

IBM dévoile la plateforme Watsonx pour alimenter des modèles de fondation de nouvelle génération pour les entreprises

- Watsonx est une nouvelle plateforme destinée aux modèles de fondation et à l'IA générative, proposant un studio, un entrepôt de données et une boîte à outils de gouvernance.
- De nouveaux produits Watson infusés des modèles de fondation et d'IA générative seront lancés pour le code, l'AIOps, le travail numérique, la sécurité et le développement durable.
- Une nouvelle collaboration avec Hugging Face va permettre d'apporter le meilleur des modèles d'IA open-source à l'entreprise sur la plateforme watsonx.
- IBM Consulting annonce un Centre d'Excellence pour l'IA générative avec plus de 1000 experts en IA prêts à mettre en œuvre la transformation des activités de leurs clients avec une IA pour les entreprises.



Armonk, N.Y., le 10 mai, 2023 : Aujourd'hui, IBM (NYSE : IBM) a annoncé, lors de sa conférence annuelle Think, [IBM watsonx](#), une nouvelle plateforme d'IA et de données qui permettra aux entreprises de mettre à l'échelle et d'accélérer l'impact de l'IA la plus avancée avec des données de confiance. Les entreprises qui se tournent aujourd'hui vers l'IA ont besoin d'accéder à une couche technologique complète qui leur permet d'entraîner, de régler et de déployer des modèles d'IA, y compris des modèles de fondation et des capacités de machine learning, à travers leur organisation avec des données de confiance, de la vitesse et de la gouvernance - le tout en un seul endroit et qui peut fonctionner dans n'importe quel environnement Cloud.

IBM a également annoncé à Think d'autres avancées prévues, notamment une offre d'infrastructure « GPU-as-a-service » conçue pour prendre en charge les applications intensives en IA, un tableau de bord alimenté par l'IA pour mesurer, suivre, gérer et aider à rendre compte des émissions de carbone dans le Cloud, ainsi qu'une nouvelle activité de service d'IBM Consulting pour watsonx et l'IA générative qui soutiendra le déploiement de l'IA chez les clients.

Avec watsonx, IBM propose un studio de développement de l'IA avec un accès à des modèles de fondation et des modèles open-source créés et entraînés par IBM, un accès à un entrepôt de données pour permettre la collecte et le tri des données d'entraînement et de réglage et une boîte à outils pour la gouvernance de l'IA dans les entreprises qui fournira un flux de travail d'IA transparent de bout en bout qui facilitera l'adaptation et la mise à l'échelle de l'IA.

*« Avec le développement des modèles de fondation, l'IA pour les entreprises est plus puissante que jamais », a déclaré **Arvind Krishna, IBM Chairman et CEO**. « Les modèles de fondation rendent le déploiement de l'IA beaucoup plus évolutif, accessible et efficace. Nous avons conçu IBM watsonx pour répondre aux besoins des entreprises, afin que les clients puissent être plus que de simples utilisateurs, ils peuvent devenir des acteurs de l'IA. Avec IBM watsonx, les clients peuvent rapidement former et déployer des capacités d'IA personnalisées dans l'ensemble de leur entreprise, tout en conservant le contrôle total de leurs données. »*

Les clients auront accès à l'ensemble des outils, à la technologie, à l'infrastructure et à l'expertise en matière de conseil pour créer leurs propres modèles d'IA - ou pour ajuster et adapter les modèles d'IA disponibles - sur leurs propres données et les déployer à grande échelle dans un environnement de confiance et plus ouvert afin de favoriser la réussite de l'entreprise. La différenciation concurrentielle et la valeur commerciale unique pourront être de plus en plus issues de la capacité d'adaptation d'un modèle d'IA aux données et aux connaissances uniques d'une entreprise.

La plateforme IBM watsonx se compose de [trois ensembles de produits uniques](#) pour répondre à ces besoins :

IBM watsonx.ai : un studio d'entreprise de nouvelle génération, qui devrait être disponible en juillet 2023, pour les créateurs d'IA afin d'entraîner, tester, régler et déployer à la fois le machine learning traditionnel et les nouvelles capacités d'IA générative alimentées par des modèles de fondation à travers une interface utilisateur ouverte et intuitive.

- Le studio d'IA fournit une gamme de modèles de fondation, d'outils d'entraînement et de réglage, ainsi qu'une infrastructure rentable qui facilite l'ensemble du cycle de vie des données et de l'IA, de la préparation des données au développement, au déploiement et à la surveillance des modèles.
- Le [studio](#) comprend également une bibliothèque de modèles de fondation qui permet aux utilisateurs d'accéder facilement à des modèles de fondation entraînés par IBM. Les modèles de fondation d'IBM utilisent un vaste ensemble de données d'entreprise, qui s'appuient sur un processus de filtrage et de nettoyage robuste et sur une lignée de données vérifiable. Ces modèles sont entraînés non seulement sur le langage, mais aussi sur une variété de modalités, y compris le code, les données chronologiques, les données tabulaires, les données géospatiales et les données sur les événements informatiques. Un premier ensemble de modèles de fondation sera mis à la disposition de certains clients en version bêta tech preview.

Voici quelques exemples de modèles :

- **fm.code** : modèles conçus pour générer automatiquement du code pour les développeurs par le biais d'une interface en langage naturel afin d'accroître la productivité des développeurs et de permettre l'automatisation de nombreuses tâches informatiques.
- **fm.NLP** : collection de langages de grande taille (LLM : large language models) pour des domaines spécifiques ou sectoriels qui utilisent des données conservées permettant d'atténuer plus facilement les biais et de les personnaliser rapidement à l'aide des données du client.
- **fm.geospatial** : **modèle** construit à partir de données climatiques et de données de télédétection pour aider les organisations à comprendre et à planifier les changements dans les schémas de catastrophes naturelles, la biodiversité, l'utilisation des sols et d'autres processus géophysiques susceptibles d'avoir un impact sur leurs activités.
- Dans le cadre d'une nouvelle collaboration entre IBM et Hugging Face, le studio watsonx.ai s'appuiera sur les bibliothèques open-source de Hugging Face et proposera des milliers de modèles et de données ouverts de Hugging Face. Cela fait partie de l'engagement d'IBM à fournir à ses clients une approche d'écosystème ouvert qui leur permet d'exploiter les meilleurs modèles et architectures pour leurs propres besoins commerciaux.

IBM watsonx.data : un entrepôt de données adapté construit sur l'architecture open lakehouse qui est optimisé pour les données régies et les applications d'IA, pris en charge par l'interrogation, la gouvernance et les formats de données ouvertes pour accéder aux données et les partager. La **solution** devrait être disponible en juillet 2023 :

- La solution peut gérer les applications à la fois sur site et à travers des environnements multi-cloud.
- Grâce à l'optimisation de l'application, avec cette solution, une organisation peut réduire les coûts de stockage de données jusqu'à 50 %.^[1]
- Watsonx.data permettra aux utilisateurs d'accéder à leur données de plus en plus volumineuses via un point d'entrée unique, tout en utilisant plusieurs moteurs d'interrogation adaptés pour découvrir des informations précieuses.
- Il fournira également des outils de gouvernance intégrés, une automatisation et des intégrations avec les bases de données et les outils existants d'une organisation afin de simplifier la mise en place et l'expérience des utilisateurs.

IBM watsonx.gouvernance : une boîte à outils de gouvernance de l'IA pour permettre des flux de travail d'IA de confiance. La solution, qui devrait être disponible plus tard cette année :

- Opérationnalise la gouvernance pour aider à atténuer le risque, le temps et le coût associés aux processus manuels et fournit la documentation nécessaire pour conduire des résultats transparents et explicables.
- Fournit les dispositifs permettant de protéger la vie privée des clients, de détecter de manière proactive les biais et les

dérives des modèles, et d'aider les organisations à respecter leurs normes éthiques.

Grâce à la plateforme watsonx, les clients sont en mesure de répondre aux attentes de leur organisation dans cinq domaines clés : l'interaction et la conversation avec les clients et les employés ; l'automatisation des flux de travail et des processus internes ; l'automatisation des processus informatiques ; la protection contre les risques ; et la réalisation d'objectifs de développement durable.

IBM prévoit également d'infuser les modèles de fondation `watsonx.ai` foundation de `watsonx.ai` dans tous ses principaux logiciels à l'avenir, par exemple :

- **Watson Code Assistant** : une solution, attendue dans le courant de l'année, qui exploite l'IA générative pour permettre aux développeurs de générer du code à l'aide d'une simple instruction en anglais.
- **AIOps Insights** : les capacités d'AI Operations (AIOps) sont améliorées par des modèles de fondation prévus pour le code et le NLP afin d'offrir une meilleure visibilité des performances dans les environnements informatiques, aidant ainsi les responsables des opérations informatiques (ITOps) et les ingénieurs de fiabilité des sites (SRE) à résoudre les incidents de manière plus rapide et plus rentable.
- **Watson Assistant et Watson Orchestrate** : les solutions de travail numérique d'IBM devraient être associées à un modèle de fondation NLP afin d'améliorer la productivité des employés et les expériences de service à la clientèle.
- **Environmental Intelligence Suite** : IBM EIS Builder Edition devrait être activée par le modèle de fondation géospatial et disponible en avant-première plus tard dans l'année, permettant aux organisations de créer des solutions sur mesure qui traitent et atténuent les risques environnementaux en fonction de leurs objectifs et besoins spécifiques.

Toujours dans le cadre de Think 2023, IBM annoncera également un certain nombre d'offres supplémentaires prévues pour favoriser l'adoption de l'IA, notamment :

- **Une nouvelle offre GPU sur IBM Cloud** : pour répondre au besoin mondial de modèles de fondation, IBM annonce de nouvelles offres GPU sur IBM Cloud, une infrastructure adaptée à l'IA conçue pour prendre en charge les tâches informatiques intensives des entreprises. Dans le courant de l'année, IBM devrait proposer une infrastructure complète de haute performance, flexible et optimisée pour l'IA, fournie en mode as a service sur IBM Cloud, pour l'entraînement et le service de modèles de fondation.
- **Un Centre d'Excellence pour l'IA générative proposé par IBM Consulting** : IBM Consulting annonce la création d'un centre d'excellence pour l'IA générative avec plus de 1 000 experts en IA générative et prévoit de créer une activité de service axée sur watsonx qui concevra et déploiera activement watsonx pour les clients. IBM Consulting a réalisé des dizaines d'engagements clients en infusant l'IA générative avec IBM Watson et un portefeuille de partenaires de l'écosystème par le biais de sa méthode éprouvée IBM Garage.

- **IBM Cloud Carbon Calculator** : un tableau de bord fondé sur l'IA pour permettre aux clients de mesurer, de suivre, de gérer et d'aider à rendre compte de leurs émissions de carbone associées à leur parcours hybride multi-cloud. Basé sur la technologie d'IBM Research, IBM Cloud Carbon Calculator devrait être disponible dans le courant de l'année. Le tableau de bord complète les solutions de développement durable existantes d'IBM avec un portefeuille complet de technologies et d'expertise, y compris la suite [IBM Envizi ESG Suite](#), [IBM Turbonomic](#), [IBM Planning Analytics](#) et [IBM LinuxONE](#), qui aident les organisations à accélérer leurs objectifs à la fois commerciaux et en matière de développement durable.
- **Une nouvelle étude révèle que l'IA générative fait partie des sept tendances qui façonnent les entreprises** : un nouveau rapport de l'IBM Institute for Business Value, intitulé Seven Bets, présente sept tendances qui ont un impact sur les entreprises aujourd'hui et décrit les sept paris qui valent la peine d'être faits pour améliorer leurs activités, y compris des idées sur les raisons pour lesquelles les entreprises devraient adopter un état d'esprit "AI-first" et comment les dirigeants peuvent capitaliser le plus efficacement possible sur les opportunités de l'IA aujourd'hui et à l'avenir, ainsi que gérer les risques accrus au sein de leurs organisations.

La retransmission de Think 2023 peut être visionnée ici : www.ibm.com/events/think.

Lors de son discours d'ouverture diffusé le mardi 09 mai, de 14h30 à 15h15, Arvind Krishna, IBM CEO and Chairman, a été rejoint sur scène par des dirigeants de Delta Airlines, Citi et Red Hat pour discuter de leurs visions pour transformer de manière créative les défis en opportunités avec le Cloud hybride et l'IA. Le replay peut être visionné ici :

<https://www.ibm.com/events/think/hybrid-cloud-and-ai/>

Le mercredi 10 mai, de 22h30 à 23h00, Darío Gil, Senior Vice President and Director of IBM Research, présentera dans le cadre d'une keynote, la façon dont l'IA générative peut être appliquée à des cas d'usage commerciaux.

Après l'événement phare d'Orlando, Think on Tour se rendra dans plus d'une douzaine de villes dans le monde. Pour obtenir des informations sur les villes confirmées, veuillez consulter le site suivant : <https://www.ibm.com/events/think/on-tour/>.

A propos d'IBM

IBM est un leader mondial du Cloud hybride et de l'IA, ainsi que des services aux entreprises, qui aide ses clients dans plus de 175 pays à capitaliser sur les connaissances issues de leurs données, à rationaliser leurs processus métier, à réduire leurs coûts et à acquérir un avantage concurrentiel dans leurs secteurs d'activité. Près de 4 000 entités gouvernementales et entreprises dans des domaines d'infrastructures critiques tels que les services financiers, les télécommunications et les soins de santé font confiance à la plateforme Cloud hybride d'IBM et à Red Hat OpenShift pour impacter leurs transformations numériques rapidement, efficacement et en toute sécurité. Les innovations révolutionnaires d'IBM en matière d'IA, d'informatique quantique, de solutions Cloud spécifiques à certains secteurs et de services aux entreprises offrent des options ouvertes et flexibles à nos clients. Tout cela est soutenu par l'engagement légendaire d'IBM en matière de confiance, de transparence, de responsabilité, d'inclusivité et de service.

Pour en savoir plus : www.ibm.com

###

Hugging Face et IBM s'associent pour démocratiser l'accès aux modèles de fondation

*« Pour bénéficier des dernières capacités de l'IA, les entreprises veulent s'appuyer sur le machine learning open-source : des modèles open-source formés sur des ensembles de données accessibles qui peuvent fonctionner dans un environnement sécurisé avec une conformité et une gouvernance des données appropriées », a déclaré **Clem Delangue, co-founder and CEO, Hugging Face**. « Nous sommes ravis de nous associer à IBM pour offrir aux développeurs une expérience de premier ordre basée sur le machine learning open-source et communautaire de Hugging Face, au sein de la nouvelle plateforme watsonx qui offre la lisibilité et la fiabilité d'IBM. »*

L'équipe PyTorch de Meta et IBM modifient une ligne de code pour améliorer massivement l'entraînement des modèles d'IA

*« Les entreprises d'aujourd'hui ne peuvent pas tirer le meilleur parti de leurs technologies d'IA et de Cloud si ces innovations ne sont pas équipées pour réussir à l'échelle », a déclaré **Damien Sereni, Engineering Director, Meta**. « Meta a bénéficié d'une collaboration de plusieurs années avec IBM pour relever ce défi majeur. Avec la sortie d'IBM watsonx, l'équipe PyTorch de Meta et IBM collaborent encore davantage pour construire une couche logicielle prête à la production pour l'entraînement de bout en bout, le réglage fin et l'inférence de modèles de fondation à grande échelle. »*

Comment le modèle de fondation IA d'IBM et de la NASA peut aider à suivre et à s'adapter au changement climatique

*« La NASA s'est engagée à partager entièrement, gratuitement et ouvertement plus de 100 PB de données et d'informations scientifiques. C'est la première fois que des modèles de fondation sont appliqués à l'immense archive de données sur les sciences de la Terre et nous travaillons avec IBM pour évaluer la pile de modèles de fondation dans le Science Managed Cloud Environment de la NASA », a déclaré **Kevin Murphy, Chief Science Data Officer à la NASA**. « La technologie d'IA de nouvelle génération peut accélérer notre analyse et notre découverte d'informations critiques relatives à la Terre et aux questions climatiques, permettant aux scientifiques d'acquérir des connaissances qui peuvent nous aider à protéger notre planète. »*

WIX s'appuie sur les langages de grande taille d'IBM pour améliorer le service à la clientèle

« Wix est fier d'exploiter les langages de grande taille d'IBM de pointe pour découvrir de nouvelles informations à partir de vastes quantités de données d'assistance à la clientèle dépouillées d'informations confidentielles et ainsi améliorer notre assistance à plus de 243 millions de clients à travers le monde », a déclaré **Yoav Abrahami, Chief Architect chez Wix and Head of Velo**.

Les plans, orientations et intentions d'IBM peuvent être modifiés ou retirés à tout moment, à la discrétion d'IBM et sans préavis. Les informations sur les futurs produits et les futures améliorations potentielles sont fournies pour donner une orientation générale des buts et objectifs d'IBM et ne doivent pas être utilisées pour prendre une décision d'achat. IBM n'est pas tenue de fournir du matériel, du code ou des fonctionnalités basés sur ces informations.

Contacts Presse :

IBM

Charlotte Maes

Tél. + 33 (0)7 86 09 83 33

charlotte.maes@ibm.com

Weber Shandwick pour IBM

Louise Weber / Jennifer Tshidibi

Tél. : + 33 (0)6 89 59 12 54 / + 33 (0)6 13 94 26 58

ibmfrance@webershandwick.com

[1] En comparant les prix publiés pour 2023 normalisés pour les heures VPC de watsonx.data à ceux de plusieurs grands fournisseurs d'entrepôts de données dans le Cloud. Les économies peuvent varier en fonction des configurations, des charges de travail et des fournisseurs.
