

**La science des données est un sport d'équipe.
Possédez-vous les compétences nécessaires pour être sélectionné au bon poste ?**

IBM

1

Introduction

2

L'analyste scientifique

3

L'ingénieur spécialiste des données

4

Le développeur

5

L'analyste métier

6

Les équipes de science des données : Les nouveaux catalyseurs du changement

7

Ressources

Introduction

Dans la vaste économie digitale d'aujourd'hui, le défi majeur de la révolution du Big data consiste à extraire la logique enfouie dans l'ensemble des informations produites. La collecte de la moindre donnée accessible est déjà toute une affaire pour une entreprise, mais comment extraire de la valeur d'un tel volume d'informations ?

Plus vous possédez de données, meilleure est la qualité de vos rapports et de vos recommandations stratégiques, n'est-ce pas ? Oui... si vous pouvez analyser ces données de manière intelligente et rapide et les exploiter pour obtenir des connaissances valables. Dans le cas contraire, vous vous exposez à des problèmes proportionnels au volume de vos données, notamment la désorganisation des informations, les soucis de stockage, les risques de sécurité,



la frustration des équipes opérationnelles et les activités pléthoriques du personnel informatique.

Face à cela, comment définir ce que vous souhaitez réellement extraire de vos données, et quelles sont celles qui vont vous permettre d'obtenir des réponses ?

Comment allez-vous transformer les indicateurs et les fichiers bruts en informations ayant une valeur réelle, et qui vont vous permettre d'en créer davantage ? Possédez-vous les compétences dont votre entreprise a besoin pour obtenir des résultats ?

Les entreprises les plus performantes créent des équipes spécialisées dans la science des données en y incorporant des compétences et des responsabilités variées au lieu de ne s'appuyer que sur quelques personnes triées sur le volet.

En pratique, les produits de données sont créés par une équipe constituée de différentes personnes. La qualité de vos analyses dépendra de l'équipe responsable de la collecte, de l'élaboration et de l'analyse des données traitées.

Dans ce cadre, qu'est-ce qu'apporte une équipe d'analystes scientifiques ?

Les ingénieurs spécialistes des données et les développeurs d'interfaces apportent les ressources nécessaires pour exploiter les résultats obtenus par la science des données en s'appuyant sur la programmation détaillée. Les analystes scientifiques utilisent des algorithmes statistiques pour repérer des modèles dans

les données. Tout au long du processus, toutes les démarches doivent être en phase avec les résultats opérationnels attendus, auxquels veille l'analyste métier. Par leur action collective, l'équipe spécialisée dans la science des données peut résoudre les défis et les problématiques d'aujourd'hui pour créer des opportunités et des possibilités inédites pour l'avenir.

Quels sont les talents et les aptitudes caractéristiques des membres d'une équipe spécialisée dans la science des données et comment se complètent-elles ? **Découvrez comment en poursuivant votre lecture.**

L'analyste scientifique

Résoudre des problèmes et répondre à des questions grâce à l'analyse des données est aujourd'hui devenu un phénomène normal, dans un environnement piloté par les données. Dans les entreprises actuelles où l'expérimentation avec le réel est omniprésente, les analystes scientifiques évoluent vers un rôle consistant à raviver, à régler et à alimenter ce moteur opérationnel.

Un analyste scientifique performant ne se contentera pas de résoudre des problèmes métier, il ciblera les problèmes les plus cruciaux pour l'entreprise

Ces spécialistes sont souvent baptisés « licornes » car ils conjuguent des talents rares, aussi bien en termes de responsabilités que de compétences : mathématiques, statistiques, expertise spécifique et communication, entre autres. Pour l'essentiel, le travail d'un analyste scientifique consiste à rechercher des modèles masqués. Pour ce faire, il applique des techniques d'analyse avancée englobant (entre autres) l'apprentissage automatique, la modélisation, les statistiques et la visualisation. La plupart du temps, les analystes scientifiques des données vont construire des modèles pour prédire des résultats ou découvrir des schémas masqués ; leur stratégie consiste à produire des connaissances exploitables pour améliorer les résultats futurs.



Les analystes scientifiques participent à chaque étape de la construction des produits de données, et ils ont donc une vue globale sur la résolution des problèmes grâce aux données. Cependant, ils ne peuvent pas être des experts dans tous les domaines, et c'est là que leur équipe peut les épauler.

Ils procèdent en permanence à des expériences en intégrant des modèles prédictifs, des règles métier et des algorithmes d'orchestration inédits dans des applications programmées pour décider de l'action la plus efficace en fonction du contexte.

Un analyste scientifique compétent explore et examine les données issues de différentes sources hétérogènes. Son rôle consiste à étudier en détail l'ensemble des données entrantes avec pour objectif de mettre en relation les nouvelles informations et les données d'historique. Il sera ainsi possible de mettre en évidence une relation

ou une tendance offrant un avantage concurrentiel crucial ou une réponse à un problème métier urgent. L'analyste scientifique ne se contente pas de collecter des données et d'établir des rapports. Il les observe sous différents angles, détermine leur signification puis préconise les approches nécessaires pour appliquer les résultats obtenus. Il doit aussi veiller à ce que ses requêtes soient correctes et doit pouvoir justifier leurs conclusions en s'appuyant sur des modèles pertinents et des données fiables. En effet, un analyste scientifique est souvent amené à proposer des recommandations aux équipes de management et de direction.



Principales compétences des analystes scientifiques

Les analystes scientifiques se distinguent par leur réelle sagacité concernant le fonctionnement de l'entreprise, mais aussi par leur capacité à communiquer les résultats obtenus à la fois aux responsables métier et informatiques. Ils peuvent ainsi influencer la manière dont une entreprise aborde un défi métier. Un analyste scientifique joue fréquemment le rôle d'intermédiaire entre le département informatique et le comité de direction. Loin de se cantonner à son expertise, il doit être capable de dialoguer avec l'un et l'autre et de comprendre la hiérarchie des données. Il doit également connaître de manière approfondie l'entreprise, mais aussi faire preuve de la conviction nécessaire pour s'appuyer fermement sur les résultats obtenus en cas de désaccord.

Les analystes scientifiques sont curieux et mènent des investigations, explorent et posent des questions. Ils procèdent à des simulations (« what-if ») et remettent en cause les hypothèses et les processus établis. En termes de compétences techniques, ces spécialistes pratiquent généralement plusieurs langages de programmation. Ils connaissent bien la gestion du Big Data et les outils d'analyse comme Apache Hadoop et Spark, et utilisent des outils pour visualiser plus facilement les données et les connaissances.

L'ingénieur spécialiste des données

Un analyste scientifique peut avoir pour responsabilité de mettre en lumière les modèles enfouis dans les données. Mais comment, selon vous, vont-ils obtenir ces données, et que se passe-t-il lorsque leurs solutions doivent s'appliquer à des milliers d'utilisateurs ou traiter des informations sensibles ? Les scientifiques ne peuvent élargir leur périmètre d'intervention sans un soutien approprié pour concrétiser leur travail.

C'est ici qu'intervient l'ingénieur spécialiste des données. Pour l'essentiel, ces ingénieurs contribuent à collecter, à organiser et à nettoyer les données que les scientifiques vont au final exploiter pour construire leur analyse. Alors que les analystes scientifiques des données extraient de la valeur des données, les ingénieurs veillent à ce que ces données circulent de manière fluide entre la source et la destination pour pouvoir les traiter.

Les ingénieurs spécialistes des données ont la responsabilité de configurer les systèmes et processus que d'autres intervenants, notamment les scientifiques, utilisent pour travailler avec les données. Ils doivent comprendre comment simplifier la circulation des données pour minimiser les retards liés aux mouvements et apporter de l'agilité aux analyses. Ils collaborent également avec les développeurs d'interfaces lors de la mise en œuvre des projets d'analyse des données.

Dans la plupart des entreprises, l'ingénieur spécialiste des données a pour responsabilité l'intégration des données, et en particulier la conception, la construction et l'évaluation des performances des « pipelines » d'absorption et d'intégration des données, qui traitent des volumes importants de données temporelles issues de différentes sources. À titre d'exemple, ces données englobent les extraits de bases de données, les journaux

Les ingénieurs spécialistes des données ont souvent pour responsabilité de poser les fondations nécessaires aux analystes ou aux scientifiques spécialistes des données pour extraire facilement les données nécessaires pour leurs évaluations et leurs expériences.

de serveur d'applications, les images numérisées, les enregistrements vocaux, les flux Twitter, les sites Web et les données relatives à l'état de santé issues de capteurs. Une fois les pipelines établis (et connectés en entrée et en sortie à ces vastes « viviers » d'informations filtrées), les analystes scientifiques peuvent en extraire des ensembles de données pertinents pour effectuer leurs analyses.

Principales compétences des ingénieurs spécialistes des données

Les ingénieurs spécialistes des données sont de purs ingénieurs qui comprennent le fonctionnement interne d'un logiciel de base de données. Leur rôle est de compiler et d'installer des systèmes de bases de données, d'écrire des requêtes complexes et de les déployer sur différentes machines, mais aussi de gérer les sauvegardes et de déployer des systèmes de reprise après incident. Par ailleurs, ils développent, construisent, testent et maintiennent des architectures, notamment pour les bases de données et les systèmes de traitement des données à grande échelle.

Les ingénieurs spécialistes des données les plus performants apprennent et réfléchissent en permanence pour identifier les nouvelles technologies qui les aideront à faire évoluer l'entreprise. Ce qui les conduit à acquérir des compétences approfondies en programmation, mais aussi à se familiariser avec les technologies basées sur l'environnement Hadoop, comme MapReduce, Hive et Pig. Ces ingénieurs ont généralement une solide expérience des technologies SQL et NoSQL, mais aussi des méthodologies et des solutions d'entrepôts de données comme ETL (extract, transform, load).

Le développeur

Bien que les développeurs n'occupent pas toujours une fonction totalement dédiée au sein d'une équipe de science des données, ils constituent un élément essentiel du puzzle. Un projet efficace permet en général de banaliser les activités de science des données pour répondre aux besoins de parties prenantes internes ou de clients externes. Les entreprises gagnent en compétitivité en s'appuyant sur les données traitées en imbriquant la science des données dans les processus métier.

Les développeurs interviennent souvent lors de la phase finale du workflow d'analyse scientifique des données. Il leur appartient de construire les applications contenant les modèles, qui permettent d'exploiter les connaissances et les données collectées auprès des autres membres de l'équipe scientifique. La démarche implique une bonne dose de temps de programmation, et cette activité peut entraîner des difficultés. En effet, faute de technologie intégrée, la démarche des développeurs peut se trouver ralentie, rendant ainsi difficile l'incorporation des analyses de l'équipe scientifique.



Les développeurs apportent les moyens d'exploiter les résultats obtenus par la science des données, permettant ainsi aux parties prenantes internes et externes à l'entreprise de tirer parti des données et des analyses dans leurs activités quotidiennes.

Autre exigence, les développeurs doivent disposer de compétences de programmation complètes. Il s'agit notamment de compétences approfondies dans la construction de services web, le développement de logiciels en lien direct avec les utilisateurs et les fonctionnalités des interfaces utilisateur. En outre, les développeurs doivent bien connaître les interfaces de programmation d'application (API) et savoir les utiliser pour intégrer différents produits et sources de données dans les applications. Bien appliquées, les compétences d'un développeur permettent de créer une application si robuste que vous ne vous rendez même pas compte du produit de données autour duquel elle est construite.

Pour créer des applications mobiles, web et d'entreprise capables de changer la donne et de révolutionner le marché, les développeurs ont besoin des données et des outils nécessaires pour donner une réalité à leurs ambitions. Un développeur doit collaborer avec le scientifique, l'ingénieur et l'analyste métier pour veiller à mettre en phase les objectifs métier et les outils d'analyse centralisés.

L'analyste métier

L'analyste métier, parfois baptisé « analyste indépendant », apporte son expertise et ses conseils métier aux équipes de science des données. Sa responsabilité consiste à appliquer ses connaissances spécifiques et de prendre des décisions étayées par des données.

L'analyste métier comprend les besoins de l'entreprise, mais ne dispose pas des compétences techniques pour développer une analyse détaillée. Les améliorations technologiques et leur compréhension plus détaillée des profils d'utilisateurs les ont

transformés en « consommateurs d'analyse », ce qui leur permet de prendre des décisions en se basant sur les connaissances extraites des données. Les technologies flexibles et conviviales leur permettent de développer certaines analyses métier sans code de programmation ni intervention des analystes scientifiques. Ce qui donne aux analystes métier davantage d'indépendance et élargit leurs possibilités d'utilisation des résultats obtenus par l'équipe scientifique dans leurs analyses courantes.



Les équipes de science des données : Les nouveaux catalyseurs du changement



Dans l'entreprise d'aujourd'hui, les scientifiques, les ingénieurs, les développeurs et les analystes métier spécialistes des données sont de véritables acteurs de la transformation. Ils sont au cœur de la vitalité de l'économie du Big Data. Loin de se contenter de concevoir de nouveaux systèmes, ils s'efforcent d'identifier de nouvelles sources de données et des méthodes inédites pour les exploiter. Il appartient bien sûr encore au service informatique de construire le système, mais les professionnels de la science des données sont aux avant-postes pour apporter aux départements de l'entreprise les moyens de collaborer pour résoudre des problèmes et accélérer l'innovation.

Les produits de données les plus performants passent inaperçus pour les utilisateurs. Ce qui suppose de réfléchir soigneusement aux interfaces, aux données capturées et à l'automatisation que permet le système d'analyse centralisé. La technologie nécessaire pour collecter et analyser des volumes importants de données métier est aujourd'hui disponible, et vous pouvez l'exploiter pour faire fructifier votre entreprise Êtes-vous prêt à faire partie d'une équipe destinée à faire face aux enjeux majeurs qu'offre la science des données ?

Former les professionnels de la science de données pour l'avenir

Pour former une équipe solide, prête à agir et opérationnelle à tout moment, vous devez améliorer vos performances en construisant vos connaissances et en développant vos compétences. IBM® fait partie des nombreuses entreprises qui apportent aux professionnels de la science des données les moyens de remplir leur mission plus rapidement et efficacement.

Les secteurs porteurs d'une dynamique de croissance nécessitent des professionnels possédant des compétences et des domaines d'expertise inédits. Soutien actif de la communauté des

analystes scientifiques, IBM prévoit de poursuivre dans cette voie au travers de forums en ligne et de formations individuelles. Par son engagement (événements, réunions, cours, contenus, contributions auprès de la communauté open source, etc.), IBM est aux côtés des spécialistes des données en activité et en devenir pour mieux les préparer à faire face aux demandes futures de traitements ultraperformants de vastes volumes de données.



Ressources

Êtes vous prêt(e) à élargir vos compétences en science des données ? Nous vous proposons des ressources pour vous accompagner :

1. Développez vos compétences en science des données avec les formations [Big Data University](#)
2. Engagez-vous dans la démarche avec les outils dont vous avez besoin en vous appuyant sur les offres [IBM Data Science Experience](#)



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

IBM Ireland registered in Ireland under company number 16226.

IBM, le logo IBM, **ibm.com** et SPSS sont des marques d'International Business Machines Corp., déposées dans de nombreux pays du monde. Les autres noms de produits et de services peuvent être des marques d'IBM ou d'autres sociétés. Une liste actualisée des marques déposées d'IBM est accessible sur le web sous la mention « Copyright and trademark information » à l'adresse www.ibm.com/legal/copytrade.shtml

Le présent document est proposé dans sa version actuelle à sa date initiale de publication et peut être modifié à tout moment par IBM. Toutes les offres ne sont pas disponibles dans chaque pays dans lequel opère IBM.

Les exemples de clients fournis ne sont mentionnés qu'à titre d'information. Les performances réelles peuvent varier selon les configurations et les conditions de fonctionnement spécifiques. Il appartient à l'utilisateur d'évaluer et de vérifier le fonctionnement d'autres produits ou programmes avec les produits et programmes d'IBM. LES INFORMATIONS CONTENUES DANS CE DOCUMENT SONT FOURNIES « EN L'ÉTAT », SANS AUCUNE GARANTIE EXPRESSE OU TACITE, NOTAMMENT SANS AUCUNE GARANTIE DE QUALITÉ MARCHANDE OU D'ADAPTATION À UN EMPLOI SPÉCIFIQUE, ET SANS AUCUNE GARANTIE OU CONDITION DE NON INFRACTION VIS-À-VIS DES LOIS. Les produits IBM bénéficient d'une garantie, conformément aux conditions générales des contrats dans le cadre desquels ils sont fournis.

© Copyright IBM Corporation 2017



Veuillez recycler