

Simplifier le développement de l'IA générative

Stations de travail Dell Precision pour l'IA
et NVIDIA AI Workbench

Commencer



Sommaire

Défis du développement de l'IA générative (GenAI).....	3
La solution : NVIDIA AI Workbench sur les stations de travail Dell Precision	4-7
Témoignage client : Invoke et Invoke Enterprise	8
NVIDIA AI Workbench et NVIDIA AI Enterprise	9
Étapes suivantes et ressources.....	10

Défis du développement de l'IA générative

Qu'il s'agisse de configurer des processeurs graphiques (GPU) et des notebooks Jupyter, de déplacer des charges applicatives ou de gérer des workflows, les défis quotidiens du développement de l'IA générative peuvent affecter l'avancement des projets. Le temps et l'expertise technique nécessaires peuvent intimider les scientifiques des données, les ingénieurs et les développeurs. Heureusement, avec les bons partenaires et outils technologiques, il est possible de surmonter les obstacles et d'accélérer le développement.

Obstacles courants à l'IA générative



Configuration complexe des processeurs graphiques



Portabilité limitée des charges applicatives



Gestion chronophage des workflows



La solution : NVIDIA AI Workbench sur les stations de travail Dell Precision prêtes pour l'IA

Avec NVIDIA AI Workbench sur les stations de travail Dell Precision, les développeurs experts et novices peuvent rapidement et facilement créer, tester et personnaliser des projets d'IA générative sur des stations de travail locales, puis les faire évoluer à la hausse ou à la baisse sur n'importe quelle plateforme. Tout peut être fait sur les stations de travail Dell Precision prêtes pour l'IA qui fournissent les hautes performances requises pour traiter de grands jeux de données et des réseaux neuronaux complexes.

NVIDIA AI Workbench

- ✓ Rationalisez la configuration
Commencez à coder et à personnaliser des modèles plus vite avec la configuration automatisée des systèmes de processeurs graphiques
- ✓ Exécutez des projets partout
Développez en local et déplacez des charges applicatives facilement pour les exécuter sur la meilleure plateforme en termes de coûts, de vitesse et d'échelle
- ✓ Accélérez les workflows d'IA
Améliorez la productivité avec l'automatisation pour la gestion des conteneurs et des versions

Stations de travail Dell Precision

- ✓ Développement local de l'IA générative
Personnalisez et développez des modèles d'IA générative localement sur des stations de travail accélérées par des processeurs graphiques
- ✓ Architecture évolutive
Évoluez de un à quatre processeurs graphiques NVIDIA RTX™ de génération Ada
- ✓ Solutions prêtes pour l'IA
Conçues et configurées pour répondre à vos besoins de charges applicatives d'IA

Rationalisez la configuration matérielle

La productivité baisse lorsque les développeurs doivent gérer le processus complexe de configuration d'une machine GPU pour le Deep Learning et l'IA générative. NVIDIA AI Workbench élimine les problèmes courants de configuration tels que l'incompatibilité des systèmes d'exploitation et des pilotes, ainsi que celle du pilote CUDA et du firmware. Les développeurs peuvent commencer à coder rapidement grâce à l'installation et à la configuration automatisées des logiciels et pilotes. NVIDIA AI Workbench aide également à garantir que les conteneurs sont configurés pour une efficacité optimale des processeurs graphiques, en fonction de votre charge applicative.



Augmentez la productivité grâce à la configuration automatisée



Éliminez les problèmes de configuration matérielle courants



Assurez la configuration appropriée des conteneurs

Exécutez des projets partout

NVIDIA AI Workbench donne aux développeurs la liberté de créer, personnaliser et collaborer facilement grâce à la mobilité fluide des charges de travail. Déplacez les projets NVIDIA AI Workbench vers le système le mieux équipé en processeurs graphiques pour optimiser les coûts, la vitesse et pouvoir évoluer en toute sécurité. Épargnez-vous le temps et les compétences techniques généralement nécessaires en la matière. Traitez facilement les changements d'environnement, de données, de modèle et d'emplacement ainsi que le contrôle des versions pendant que NVIDIA AI Workbench gère en toute sécurité l'authentification et les secrets requis pour permettre aux utilisateurs d'accéder aux différentes ressources.



Transférez les charges applicatives facilement



Déplacez les informations d'identification et les secrets en toute sécurité



Traitez les changements et le contrôle des versions en toute simplicité

Accélérez les workflows d'IA



Automatisez l'installation et la gestion de projets



Suivez les dépendances de projet automatiquement



Simplifiez la gestion de projet avec une UI et une CLI intuitives

Avec NVIDIA AI Workbench, les développeurs experts et novices peuvent rapidement identifier, installer et gérer les éléments des projets d'IA et de Deep Learning. Boostez la productivité avec l'installation automatisée de plusieurs ressources de projet. Procédez facilement au suivi des dépendances de projet pour un modèle et un GPU donnés afin que les développeurs puissent se concentrer sur l'exécution. NVIDIA AI Workbench simplifie la gestion des workflows grâce à une interface utilisateur (UI) intuitive et à une puissante interface de ligne de commande (CLI).

Voir des exemples de workflows sur GitHub :

- ➔ [Projet Hybrid RAG](#)
- ➔ [Projet NIM Anywhere](#)
- ➔ [Projet Stable Diffusion](#)
- ➔ [Projet Fine-Tuning](#)

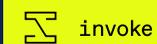
Témoignage client : Invoke Enterprise

Invoke a utilisé un projet NVIDIA AI Workbench pour permettre à des entreprises d'entraîner des modèles d'images d'IA de manière sécurisée et collaborative. Invoke a déployé Invoke Enterprise et NVIDIA AI Workbench sur site, en exécutant le calcul IA localement sur des [stations de travail mobiles Dell Precision 7780](#) équipées de processeurs graphiques NVIDIA RTX™ 5000 de génération Ada.

« Le déploiement de modèles d'IA générative sur des stations de travail Dell Precision prêtes pour l'IA, plus particulièrement la Precision 7780, a établi une nouvelle norme en matière de performances. Le matériel plus ancien peut avoir du mal à traiter le calcul IA, mais l'exécution d'Invoke sur la 7780 est simple, puissante et prend en charge l'ensemble de nos workflows d'IA les plus avancés. »

— Kent Keirse, fondateur et PDG, Invoke

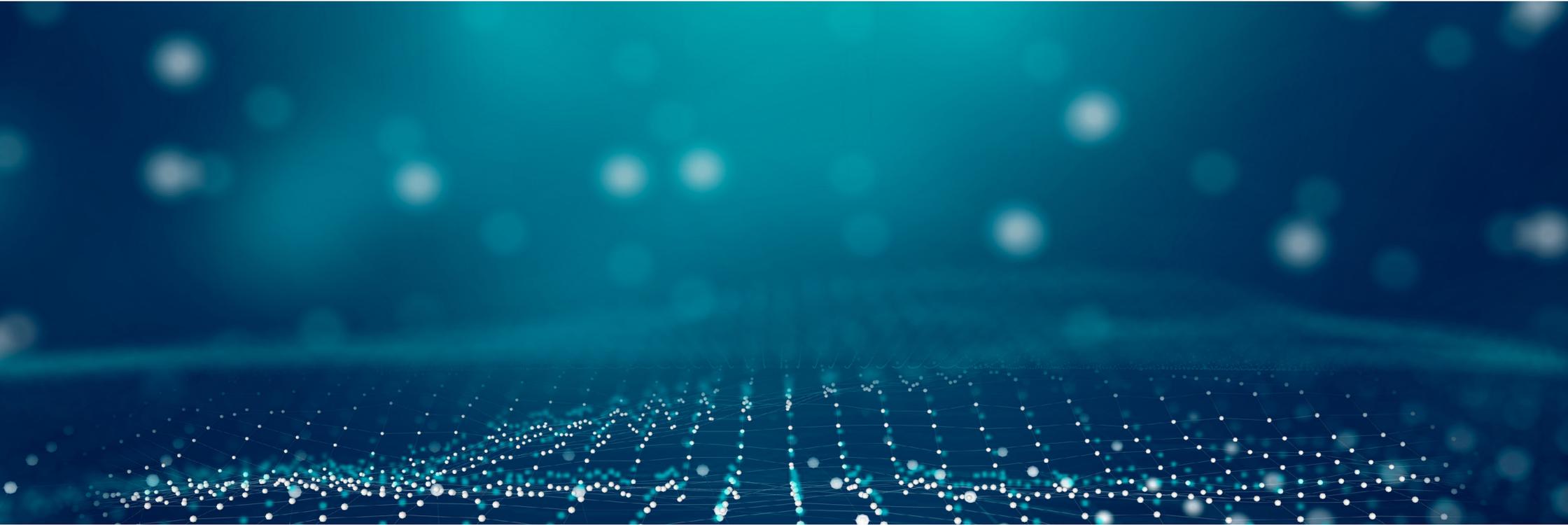
[Lire l'étude de cas](#)





NVIDIA AI Workbench et NVIDIA AI Enterprise

Combinez NVIDIA AI Workbench avec NVIDIA AI Enterprise pour accéder à une plateforme logicielle Cloud native de bout en bout qui accélère les pipelines de science des données et rationalise le développement et le déploiement d'applications d'IA de production. NVIDIA AI Workbench est une solution gratuite, accompagnée d'un support d'entreprise pour les détenteurs d'une licence NVIDIA AI Enterprise, notamment de fonctions de sécurité, de NIM et de branches de production.



Étapes suivantes et ressources

Lancez-vous

- ➔ Téléchargez AI Workbench gratuitement
- ➔ Trouvez des configurations de stations de travail prêtes pour l'IA
- ➔ Contactez-nous pour acheter NVIDIA AI Enterprise via Dell
- ➔ Essayez NVIDIA AI Enterprise gratuitement pendant 90 jours

En savoir plus

- ➔ Blog : Harnessing NVIDIA Tools for Developers with Precision AI-Ready Workstations
- ➔ Blog : Decoding How NVIDIA AI Workbench Powers App Development
- ➔ Blog : Optimize AI Model Performance and Maintain Data Privacy with Hybrid RAG
- ➔ Blog : Develop and Deploy Scalable Generative AI Models Seamlessly with NVIDIA AI Workbench