

## Mise en œuvre de la plateforme d'IA

Surmontez la complexité des solutions d'IA pour accélérer la création de valeur



La puissance transformatrice de l'IA et de l'IA générative appelle une action rapide. De nombreuses organisations sont forcées de composer avec des infrastructures IT obsolètes, des modèles opérationnels inefficaces et des déficits de compétences croissants alors même que les charges applicatives d'IA consomment plus de ressources et gagnent en complexité. Il est absolument primordial d'être accompagné par des experts techniques afin de mettre en œuvre et d'intégrer la technologie sans mettre en péril la rapidité, la rentabilité et l'évolutivité.

Il est tout à fait possible d'acquérir des solutions personnalisées, sécurisées et fiables qui répondent aux cas d'utilisation et aux objectifs métiers, notamment grâce à des prestataires de services en qui vous avez confiance. En surmontant rapidement ces défis, votre organisation libérera tout le potentiel de l'IA, générera une valeur mesurable et réussira à long terme.

92 %

des organisations s'attendent à ce que le pourcentage de collaborateurs ayant accès à l'IA augmente au cours de l'année à venir<sup>1</sup>

50 %

n'ont pas suffisamment confiance dans leur capacité à rechercher, sélectionner, déployer et gérer des solutions d'IA pour répondre aux besoins de leur entreprise<sup>2</sup>

84 %

des organisations préfèrent externaliser/collaborer avec un prestataire de services pour mettre en œuvre des solutions d'IA/IA générative<sup>3</sup>



Dell Services relève toute une palette de défis liés à la conception, à la mise en œuvre et à la gestion de solutions d'IA pour améliorer les performances, l'évolutivité et l'efficacité. Fiez-vous à des experts qui évaluent vos besoins et vous apportent les solutions qu'il vous faut.

- ✓ Opérations rationalisées
- ✓ Meilleure adoption
- ✓ Retour sur investissement maximisé
- ✓ Flexibilité en termes d'évolutivité

## Complexité de la mise en œuvre

## Impact sur le retour sur investissement

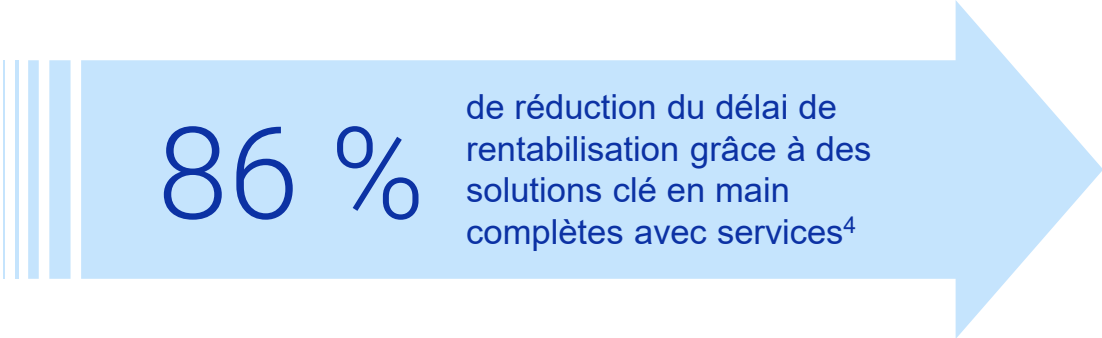


### Raccourcissez le délai de rentabilisation des solutions sur site

- ✓ Adapté à vos besoins métiers
- ✓ Optimisé pour la performance
- ✓ Évolutif à souhait
- ✓ Meilleure adoption et utilisation
- ✓ Opérations rationalisées grâce à l'orchestration et à l'automatisation
- ✓ Générations améliorées à partir des données intégrées
- ✓ Résultats optimisés grâce à l'expertise de l'écosystème ouvert

Les experts Dell Services vous aident à simplifier la mise en œuvre et à disposer de solutions d'IA à toute épreuve. Configuration de l'infrastructure, sélection, intégration et optimisation de la plateforme, efficacité opérationnelle, et toujours une marge d'adaptation.

Focus Areas		
Use Cases		
Application Infrastructure	Training Frameworks	Inference
	Foundational Models / Model Training	
	MLOps / LLMOps	
Software Infrastructure	Workload Manager	
	Driver / Compiler / Toolchain	
	Operating Systems	
	Deployment Tools	
Hardware Infrastructure	Infrastructure from device to data center Dell compute, storage and networking	
		Partner GPUs





## Conception et planification

- Identification des habitudes d'utilisation de l'IA, stratégie d'optimisation des processeurs graphiques/ du réseau, composants de l'infrastructure d'IA pour la sélection du fournisseur, considérations relatives aux performances, etc.
- Étude des besoins en matière de préparation et de conformité des données
- Évaluation complète du niveau de préparation du datacenter pour les fonctionnalités d'IA
- Vérification de l'alignement entre les solutions et les objectifs et adaptation à la production et à plusieurs cas d'utilisation
- Atelier pour définir les besoins spécifiques en matière d'intégration



## Déploiement de l'infrastructure

- Montage et intégration de racks en usine
- Connectivité entre les racks et câblage des racks de processeurs graphiques et des commutateurs
- Gestion de l'alimentation et connectivité du système de refroidissement (liquide ou air)
- Montage du cluster et tests et validation du rack complet
- Configuration et tests du fabric réseau
- Tests d'acceptation de l'ensemble du cluster (HPL, NCCL, etc.)
- Transfert de connaissances sur le matériel déployé



## Implémentation de la plateforme

Mise en œuvre et configuration de l'infrastructure logicielle sélectionnée

- Mise en œuvre et démonstration du serveur d'inférence avec un LLM pré-entraîné
- Création d'un exemple de charge applicative d'entraînement de modèle (réglage fin) pour tester l'environnement en fonctionnement
- Proposition d'outils/ de configurations supplémentaires en fonction des résultats souhaités
- Scripts de test pour tester le fonctionnement

Plan d'adoption de la plateforme d'IA



## Adoption et évolutivité

- Intégration avec les plateformes de stockage, d'observabilité et de journalisation ; amélioration des performances système avec surveillance pour une allocation optimale des processeurs graphiques
- Mise au point des procédures opérationnelles et des workflows pour soutenir l'amélioration des opérations de production
- Développement avec des outils conteneurisés, le réglage fin des modèles et le dépannage
- Acquisition de compétences grâce à une expertise en la matière pour la mise à l'échelle des charges applicatives d'IA
- Amélioration de l'utilisation avec des experts délégués sur site dédiés
- Rationalisation des opérations avec des solutions entièrement gérées

Jour 0

Jour 1

Jour 2+

Vous ne savez pas par où commencer ? Nous vous proposons un [Accelerator Workshop](#) ou un pilote de cas d'utilisation pour faire vos premiers pas. Contactez un représentant Dell dès aujourd'hui pour en savoir plus

1. ESG, « There's no time to waste: Confidently moving ahead with Enterprise AI », Mike Leone, novembre 2024

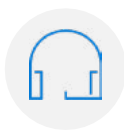
2. ESG, « The AI Maturity Playbook: Driving Business Impact and Innovation », Mike Leone, réalisé à la demande de Dell Technologies, les résultats réels peuvent varier

3. Étude réalisée par Forrester Consulting à la demande de Dell Services, septembre 2024

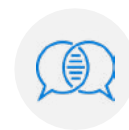
4. Estimation basée sur une analyse réalisée par Dell en mai 2024 comparant le temps de configuration d'un cluster Kubernetes® à deux nœuds pour un LLM à usage général utilisant des scripts automatisés au délai de déploiement manuel d'une conception commune. Le temps de configuration inclut uniquement l'installation de la base. Le temps de configuration réel varie en fonction de la configuration de la solution.



Explorer les [services IA Dell](#)



[Contacter](#) un expert  
Dell Technologies



Participez à la conversation  
avec #DellTechnologies