



INNOVATION AVEC L'IA

ALLER DE L'AVANT AVEC DATAOPS ET MLOPS



ENQUÊTE THÉMATIQUE

Auteurs

Alexandre Rodé
Timm Grosser

Cette étude a été préparée par BARC, une société d'analyse de marché indépendante. Il est disponible gratuitement grâce à la générosité de DataRobot, Domino Data Lab et ONE LOGIC.



DOMINO
DATA LAB

DataRobot



ONE LOGIC



CONTENU

Avant-propos.....	2
Résumé de gestion.....	3
Applications ML hautes performances: tout sauf la norme	6
DataOps et MLOps sont des approches prometteuses pour une plus grande efficacité en ML.....	9
Un impact positif indiscutable.....	dix
Le choix du logiciel est important.....	15
Les entreprises prennent de l'avance avec DataOps et MLOps, mais doivent veiller à embarquer tout le monde	18
Démographie	20
Profil BARC	22
Profil des parrains	23
Auteurs	26

AVANT-PROPOS

L'intelligence artificielle (IA), l'apprentissage automatique (ML) et la science des données font partie d'un grand groupe de mots à la mode associés à des percées majeures dans la façon dont les affaires sont menées et à de meilleurs résultats. Dans cette étude, nous examinerons les mécanismes à l'origine de ces réussites et les méthodes qui entrent en jeu pour les concrétiser. En bref, nous nous concentrerons sur les concepts de DataOps et de MLOps et leur impact sur l'application du ML.

Nous évaluerons en détail la perception et l'adoption de ces concepts ainsi que leur contribution au succès de l'IA, ou plus spécifiquement, des applications ML dans les entreprises. Nous aborderons donc le périmètre fonctionnel du DataOps étendu par MLOps sans différencier les deux. Il s'agit d'une étude exploratoire qui donne un aperçu et révèle les interrelations.

Parce qu'il existe des défis conceptuels liés à DataOps et MLOps, nous aimerions d'abord partager notre compréhension de ces deux concepts. Cela illustrera pourquoi nous ne les distinguons pas tout au long de cette étude.

DataOps et MLOps : seulement une transition progressive de l'un à l'autre En raison du grand chevauchement fonctionnel et des diverses interprétations, le marquage des frontières entre DataOps et MLOps est arbitraire et ne produit jamais de résultats satisfaisants. Par conséquent, nous considérerons ces concepts comme un seul pour les besoins de cette étude.

DataOps se concentre sur la réalisation d'un flux gérable, maintenable et automatisé de données de qualité garantie. L'objectif est d'atteindre la transparence

concernant toutes les interdépendances entre les systèmes impliqués le long d'un pipeline de données de bout en bout. C'est un concept qui favorise la collaboration entre experts et rend le processus de développement de produits de données plus agile et efficace. Il est pertinent pour toutes sortes de produits de données dont le succès dépend de leur mise à jour.

Cela s'applique également sans exception aux modèles d'apprentissage automatique (ML), qui sont un type particulier de produit de données. Ainsi, le développement et le déploiement de modèles ML s'accompagnent d'exigences particulières telles que le recyclage, les tests, le suivi des métriques de performance pertinentes et le versionnage des différentes configurations de modèles ainsi que la gestion générale du code. Le maintien de processus efficaces et automatisés pour la satisfaction de ces exigences particulières est pris en charge par les MLOps.

Par conséquent, nous considérons MLOps comme une extension fonctionnelle de DataOps. Les deux concepts dérivent des principes DevOps et pourraient être considérés comme une extension de ceux-ci avec un accent particulier sur la qualité, la validité et la fiabilité des données en tant que support de la logique.

Nous considérons DataOps, MLOps et DevOps comme des concepts de conception de processus qui garantissent le fonctionnement sans faille des logiciels et des produits de données. Les outils et solutions dans ce domaine fournissent généralement des fonctions et des cadres techniques pour le support de ces processus.

Alexander Rode et Timm Grosser

Wurtzbourg, mai 2022

RÉSUMÉ DE GESTION



01 LES APPLICATIONS ML HAUTES PERFORMANCES SONT TOUT MAIS LA NORME DANS LES ENTREPRISES D'AUJOURD'HUI

La plupart des entreprises commencent tout juste à exploiter les avantages du ML pour elles-mêmes. 55 % des entreprises représentées dans cette enquête n'ont pas encore déployé de modèle ML et seulement 10 % se considèrent avancées dans ce domaine.

Dès que les modèles ML sont prêts à être transférés en production, les choses commencent vraiment à se compliquer. Le déploiement est un obstacle difficile à surmonter pour de nombreuses entreprises.

02 DATAOPS ET MLOPS SONT GÉNÉRALEMENT RECONNUS DES CONCEPTS POUR FAIRE FACE AUX DÉFIS TYPIQUES DE ML

Ces concepts sont à la hauteur de leur réputation et apportent des solutions efficaces à de nombreux défis (par exemple, collaboration efficace, documentation, surveillance, automatisation).

Les entreprises familiarisées avec les sujets liés aux DataOps et aux MLOps ont des attentes plus réalistes quant à ce qu'elles peuvent réaliser avec l'apprentissage automatique et peuvent mieux planifier les projets de ML. Les utilisateurs de données/MLOps sont 3,5 fois moins susceptibles d'être confrontés à une complexité écrasante.

L'introduction de Data/MLOps permet aux modèles ML d'être déployés plus rapidement et plus efficacement tout en garantissant la qualité des opérations. Les adopteurs de données/MLOps sont 4,2 fois plus susceptibles de pouvoir se déployer rapidement (en quelques semaines ou quelques jours).



RECOMMANDATION

Développer des modèles ML est la partie la plus facile. Informez-vous tôt sur les défis du déploiement en vous renseignant sur DataOps et MLOps. Cela vous aidera à vous préparer ainsi qu'à éviter les contretemps et les mauvaises surprises.



RECOMMANDATION

Avant de vous fier aux résultats et au bon fonctionnement des modèles ML en production, assurez-vous d'avoir tout sous contrôle en cas de problème. DataOps et MLOps peuvent vous donner un bon guide pour comprendre ce qui peut mal tourner, comment éviter les erreurs et comment réagir rapidement en cas de besoin.

De cette façon, vous pouvez garantir l'application sûre du ML et garantir son acceptation.

RÉSUMÉ DE GESTION



RECOMMANDATION

Vous pouvez utiliser l'open source tant que vous pouvez gérer la complexité. Les outils commerciaux, en particulier les solutions de plate-forme, peuvent vous aider à mieux faire face à la complexité et à déployer plus rapidement. Basez votre sélection de logiciels sur les exigences actuelles et futures (par exemple, surveillance et documentation). En termes d'exigences futures, vous pouvez également vous appuyer sur les concepts de DataOps et de MLOps.



RECOMMANDATION

N'oubliez pas de renforcer le soutien au ML et à la science des données dans toute l'entreprise. La mise en œuvre technique et procédurale de DataOps et MLOps est une exigence importante pour une application réussie du ML, mais la résistance et la peur déraisonnable parmi les employés peuvent constituer un obstacle insurmontable au progrès.

03 LE BON OUTILLAGE ML PEUT AVOIR UN IMPACT SIGNIFICATIF SUR LE SUCCÈS DU ML ET AIDE À FAIRE DES PROCESSUS DE DONNÉES/MLOPS PLUS FACILE À METTRE EN ŒUVRE

Les piles d'outils ML de la plupart des entreprises sont dominées par des solutions open source, tandis que les solutions de plate-forme sont encore relativement rarement utilisées.

L'effet de réduction de la complexité et d'augmentation de l'efficacité des outils commerciaux et des solutions de plate-forme prend tout son sens une fois les modèles ML déployés. Les utilisateurs pratiquant l'apprentissage automatique d'outils commerciaux sont 8,25 fois moins susceptibles de déclarer être submergés par la complexité que les utilisateurs de solutions open source.

Parmi les praticiens du domaine du ML, les utilisateurs d'outils open source sont plus fréquemment aux prises avec une complexité écrasante et sont moins susceptibles de pouvoir se déployer rapidement.

04 BÉNÉFICIEZ DES DATAOPS ET MLOPS TOUT EN AUGMENTANT SOUTIEN DANS TOUTE L'ENTREPRISE

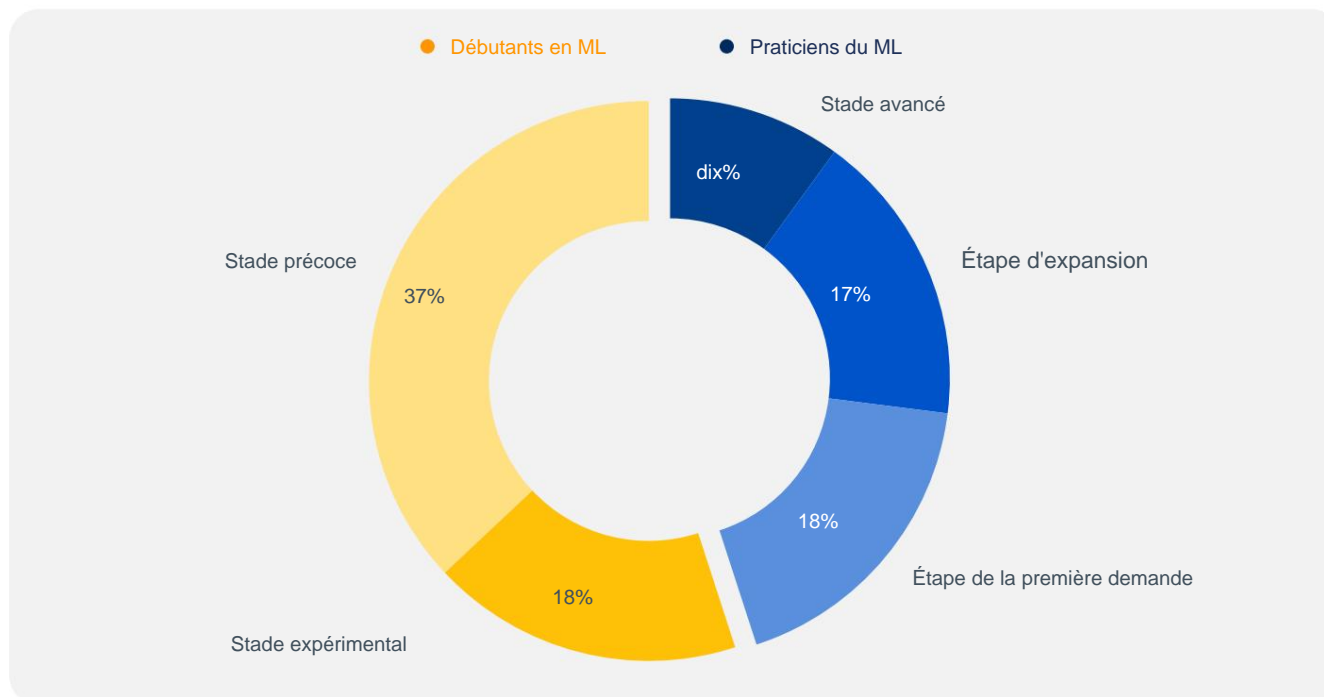
Les avantages obtenus grâce aux Data/MLOps influencent également positivement le succès d'autres mesures telles que l'embauche d'experts et l'utilisation de nouveaux outils, plateformes et infrastructures.

Les adopteurs de Data/MLOps semblent se concentrer davantage sur les améliorations techniques et procédurales que sur les mesures visant à ancrer un état d'esprit axé sur les données dans toute l'entreprise en augmentant généralement les compétences en matière de données.

01 APPLICATIONS ML HAUTES PERFORMANCES : TOUT SAUF LA NORME



DÉPLOYÉS OU NON DÉPLOYÉS : DIFFÉRENCES ENTRE LES PRATICIENS ML ET LES DÉBUTANTS COMPTENT



Veuillez décrire le niveau d'avancement de votre entreprise dans le domaine de l'apprentissage automatique (n=248)

Les résultats de notre enquête sont surprenants et quelque peu trompeurs. Malgré le battage médiatique et les nombreuses réussites sur une longue période, la majorité des entreprises semble être bloquée dans une phase précoce de l'avancement du ML.

Plus de la moitié des participants à notre enquête (55%) n'ont encore opérationnalisé aucun modèle de ML, et ceux qui n'ont même pas commencé à développer des modèles de ML constituent le groupe le plus important (37%). Ce résultat reflète une réalité où

de nombreuses entreprises découvrent seulement maintenant le potentiel du ML.

Des informations intéressantes sont fournies par une vision différenciée de ceux qui utilisent déjà des modèles ML et de ceux qui n'ont pas encore mis en œuvre de modèles ML.

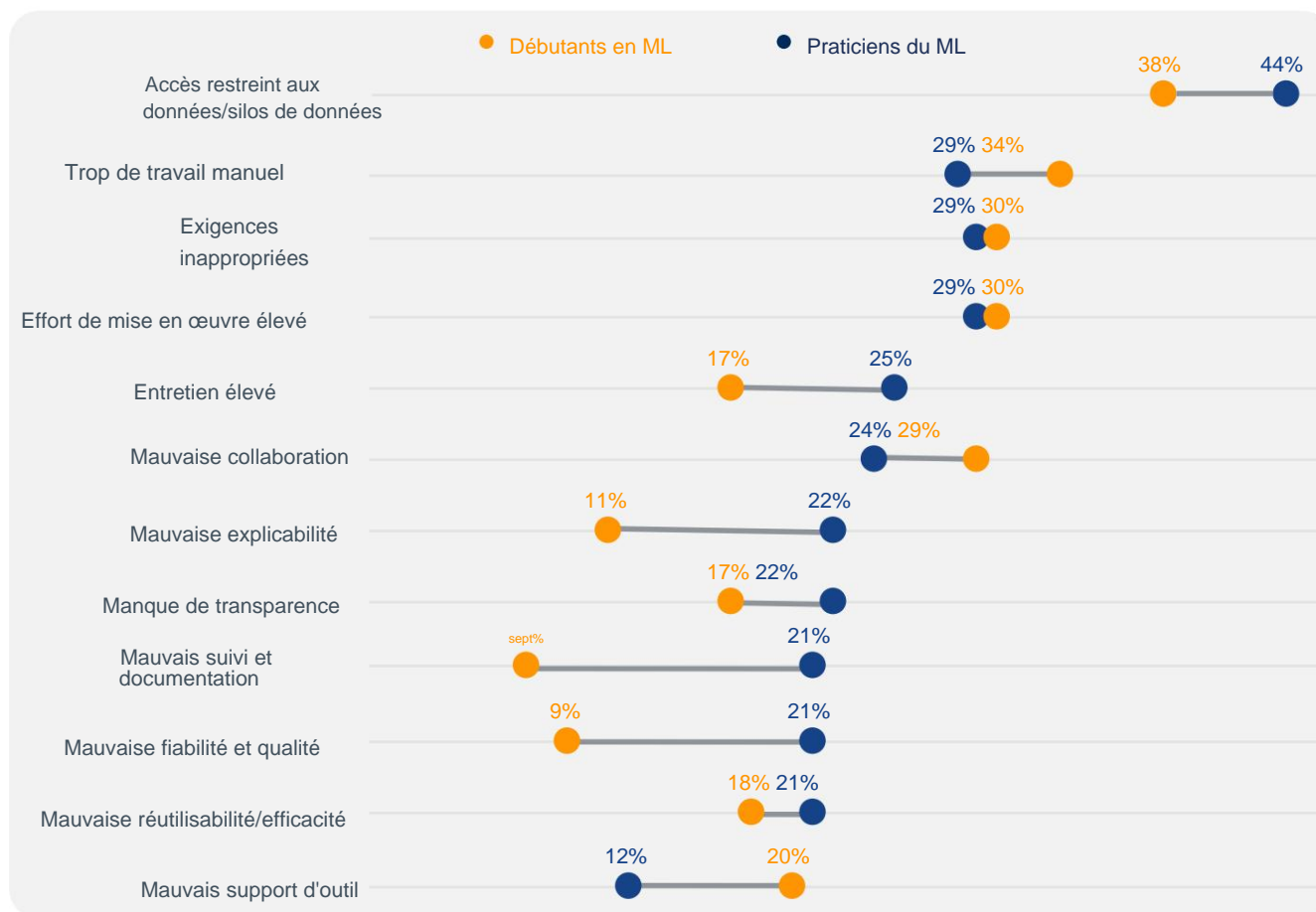
Les défis changent avec l'échelle et la complexité des initiatives de ML. La vue différenciée nous donne une idée de ce à quoi il faut prêter attention dans un premier temps, à court terme, mais aussi des défis qui peuvent survenir à moyen et long terme.

Par conséquent, nous différencierons régulièrement les entreprises qui ont déjà déployé au moins un modèle ML (praticiens ML) et celles qui ne l'ont pas fait (débutants ML).

01 APPLICATIONS ML HAUTES PERFORMANCES: TOUT SAUF LA NORME



CHAQUE NIVEAU DE ML DEVIENT PLUS DIFFICILE: PLUS VOUS AVEZ PLUS LOIN, PLUS VOUS AVEZ DE DÉFIS DOIT SURMONTER



Quels sont les principaux problèmes auxquels votre entreprise est confrontée lors du développement et du déploiement de modèles de ML? (n=244)

En moyenne, les praticiens du BC ont indiqué qu'ils sont actuellement confrontés à 3,5 des 12 problèmes disponibles pour la sélection dans l'enquête. Ce taux dépasse celui des débutants de plus de 20 %.

Vraisemblablement, les praticiens font face à plus de défis que les débutants en général. Cela souligne l'hypothèse selon laquelle des solutions de ML bien gérées et performantes, souvent présentées comme des projets phares, ne sont pas du tout la norme dans le ML appliqué et la plupart des entreprises ont du mal à les fournir. Les praticiens ont déjà structuré et professionnalisé leur travail et considèrent l'effort manuel, la collaboration et l'utilisation d'outils comme des problèmes beaucoup moins fréquemment que les débutants.

Néanmoins, nous constatons des problèmes de fonctionnement nettement plus importants. Il s'agit notamment des défis de pouvoir s'appuyer sur les travaux antérieurs, d'établir la transparence en ce qui concerne les interdépendances, de maintenir les applications en cours d'exécution et d'assurer la surveillance et la documentation des artefacts importants ainsi que la fiabilité, la qualité et l'explicabilité des résultats générés.

01 APPLICATIONS ML HAUTES PERFORMANCES: TOUT SAUF LA NORME



Ce constat rappelle d'autres projets orientés code, comme le chargement d'un entrepôt de données.

Malgré les meilleures intentions de documenter correctement le code ETL, de conserver les métadonnées et de créer du code réutilisable à l'aide de fonctions standard, la réalité est souvent différente. Cela s'explique généralement par les pressions exercées par des activités quotidiennes flexibles associées à la complexité croissante du paysage des données et à l'augmentation des exigences dynamiques.

Soit le temps manque, soit les structures pour

la création de code durable et le fonctionnement efficace ne sont pas en place. Cette image est similaire dans le domaine de ML. Mais dans ce cas, nous avons la chance de faire meilleur.

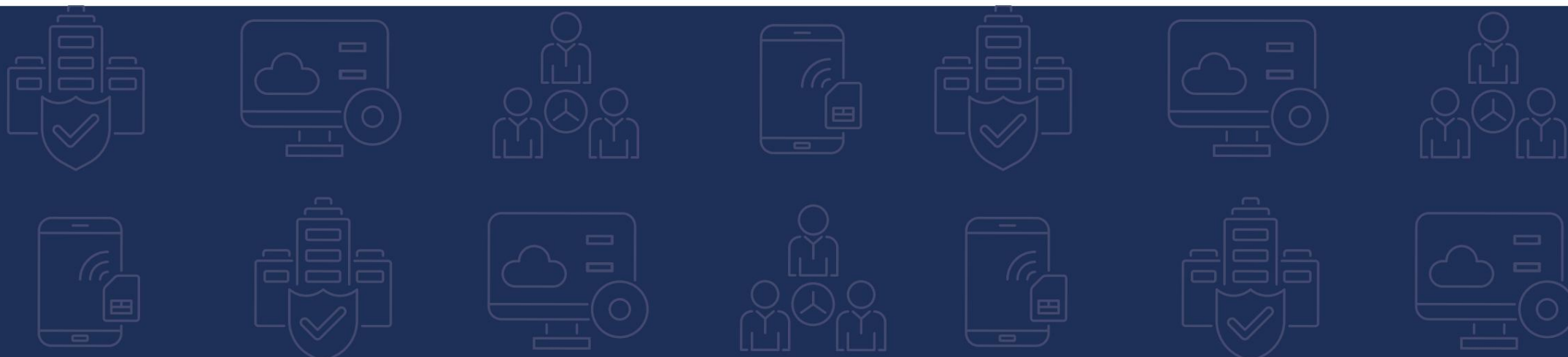
Le problème des silos de données mérite également d'être souligné. Les silos de données représentent l'un des plus grands défis de notre époque. Cela vaut également dans le domaine du ML, pour les praticiens (44%) encore plus que pour les débutants (38%). De notre point de vue, ce défi peut être relevé à l'aide d'une stratégie de données globale et d'un changement culturel concernant le traitement des données.

Comme nous le verrons plus loin, la sous-représentation d'autres problèmes non mentionnés ici est plutôt un indicateur d'une méconnaissance que de l'absence de problèmes.

Pour **44%** des praticiens du ML

et **38%** des débutants en ML,

l'accès restreint aux silos de données est le plus grand défi lors du développement et du déploiement de modèles ML.

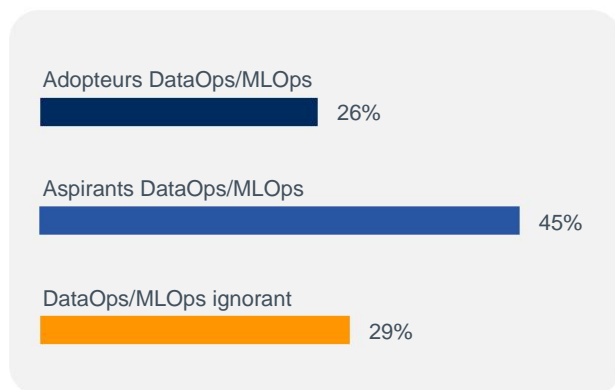




02 DATAOPS ET MLOPS SONT PROMETTEURS

APPROCHES POUR UNE PLUS GRANDE EFFICACITÉ EN ML

IMPORTANCE DES DATAOPS ET MLOPS LARGEMENT RECONNUE



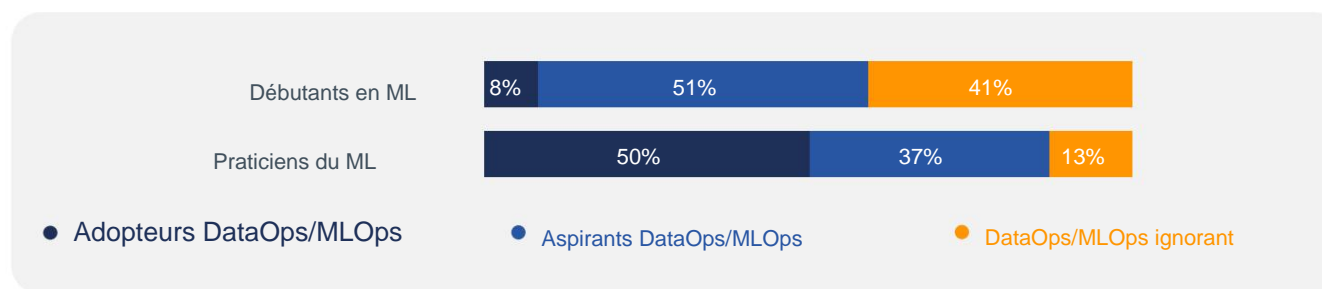
Statut d'adoption de DataOps/MLOps (n=228)

Développer et déployer avec succès le ML est un défi et sa complexité est souvent sous-estimée, comme l'ont indiqué 68 % de nos répondants. Il est largement admis que la mise en œuvre de DataOps et de MLOps permet aux entreprises de faire face à ces défis complexes et d'augmenter leur efficacité.

Comme la majorité des entreprises représentées dans cette enquête n'ont même pas encore déployé de modèles ML opérationnalisés, il n'est pas surprenant que la proportion de celles qui adoptent déjà Data/MLOps ne soit pas très importante à 26 % (voir graphique).

Étant donné que le pourcentage de ces entreprises qui envisagent de travailler avec Data/MLOps à l'avenir (aspirants) est très élevé à 45%, on peut conclure que DataOps et MLOps sont largement perçus comme des approches prometteuses qui n'ont été adoptées que par quelques pionniers. à ce jour.

Même parmi les débutants en ML, la majorité des entreprises prévoient au moins de travailler avec Data/MLOps (51 %) et un petit pourcentage ont même commencé à l'implémenter (8 %). L'importance de DataOps semble être largement reconnue à tous les niveaux de maturité ML.



Adoption de DataOps/MLOps par les praticiens du ML et les débutants (n=228)

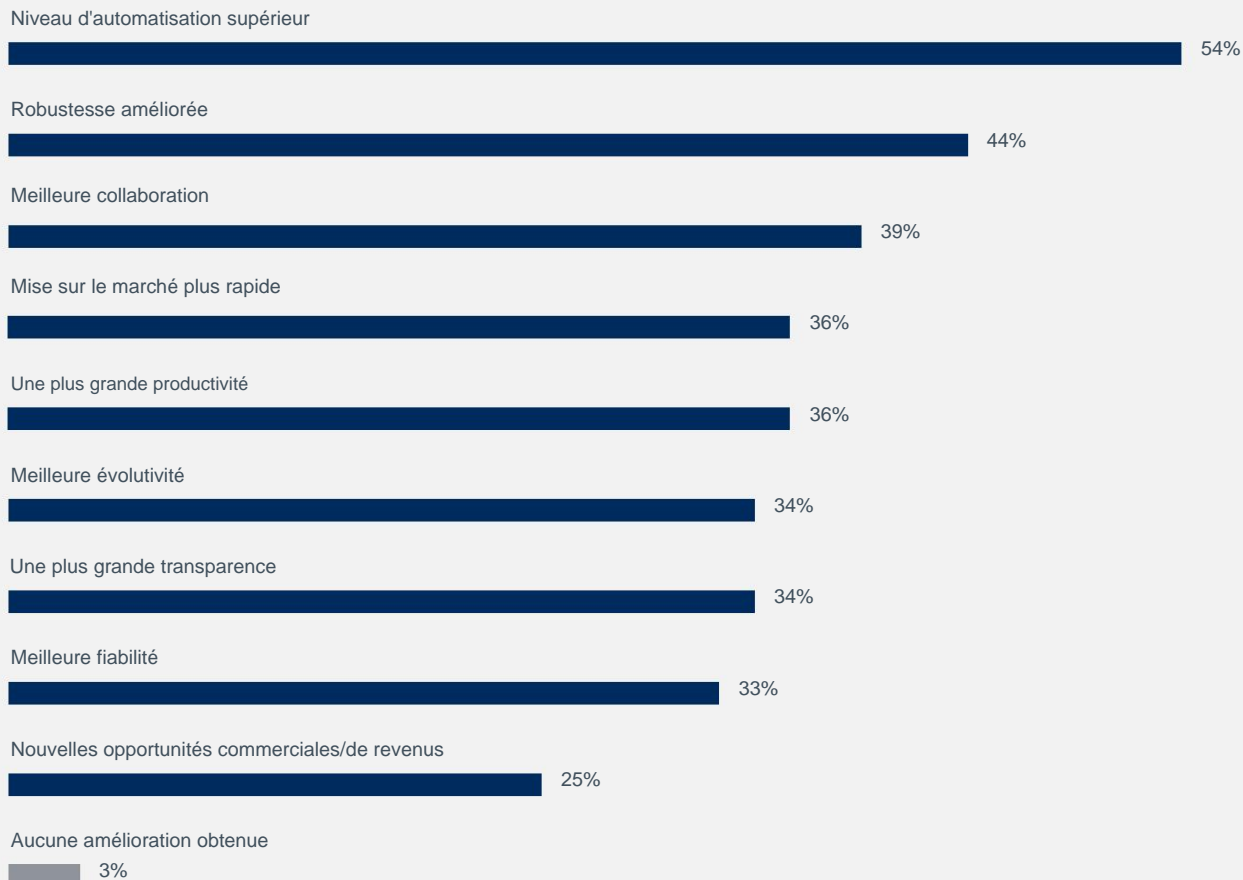
Cette impression devient encore plus apparente lorsque l'on différencie les praticiens du ML des débutants (voir graphique). Les praticiens du ML ont pour la plupart commencé à adopter DataOps et MLOps (50%) ou évaluent les étapes à suivre pour le faire (37%). Il semble qu'il n'y ait aucun moyen de contourner Data/MLOps pour faire face aux complexités découlant du déploiement du ML.

Mais pourquoi devrait-on se soucier des DataOps et des MLOps avant d'avoir opérationnalisé des modèles de ML? Il convient certainement d'aborder les objectifs de ces concepts en amont afin de prendre conscience des nombreux problèmes qui peuvent survenir lors du déploiement du ML. Commencer à déployer et à s'appuyer sur des modèles ML qui ne sont pas soutenus par des précautions adéquates dans le pire des cas peut entraîner une déception, voire un désastre.



03 IMPACT POSITIF INCONTESTABLE

LARGE GAMME D'AMÉLIORATIONS OBTENUES



Lorsqu'on leur a demandé quelles mesures avaient le plus augmenté succès grâce à l'utilisation de ML, DataOps et MLOps Les utilisateurs ont principalement cité l'intégration stratégique de ces mêmes concepts (59 %). 97 % d'entre eux ont déclaré avoir réalisé des améliorations significatives grâce à la mise en œuvre de Data/MLOps. Les avantages en termes d'automatisation, de robustesse de l'application et de collaboration arrivent en tête de liste des réalisations grâce à Data/

MLOps.

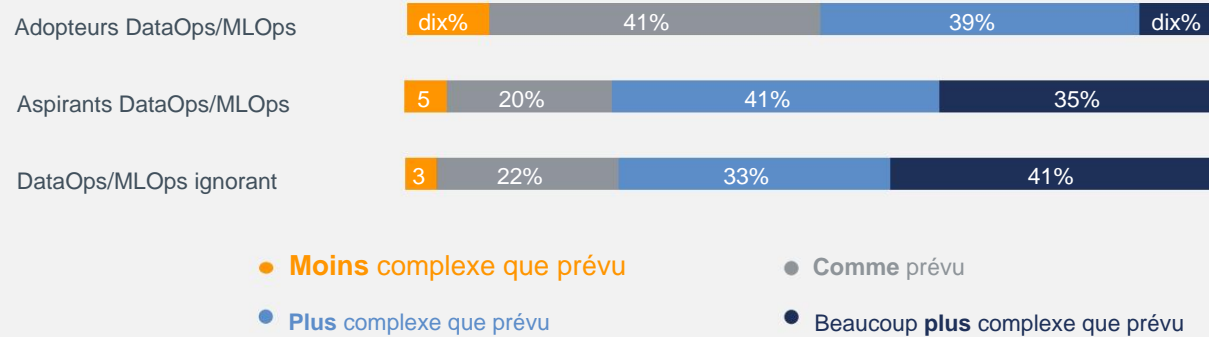
Les chiffres du graphique indiquent que s'engager dans DataOps et MLOps est une bonne idée.

Quelles améliorations significatives votre entreprise a-t-elle réalisées avec l'introduction de MLOps/DataOps? (n=61)



03 IMPACT POSITIF INCONTESTABLE

LA FAMILIARITÉ AVEC DATAOPS ET MLOPS CONDUIT À DES ATTENTES PLUS RÉALISTES



Comparaison de l'évaluation de la complexité, par statut de mise en œuvre DataOps/MLOps (n=218)

Lorsqu'on leur a demandé si leurs attentes initiales concernant l'impact positif du ML avaient été satisfaites, 53% des utilisateurs de Data/MLOps, 45% des aspirants et seulement 31% de ceux qui n'étaient pas au courant de Data/MLOps a accepté. En effet, les « inconscients » ont affiché des taux plus élevés à la fois de dépassement des attentes et de déception. Ces résultats suggèrent que les entreprises familiarisées avec les sujets liés au DataOps et au MLOps ont des attentes plus réalistes quant à ce qu'elles peuvent réaliser avec l'apprentissage automatique. Comme nous le verrons, ce constat sera confirmé lors de l'évaluation de la complexité attendue.

Les adopteurs de ces concepts sont moins susceptibles d'être surpris par des complexités imprévues : 41 % ont décrit le niveau de complexité « comme prévu ».

Il existe une plus grande différence de complexité perçue entre les entreprises qui ont déjà adopté DataOps/MLOps et celles qui envisagent de le faire. Étant donné qu'environ les trois quarts de ce dernier groupe (76%) déclarent avoir sous-estimé la complexité de l'utilisation du ML, il est raisonnable de supposer qu'ils ne font que

a commencé à explorer DataOps et MLOps après avoir été submergé par la complexité de l'IA appliquée.

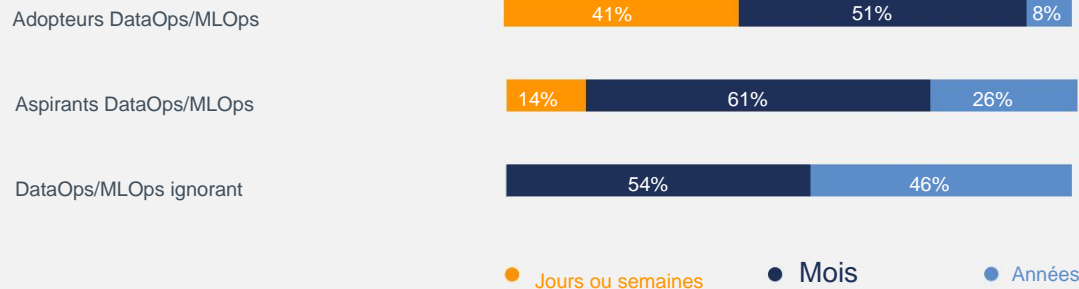
Les approches pragmatiques du ML ne posent généralement pas de problème dans les premières étapes des tests et du déploiement, mais la prudence est recommandée dès que des décisions et des fonctions importantes dépendent du bon fonctionnement des modèles de ML. La négligence ici peut conduire à une méfiance générale et au rejet des résultats du ML.

Plus de **70%** des DataOps/MLOps les aspirants et les ignorants ont sous-estimé la complexité de l'adoption de l'apprentissage automatique.



03 IMPACT POSITIF INCONTESTABLE

DATAOPS ET MLOPS POUR DES DÉPLOIEMENTS PLUS RAPIDES



Délai entre le développement et le déploiement par les praticiens ML, par statut d'adoption DataOps/MLOps (n=102)

Nos résultats prouvent que les adopteurs de DataOps et de MLOps obtiennent des délais de mise sur le marché plus rapides, une productivité plus élevée, une meilleure évolutivité et des niveaux d'automatisation plus élevés. Ce sont toutes des mesures d'efficacité et de rapidité améliorées dans la livraison des résultats de ML.

L'examen du délai entre le développement et le déploiement semble confirmer ces résultats (voir graphique).

Les utilisateurs de Data/MLOps sont nettement moins susceptibles de consacrer des « années » à ce processus (8 % contre 26 %/46 %). Alors que les déploiements prenant des "mois" semblent être la norme et ne

ne montrent pas beaucoup de variation en termes d'adoption de données/MLOps, des déploiements plus rapides d'une durée de seulement « ysemainesy » ou « yjoursy » sont courants parmi les adoptants, mais sont inconnus parmi ceux qui ne le savent pas.

Concernant les débutants, les entreprises qui n'ont pas encore mis en production des modèles ML mais envisagent au moins de mettre en place DataOps/

Les MLOps sont généralement moins susceptibles de s'attendre à des délais de mise en œuvre très longs pouvant durer des années.

DataOps et MLOps facilitent non seulement des délais de mise en œuvre plus courts, mais forgent également le

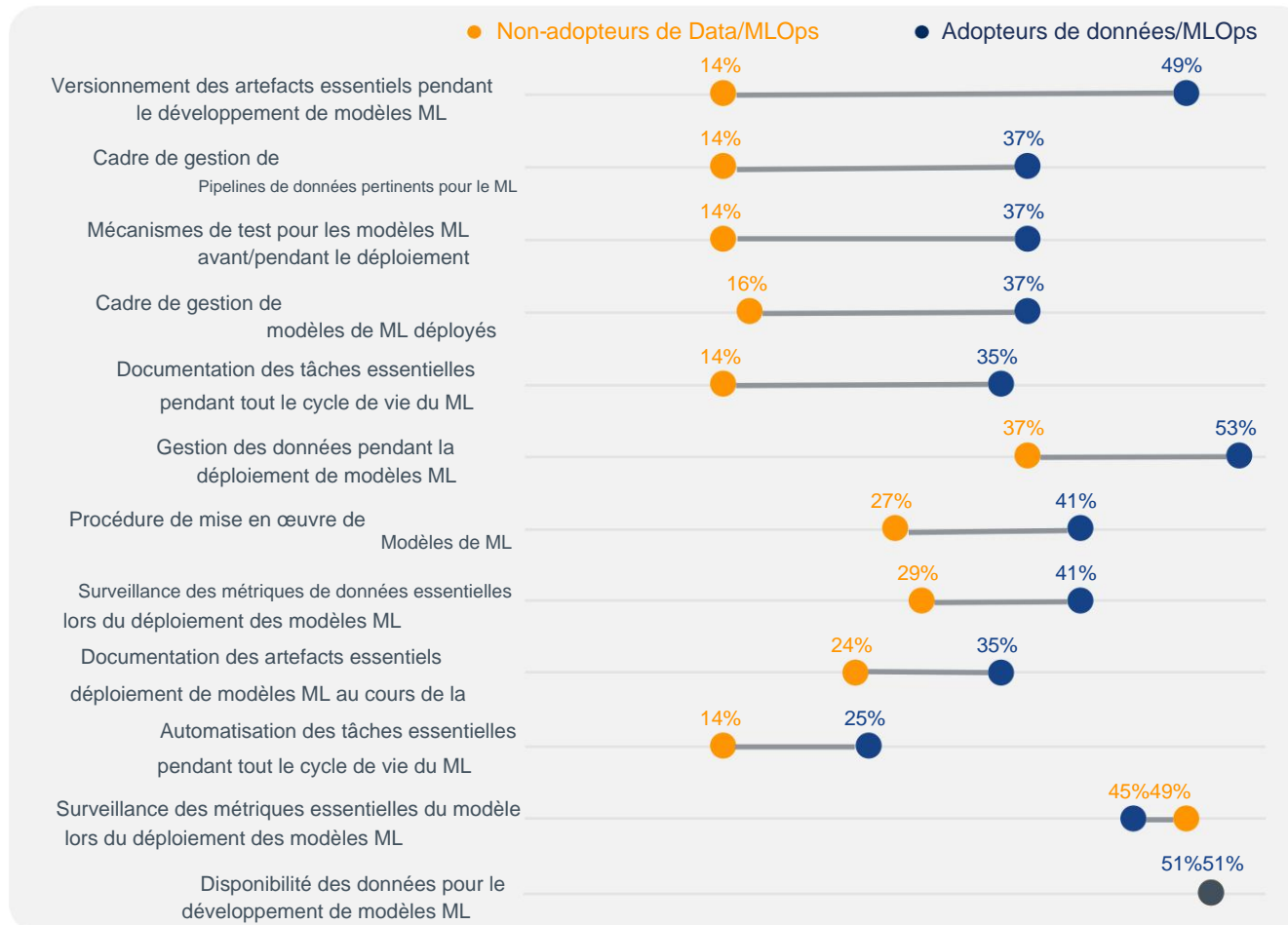
confiance pour être en mesure de livrer plus rapidement lorsque le déploiement est à l'ordre du jour. Ce dernier effet semble être réalisable en se familiarisant simplement avec les concepts et leurs implications.

Aucun des répondants qui ne connaissaient pas DataOps/ MLOps n'a passé moins de « ymois » sur le processus, du développement d'un modèle d'apprentissage automatique à son déploiement.



03 IMPACT POSITIF INCONTESTABLE

DATAOPS ET MLOPS COMME CONCEPTS POUR FAIRE FACE AUX DÉFIS DU ML



Défis pour lesquels les entreprises ont déjà une bonne solution, par les praticiens du ML, par le statut d'adoption de DataOps/MLOps (n=102)

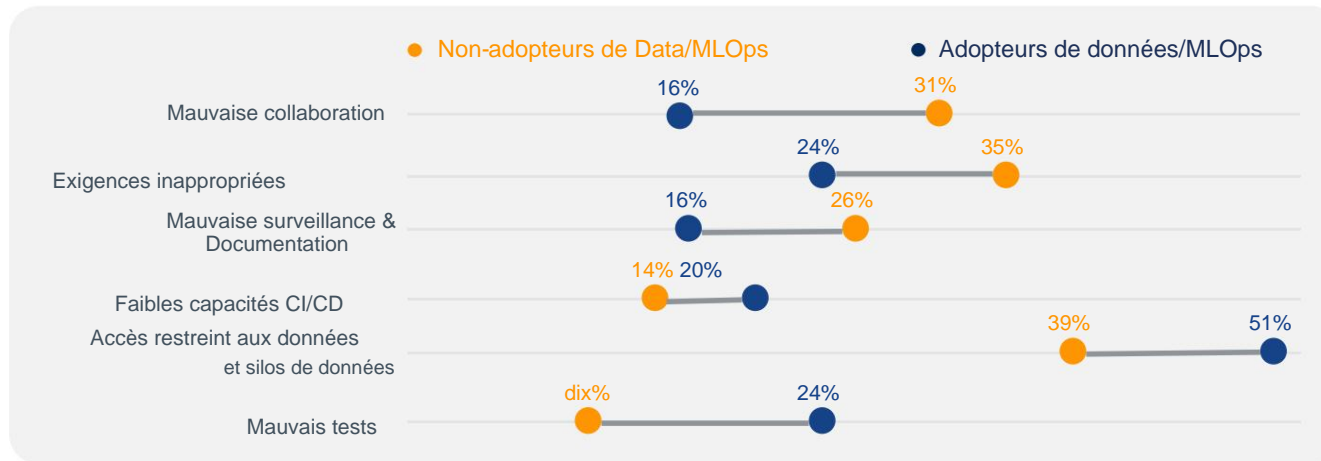
Les adopteurs de Data/MLOps sont mieux lotis à presque tous les égards en ce qui concerne les indicateurs d'une bonne solution pour relever les défis typiques du ML (voir graphique). En moyenne, les adoptants ont sélectionné plus de deux fois plus d'options de réponse que les non-adoptants (4,5 contre 2,2).

Les adoptants réussissent particulièrement bien dans les domaines de la gestion des versions et de la documentation, en fournissant des cadres de gestion et des tests. Dans le cas du suivi des métriques essentielles du modèle, une proportion plus faible d'adoptants par rapport aux non-adoptants affirment disposer d'une bonne solution à cette fin. Ce résultat peut probablement être attribué aux adopteurs ayant des exigences plus élevées en termes de qualité de modèle.



03 IMPACT POSITIF INCONTESTABLE

DATAOPS ET MLOPS COMME CONCEPTS POUR FAIRE FACE AUX PROBLÈMES DE ML



Problèmes rencontrés par les praticiens du ML, par statut d'adoption DataOps/MLOps (n=102)

À leur tour, nous nous attendrions également à ce que les adopteurs de Data/MLOps rencontrent moins de problèmes lors du déploiement de modèles ML. En effet, c'est le cas, mais l'avantage d'adopter Data/MLOps n'est pas aussi important que nous l'imaginons. En sélectionnant en moyenne 3,4 des 16 options présentées dans notre enquête, les adopteurs de Data/MLOps ne diffèrent pas beaucoup de ceux qui déploient le ML sans mettre en œuvre ces concepts (3,8 sur 16). Outre cette différence quantitative marginale, des variations significatives concernant le comportement de réponse peuvent être observées (voir graphique).

L'effet le plus significatif peut être identifié avec la collaboration. Les adopteurs de MLOps semblent généralement mieux placés pour surmonter les obstacles qui entravent une coopération efficace entre les parties prenantes. La nécessité de surmonter ce problème est apparemment l'un des principaux moteurs de l'adoption de Data/MLOps. Ceci est souligné par une comparaison des rôles impliqués dans le développement et le déploiement du ML. Le nombre moyen de rôles différents impliqués pour les adoptants est de 4,2, ce qui est nettement supérieur aux 3,0 pour les praticiens ML n'utilisant pas Data/MLOps.

Outre les améliorations procédurales dans les domaines de la surveillance et de la documentation, la réduction des efforts manuels et la garantie de la fiabilité et de la qualité, les utilisateurs de Data/MLOps sont beaucoup moins susceptibles d'avoir des problèmes avec des exigences irréalistes des départements et de la direction. Vraisemblablement, cette observation est étroitement liée à la meilleure gestion des attentes des adopteurs Data/MLOps décrite ci-dessus. Lorsque des attentes réalistes sur ce qui est possible et réalisable prévalent, il est plus facile de faire face à des exigences excessives.

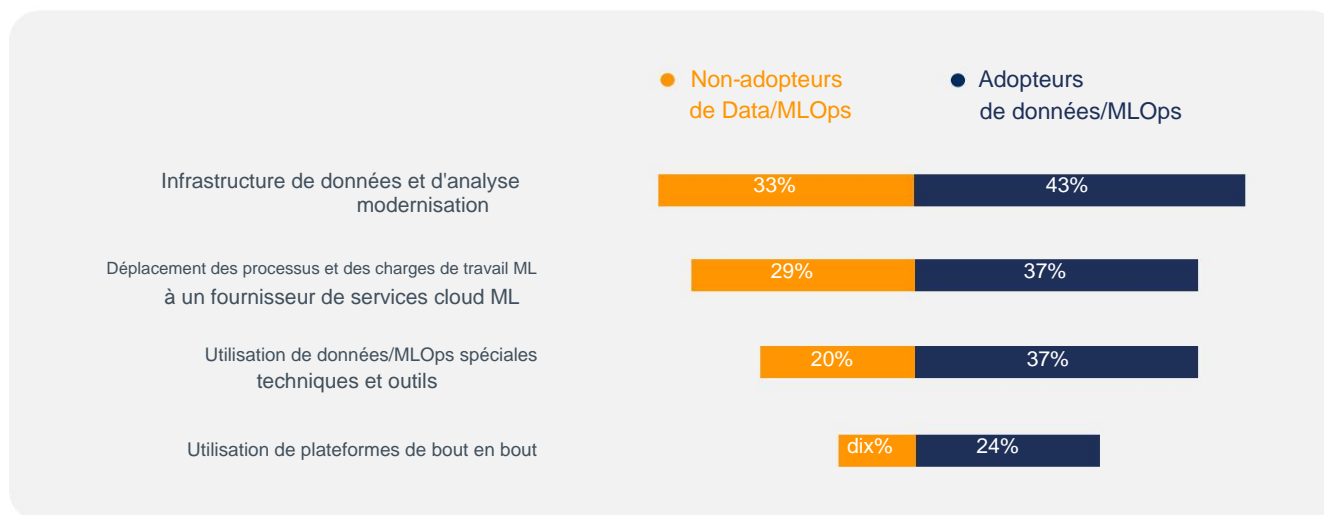
La fréquence plus élevée de problèmes liés à de faibles capacités de CI/CD et de test indique très probablement une plus grande prise de conscience de ces défis d'une part, et aussi qu'ils semblent être des défis en aval qui pourraient bien être plus difficiles à résoudre.

Cependant, la gestion des accès restreints aux données et des problèmes de silos de données semble sortir du cadre de la plupart des initiatives Data/MLOps, car ce problème est encore plus répandu parmi les utilisateurs.



04 QUESTIONS DE SÉLECTION DE LOGICIEL

MALGRÉ LES AVANTAGES DES OUTILS COMMERCIAUX, L'OPEN SOURCE DOMINE TOUJOURS



Mesures réussies liées aux outils et à l'infrastructure prises par les praticiens du ML, par statut d'adoption de DataOps/MLOps (n=102)

L'adoption de bonnes procédures DataOps et MLOps est généralement indépendante de la nature des outils utilisés. Néanmoins, les outils peuvent avoir un impact important. Le graphique montre que les adopteurs de DataOps et de MLOps ont déclaré beaucoup plus souvent qu'ils ont obtenu des succès grâce à des mesures liées aux outils et à l'infrastructure. UN comparaison générale de l'impact de l'utilisation d'outils open source, auto-développés et commerciaux,

notamment les plateformes de solutions, éclaire la pertinence du choix des logiciels.

En termes de pile d'outils pour réaliser des applications d'IA, les utilisateurs de Data/MLOps ne diffèrent pas significativement des autres. Globalement, la moitié des entreprises interrogées ont misé sur des outils open source, près d'un tiers (31%) utilisent des outils commerciaux et 19% ont construit leurs propres outils.

La prédominance de l'open source a de multiples raisons:

- La communauté open source est un moteur majeur de l'innovation
- La communauté open source est vivante et active • Les solutions open source sont gratuites et disponibles instantanément
- L'open source est principalement utilisé pour le transfert de connaissances dans les établissements d'enseignement

D'autre part, le marché des outils commerciaux dans le domaine de l'IA est loin d'être consolidé et aucun outil standard n'a encore été établi.

Le marché est très concurrentiel et les départements marketing proposent constamment de nouveaux mots à la mode pour décrire leurs outils, alors choisir des outils commerciaux adaptés est tout sauf facile.

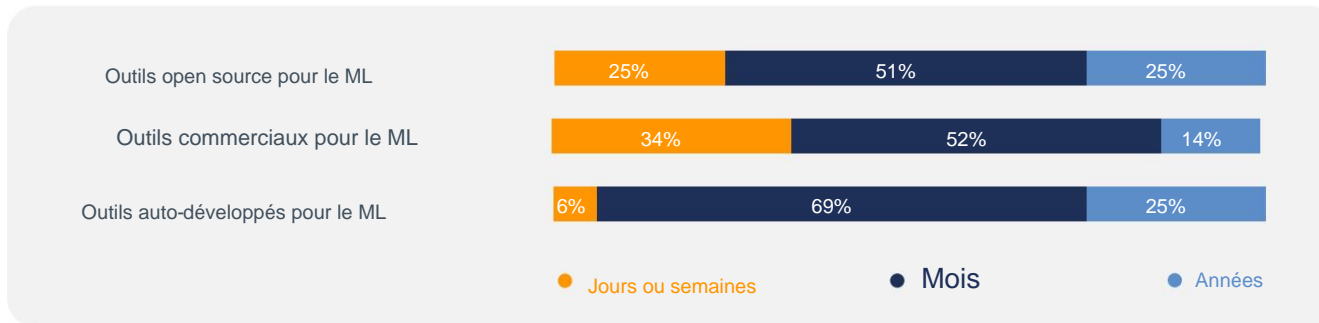
Néanmoins, cela peut en valoir la peine.

De nombreux fournisseurs sont de plus en plus capables de rendre les fonctionnalités des solutions open source plus conviviales en réduisant la complexité, tout en intégrant les nombreux avantages de l'open source grâce à la conception d'outils ouverts, adaptables et compatibles.



04 QUESTIONS DE SÉLECTION DE LOGICIEL

IMPACT POSITIF DES OUTILS COMMERCIAUX OÙ LE ML EST OPÉRATIONNALISÉ



Durées de déploiement par les praticiens du ML, par type d'outils de ML (n=102)

Une comparaison de la durée des déploiements des praticiens du ML (voir graphique) révèle un avantage comparatif dans l'utilisation d'outils commerciaux. Leurs utilisateurs sont plus susceptibles de pouvoir se déployer en de courtes périodes de jours ou de semaines, et ils sont moins susceptibles de mettre des années à se déployer. Les entreprises qui travaillent avec des développements internes s'en tirent le moins bien dans cette comparaison.

Cependant, les débutants ne s'attendent pas à des temps de mise en œuvre plus courts lorsqu'ils travaillent avec des outils commerciaux. À cet égard, les débutants qui utilisent des outils open source ou développés par eux-mêmes sont plus confiants. Une raison possible à cela pourrait être qu'ils ne prennent pas en compte les considérations opérationnelles

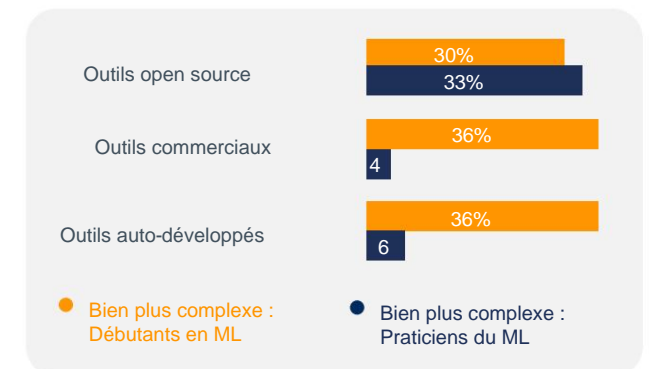
assez sérieusement lors de l'achat d'outils ML.

Les outils commerciaux pour le développement de modèles ML doivent donc toujours être vérifiés pour leur adéquation aux exigences d'opérationnalisation des modèles (par exemple, fournissent-ils des interfaces et des fonctions adéquates?).

Une autre distinction concernant l'utilisation des outils entre les débutants et les praticiens est évidente en ce qui concerne la complexité perçue. Bien que le type d'outil ne fasse pas beaucoup de différence pour les débutants, les praticiens utilisant des outils commerciaux et auto-développés sont beaucoup moins susceptibles d'être pris au dépourvu par une complexité excessive lorsqu'ils répondent à leurs exigences en matière d'apprentissage automatique.

Les chiffres presque égaux pour les utilisateurs d'outils open source indiquent que les praticiens du ML peuvent mieux réduire la complexité avec des outils commerciaux qu'avec des outils open source.

Les utilisateurs d'outils auto-développés sont également plus performants que les utilisateurs open source. Cependant, l'utilisation d'outils auto-développés ne semble pas être une pratique courante, avec seulement 16 % des praticiens du ML déclarant les utiliser. Nous observons également souvent que les utilisateurs d'outils ML auto-développés ont tendance à passer à des outils commerciaux à mesure que leur niveau de maturité ML augmente.

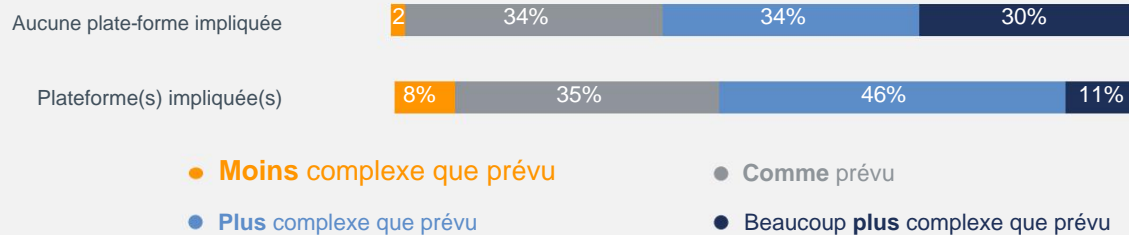


Piles d'outils comparées par complexité réelle par rapport à la complexité attendue, par niveau d'avancement (n = 100)



04 QUESTIONS DE SÉLECTION DE LOGICIEL

SOLUTIONS DE PLATEFORME POUR UNE COMPLEXITÉ RÉDUITE



L'impact de l'utilisation de la plateforme sur la complexité perçue par les praticiens du ML (n=87)

Une autre option potentiellement intéressante pour réduire la complexité consiste à utiliser des plates-formes qui combinent les fonctionnalités de plusieurs solutions ou même couvrent l'ensemble du processus, du développement au déploiement (de bout en bout). La plupart des entreprises combinent plusieurs solutions pour réaliser le développement et

déploiement (53 %). La deuxième option la plus populaire consiste à exécuter des plates-formes distinctes pour le développement et le déploiement (23%). L'utilisation d'une plate-forme pour l'un ou l'autre ne représente que 17% de nos données, tandis que les plates-formes de bout en bout ne sont utilisées que par 6% des répondants à cette enquête.

Pour presque **1 entreprise sur 3** sans plate-forme, les projets ML sont bien plus complexes que prévu.

Ceci n'est vrai que pour **1 sur 10** entreprises avec une plate-forme.

Il n'y a pas de différences majeures entre les praticiens du ML et les débutants concernant l'utilisation générale des plates-formes. Étonnamment, 43% des entreprises qui n'ont pas encore déployé de modèle ML (débutants) déclarent que leurs exigences de déploiement ML sont déjà couvertes par une forme de plate-forme.

Les praticiens du ML dans les entreprises équipées d'une sorte de plate-forme sont beaucoup moins susceptibles de rencontrer "beaucoup plus" de complexité que prévu. Vraisemblablement, l'utilisation de plates-formes peut aider à réduire la complexité.

Bien sûr, la mise en place de DataOps et de MLOps va bien au-delà de la sélection de logiciels, mais l'impact d'outils bien choisis n'est pas négligeable. Ils peuvent apporter une grande contribution en fournissant des cadres pour la définition des processus ainsi qu'en assistant techniquement et en facilitant l'accomplissement de tâches cruciales.

Généralement, il est logique pour les débutants de trouver leur chemin et d'explorer de nouvelles fonctions avec des versions open source ou gratuites d'outils commerciaux avant d'acheter des licences coûteuses. Encore une fois, de notre point de vue, étudier les principes de DataOps et MLOps et apprendre à les appliquer dans la pratique peut faciliter l'identification des exigences importantes et aider à éviter les erreurs et les revers courants sur la voie de déploiement ML réussi.

05 LES ENTREPRISES PRENNENT DE L'AVANCE AVEC DATAOPS ET MLOPS, MAIS DEVRAIT FAIRE ATTENTION À EMMENER TOUT LE MONDE À BORD



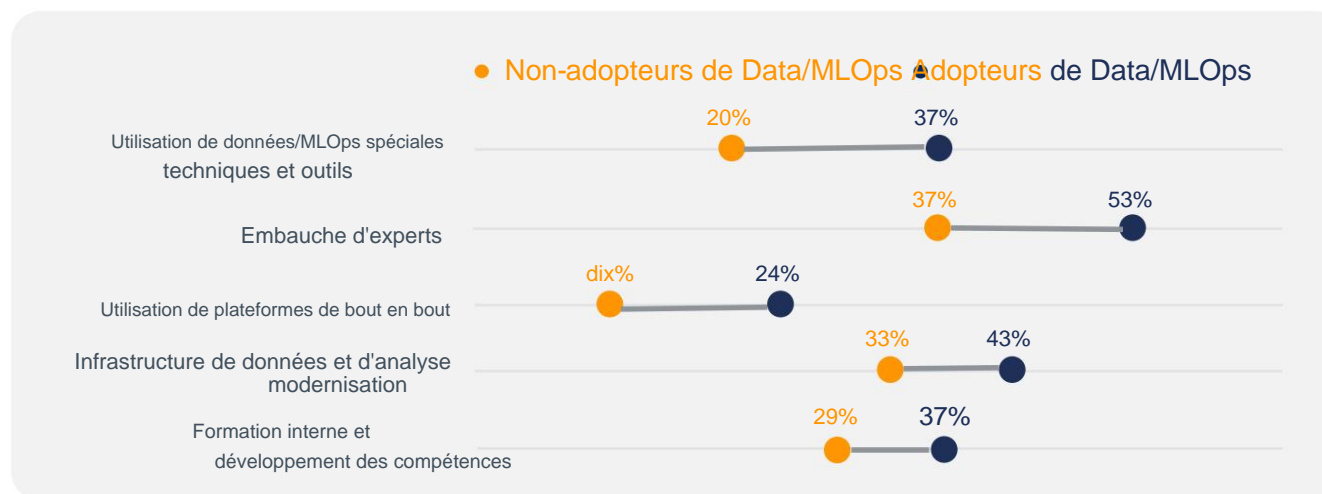
MESURES ET OBSTACLES SUR LA VOIE VERS L'UTILISATION RÉUSSIE DU ML

Près de 2 adoptants sur 3 envisagent l'intégration stratégique de DataOps/MLOps pour être très propice au succès avec ML.

Nous avons déjà souligné que DataOps et MLOps sont des concepts précieux pour des initiatives de ML réussies et efficaces. En effet, les données d'enquête sur le succès de 12 mesures individuelles corroborent cette constatation. 59% des utilisateurs de Data/MLOps considèrent que l'intégration de ces concepts est très propice au succès avec le ML.

Cependant, il existe également d'autres mesures prometteuses. En moyenne, les adoptants étaient 52% plus susceptibles de déclarer qu'une mesure donnée avait contribué à accroître le succès. En regardant les plus grandes différences entre les adoptants et les non-adoptants de Data/MLOps dans le graphique, il est raisonnable de supposer que la mise en œuvre de Data/MLOps a également un impact positif sur le succès d'autres

les mesures.



Comparaison des mesures efficaces prises par les praticiens du ML, par statut d'adoption DataOps/MLOps (n=102)

Les adopteurs de DataOps et de MLOps semblent également mieux à même d'attirer des experts compétents sur le marché du travail qui peuvent générer de la valeur ajoutée. Ce n'est pas une surprise, car ils peuvent généralement mieux assurer une atmosphère de travail professionnelle car ils ont déjà fait des efforts pour faire fonctionner la collaboration et établir des cadres de gestion pour le ML. Néanmoins, le manque de personnel, en particulier dans le domaine de l'analyse avancée, reste également un défi pour les adopteurs de Data/MLOps.

Cependant, on peut affirmer qu'ils ont une plus grande capacité à profiter de l'utilisation des nouvelles technologies et des infrastructures modernisées, car ils sont probablement plus capables de sélectionner les outils adaptés et de profiter de leur éventail de fonctions. Cette hypothèse est étayée par notre constatation selon laquelle les adoptants sont beaucoup moins susceptibles de voir un manque d'orientation comme un défi (voir le graphique à la page suivante).

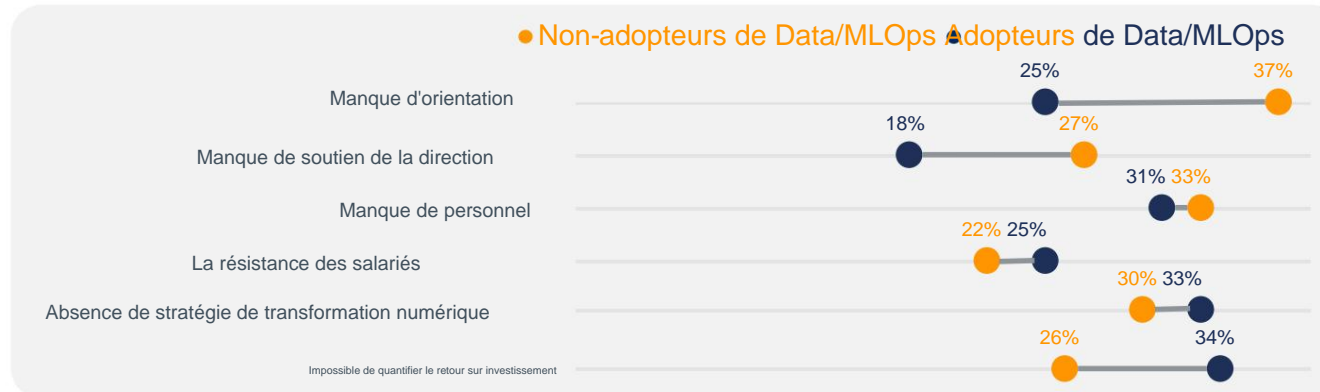
05 LES ENTREPRISES PRENNENT DE L'AVANCE AVEC DATAOPS ET MLOPS, MAIS DEVRAIT FAIRE ATTENTION À EMMENER TOUT LE MONDE À BORD



Un autre différenciateur important pour les adopteurs de données/MLOps est qu'ils sont moins susceptibles de percevoir un manque de soutien de la direction comme un défi.

En règle générale, le soutien de la direction est très important dans les phases initiales de l'introduction du ML. La conviction, ou du moins un acte de foi significatif de la part des décideurs, est essentielle à toute initiative visant à faire fonctionner le ML dans les organisations. On s'attendrait à ce que la question du soutien de la direction se dissipe progressivement à mesure que ces initiatives commencent à avoir un impact positif. Nous supposons que la plupart des adoptants qui ont répondu à notre enquête ont déjà surmonté cet obstacle avec succès.

Cependant, parmi les nombreux avantages que les utilisateurs de Data/MLOps ont obtenus, des aspects problématiques émergent qui ont une corrélation frappante. Les adopteurs se concentrent moins sur les mesures visant à accroître l'approbation du ML dans le contexte plus large de l'entreprise et déploient moins d'efforts pour renforcer la collaboration entre les rôles, comme le montre le graphique. Il semble que les adoptants



Facteurs qui entravent les avantages du ML, par statut d'adoption des données/MLOps (n=228)

des données/MLOps ont tendance à être plus concernés par les questions techniques et spécialisées que par la communication dans le contexte plus large de l'entreprise.

Cela se reflète également dans plusieurs facteurs préventifs qui ont été cités par les adoptants plus souvent que quiconque. À cet égard, ne pas être en mesure de communiquer le retour sur investissement des projets aux parties prenantes peut être interprété comme un manque de compétences majeur. De même, une incapacité des parties prenantes à bien comprendre les

le but des projets ML peut être problématique. Dans les deux cas, cela augmente le risque que des projets importants échouent ou ne soient pas autorisés en premier lieu en raison d'un manque d'acceptation et de consentement.

Les entreprises où la majorité des employés n'ont pas une compréhension de base des opportunités et des possibilités de l'analyse avancée ont également du mal à établir des stratégies durables de transformation numérique car ce manque d'informations

se manifeste généralement par une résistance aux nouvelles approches. L'adoption de DataOps/MLOps doit donc toujours être intégrée dans une initiative à l'échelle de l'entreprise pour accroître généralement la littératie des données et établir une culture des données dans toute l'organisation.

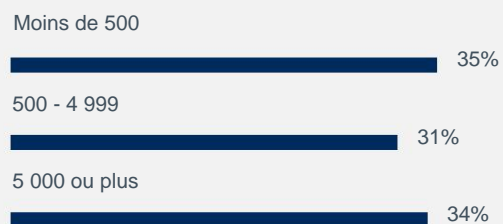


Comparaison des mesures efficaces par les praticiens du ML, par statut d'adoption DataOps/MLOps (n=102)

DÉMOGRAPHIE

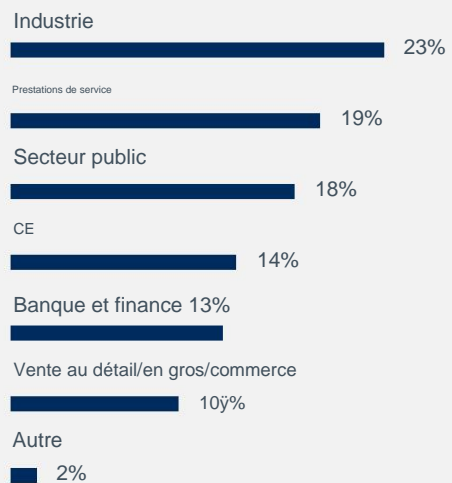


TAILLE DE L'ENTREPRISE



Combien d'employés votre entreprise compte-t-elle?
(n=248)

SECTEUR DE L'INDUSTRIE



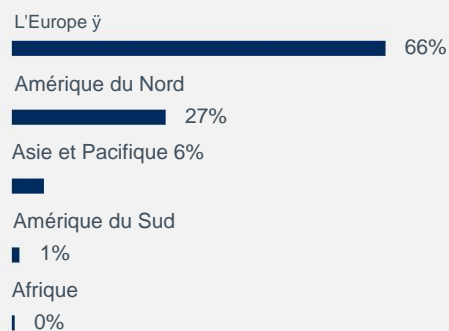
Lequel des énoncés suivants décrit le mieux le
secteur d'activité de votre organisation? (n=248)

APERÇU

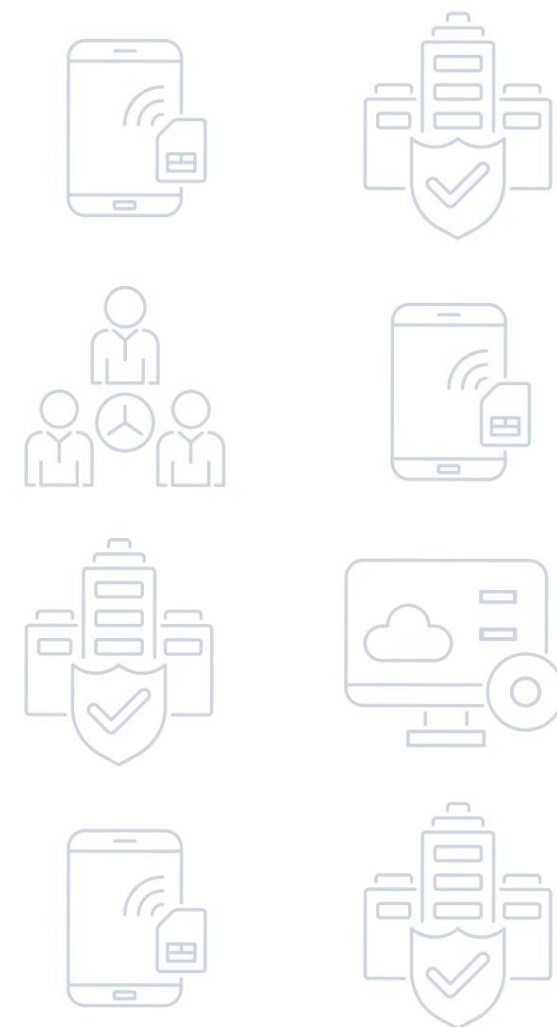
Cette étude est basée sur une enquête mondiale en ligne menée de février à mars 2022. Nous avons fait la promotion de l'enquête auprès du panel BARC et via nos divers canaux de communication. La

la majorité des participants viennent d'Europe. Ils représentent des entreprises de différentes tailles et de diverses industries, ainsi que diverses approches et niveaux de progrès dans l'application du ML.

RÉGION



Dans quelle région êtes-vous situé ? (n=163)



DÉMOGRAPHIE

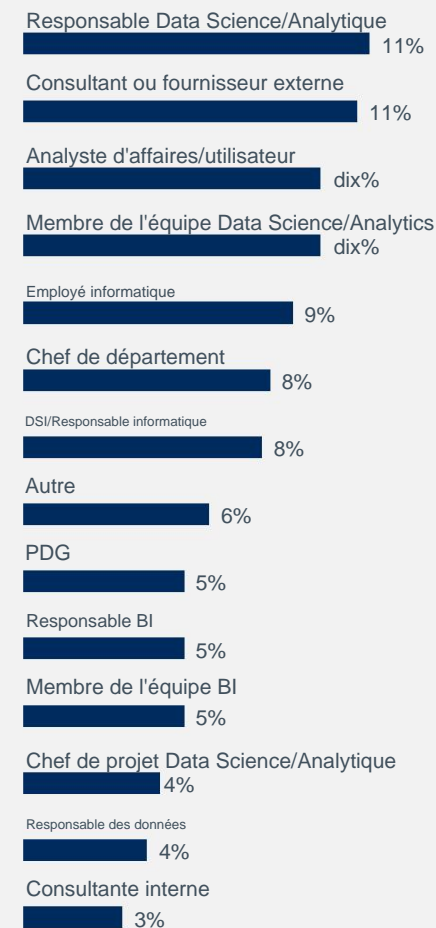


QUELLES DONNÉES AVONS-NOUS UTILISÉES?

Après l'assurance qualité et le nettoyage des données, notre échantillon contient des données provenant de 248 questionnaires remplis. Les tailles d'entreprises allant de « moins de 500 » à « 5000 employés ou plus » sont représentées de manière égale et proviennent de divers secteurs d'activité secteurs.

Les répondants occupent divers postes au sein de leur entreprise, comme le montre le graphique. Les participants ayant une formation en analyse, comme la science des données ou l'intelligence d'affaires, représentent la plus grande part. Les consultants et fournisseurs externes représentent le deuxième segment le plus important de l'enquête avec 11%. Chacun des répondants de ce groupe a rempli le questionnaire au nom d'un de ses clients.

POSTE DANS L'ENTREPRISE



Quel est votre rôle dans l'entreprise ? (n=248)

BARC – FAIRE DES LEADERS NUMÉRIQUES

BARC – APPLICATION D'ENTREPRISE CENTRE DE RECHERCHE

BARC (Business Application Research Center) est l'un des principaux cabinets d'analyse en Europe pour les logiciels d'entreprise, se concentrant sur les domaines des données, de l'informatique décisionnelle (BI) et de l'analyse, de la gestion de contenu d'entreprise (ECM), de la gestion de la relation client (CRM) et de la planification des ressources d'entreprise. (ERP).

Notre passion est d'aider les organisations à devenir les entreprises numériques de demain. Nous y parvenons en utilisant la technologie pour repenser le monde, en faisant confiance aux décisions basées sur les données et en optimisant et en numérisant les processus. Il s'agit de trouver les bons outils et de les utiliser de manière à donner à votre entreprise le meilleur avantage possible.

Ce mélange unique de connaissances, d'échange d'informations et d'indépendance distingue nos services dans les domaines de la recherche, de l'événementiel et du conseil.

Rechercher

Nos études BARC sont basées sur le marché intérieur recherche, tests de logiciels et commentaires d'analystes, vous donnant la sécurité de prendre les bonnes décisions.

Notre recherche indépendante met l'accent sur les développements du marché, met les logiciels et les fournisseurs à l'épreuve et donne aux utilisateurs un espace pour exprimer leurs opinions.

Événements

Les décideurs et les leaders de l'industrie informatique se réunissent lors des événements BARC. Les séminaires BARC en petits groupes, les webinaires en ligne et les conférences avec plus de 1 000 participants par an offrent inspiration et interactivité. Grâce à des échanges avec des pairs et à un aperçu des tendances actuelles et des développements du marché, vous recevrez un nouvel élan pour faire avancer votre entreprise.

Consultant

Lors d'ateliers d'experts confidentiels, de coaching et de consultations internes, nous transformons les besoins de votre entreprise en décisions pérennes. Nous vous fournissons des concepts holistiques réussis qui vous permettent d'utiliser correctement les bonnes informations.

Notre accompagnement de projet couvre toutes les étapes de l'utilisation réussie d'un logiciel.



Allemagne

BARC GmbH
Place de Berlin 7
D-97080 Wurtzbourg
+49 931 880 6510

www.barc.de

L'Autriche

BARC GmbH
Hirschstettner Straße 19 / I / IS314
A-1200 Vienne
+43 660 6366870

Suisse

BARC Suisse SA
Täferstraße 22a
CH-5405 Baden-Dättwil
+41 56 470 94 34

Reste du monde

+44 1536 772 451

www.barc-research.com

PROFIL DU COMMANDITAIRE : ROBOT DE DONNÉES

À PROPOS DU ROBOT INFORMATIQUE

DataRobot AI Cloud est la prochaine génération d'IA.

La vision AI Cloud de DataRobot est de rassembler tous les types de données, tous les utilisateurs et tous les environnements pour fournir des informations commerciales essentielles à chaque organisation.

DataRobot bénéficie de la confiance de clients mondiaux de tous les secteurs, dont un tiers du Fortune 50, fournissant plus d'un billion de prédictions pour les principales entreprises du monde entier.

DataRobot propose également DataRobot AI Cloud pour

Industries, une extension de notre plate-forme AI Cloud avec des capacités et une expertise uniques adaptées pour répondre aux besoins de croissance, de risque et d'échelle dans les principales industries. DataRobot a livré plus d'un

millions de projets clients actifs pour des organisations telles que Mars, McLaren Racing et Lenovo, résultant en une vaste bibliothèque de meilleures pratiques pour appliquer l'IA afin d'améliorer les soins aux patients, la gestion de la chaîne d'approvisionnement, la détection des fraudes et plus encore. DataRobot AI Cloud pour

Industries étend cette expertise pour accélérer l'impact de l'IA pour toutes les organisations bancaires, de vente au détail, de fabrication et de soins de santé.

DataRobot AI Cloud est conçu pour toutes les organisations de tous les secteurs. Des organisations gouvernementales, commerciales et à but non lucratif du monde entier, et a aidé des organisations de tous les secteurs à exploiter le pouvoir de transformation de l'IA. De la restauration de la résilience de la chaîne d'approvisionnement à l'accélération du traitement et de la prévention d'innombrables maladies en passant par la lutte contre le changement climatique, tout en stimulant l'innovation pour alimenter les économies.

En savoir plus sur [datarobot.com](https://www.datarobot.com).

DataRobot

Robot de données

Contact:

Lauren Sanborn

presse@datarobot.com

+1-617-765-4500

www.datarobot.com

PROFIL SPONSOR : DOMINO DATA LAB

À PROPOS DU LABORATOIRE DE DONNÉES DOMINO

Domino propulse les entreprises axées sur les modèles avec sa plate-forme Enterprise MLOps leader qui accélère le développement et le déploiement du travail de science des données tout en augmentant la collaboration et la gouvernance. Plus de 20% des entreprises du Fortune 100 comptent sur Domino pour les aider à faire évoluer la science des données, ce qui en fait un avantage concurrentiel. Fondée en 2013, Domino est soutenue par Sequoia Capital et d'autres investisseurs de premier plan.



DOMINO
DATA LAB

LABORATOIRE DE DONNÉES DOMINO

Cote est

Gare de King's Cross
Londres
N1C 4AX

Contact : Kathryn McDonagh
Kathryn.mcdonagh@dominodatalab.com

PROFIL PARRAIN : UNE LOGIQUE

À PROPOS D'UNE LOGIQUE

Fondée en 2013, ONE LOGIC est une entreprise leader dans l'utilisation de l'intelligence artificielle (IA) avec des produits de données clés en main (applications de bout en bout). Avec une vaste expérience dans les projets de science des données, ONE LOGIC est un expert pour relever les défis spécifiques à l'industrie avec des solutions ciblées. L'objectif de l'entreprise est de permettre à ses clients de monétiser rapidement et durablement leurs données, de digitaliser leurs modèles économiques et de générer de manière autonome de la valeur ajoutée. Il le fait grâce à une combinaison unique de produits de données clés en main qui transforment les données en actifs d'entreprise à long terme et à l'assistance d'experts de spécialistes expérimentés de la science des données et de l'IA. Cela ouvre la voie aux utilisateurs à n'importe quelle étape de leur parcours professionnel pour développer des produits de données de haute qualité, productifs et évolutifs et des modèles commerciaux innovants qui créent de la valeur et le faire à un rythme beaucoup plus rapide.

Notre technologie habilitante pour cela est la cartographie ONE DATA. Cela automatise les processus et crée systématiquement les conditions idéales pour l'utilisation des données en fusionnant des sources de données hétérogènes via l'IA. ONE LOGIC propose également des produits de données clés en main

qui analysent et relient les sources de données via l'IA pour relever les plus grands défis auxquels sont confrontées les entreprises aujourd'hui.

En tant que fournisseur agile de produits de données basés sur l'IA, ONE LOGIC s'appuie sur les dernières technologies et algorithmes d'apprentissage automatique, d'apprentissage en profondeur et d'IA. Il favorise également des partenariats avec des entreprises et des universités pour poursuivre la recherche et le développement intensifs de produits de données innovants. Notre objectif? Pour aider les entreprises à se numériser rapidement et complètement. La durabilité est toujours une priorité absolue. Chez ONE LOGIC, nous nous appuyons sur notre vaste expertise en science des données pour créer des produits de données efficaces qui ajoutent une valeur tangible et ont un impact omniprésent sur les entreprises et la société.

Entreprise dirigée par son propriétaire, le conseil d'administration de ONE LOGIC est composé du fondateur de l'entreprise, le Dr Andreas Böhm, du Dr Stefan Roskos, de Christian Aumüller et du professeur Dr Andreas Peifer (ppa.). En avril 2022, l'effectif comptait près de 300 personnes et continuait de croître, avec des équipes basées dans les bureaux de ONE LOGIC à Passau, Munich, Francfort et Zurich.

Nous ciblons principalement les secteurs de la fabrication, de l'automobile, de la vente au détail, de la pharmacie et de la biotechnologie.



ONE LOGIC

ONE LOGIC Passau

Kapuzinerstrasse 2c
94032 Passau

Allemagne
+49 851 225 906 0

UNE LOGIQUE Munich

Prinzregentenstrasse 50
80538 Munich

Allemagne
+49 89 954 592 70

ONE LOGIC Francfort

Usine d'Eschenheim 1
60316 Francfort/Main

Allemagne
+49 69 299 5703 10

ONE LOGIC Zürich

Europaallee 41
8004 Zürich

Suisse
+41 44 214 6295

Contact:

contact@onelogic.de

www.onelogic.ai

AUTEURS



ALEXANDRE RODE
ANALYSTE DATA & ANALYTICS / DATA
SCIENTIFIQUE

Alexander Rode est analyste pour les données et l'analyse et scientifique des données au Business Application Research Center (BARC). Il conseille les entreprises sur l'identification des cas d'utilisation pour l'analyse de données et la sélection d'outils pour l'analyse avancée, ainsi que sur la réalisation de preuves de concept dans le domaine de l'analyse avancée et du coaching en science des données.

Alexander soutient également la préparation d'études de marché BARC et d'articles de recherche, prend la parole lors de conférences et dirige BARC et en interne séminaires.



TIMM GRAND
ANALYSTE SÉNIOR

En tant qu'analyste principal, Timm Grosser a conseillé des entreprises nationales et internationales de différentes tailles et industries dans les domaines de

BI, data management et analytics depuis plus de 10 ans. Durant son mandat de consultant, il a conçu de nombreuses solutions en stratégie BI/big data, organisation, architecture et sélection d'outils avec les clients et dans le BARC

laboratoire d'essai. Il est un conférencier fréquent lors de conférences et de séminaires ainsi que l'auteur de nombreux articles sur l'industrie et études de marché.



Allemagne

BARC GmbH
Place de Berlin 7

D-97080 Wurtzbourg
+49 931 880 6510

L'Autriche

BARC GmbH
Hirschstettner Straße 19
/ 1 / IS314

A-1220 Vienne
+43 660 6366870

Suisse

BARC Suisse SA
Täferstraße 22a

CH-5405 Baden-Dättwil
+41 56 470 94 34

Reste du monde

www.barc-research.com
+44 1536 772 451