

IBM et VMware aident les entreprises à adopter l'IA générative avec watsonx sur site (« on premises »)



Par Tarun Chopra, Vice President, Product Management, Data & AI, IBM et Chris Wolf, Vice President, VMware AI Labs, le 07 novembre 2023 : IBM et VMware collaborent pour mettre à disposition IBM watsonx sur des environnements sur site, sur la solution VMware Private AI et Redhat OpenShift, et ce afin d'aider à mettre en place des fonctionnalités d'IA générative rapides et transparentes.

En 2023, nous avons vu l'adoption de l'IA se généraliser, les entreprises cherchant à conduire des innovations transformatrices. L'IA générative, en particulier, a capté l'imagination des organisations en tant que moyen de générer un nouveau contenu basé sur une variété d'informations. Les cas d'usage et les tâches de l'IA générative sont très variés - de la génération de codes à l'expérience de résolution des centres de contact, à l'automatisation des opérations informatiques, à la recherche d'informations et à la création de contenu.

Les données sont au cœur de tous les cas d'usage de l'IA. Plus important encore, l'accès aux données nécessaires et leur traitement pour obtenir les meilleurs résultats des modèles d'IA. Les entreprises doivent être en mesure d'apporter la capacité de calcul et les modèles d'IA là où les données de l'entreprise sont créées, traitées et consommées - pour soutenir les cas d'usage de l'IA – et ce pour des cas d'usage d'applications traditionnelles d'IA/ML (Machine Learning) et l'IA générative. Les entreprises doivent également avoir accès à une suite technologique complète qui réponde aux exigences traditionnelles des entreprises, telles que l'utilisation du confidential computing (informatique confidentielle) et la capacité d'entraîner, de régler (« tune ») et de déployer des modèles d'IA en un seul endroit qui peut s'exécuter dans des environnements Cloud.

C'est la promesse de Private AI, une approche architecturale qui vise à équilibrer les gains commerciaux de l'IA avec les besoins de confidentialité et de conformité de l'organisation. Elle se concentre sur la plateforme et l'architecture d'infrastructure construites pour soutenir l'IA, qui peut être déployée dans des Clouds publics, des Clouds privés virtuels, des datacenters et

des sites périphériques (edge).

VMware Private AI et IBM watsonx apportent de la flexibilité à l'IA générative

IBM et VMware offrent aux entreprises clientes la flexibilité nécessaire pour mettre en œuvre des cas d'usage d'IA générative là où se trouvent leurs opérations critiques. C'est pourquoi IBM et VMware associent VMware Cloud Foundation, Red Hat OpenShift et la [plateforme de données et d'IA IBM watsonx](#). Cette combinaison permettra aux entreprises d'accéder à IBM watsonx dans des environnements privés, sur site, sur du IaaS (Infrastructure as a Service), ainsi que dans le Cloud hybride avec les offres SaaS de watsonx sur IBM Cloud. Les clients communs pourront utiliser les capacités d'IA d'IBM watsonx, de gestion des données et, dans un avenir proche, ses capacités de gouvernance et ce, dans les Clouds publics, les Clouds privés virtuels, les datacenters et les sites périphériques (edge), ce qui contribuera à mettre en place des modèles d'IA transparents et explicables.

VMware Private AI est disponible dans le cadre d'un déploiement hybride multi-cloud. En utilisant Red Hat OpenShift, les clients peuvent désormais utiliser watsonx sur site pour les applications déployées sur VMware. Les clients communs pourront accéder à l'ensemble de la plateforme watsonx, qui propose des solutions pour aider les clients à accélérer et à mettre à l'échelle l'IA. Les clients peuvent éventuellement utiliser IBM Cloud Satellite pour automatiser et simplifier le déploiement et les opérations du jour 2 (day 2 operation) de leurs environnements VMware Private AI avec OpenShift.

Un partenariat déterminé à offrir la valeur du multi-cloud

Cette collaboration est née du Joint Innovation Lab d'IBM et de VMware, un moteur de co-innovation doté d'une équipe d'ingénieurs dédiée qui a pour mission de s'appuyer sur les atouts des deux sociétés et de résoudre les problèmes des clients en mettant de nouvelles offres sur le marché. Le laboratoire, annoncé en 2018, a depuis réalisé de nombreux projets axés sur le Cloud hybride et l'IA.

VMware et IBM prévoient de développer une architecture de référence validée pour permettre à leurs clients communs d'exploiter des environnements Cloud hybride et sur site, soumis à une réglementation pour l'entraînement et le réglage fin (fine tuning) des modèles de ML - y compris les modèles de langage de grande taille (LLM) pour les cas d'usage de l'IA générative. Par exemple, le transfert de données d'une banque ou d'une institution financière d'un pays vers un fournisseur de Cloud (hyperscaler) pourrait être interdit en raison de restrictions réglementaires ou parce que les données sont trop volumineuses et qu'il serait impossible d'essayer de les transférer d'un endroit à un autre.

IBM Consulting apporte à ses clients son expertise en matière de solutions spécifiques VMware et d'IA générative

Au début de cette année, IBM Consulting a créé un centre d'excellence pour l'IA générative et compte désormais plus de 1 000 consultants spécialisés dans l'IA générative qui s'engagent auprès d'un ensemble mondial de clients pour augmenter la productivité dans les opérations informatiques et les processus métier fondamentaux tels que les RH ou le marketing, afin d'améliorer leurs expériences client et de créer de nouveaux modèles d'entreprise.

Ceci, combiné à la forte expertise VMware d'IBM Consulting et à ses capacités de service, permettra d'accélérer les transformations commerciales de nos clients grâce à l'IA d'entreprise sur l'architecture de référence VMware Private AI.

En outre, pour les clients qui cherchent à moderniser et à transformer leurs applications, IBM Consulting prévoit d'intégrer les services IBM watsonx et VMware Private AI dans son service exclusif [IBM Consulting Cloud Accelerator](#) afin d'accélérer le processus de transformation du Cloud. Une fois lancée, cette intégration facilitera la rétro-ingénierie, la génération de code et la conversion de code tout en aidant à gérer les opérations du jour 2 (day 2 operation) et au-delà pour des services de gestion de Cloud hybride transparents de la part d'IBM Consulting.

Favoriser l'avenir de l'IA générative des clients

Le partenariat stratégique entre IBM et VMware vise à permettre à leurs clients communs d'adopter le Cloud hybride grâce à des modèles de consommation flexibles et de moderniser leurs applications critiques. Les clients peuvent également bénéficier de la présence d'IBM Consulting en tant que partenaire stratégique dans le cadre de ce projet. En ayant le choix du moment, de l'endroit et de la manière d'intégrer les technologies d'IA générative à VMware Cloud Foundation, les entreprises seront habilitées pour entraîner et déployer rapidement des fonctionnalités d'IA personnalisées au sein de leur entreprise, tout en conservant le contrôle de leurs données. Elles bénéficient également d'un guide puissant pour exploiter les meilleures innovations des offres logicielles de VMware et ce, de manière unifiée afin de fournir un environnement cohérent intégré aux données et aux capacités d'IA apportées par la technologie des partenaires d'IBM Cloud.

Avertissement : VMware et IBM ne garantissent pas que les services annoncés en avant-première ou en version bêta seront disponibles à une date ultérieure. Les informations contenues dans ce communiqué de presse sont données à titre indicatif uniquement et ne peuvent être intégrées dans aucun contrat. Cet article peut contenir des hyperliens vers des sites Web autres que ceux de VMware ou d'IBM, créés et gérés par des tiers qui sont seuls responsables du contenu de ces sites.

Contacts Presse :

Weber Shandwick pour IBM

IBM

Gaëlle Dussutour

Tél. : + 33 (0)6 74 98 26 92

dusga@fr.ibm.com

Louise Weber

Tél. : + 33 (0)6 89 59 12 54

ibmfrance@webershandwick.com
