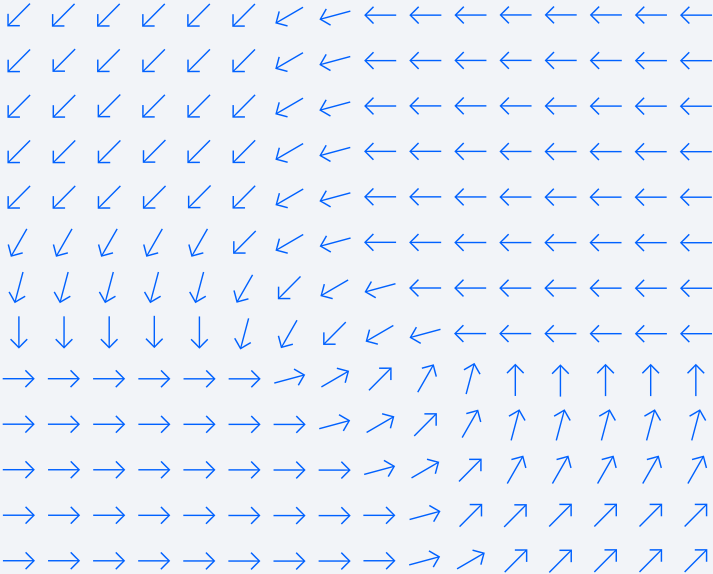




Pratiques MLOps
et IA fiable ↪
pour les
responsables des
données



Sommaire



01

Introduction : une approche des pratiques MLOps et de l'IA fiable basée sur la data fabric

02

Pourquoi mettre en place des pratiques MLOps et une IA fiable ?

03

Les éléments constitutifs des MLOps et d'une IA fiable

04

Data fabric : une approche holistique

05

Exemples de réussite de mise en œuvre de pratiques MLOps et d'une IA fiable

06

Éléments à prendre en considération

07

Créez votre solution idéale de pratiques MLOps et d'IA fiable

01

Introduction : une approche des pratiques MLOps et de l'IA fiable basée sur la data fabric

Maintenant que la confiance est devenue une valeur essentielle aux yeux des clients, chaque organisation se doit d'utiliser une IA éthique et explicable, respectant les droits individuels, la vie privée et les principes de non-discrimination.

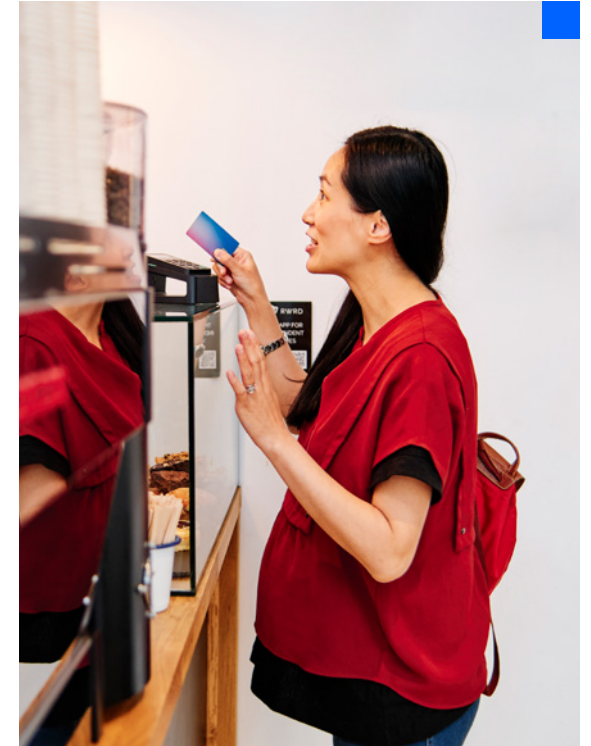
Lorsque l'IA est fiable, les revenus et la satisfaction des clients augmentent, les délais de mise sur le marché diminuent et le positionnement vis-à-vis de la concurrence s'améliore. À l'inverse, lorsque l'IA est négligée, la confiance envers l'organisation peut s'altérer et se traduire par des audits ratés ainsi que par des sanctions réglementaires entraînant une baisse de la réputation de la marque et de ses revenus.

Le succès de la construction, du déploiement et de la gestion des modèles d'IA/de ML repose sur des données fiables et des outils et processus automatisés de science des

données. Cela nécessite une plateforme technologique capable d'harmoniser de nombreux types et sources de données dans des environnements multicloud hybrides. L'architecture data fabric est une approche technologique qui garantit que des données de qualité sont accessibles aux bonnes personnes au moment opportun, quelle que soit leur localisation.

La data fabric constitue une assise solide pour les pratiques MLOps et une IA fiable. Poursuivez votre lecture pour en savoir plus ou essayez cette technologie vous-même avec notre [essai de MLOps et d'une IA fiable](#) gratuit.

En outre, la data fabric permet l'intégration de données multicloud, la gouvernance et la conformité des données, ainsi qu'une « intelligence client » à 360 degrés (ces cas d'application sont abordés dans d'autres eBooks).



02

Pourquoi mettre en place des pratiques MLOps et une IA fiable ?

Une IA bien préparée, exécutée et contrôlée, conçue pour atténuer les risques et produire les résultats analytiques souhaités, nécessite d'instaurer la confiance dans les données, les modèles et les processus.

Confiance envers les données

Pour disposer d'une IA solide, produisant des résultats dignes de confiance, il faut une connexion à des données précises, de grande qualité et prêtes à être utilisées en libre-service par des parties prenantes autorisées. La solidité du modèle d'IA dépend de sa capacité à agréger des données structurées et non structurées provenant de sources internes et externes disparates, depuis des emplacements locaux ou des clouds publics ou privés. L'efficacité de la collecte et de l'utilisation des données dépend de l'équité apportée à la formation des données, à leur traçabilité et à la garantie de la confidentialité lorsqu'une analyse en libre-service est proposée à plusieurs types d'utilisateurs.

Confiance envers les modèles

Les outils de science des données intégrés et automatisés par MLOps permettent de rendre opérationnels la mise en place, le déploiement et le suivi des modèles d'IA et d'en garantir la transparence et la responsabilité à chaque étape du cycle de vie. Les pratiques MLOps renforcent l'efficacité de l'intégration, du service et du déploiement en continu des flux de travail afin d'atténuer les biais, les risques et les écarts pour parvenir à des décisions plus précises basées sur les données. Certaines mises en œuvre uniques des MLOps apportent également équité, compréhension et robustesse au processus de modélisation de l'IA.

Confiance envers les processus

Tout au long du cycle de vie du modèle, de sa création, à son déploiement, sa gestion et son suivi, l'absence de processus automatisés peut conduire à des incohérences, à de l'inefficacité ou à un manque de transparence. La gouvernance de l'IA fournit une automatisation qui permet de mettre en place des processus cohérents et reproductibles qui réduisent les délais de production, améliorent la transparence des modèles, assurent la traçabilité et permettent l'extensibilité de l'IA.

L'IDC prévoit que, d'ici 2025, 60 % des entreprises auront rendu opérationnels leurs flux de travail de ML grâce à des fonctionnalités MLOps/ModelOps et intégré l'IA dans les opérations de leur infrastructure informatique grâce à des capacités d'AIOps.¹

03

Les éléments constitutifs des MLOps et d'une IA fiable

N°1

Collecte des données et accès aux données

La qualité des modèles et des algorithmes dépend de la qualité des données utilisées pour les créer. Si les données sont incomplètes ou partiales, les algorithmes et les résultats analytiques peuvent s'en trouver faussés. Il est essentiel de trouver une méthode permettant de relier ces sources entre elles.

Les scientifiques des données, les analystes et les développeurs ont besoin d'un accès en libre-service aux données les plus pertinentes pour leur projet. Sans les outils appropriés, il sera difficile de mettre en place des contrôles de la confidentialité pour plusieurs profils, de fournir un accès en temps réel aux « bons » utilisateurs de données et d'avoir une traçabilité des données.

N°2

Création et déploiement de modèles

Pour les scientifiques des données, le nettoyage et la préparation des données sont des processus manuels fastidieux et peu agréables. Beaucoup préfèrent consacrer leur temps à l'extraction ou à la modélisation de données. Une solution permettant d'automatiser ces tâches est essentielle.

Les scientifiques des données se plaignent souvent du manque d'outils intégrés pour la création, le déploiement, la mise à l'échelle et la formation des modèles. Utiliser plusieurs outils indépendants s'avère fastidieux, surtout lorsque ces outils ne s'accompagnent pas d'une documentation, de FAQ et de cas d'utilisation. L'absence d'outils intégrés peut être une source de problèmes de gouvernance et de conformité. Une fois le modèle créé, la collaboration est impérative entre le concepteur du modèle et les autres équipes, notamment celle de l'ingénierie logicielle, les développeurs d'IA et les analystes commerciaux. Élaborer une structure automatisée permettant cette collaboration au moment opportun revêt une importance essentielle.

N°3

Suivi et gestion des modèles

Une fois un modèle déployé, il faut impérativement en évaluer et en surveiller les performances. La dégradation du modèle, des écarts, des biais et d'autres facteurs peuvent entraîner des changements. L'utilisation de processus manuels pour détecter ces changements peut conduire à des erreurs coûteuses et à un défaut de gouvernance entraînant des transgressions réglementaires et une perte de confiance chez les clients.

L'automatisation des processus de surveillance et de requalification des modèles, mais aussi de la collecte des faits relatifs aux modèles tout au long de leur cycle de vie, renforce la cohérence et la transparence et permet de répondre automatiquement aux exigences réglementaires et au besoin d'équité des modèles.

TechTarget a calculé que 83 % des organisations ont accru le budget consacré à l'IA et que le nombre moyen de scientifiques des données a augmenté de 76 %. Malgré cela, le temps nécessaire au déploiement d'un modèle augmente également, 64 % des organisations devant y consacrer un mois ou plus.²

Data fabric : une approche holistique

Dans les sections précédentes, nous avons évoqué les éléments nécessaires pour instaurer la confiance envers les données, les modèles et les processus, ainsi que les défis auxquels de nombreuses organisations sont confrontées lorsqu'elles ne disposent pas d'une automatisation adaptée à chaque étape du cycle de vie de l'IA. Nous avons décrit les caractéristiques technologiques spécifiques requises à chaque étape pour mettre en place une IA bien préparée, exécutée et contrôlée.

Ces fonctionnalités sont les suivantes :

- Un moyen d'intégrer des données de divers types, issues de sources variées, à des déploiements différents
- Un accès en libre-service doté de contrôles de la confidentialité et d'un système de traçabilité des données
- Une automatisation de la création, du déploiement, de l'évolutivité, de la formation et du contrôle des modèles
- Une gouvernance automatisée pour garantir la qualité des données et la conformité aux réglementations

Pour mettre en place tous ces éléments dans l'entreprise et surmonter les problèmes de complexité des données, les organisations ont besoin d'une stratégie et d'une architecture de données intégrées. C'est là qu'intervient l'architecture data fabric.

La data fabric est une approche architecturale qui simplifie l'accès aux données d'une organisation et facilite la consommation de données en libre-service. Il rassemble les fonctionnalités citées plus haut dans une architecture unifiée, évitant ainsi le coût et la complexité de l'intégration d'une multitude de solutions spécifiques. Au lieu d'un ensemble fragmenté d'outils accolés les uns aux autres, la data fabric offre une solution unique et holistique, conçue pour fonctionner en toute transparence.

La data fabric connecte, gouverne et protège vos silos de données répartis dans un environnement de cloud hybride et peut vous aider à concrétiser les promesses de votre stratégie en matière de données.

Gartner prévoit que d'ici 2023, les entreprises utilisant des architectures data fabric connecteront, optimiseront et automatiseront de manière dynamique les processus de gestion de données et réduiront de 30 % les délais de mise à disposition des données intégrées.³

Dans la section suivante, vous verrez comment d'autres organisations ont déployé avec succès une approche de data fabric pour les MLOps et une IA fiable.

05

Exemples de réussite de mise en œuvre de pratiques MLOps et d'une IA fiable

Soins de santé : Penn Medicine ↻

Que ce soit pour accélérer la recherche, aider les médecins à prendre de meilleures décisions ou améliorer l'expérience des patients grâce à une médecine personnalisée, l'IA peut transformer le secteur de la santé.

Penn Medicine est un centre médical universitaire situé à Philadelphie. Sa réputation est mondiale et ses hôpitaux sont classés parmi les meilleurs des États-Unis. Le centre avait besoin d'un meilleur moyen de prévoir l'hypertension intracrânienne (crises d'HIC) chez les patients souffrant de lésions cérébrales traumatiques graves, afin de détecter plus rapidement cette affection et d'appliquer un traitement préventif.

Les neurochirurgiens de Penn Medicine utilisent diverses séries chronologiques de données physiologiques pour prendre des décisions thérapeutiques, mais le suivi des crises d'HIC est un processus manuel et laborieux.

L'équipe Data Science and AI Elite d'IBM a conçu un modèle de machine learning permettant de prévoir 20 à 30 minutes à l'avance l'augmentation de l'hypertension intracrânienne (crises d'HIC), un délai qui laisse suffisamment de temps aux neurochirurgiens pour intervenir et, potentiellement, prévenir ces crises. Cette solution axée sur les données a utilisé le machine learning pour découvrir des modèles cachés dans les séries chronologiques de données physiologiques. Le modèle pourrait permettre une prévision continue, à intervalles réguliers. Son intégration aux processus de travail cliniques pourrait également améliorer les résultats pour les patients victimes de lésions cérébrales traumatiques graves.



05

Exemples de réussite de mise en œuvre de pratiques MLOps et d'une IA fiable

Fabrication : ABB ↩

L'imprévisibilité de la demande client, les défaillances chroniques de la chaîne d'approvisionnement et les pénuries de main-d'œuvre obligent les organisations à réévaluer leurs processus de fabrication. L'accélération de la technologie numérique basée sur l'IA promet d'accroître l'efficacité opérationnelle à grande échelle, d'augmenter la productivité et d'améliorer la gestion de la chaîne d'approvisionnement. ABB est une multinationale dont le siège social se trouve en Suisse et qui intervient dans les domaines de la robotique, de l'énergie, des équipements électriques et des technologies d'automatisation. Elle souhaitait exploiter ses données historiques en matière de spécifications de tuyaux pour réaliser des analyses avancées et ainsi préserver, entretenir, mais aussi élargir et partager, les connaissances de ses ingénieurs commerciaux partant à la retraite.

L'entreprise avait besoin d'un processus complet offrant les capacités suivantes :

- Possibilité d'exploiter des données en dépit de leur formatage inadéquat
- Utilisation du machine learning pour l'analyse
- Utilisation des résultats pour avoir un impact réel sur les processus métiers

La création d'un pipeline de bout en bout a permis d'obtenir des informations clés sur des techniques d'analyses avancées, telles que l'exploration d'associations et le machine learning non supervisé, pour révéler des modèles cachés.

La mise en œuvre de technologies d'IA aide ces entreprises traditionnelles à adopter des canaux numériques offrant des options mobiles et de libre-service. En encapsulant les résultats dans un prototype Web, ABB a pu considérer le développement éventuel d'une application interne utile pour ses ingénieurs commerciaux.

ABB a ainsi pu :

- Réduire la complexité et bénéficier d'une infrastructure de données transparente
- Réduire les autres coûts et les ressources pour former ses nouveaux ingénieurs commerciaux grâce à une nouvelle approche basée sur les données
- Gagner du temps dans la création des spécifications des tuyaux, un processus qui aurait dû prendre plusieurs mois et non quelques jours
- Effectuer pour la première fois des analyses avancées sur son historique de données de spécifications des tuyaux



05

Exemples de réussite de mise en œuvre de pratiques MLOps et d'une IA fiable

Services financiers : NatWest Group ↻

Les clients actuels attendent des expériences exceptionnelles et fluides tous les domaines. Pour créer un avantage concurrentiel durable et maintenir la confiance, les agences bancaires, les caisses d'épargne et de crédit et les autres établissements financiers doivent comprendre les besoins actuels et futurs de leurs clients et savoir comment y répondre.

Pour de nombreuses personnes, devenir propriétaire de leur logement est la concrétisation d'un rêve. Un prêt hypothécaire accordé par une banque est souvent indispensable à la réalisation de ce rêve. L'émission et l'obtention d'un prêt hypothécaire peuvent toutefois se compliquer au fur et à mesure de l'évolution des réglementations, des produits et des processus. Les banques doivent être en mesure d'accéder à des informations précises sur leurs politiques et de les appliquer aux besoins uniques de chaque client en temps réel, tout au long du processus d'achat d'un logement.

La Royal Bank of Scotland (RBS), aujourd'hui Natwest Group, résout ce problème avec un programme numérique d'aide aux prêts hypothécaires. Elle a créé une plateforme cloud

alimentée par l'IA qui permet à ses employés des centres d'appels pour les prêts d'apporter une aide numérique en temps réel aux acheteurs de logement. La plateforme, baptisée « Marge », est un nouvel atout précieux pour la transformation numérique de RBS.

Elle a été volontairement personnifiée en tant que membre de l'équipe de RBS et possède même sa propre personnalité évolutive. Marge a été créée directement dans le cloud et intégrée aux structures de données existantes de RBS. Elle a accès à de nouvelles données ajoutées chaque minute par le biais des mises à jour de contenu et des interactions avec les clients. Pour les employés du centre d'appel des prêts de RBS, elle constitue un point d'accès unique à une assistance numérique intelligente sur les prêts hypothécaires.

Lorsqu'ils sont au téléphone avec un client, les conseillers de RBS tapent des mots clés sur une console et Marge leur fournit une assistance numérique rapide en matière de prêt hypothécaire. Grâce à la technologie d'entreprise cognitive à leur disposition, ils peuvent aider les clients, anciens ou nouveaux, qui achètent un logement.



Depuis la mise en service de son outil numérique d'aide au prêt hypothécaire, RBS a enregistré une amélioration de 20 % du taux de satisfaction des clients et une diminution de 10 % de la durée des appels. Au fur et à mesure de l'évolution de Marge, les employés de la Royal Bank of Scotland auront les moyens de faire un excellent travail pendant la transformation numérique.

« Le secteur du crédit hypothécaire change en permanence. Changements de réglementation, changements de produits, changements de processus. Il est impératif que le client dispose de l'aide et des informations dont il a besoin pour pouvoir avancer dans son projet d'achat immobilier. »

MaryAnn Fleming

Responsable des services d'achat immobilier

Royal Bank of Scotland

06

Éléments à prendre en considération

IBM propose une plateforme de science des données complète, à laquelle les utilisateurs peuvent accéder en libre-service, et qui intègre des outils mis à jour et automatisés pour créer des modèles et les passer en production de manière continue.

Avantages :

- Regroupement des données et des services d'IA sur une même plateforme intégrée
- Intégration des données améliorée avec l'automatisation des tâches d'ingénierie
- Automatisation de la gouvernance, de la protection des données et de la sécurité grâce aux métadonnées actives
- Outils de collaboration à disposition
- Outils dédiés à la gestion des risques liés aux modèles d'IA

L'approche data fabric d'IBM repose sur une plateforme de données natives du cloud avec IA destinée à améliorer la productivité et à réduire la complexité en connectant des données en silos, des sources et des charges de travail de tous types dans un environnement de cloud hybride offrant un accès en libre-service et des modèles d'IA.

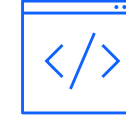


IBM Cloud Pak for Data

IBM Cloud Pak® for Data est une plateforme construite spécifiquement pour une architecture data fabric et visant à prévoir les résultats plus rapidement pour vous permettre de collecter, d'organiser et d'analyser vos données, quel que soit leur emplacement. La plateforme permet ainsi d'améliorer la productivité et de réduire la complexité en construisant une architecture data fabric qui relie les données en silo distribuées dans un environnement de cloud hybride.

Rendez les opérations technologiques plus agiles et plus efficaces tout en réduisant les coûts et les risques liés à la combinaison des opérations d'IA et de DevOps. Avec à son vaste écosystème, IBM Cloud Pak for Data vous aide à étendre vos investissements libres ou propriétaires et à exploiter les dernières innovations d'IBM en matière d'IA. Tout en développant votre ingénierie d'IA, vous modernisez votre architecture d'information pour qu'elle soit pérenne.

[En savoir plus sur IBM Cloud Pak for Data →](#)



IBM Watson Studio

IBM Watson® Studio réunit des outils permettant d'explorer les données, de créer des modèles (visuellement ou à l'aide de code) et de déployer et surveiller ces modèles en bénéficiant d'explications et d'équité tout au long du cycle de vie. Watson Studio utilise les pratiques MLOps pour simplifier la production de modèles à partir de n'importe quel outil. Il permet de former à nouveau des modèles de manière automatique, favorise la transparence et contrôle la précision et les biais des modèles dans la durée.

En tant que composante intégrale de Watson Studio, IBM Factsheets enregistre la manière dont chaque modèle est développé et déployé. Une transparence accrue fournit des informations aux utilisateurs d'IA afin qu'ils comprennent plus aisément comment un modèle ou un service a été créé et qu'ils déterminent s'il est adapté à une situation ou à un besoin spécifique.

[En savoir plus sur IBM Factsheets →](#)

[En savoir plus sur IBM Watson Studio →](#)

06

Éléments à prendre en considération



IBM Watson Knowledge Catalog

IBM Watson® Knowledge Catalog offre un catalogage intelligent, avec collecte automatique des métadonnées et gestion des politiques pour veiller à ce que les détails d'un modèle soient automatiquement collectés et stockés pour une transparence et une répétabilité maximales. La solution garantit que les modèles sont impartiaux, qu'ils ne sont pas biaisés, qu'ils sont explicables et qu'ils s'adaptent à l'évolution des paramètres du modèle.

[En savoir plus sur IBM Watson Knowledge Catalog →](#)



IBM OpenPages with Watson

IBM OpenPages with Watson vous aide à identifier, gérer et surveiller le risque et la conformité réglementaire tout au long du cycle de développement du modèle, et à produire des rapports, sans devoir modifier les exigences des outils d'IA/de ML utilisés aujourd'hui.

[En savoir plus sur IBM OpenPages →](#)



IBM propose également des boîtes à outils de gouvernance d'IA en open source. [AI Fairness 360](#) facilite l'examen, le signalement et l'atténuation des biais dans les modèles tout au long du cycle de vie des applications d'IA. [AI Explainability 360](#) inclut des mesures expliquant les processus et les prises de décisions d'un modèle.

07

Créez votre solution idéale de pratiques MLOps et d'IA fiable

Si vous voulez avoir davantage confiance en votre IA, nous vous encourageons à profiter de quelques-unes de nos ressources gratuites. La plus importante de ces ressources est un essai des pratiques de [MLOps et de l'IA fiable](#) qui vous permettra d'acquérir une expérience concrète d'une architecture data fabric conçue dans ce but. Vous pouvez également consulter les informations sur les pratiques de [MLOps et l'IA fiable](#) disponibles sur notre site Web pour en savoir encore plus sur leurs avantages. Enfin, vous pouvez contacter l'un de nos spécialistes en [prenant rendez-vous en ligne](#), en parlant avec votre représentant IBM ou en vous adressant à l'un de nos partenaires commerciaux.

Consultez les trois autres eBooks sur les cas d'utilisation d'une architecture data fabric :

- [Intégration de données multicloud](#)
- [Gouvernance des données et confidentialité](#)
- [Vue client à 360°](#)





© Copyright IBM Corporation 2022

Compagnie IBM France
17 avenue de l'Europe
92275 Bois-Colombes Cedex

Produit aux États-Unis d'Amérique
Mai 2022

IBM, le logo IBM et ibm.com sont des marques d'International Business Machines Corp. enregistrées dans de nombreux pays. Les autres noms de produits et de services peuvent être des marques d'IBM ou d'autres sociétés. La liste actualisée de toutes les marques d'IBM est disponible sur la page Web « Copyright and trademark information » à l'adresse [ibm.com/legal/copytrade.shtml](https://www.ibm.com/legal/copytrade.shtml).

L'information contenue dans ce document était à jour à la date de sa publication initiale et peut être modifiée sans préavis par IBM. Les offres mentionnées dans le présent document ne sont pas toutes disponibles dans tous les pays où la société IBM est présente.

Les données de performance et les exemples de clients rapportés dans ce document sont présentés à titre d'illustration uniquement. Les résultats de performance réels peuvent varier en fonction des configurations et des conditions de fonctionnement spécifiques.

LES INFORMATIONS CONTENUES DANS LE PRÉSENT DOCUMENT SONT FOURNIES « EN L'ÉTAT », SANS AUCUNE GARANTIE EXPLICITE OU IMPLICITE. IBM DÉCLINE NOTAMMENT TOUTE RESPONSABILITÉ RELATIVE À CES INFORMATIONS EN CAS DE CONTREFAÇON AINSI QU'EN CAS DE DÉFAUT D'APTITUDE À L'EXÉCUTION D'UN TRAVAIL DONNÉ. Les produits IBM sont garantis conformément aux dispositions des contrats qui régissent leur utilisation.

- 01 IDC Future Scape: Worldwide Artificial Intelligence and Automation 2022 Predictions—European Implications
IDC, Jan 14, 2022, <https://www.idc.com/getdoc.jsp?containerId=EUR148675522>
- 02 TechTarget, “Addressing 3 infrastructure issues that challenge AI adoption,” May 14, 2021
- 03 Magic Quadrant for Data Integration Tools, Gartner, 9 March 2021.