionnement, les TI et les équipes de données plus tôt et plus concrètement que n'importe quel programme de données d'entreprise. L'IA convertit un problème de qualité ouvert en un exercice de comblement d'écart circonscrit avec une ligne d'arrivée.

Approches architecturales face à la fragmentation des données

Les organisations de pointe répondent à la fragmentation des données par des approches architecturales pragmatiques qui ne nécessitent pas de refontes d'entreprise pluriannuelles. Les trames de données de la chaîne d'approvisionnement, c'est-à-dire des couches logiques unifiées qui s'appliquent aux systèmes cloisonnés existants, permettent aux agents IA d'accéder aux données des progiciels de gestion intégrés, des systèmes de gestion d'entrepôts, des systèmes de gestion de la formation et de sources externes sans devoir procéder à une migration des données. Kinaxis et Databricks se sont associés pour construire une trame de données de chaîne d'approvisionnement reliant les moteurs de planification directement à l'infrastructure d'un lac de données. Blue Yonder et Snowflake offrent une intégration comparable pour les données de logistique et de gestion optimale. SAP Business Data Cloud met en place une architecture similaire au sein de l'écosystème SAP²³. Ces initiatives témoignent d'une reconnaissance pragmatique : le déploiement de l'IA ne peut attendre que la consolidation des systèmes soit achevée.

6. Gouvernance : passer de la gestion des modèles à la délégation des décisions

- Avec le passage de l'IA du conseil à l'action, la gouvernance doit évoluer de la supervision des modèles à la délégation des décisions.
- L'échelle d'autonomie offre un cadre pratique pour gagner et étendre la confiance de manière progressive.
- La réglementation n'est plus théorique : la Loi européenne sur l'intelligence artificielle entrera pleinement en vigueur en août 2026.

Le défi de gouvernance de l'IA agentique dans le domaine de la chaîne d'approvisionnement diffère fondamentalement de celui de la gouvernance des modèles prédictifs ou des assistants d'IA générative. Les modèles prédictifs produisent des prévisions; gouverner signifie surveiller la précision. L'IA générative produit du contenu; gouverner signifie vérifier les hallucinations et les biais. L'IA agentique prend des décisions; gouverner les agents signifie définir quelles décisions ils peuvent prendre, dans quelles conditions et selon quels mécanismes d'escalade et quelles pistes de vérification. Les modèles de gouvernance les plus efficaces ne sont pas des fonctions de contrôle qui ralentissent le déploiement. Ce sont des mécanismes d'accélération décisionnelle : des mécanismes d'escalade clairs, des balises prédéfinies et une gestion rapide des exceptions qui donnent aux organisations la confiance de déléguer rapidement.

L'échelle d'autonomie est un cadre pratique pour gérer cette transition. Plutôt que de traiter l'autonomie comme binaire, soit contrôlée par l'humain soit pleinement autonome, l'échelle définit trois niveaux gradués permettant aux organisations de gagner en autonomie de manière progressive, sur la base de performances démontrées et de risques circonscrits.

Chaque agent devrait d'abord faire intervenir l'humain. L'autonomie se mérite à mesure que la fiabilité est démontrée dans des conditions définies. Elle n'est pas accordée par défaut. L'objectif est de gagner rapidement en autonomie pour les cas d'utilisation performants, plutôt que d'accorder une autonomie généralisée du jour au lendemain.

Figure 4 : Échelle d'autonomie : gouvernance par type de décision

Débuter par le contrôle. Déléguer au fur et à mesure que la confiance s'établit.

	APPLICATION	SEUIL	EXEMPLES
<p>INTERVENTION HUMAINE SYSTÉMATIQUE</p> <p>Supervision complète</p> <p>L'agent recommande; l'humain approuve avant toute action. <i>Loi européenne sur l'intelligence artificielle, risque élevé : supervision humaine obligatoire</i></p>	<p>Décisions irréversibles ou à enjeux élevés, exposition financière significative, sécurité ou réglementation.</p>	<p>Approbation de la haute direction requise. Aucun seuil d'approbation automatique.</p>	<p>Bons de commande importants. Contrats avec des fournisseurs. Réacheminement critique pour la sécurité.</p>
<p>INTERVENTION HUMAINE PONCTUELLE</p> <p>Exécution supervisée</p> <p>L'agent exécute; l'humain surveille rétrospectivement et peut intervenir.</p>	<p>Décisions partiellement réversibles, incidence financière modérée, décisions fondées sur des schémas récurrents.</p>	<p>Révision du gestionnaire : plus de 25 000 \$. Vérification rétrospective.</p>	<p>Déclencheurs de réapprovisionnement. Planification de rendez-vous. Optimisation des itinéraires.</p>
<p>SANS INTERVENTION HUMAINE</p> <p>Autonomie gagnée</p> <p>Exécution pleinement autonome; supervision humaine uniquement en cas de violation de politique.</p>	<p>Généralement réversible, faible exposition financière, étendue restreinte.</p>	<p>Approbation automatique : moins de 5 000 \$. Surveillance par journalisation uniquement.</p>	<p>Rapprochement de factures. Mises à jour de statut. Enrichissement de données. Classification du chargement partiel.</p>

L'autonomie est toujours réversible. Tout agent peut être ramené à un niveau de supervision supérieur à tout moment.

Source : Analyse de Deloitte, 2026. Les seuils financiers sont illustratifs. Loi européenne sur l'intelligence artificielle (Union européenne), 2024/1689.

Trois niveaux d'autonomie

Intervention humaine systématique : supervision complète

Au niveau de supervision le plus élevé, chaque action significative de l'agent requiert une approbation humaine explicite avant exécution. Ce niveau convient aux décisions à haute valeur, irréversibles ou inédites : bons de commande importants, engagements contractuels stratégiques avec les fournisseurs, réacheminements logistiques critiques pour la sécurité et toute décision créant une exposition financière ou réputationnelle significative.

Intervention humaine ponctuelle : autonomie supervisée

Au niveau de supervision intermédiaire, les agents prennent des décisions courantes selon des paramètres définis, tandis que les humains surveillent le comportement rétrospectivement et interviennent lorsque des exceptions surviennent. Un agent pourrait approuver de manière autonome des achats allant jusqu'à 5 000 \$ auprès d'un fournisseur vérifié, tandis que tout montant supérieur à 25 000 \$ exigerait l'approbation d'un cadre supérieur.

Sans intervention humaine : autonomie gagnée

Au niveau d'autonomie le plus élevé, les agents opèrent en toute indépendance dans un créneau d'action étroitement défini, selon une feuille de route précise. La confiance se gagne par la performance, elle n'est pas décrétée par politique.

Gouvernance en pratique : contexte réglementaire

La *Loi européenne sur l'intelligence artificielle* de l'Union européenne est la première réglementation exhaustive en matière d'IA au monde. Sa classification à haut risque s'applique principalement aux systèmes d'IA utilisés dans la gestion du trafic routier et dans la fourniture d'infrastructures critiques (annexe III). Pour les chaînes d'approvisionnement de produits de consommation et de détail, l'exposition est généralement indirecte, mais réelle : les fournisseurs et les prestataires de services logistiques exécutant des activités en Europe seront soumis aux exigences de la loi, et les cadres de gouvernance devront intégrer les obligations de conformité.

Les obligations pour les systèmes à haut risque comprennent : 1) la gestion continue des risques tout au long du cycle de vie de l'IA; 2) une gouvernance stricte des données d'entraînement; 3) une documentation technique complète; 4) une journalisation automatique pour la traçabilité; 5) des mécanismes de supervision humaine, y compris la capacité d'intervenir ou d'arrêter le système; et 6) des normes appropriées de précision, de robustesse et de cybersécurité⁵. Les sanctions pour les violations graves atteignent 35 millions d'euros ou 7 % du chiffre d'affaires annuel mondial. Les exigences relatives aux risques élevés entrent pleinement en vigueur le 2 août 2026²⁴.

Pour les dirigeants de la chaîne d'approvisionnement au Canada, le portrait réglementaire national n'est pas encore clairement défini. La *Loi sur l'intelligence artificielle et les données (LIAD)* du Canada est morte au feuillet en janvier 2025, et aucune législation fédérale de remplacement sur l'IA n'a été présentée⁴². Des cadres provinciaux comblent partiellement le vide (la Loi 25 du Québec, les modifications à la Loi sur les normes d'emploi de l'Ontario), mais il s'agit de mesures restreintes, et non d'une gouvernance exhaustive de l'IA³⁸.

L'orientation de la réglementation en matière de protection des données reste instructive. Le *Règlement général sur la protection des données (RGPD)* a établi la norme; les cadres nord-américains subséquents ont adopté et adapté les principes fondamentaux. La gouvernance de l'IA suit la même trajectoire. Les organisations qui alignent dès aujourd'hui leurs cadres de gouvernance internes sur les principes de la *Loi européenne sur l'intelligence artificielle* seront mieux positionnées, quel que soit le prochain territoire à légiférer. La réglementation évolue, mais l'orientation est déjà toute tracée.

La gouvernance n'est pas une contrainte sur la valeur de l'IA. C'est le mécanisme par lequel la valeur est pérennisée et la confiance, méritée.

7. Mesurer ce qui compte : « Clarté, intention et transparence »

- L'écart entre l'aspiration en matière d'IA et la valeur de l'IA n'est pas un écart technologique. C'est un écart de mesure.
- La méthodologie « Clarté, intention et transparence » est une approche structurée pour convertir l'investissement en IA en résultats financiers vérifiables.
- La valeur éprouvée finance le prochain cas d'utilisation. Le cadre contient son propre volant d'inertie.

Dans l'ensemble, 74 % des organisations aspirent à utiliser l'IA pour la croissance des revenus, mais seulement 20 % y sont parvenues²⁵. L'écart ne s'explique pas par des limitations technologiques ou la qualité des données; il s'explique par l'absence d'une approche structurée pour définir, concevoir et suivre la valeur des investissements en IA. La recherche récente a constaté que les entreprises performantes en IA étaient près de trois fois plus susceptibles de modifier significativement leurs flux de travail parallèlement au déploiement de l'IA, confirmant que la réalisation de valeur exige un changement organisationnel, et pas seulement technologique³⁹.

74 %/20 %

aspirent à la croissance des revenus par l'IA / l'ont atteinte

Le cadre « Clarté, intention et transparence » répond directement à cet écart. Ce n'est ni un outil de rapport ni un tableau de bord; c'est une méthodologie pour intégrer la valeur à chaque phase d'un déploiement d'IA, de l'idéation initiale aux opérations courantes. Le cadre contient un mécanisme de rétroaction intégré : la valeur éprouvée d'un cas d'utilisation finance et justifie le suivant, créant un cycle autofinancé qui soutient l'élan au-delà du projet pilote initial.

Clarté : définition de la valeur avant la conception

La clarté consiste à définir, à décomposer et à chiffrer en dollars la valeur avant tout développement. Elle répond à quatre questions séquentielles :

Catalyseur de valeur : quel résultat d'affaires spécifique ce cas d'utilisation vise-t-il? Un bon exemple serait de « réduire le fonds de roulement immobilisé dans les stocks de 15 % », plutôt que d'« améliorer l'efficacité de la chaîne d'approvisionnement ».

Mécanisme : par quels leviers opérationnels l'IA produit-elle ce résultat? Pour la réduction des stocks, les leviers comprennent l'optimisation des stocks de sécurité, les changements de politique des stocks cycliques, et l'identification des excédents et de l'obsolescence.

Incidence : quelle est l'incidence financière quantifiée de chaque levier? Les agents de soumission de C.H. Robinson ciblent un mécanisme spécifique : chaque retard de soumission augmente les coûts au comptant de 23 % à 25 %.

Chiffrage en dollars : quelle est la valeur financière annuelle, exprimée dans une devise que le directeur financier reconnaît? Pour une entreprise avec 300 millions de dollars de fonds de roulement, une réduction de 15 % libère 45 millions de dollars : un calcul plutôt élémentaire.

Pour illustrer le fonctionnement de l'aspect « clarté » en pratique, considérons la capacité de rééquilibrage autonome des stocks de Walmart selon les quatre questions. Le *catalyseur de valeur* est la réduction du gaspillage et du fonds de roulement immobilisé dans des stocks mal alloués. Le *mécanisme* est un agent qui identifie de manière autonome les surplus dans un emplacement et les redirige vers les magasins à plus forte demande par le biais de transferts intermagasins. L'*incidence* se traduit par une réduction des démarquages, une amélioration de la disponibilité sans croissance proportionnelle des stocks et une correction plus rapide des déséquilibres. Le *chiffrage en dollars* s'exprime dans les résultats financiers de Walmart : une croissance des ventes de 5 % avec une croissance des stocks de seulement 2,6 %, ce qui correspond à une précision de gestion des stocks alimentée par l'IA libérant des capitaux du fonds de roulement à grande échelle¹⁷.

Intention : conception axée sur la valeur, et non sur les fonctionnalités

L'intention traduit la clarté en une approche d'exécution. Elle se décline en trois phases.

Conception : décomposer le cas d'utilisation en conception fonctionnelle, liée à l'expérience utilisateur et technique, chaque composante étant directement alignée sur le catalyseur de valeur.

Planification : catégoriser les fonctionnalités comme fondamentales ou enrichies. Prioriser ce qui génère de la valeur directe par rapport à ce qui est complexe à construire.

Mise à l'échelle : approfondir (capacités avancées), aller loin (utilisateurs/régions supplémentaires), élargir (flux de travail adjacents).

Un fabricant alimentaire a utilisé cette approche pour la planification du déploiement et l'équilibrage mondial des stocks. Le catalyseur de valeur était clair : s'assurer que les bons produits se trouvent dans le bon centre de distribution au bon moment. La première version était calibrée pour générer un retour annuel de un million de dollars, une cible ciblée et réfutable²⁶.

Transparence : suivi, communication et pérennité

Le mécanisme de suivi doit être défini avant le déploiement, pas après. Le mécanisme de communication est tout aussi important. La valeur suivie, mais non communiquée ne crée aucun élan organisationnel. Des revues de gouvernance hebdomadaires ou mensuelles qui incluent le suivi de la valeur de l'IA renforcent la confiance institutionnelle requise pour approuver les investissements subséquents.

Figure 5 : Le cadre de valeur : clarté × intention × transparence

Trois approches qui doivent être appliquées simultanément pour que l'investissement en IA devienne cumulatif.

CLARTÉ	INTENTION	TRANSPARENCE
<p><i>Définir la valeur avant le développement</i></p> <p>Catalyseur de valeur</p> <p>Quel indicateur opérationnel est amélioré</p> <p>Mécanisme</p> <p>Comment l'IA modifie le flux de travail</p> <p>Incidence</p> <p>Résultat mesurable</p> <p>Chiffrage en dollars</p> <p>Valeur financière quantifiée</p>	<p><i>Intégrer la valeur dans le déploiement</i></p> <p>Conception</p> <p>Intégration au flux de travail, refonte des rôles</p> <p>Planification</p> <p>Déploiement par phases, références, critères de succès</p> <p>Mise à l'échelle</p> <p>Déclencheurs d'expansion, adoption interfonctionnelle</p>	<p><i>Suivre la valeur et communiquer les résultats</i></p> <p>Suivi</p> <p>Mesure continue au niveau du flux de travail</p> <p>Communication</p> <p>Rapport à la direction, financement du prochain cas</p>
<p>ÉCHEC</p> <p><i>Investissement sans hypothèse de valeur réfutable</i></p>	<p>ÉCHEC</p> <p><i>Déploiement techniquement fonctionnel, mais n'atteignant jamais le flux de travail</i></p>	<p>ÉCHEC</p> <p><i>Valeur réalisée, mais jamais mesurée ni communiquée</i></p>

La valeur éprouvée finance le prochain cas d'utilisation

Source : Analyse de Deloitte, 2026.

Les organisations qui suivent et prouvent la valeur d'un cas d'utilisation de l'IA créent le fondement financier et institutionnel pour le suivant. Ce cycle autofinancé est le mécanisme qui distingue les organisations avec un seul projet pilote réussi de celles qui disposent d'un modèle opérationnel de mise à l'échelle de l'IA.

Le cadre « Clarté, intention et transparence » impose une question essentielle : qui est responsable de la valeur? Dans la plupart des organisations, les initiatives d'IA sont financées par les TI et mesurées par les équipes de technologies, créant une déconnexion structurelle entre ceux qui déploient la technologie et ceux qui sont responsables des résultats d'affaires. Les organisations de premier plan confient la responsabilité de la valeur à la fonction de gestion, et non aux TI. Le dirigeant de la chaîne d'approvisionnement qui bénéficie du flux de travail automatisé est responsable du cas de valeur, suit les indicateurs et rapporte les résultats. Les TI facilitent le déploiement; la fonction de gestion assume le résultat.

Quand la valeur se matérialise

Selon notre expérience, le délai de réalisation de la valeur de l'IA dans le domaine de la chaîne d'approvisionnement suit un schéma plus prévisible que ce à quoi la plupart des organisations s'attendent, à condition que la méthodologie « Clarté, intention et transparence » soit adoptée dès le départ.

Les agents à programmation schématisée et les capacités intégrées aux plateformes livrent généralement des améliorations opérationnelles mesurables dans un délai de 4 à 12 semaines après le déploiement. Les agents sur mesure ciblant des flux de travail plus complexes nécessitent généralement de 3 à 6 mois pour atteindre une performance stable en contexte de production. La transformation du modèle opérationnel à l'échelle de l'entreprise est un parcours de 12 à 24 mois, la valeur se composant à mesure que chaque flux de travail successif est automatisé.

En somme : la rapidité de réalisation de la valeur dépend de la rapidité avec laquelle les décisions décrites à la section 9 sont prises, et non de la maturité de la technologie.

8. La divergence concurrentielle

— *L'écart entre les leaders de l'IA dans le domaine de la chaîne d'approvisionnement et le reste du marché est mesurable, cumulatif et grandissant.*

— *Les leaders ne créent pas plus d'IA. Ils utilisent l'IA différemment.*

Les organisations de la chaîne d'approvisionnement les plus performantes investissent dans l'IA à un rythme plus de deux fois supérieur à celui des moins performantes, et l'écart se creuse²⁷. Un sondage mondial de Deloitte mené auprès des chefs de l'approvisionnement en 2025 a constaté que 96 % des leaders numériques avaient atteint ou dépassé leurs objectifs de réduction des coûts, contre 80 % des autres organisations⁴⁰. Pourtant, plus de 80 % des entreprises ne rapportent aucune incidence mesurable de l'IA sur la productivité malgré une adoption généralisée⁴¹. L'écart s'explique par la *manière* dont l'IA est déployée, et non par son adoption ou non.

Figure 6 : Divergence concurrentielle : les leaders par rapport aux autres intervenants du marché

Ce qui distingue les 25 % qui rapportent une valeur transformatrice de l'IA des 75 % qui sont toujours à l'étape de l'expérimentation.

LEADERS	DIFFÉRENCIATEUR	AUTRES INTERVENANTS DU MARCHÉ
Déploient l'IA dans des flux de travail et des structures de rôles repensés	Travail repensé	Déploient l'IA dans des flux de travail et des définitions de rôles inchangés
Suivent et prouvent la valeur de chaque flux de travail	Évaluation par flux de travail	Suivent la valeur au niveau du programme ou du service, si tant est qu'elle soit suivie
Bâtissent des équipes interdisciplinaires : chaîne d'approvisionnement, données et IA	Intégration interfonctionnelle	Maintiennent les initiatives d'IA au sein des TI ou des fonctions numériques
Investissent dans le perfectionnement à grande échelle à chaque déploiement	Investissement dans le perfectionnement	Cèdent le perfectionnement aux RH ou à l'initiative individuelle
S'engagent dans une architecture de production dès le départ	Engagement en production	Restent à l'étape du projet pilote ou de la validation de principe sans définir de parcours vers la production

Chaque différenciateur est une décision. L'écart se creuse à chaque cycle de planification.

Source : Analyse de Deloitte fondée sur des déploiements d'entreprise, la couverture d'analystes et des missions auprès de clients, 2026.

La figure 6 résume les cinq comportements qui distinguent systématiquement les leaders des autres intervenants du marché. Chaque différenciateur est une décision de leadership, et non un choix technologique. Les leaders repensent le travail plutôt que de superposer l'IA aux processus existants. Ils mesurent la valeur au niveau du flux de travail plutôt qu'au niveau du programme. Ils bâtissent des équipes interdisciplinaires combinant expertise en chaîne d'approvisionnement, données et IA. Ils investissent dans le perfectionnement à grande échelle à chaque déploiement. Ils s'engagent dans une architecture de production dès le départ plutôt que de rester indéfiniment à l'étape du projet pilote.

Cette tendance a des précédents. Aussi loin que l'an 2000, les ordinateurs personnels produisaient des gains de productivité marginaux qui étaient plus que neutralisés par les dépenses en TI. La transformation est survenue lorsque les entreprises ont repensé leurs processus d'affaires autour des capacités numériques, donnant naissance aux entreprises axées sur le numérique qui ont dominé les deux décennies suivantes. L'IA agentique se trouve au même point d'inflexion : la technologie fonctionne, mais les rendements ne deviennent cumulatifs que lorsque le travail lui-même est repensé autour d'elle.

La fenêtre de parité concurrentielle se rétrécit de manière très précise. Les premiers adoptants accumulent des données d'entraînement exclusives, des protocoles décisionnels affinés et une compétence institutionnelle en gouvernance que les entrants tardifs ne peuvent pas acquérir rapidement. Une entreprise qui opère un réapprovisionnement autonome depuis 18 mois n'a pas simplement automatisé un flux de travail : elle a construit une boucle de rétroaction qui a entraîné

ses agents sur son réseau, sa base de fournisseurs et ses schémas de demande. Ce savoir institutionnel ne peut pas être reproduit en achetant la même plateforme un an plus tard. C.H. Robinson en est l'illustration directe : une amélioration de la productivité de 30 % la première année et un objectif supplémentaire de 15 % la deuxième. Ce rendement cumulatif reflète l'accumulation de données d'entraînement des agents et l'affinement des protocoles opérationnels, et non simplement la poursuite du déploiement¹³. La même dynamique s'applique à la gouvernance : les organisations qui ont appliqué des cadres d'autonomie graduée pour plusieurs cas d'utilisation ont développé les réflexes opérationnels, les protocoles d'escalade et la confiance de la direction nécessaires pour déployer le prochain agent plus rapidement. Même si les retardataires accélèrent, l'effet cumulatif signifie que l'écart entre les leaders et les autres intervenants du marché ne se comble pas, il s'élargit. Chaque cycle de planification au cours duquel une décision est reportée est un cycle au cours duquel les leaders étendent leur avantage.

La fenêtre de parité concurrentielle en IA dans le domaine de la chaîne d'approvisionnement se rétrécit. L'IA agentique crée des avantages autoreforçants : données exclusives, protocoles affinés et compétence institutionnelle en gouvernance. C'est une priorité stratégique au niveau du conseil d'administration, et non une initiative technologique gérée au sein des TI.

9. Le programme des chefs de la chaîne d'approvisionnement : cinq décisions qui ne peuvent plus attendre

- Cette section est la plus incisive et la plus dense du document. Elle est rédigée pour les dirigeants de la chaîne d'approvisionnement qui ont lu les preuves et se demandent : que dois-je faire lundi matin?
- Cinq décisions, en séquence. Chacune a un responsable clairement désigné, une étendue définie et un test d'accomplissement.
- Il ne s'agit pas d'objectifs ambitieux. Ce sont des engagements concrets et exécutables.

Les sections précédentes ont établi que les conditions empêchant le passage à la production ont grandement changé et que trois des quatre obstacles structurels ont été surmontés. Ce qui reste, c'est l'ensemble des décisions que seuls les dirigeants de la chaîne d'approvisionnement peuvent prendre.

Décision n° 1 : Définir le périmètre de données du premier flux de travail autonome

La première décision est délibérément restreinte : sélectionner un flux de travail à automatiser et définir précisément quelles données il requiert. Il s'agit de déterminer en quoi consistent les données de qualité décisionnelle pour ce flux de travail spécifique et établir le jeu de données minimal viable : la plus petite collection de champs, selon l'exactitude et la latence requises, nécessaire pour automatiser

la première boucle décisionnelle. Il faut commencer par la priorisation des cas d'utilisation, mais reconnaître que les capacités de l'IA agentique et les paradigmes de développement à programmation schématisée peuvent avoir modifié les hypothèses de faisabilité depuis la dernière évaluation.

Les critères de sélection devraient mettre l'accent sur la préparation des données, la réversibilité des décisions et le caractère significatif sur le plan financier, et non sur l'ambition technique. C.H. Robinson a commencé par la soumission de chargements complets : haut volume, données propres, incidence financière claire. Walmart a commencé par le rééquilibrage des stocks.

La décision n° 1 est prise quand les conditions suivantes sont remplies :

- ✓ La priorisation des cas d'utilisation a été actualisée en fonction des capacités agentiques et de programmation schématisée courantes.
- ✓ Un seul flux de travail est sélectionné avec une justification documentée.
- ✓ Le jeu de données minimal viable est défini : champs spécifiques, seuils d'exactitude, exigences de latence.
- ✓ L'évaluation des écarts de données est terminée et la remédiation est cadrée et dotée de ressources.

Décision n° 2 : Établir la politique d'architecture avant l'approvisionnement

Avant d'évaluer des fournisseurs ou d'approuver des projets de développement, il faut établir une politique organisationnelle claire sur les cas d'intégration, de conception à programmation schématisée et de développement sur mesure. La politique devrait inclure un point de contrôle de gouvernance : toute proposition de développement sur mesure d'agents doit démontrer pourquoi les options intégrées et à programmation schématisée sont insuffisantes.

La décision n° 2 est prise quand les conditions suivantes sont remplies :

- ✓ Une politique d'architecture écrite définit les critères pour l'intégration, la programmation schématisée et le développement sur mesure.
- ✓ Un point de contrôle de gouvernance exige une justification avant l'approbation du développement sur mesure.
- ✓ Les évaluations d'approvisionnement renvoient à la politique. Il ne s'agit pas de préférences techniques ponctuelles.

Décision n° 3 : Établir la politique d'autonomie avant le déploiement

Avant qu'un agent n'atteigne la production, il faut définir la politique de l'organisation en matière de délégation décisionnelle. L'échelle d'autonomie fournit le cadre. Chaque agent commence par faire intervenir l'humain. La politique devrait préciser les critères de progression entre les niveaux.

La décision n° 3 est prise quand les conditions suivantes sont remplies :

- ✓ Une politique d'autonomie écrite définit les niveaux, les seuils et les critères de progression.
- ✓ Chaque déploiement d'agent spécifie son niveau d'autonomie initial et ses mécanismes d'escalade.
- ✓ La politique tient compte des exigences de la Loi européenne sur l'intelligence artificielle pour les organisations ayant une exposition européenne.

Décision n° 4 : Bâtir un cas de valeur réfutable avant le financement

L'écart entre aspiration et réalisation (74 % contre 20 %) dans la croissance des revenus attribuable à l'IA n'est pas un échec technologique; c'est un échec dans l'approche à l'égard de la valeur²⁵. Avant d'approuver le financement de toute initiative d'IA, il faut un cas de valeur réfutable : un résultat financier précis, mesuré par rapport à une référence définie et accompagné d'un calendrier d'évaluation.

La décision n° 4 est prise quand les conditions suivantes sont remplies :

- ✓ Chaque initiative d'IA financée dispose d'un cas de valeur documenté et d'un catalyseur, d'un mécanisme, d'une incidence et d'un plan de mesure.
- ✓ Les indicateurs de référence sont saisis avant le déploiement, pas après.
- ✓ Les revues de gouvernance incluent le suivi de la valeur de l'IA parallèlement au suivi des indicateurs clés de performance opérationnelle.

Décision n° 5 : Repenser le travail avant la mise à l'échelle

C'est la décision que la plupart des organisations n'ont pas encore prise. Avant de mettre à l'échelle tout déploiement d'IA, il faut cartographier les rôles, les responsabilités, les pouvoirs décisionnels et les indicateurs de performance qui changeront. La refonte doit être bien définie : quelles tâches seront remplacées, où ira la capacité libérée, quelles nouvelles responsabilités l'IA créera-t-elle pour les travailleurs?

Les implications pour la main-d'œuvre sont réelles. Une recherche de Gartner indique que 51 % des dirigeants de la chaîne d'approvisionnement s'attendent à ce que l'IA agentique entraîne des réductions globales de l'effectif, 55 % s'attendent à des réductions des embauches de début de carrière et 86 % conviennent que des processus entièrement nouveaux liés aux talents sont nécessaires²⁸.

Figure 7 : Transformation des rôles des chefs de la chaîne d'approvisionnement : de la pyramide au diamant

L'IA modifie ce que font les équipes de la chaîne d'approvisionnement et la manière dont la direction est structurée.

AUJOURD'HUI	TRANSITION	AVENIR
<p>Pyramide traditionnelle</p> <p>Leadership stratégique Gestion et planification Exécution opérationnelle</p>	<p>Où en sont la plupart des organisations</p> <p>Moins de dirigeants Niveau intermédiaire en expansion Base opérationnelle en contraction</p>	<p>Moins de dirigeants, mandat élargi</p> <p>Orchestrateurs stratégiques Noyau augmenté par l'IA Spécialistes opérationnels</p>
<p>51 %</p> <p>des dirigeants s'attendent à des réductions globales de l'effectif</p>	<p>55 %</p> <p>des dirigeants s'attendent à des réductions des embauches de début de carrière</p>	<p>86 %</p> <p>conviennent que de nouveaux processus liés aux talents sont nécessaires</p>

Les réductions d'effectifs reflètent la transformation et le redéploiement des rôles, pas nécessairement la réduction nette des postes.

Source : Gartner, Future of Supply Chain 2026 (n = 509), analyse de Deloitte, 2026.

La décision n° 5 est prise quand les conditions suivantes sont remplies :

- ✓ La cartographie du flux de travail est terminée pour le premier flux automatisé.
- ✓ La capacité libérée est explicitement redirigée vers des activités à plus haute valeur.
- ✓ Les indicateurs de performance sont mis à jour pour refléter le modèle opérationnel augmenté par l'IA.
- ✓ Un plan de transition de la main-d'œuvre aborde le perfectionnement, l'évolution des rôles et les nouvelles exigences de compétences.

Les conditions qui entravaient autrefois la mise en production ont changé. Ce qui reste, ce sont des décisions que seuls les dirigeants peuvent prendre. Une organisation est prête à passer du projet pilote à la mise en production lorsque quatre conditions sont réunies : un cadre de gouvernance qui permet une délégation rapide plutôt qu'une approbation lente est en place; c'est la fonction de gestion qui est responsable de la valeur, et non les TI, et celle-ci est mesurée au niveau du flux de travail; au moins un flux de travail a été repensé autour de l'IA plutôt qu'augmenté par celle-ci; et la direction peut expliquer pourquoi le coût concurrentiel du report n'est plus acceptable. Dans l'avenir, les organisations qui agiront le plus vite cesseront de raisonner selon des cas d'utilisation individuels et commenceront à concevoir des systèmes d'agents coordonnés qui gèrent la demande, l'approvisionnement, la tarification et le risque comme un réseau intégré. Cette transition se concrétisera plus vite que ce à quoi s'attendent la plupart des dirigeants, et les décisions présentées dans ce document détermineront si votre organisation sera prête.

Poursuivre la conversation

Ce document reflète notre perspective actuelle sur la manière dont l'IA redéfinit les opérations de la chaîne d'approvisionnement. Toutefois, le domaine évolue rapidement, et les décisions exposées ici gagnent à être éclairées par le dialogue autant que par l'analyse.

Si votre organisation amorce le passage du projet pilote à la mise en production et que vous vous intéressez à l'application des cadres énoncés à votre contexte particulier, nous serons heureux d'en discuter avec vous.



Sarthak Pany

Associé, Chaîne d'approvisionnement et opérations réseau

Deloitte Canada | Toronto

spany@deloitte.ca

ca.linkedin.com/in/sarthakpany



Elizabeth Baker

Associée, Chaîne d'approvisionnement et opérations réseau

Deloitte Canada | Montréal

elizabethbaker@deloitte.ca

ca.linkedin.com/in/elizabeth-baker-88597b19



Mahendra Dedasaniya

Associé et leader, Chaîne d'approvisionnement et opérations réseau

Deloitte Canada | Toronto

mddedasaniya@deloitte.ca

ca.linkedin.com/in/mahendra-dedasaniya-79457b32



Nihar Dalmia

Associé, IA et données

Deloitte Canada | Toronto

nidalmia@deloitte.ca

ca.linkedin.com/in/nihardalmia

Notes de fin

1. Utilisation de Microsoft Copilot Studio par CSX : Microsoft Copilot Studio, références clients, 2024-2025.
2. Performance de l'agent de soumission de C.H. Robinson et incidence sur la tarification au comptant : communiqué de presse de C.H. Robinson, mai 2024; Mark Albrecht, vice-président, IA, C.H. Robinson, 2025.
3. Gestion autonome des stocks de Walmart : blogue des dirigeants de Walmart, juillet 2025.
4. Cost of a Data Breach Report 2025 : 63 % des organisations n'avaient pas de politique de gouvernance de l'IA.
5. Loi européenne sur l'intelligence artificielle (Union européenne) 2024/1689. Le chapitre III (articles 9 à 15) détaille les obligations liées aux risques élevés. L'annexe III définit les catégories de risques élevés, y compris les composants de sécurité des infrastructures critiques. Commission européenne, <https://digital-strategy.ec.europa.eu/fr>.
6. Proportion d'organisations ayant déployé l'IA sans refonte du travail : Deloitte, State of AI in the Enterprise, 7^e édition, 2025-2026.
7. Paradoxe de productivité de l'IA générative dans la chaîne d'approvisionnement : communiqué de presse de Gartner, 5 février 2025. N = 265 répondants mondiaux, août 2024, gartner.com/en/newsroom/press-releases/2025-02-05-gartner-survey-supply-chain-genai.
8. Parc de plus de 30 agents d'IA de C.H. Robinson : communiqué de presse de C.H. Robinson, 31 octobre 2024.
9. Plus de 3 millions de tâches d'expédition traitées par C.H. Robinson : communiqué de presse de C.H. Robinson, avril 2025.
10. Performance de l'agent de soumission (2 min 13 s, incidence de 23 % à 25 % sur les coûts au comptant) : communiqué de presse de C.H. Robinson, mai 2024.
11. Agent de commandes (5 500 commandes quotidiennes de chargements complets, 600 heures économisées) : communiqué de presse de C.H. Robinson, octobre 2024.
12. Gestion des exceptions de chargement partiel (95 % d'automatisation, 42 % de réduction des déplacements inutiles des transporteurs) : renseignements fournis par la direction de C.H. Robinson, 2024-2025.
13. Hausse de productivité de 30 % chez C.H. Robinson : Arun Rajan, chef de la stratégie et de l'innovation, avril 2025. Récit client de Microsoft Azure AI.
14. Architecture systémique de l'IA chez Walmart : blogue mondial des technologies de Walmart, décembre 2025.
15. Stratégie d'automatisation et objectifs de coûts unitaires de Walmart : relations avec les investisseurs de Walmart, résultats du T4 de l'ex. 2026, février 2026. Amélioration cible de 20 % des coûts unitaires; 60 % des magasins américains recevant du fret de centres de distribution automatisés.
16. Négociation avec les fournisseurs de Pactum AI (68 % d'accord, 1,5 % de réduction des coûts, expansion au transport) : Logistics Viewpoints, mars 2025; renseignements fournis par la direction de Walmart, 2025-2026.
17. Croissance des ventes de 5 % vs croissance des stocks de 2,6 % chez Walmart : résultats financiers de l'ex. 2026 de Walmart, février 2026.
18. Optimisation des itinéraires de Walmart (retranchement de près de 50 millions de kilomètres) : direction de Walmart, 2024.
19. Définition de l'IA agentique de Gartner : glossaire des TI et prévisions de Gartner, 2025.
20. Prédiction de Gartner : le tiers des logiciels d'entreprise intégreront l'IA agentique d'ici 2028 : symposium et exposition sur la chaîne d'approvisionnement de Gartner, 2025.
21. Agents SAP Joule : SAP TechEd, octobre 2024. Agents Kinaxis Maestro : communiqué de presse de Kinaxis, juin 2025. Agents opérationnels Blue Yonder : Blue Yonder ICON, mai 2025.
22. Pile technologique de C.H. Robinson (Azure AI Foundry, LangChain) : récit utilisateur de Microsoft Azure AI, 2024-2025.
23. Partenariat Kinaxis-Databricks pour la trame de données de la chaîne d'approvisionnement : communiqué de presse conjoint, mars 2025. Intégration de données logistiques Blue Yonder-Snowflake : communiqué de presse conjoint, janvier 2025. SAP Business Data Cloud : SAP Sapphire, mai 2025.
24. Calendrier d'application et sanctions de la Loi européenne sur l'intelligence artificielle : Commission européenne, <https://digital-strategy.ec.europa.eu/fr>. Obligations pour les risques élevés applicables le 2 août 2026.

25. 74 % des organisations aspirent à la croissance des revenus par l'IA, 20 % l'ont atteinte : Deloitte, *State of AI in the Enterprise*, 7^e édition, 2025-2026.
26. Exemple de retour de 1 M\$ chez un fabricant alimentaire : Elizabeth Baker, Deloitte, présentation NISCL, février 2026.
27. Les organisations de la chaîne d'approvisionnement les plus performantes investissent 2 fois plus dans l'intelligence artificielle et l'apprentissage machine : Gartner, *Supply Chain Executive Report: Future of Supply Chain 2024*, communiqué de presse, 20 février 2024. N = 818 professionnels, août-oct. 2023, gartner.com/en/newsroom/press-releases/2024-02-20-gartner-top-supply-chain-organizations.
28. Projections de Gartner sur la main-d'œuvre : *Gartner Future of Supply Chain 2026*, n = 509, février 2026.
29. Sources de la figure 1 : résultats du T4 de l'ex. 2026 de Walmart (févr. 2026); Anaplan PlanIQ; Domino's Pizza UK/Microsoft Dynamics 365 (2023); Coca-Cola/Blue Yonder (2025); rapports sur la durabilité de Cargill (2025); ServiceNow/Aptiv (2025); Zycus Merlin (2025); SAP/Scoutbee (2025); documentation sur la plateforme Coupa; Bosch/SAP industry cloud (2025); GE/Siemens MindSphere; transformation numérique du système de production de Toyota (2025); Siemens/Microsoft Industry AI (2025); plateforme Blue Yonder; plateforme de visibilité FourKites; Amazon Robotics (2025); UPS ORION; Turvo/Trimble (2025).
30. « Wally » de Walmart, IA agentique pour les marchands : *blogue mondial des technologies de Walmart*, 2025; renseignements fournis par la direction.
31. Utilisation non autorisée d'outils d'IA et lacunes de maturité en gouvernance (70 % des entreprises reconnaissent que la gouvernance n'est pas pleinement optimisée) : *Forbes Tech Council*, novembre 2025; *Acuvity AI State of AI Security Report*, octobre 2025.
32. « Taxe de l'IA sur la productivité » (40 % des gains de temps perdus en reprise du travail) : *Workday AI Productivity Paradox Report*, janvier 2026.
33. Étude de METR sur les développeurs (19 % plus lents, perception de 20 % plus rapides) : Joel Becker, Nate Rush, Elizabeth Barnes et David Rein, « *Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity* », METR, 10 juillet 2025. Essai contrôlé randomisé, 16 développeurs, 246 tâches. Étude en génie logiciel; citée comme preuve intersectorielle du paradoxe de productivité. arXiv:2507.09089, metr.org/blog/2025-07-10-early-2025-ai-experienced-os-dev-study.
34. Goulot d'étranglement de la révision de l'IA (21 % plus de tâches, 98 % plus de demandes de révision, 91 % plus de temps de révision) : Faros AI, « *The AI Productivity Paradox Report 2025* », juin 2025. Télémétrie de plus de 10 000 développeurs dans 1 255 équipes. Étude en génie logiciel; citée comme preuve intersectorielle du paradoxe de productivité, faros.ai/blog/ai-software-engineering.
35. Élan de formation en IA des entreprises (plus de 60 % des dirigeants ont suivi une formation en IA, 73 % encouragent les équipes à utiliser l'IA) : assemblée générale, « *Corporate Leaders Are Increasingly Upskilling for AI* », 30 septembre 2025. N = 651 vice-présidents/directeurs, États-Unis et Royaume-Uni. Couverture : *CIO Dive*, 3 octobre 2025, generalassemb.ly/blog/leaders-are-gaining-ground-on-ai-skills.
36. Camions autonomes Gatik pour Walmart : communiqué de presse de Gatik, début 2026. Premières opérations commerciales de transport intermédiaire entièrement sans conducteur en Amérique du Nord.
37. Dette technique et rendement du capital investi de l'IA : *Institute for Business Value*, 2025. 86 % des cadres citent la dette technique comme contrainte de l'IA; réduction de 18 % à 29 % du rendement du capital investi.
38. Loi 25 du Québec (obligations en matière de prise de décisions automatisée, en vigueur); modification de la Loi sur les normes d'emploi de l'Ontario (divulgaration de l'IA dans l'embauche, en vigueur depuis janvier 2026).
39. Les entreprises performantes en IA sont significativement plus susceptibles de repenser les flux de travail autour de l'IA plutôt que de superposer l'IA aux processus existants : Deloitte, « *Bridging the AI Value Gap: Are Team Dynamics the Missing Link?* », *Deloitte Insights*, mars 2026, deloitte.com/us/en/insights/topics/talent/ai-roi-and-team-structure.
40. Sondage mondial Deloitte auprès des chefs de l'approvisionnement 2025 : 96 % des leaders numériques ont atteint leurs objectifs de réduction des coûts par rapport à 80 % de leurs pairs.
41. Étude du NBER (80 %+ des entreprises ne rapportent aucune incidence de productivité mesurable de l'IA) : *National Bureau of Economic Research*, février 2026.
42. La LIAD du Canada (partie du projet de loi C-27) est morte au feuillet en janvier 2025. Aucune législation fédérale de remplacement sur l'IA n'a été présentée au début de 2026. Les mesures provinciales (Loi 25 du Québec, divulgation de l'IA dans la Loi sur les normes d'emploi de l'Ontario) offrent une couverture partielle. *Parlement du Canada*, 2025-2026.

43. Deloitte fournit des services professionnels, y compris d'audit et de certification, à certaines entités mentionnées dans cette publication. Les références à des fournisseurs, plateformes ou produits spécifiques sont fournies à titre illustratif uniquement et ne constituent ni une recommandation, ni une approbation, ni un traitement préférentiel. Les décisions de sélection de fournisseurs devraient être fondées sur les exigences propres à chaque organisation et une évaluation indépendante.

44. Lancement de l'Intelligent Control Tower™ de FourKites et capacités agentiques de la chaîne d'approvisionnement : communiqué de presse de FourKites, 15 janvier 2025, fourkites.com/press/fourkites-introduces-intelligent-control-tower.

45. Plateforme d'intelligence des processus de Celonis et suite d'agents de la chaîne d'approvisionnement d'AgentC : communiqué de presse de Celonis, 13 mai 2025, celonis.com/news/press/celonis-showcases-latest-innovations. Plateforme : celonis.com/solutions/supply-chain-transformation.

46. Prédiction d'annulation de projets d'IA agentique et évaluation de la maturité des fournisseurs de Gartner : communiqué de presse de Gartner, 25 juin 2025, gartner.com/en/newsroom/press-releases/2025-06-25-gartner-predicts-over-40-percent.

47. Prédiction de Gartner : 50 % des solutions de gestion de la chaîne d'approvisionnement interfonctionnelle incluront des capacités d'IA agentique d'ici 2030. Communiqué de presse de Gartner, 21 mai 2025.

48. Robertson et al., « Resilient by Design: The Agentic Supply Chain », Deloitte Insights, mars 2026. Henderson et al., « From Vision to Value: A Road Map for Enterprise Transformation in Manufacturing with Agentic AI », Deloitte Insights, septembre 2025.

À propos de Deloitte Canada

Chez Deloitte, notre raison d'être est d'avoir une influence marquante. Nous existons pour inspirer et aider nos gens, nos organisations, nos collectivités et nos pays à prospérer en créant un avenir meilleur. Notre travail soutient une société prospère où les gens peuvent s'épanouir et saisir des occasions. Il renforce la confiance des consommateurs et des entreprises, aide les organisations à trouver des moyens créatifs de déployer des capitaux, habilite des institutions sociales et économiques justes, fiables et efficaces, et permet à nos amis, à nos familles et à nos collectivités de profiter de la qualité de vie qui accompagne un avenir durable. Étant le plus grand cabinet de services professionnels détenu et exploité à 100 % par des Canadiens dans notre pays, nous sommes fiers de travailler aux côtés de nos clients pour avoir une influence positive sur tous les Canadiens.

Deloitte offre des services de premier plan dans les domaines de la consultation, de la fiscalité et des services juridiques, des conseils financiers, de l'audit et de la certification ainsi que des conseils en gestion des risques à près de 90 % des sociétés du palmarès Fortune Global 500™ et à des milliers de sociétés fermées. Nous réunissons des compétences, des perspectives et des services de classe mondiale pour aborder les enjeux d'affaires les plus complexes de nos clients.

Deloitte S.E.N.C.R.L./s.r.l., société à responsabilité limitée constituée en vertu des lois de l'Ontario, est le cabinet membre canadien de Deloitte Touche Tohmatsu Limited. Deloitte désigne une ou plusieurs entités parmi Deloitte Touche Tohmatsu Limited, société fermée à responsabilité limitée par garanties du Royaume-Uni, ainsi que son réseau de cabinets membres dont chacun constitue une entité juridique distincte et indépendante. Pour obtenir une description détaillée de la structure juridique de Deloitte S.E.N.C.R.L./s.r.l. et de ses filiales, veuillez consulter www.deloitte.com/ca/apropos.

Pour en apprendre plus sur Deloitte Canada, veuillez nous suivre sur LinkedIn, X, Instagram ou Facebook.

© Deloitte S.E.N.C.R.L./s.r.l. et ses sociétés affiliées.