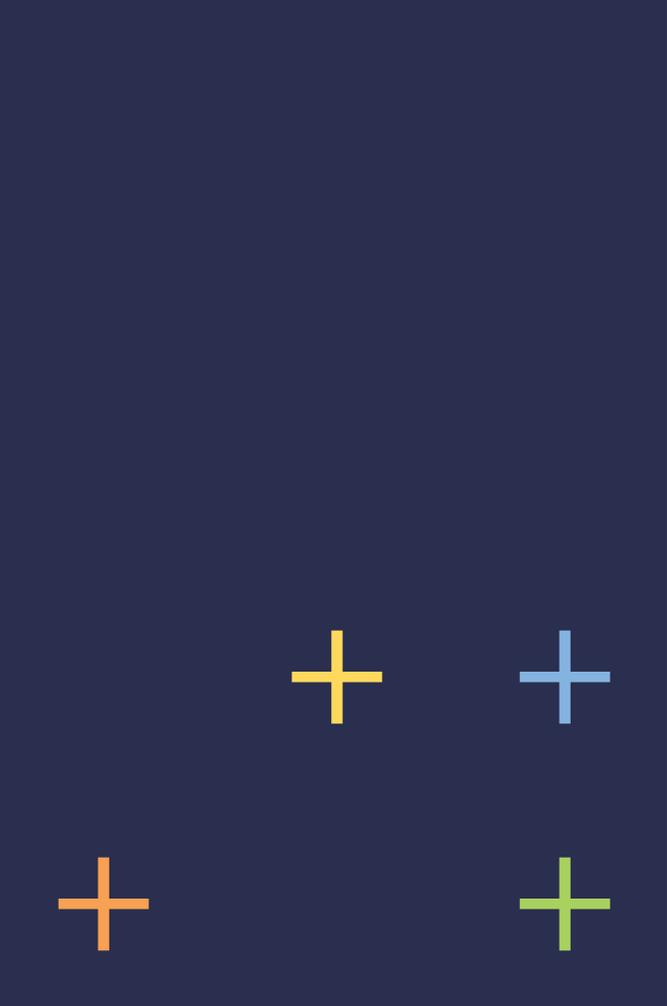


2019

Les tendances de la BI



Sommaire

- 01 L'avènement de l'IA transparente**

Les entreprises s'appuient de plus en plus sur les modèles d'intelligence artificielle et de machine learning. Mais comment s'assurer de leur fiabilité ?
- 02 Le langage naturel humanise vos données**

Les avancées des systèmes de traitement du langage naturel permettent à chacun d'interroger naturellement les données.
- 03 Les analyses exploitables contextualisent les données**

Les plates-formes BI évoluent pour placer les données au cœur des actions des utilisateurs.
- 04 La collaboration autour des données a un impact positif important sur le plan social**

Les efforts concentrés des organismes publics et privés renforcent le mouvement Data for good.
- 05 Les codes de déontologie se mettent à jour sur les questions relatives aux données**

Les responsables prédisent l'avenir des pratiques déontologiques en matière de données à la lumière des réglementations telles que le RGPD.



- 06 **La gestion des données converge avec les plates-formes BI modernes**
La curation gouvernée des données comble le fossé entre les données et les unités métier.
- 07 **La mise en récit des données est le nouveau langage des entreprises**
La recherche et la communication d'informations exploitables constituent désormais un sport d'équipe.
- 08 **Les entreprises vont plus loin en matière d'adoption de l'analytique**
Que se passe-t-il lorsque les leaders se focalisent moins sur l'adoption et davantage sur l'engagement ?
- 09 **La démocratie des données élève le rôle du data scientist**
Les data scientists développent des soft skills pour faire évoluer leur entreprise.
- 10 **L'accélération de la migration de données dans le cloud favorise l'adoption de la BI moderne**
Les données passent au cloud plus vite que jamais, obligeant les entreprises à repenser leur stratégie de gestion des données.

01

L'avènement de l'IA transparente

Les entreprises s'appuient de plus en plus sur les modèles d'intelligence artificielle et de machine learning. Mais comment s'assurer de leur fiabilité ?

L'intelligence artificielle (IA) nous promet d'accroître la compréhension humaine en automatisant la prise de décision. Josh Parenteau, directeur de la veille stratégique chez Tableau, estime qu'à l'avenir, l'intelligence artificielle et le machine learning offriront une nouvelle perspective, « en aidant à dégager des informations exploitables qu'on n'avait encore jamais décelées ». [Une étude Gartner révèle](#) que d'ici 2020, « 85 % des DSI mettront à l'essai des programmes d'IA via plusieurs approches de sourcing :

achat, développement ou externalisation ». Mais, alors que les entreprises s'appuient de plus en plus sur les modèles de machine learning, comment peut-on s'assurer de leur fiabilité ?

De nombreuses applications de machine learning ne permettent pas pour l'instant de « regarder sous le capot » pour comprendre les algorithmes ou la logique ayant mené aux recommandations ou à la prise de décision. Quand elles envisagent une adoption à plus grande

85 %
des DSI mettront
à l'essai des
programmes
d'IA via plusieurs
approches de
sourçage : achat,
développement ou
externalisation.

échelle, les entreprises qui pilotent des programmes d'IA émettent donc des doutes tout à fait légitimes. Comme l'explique Adrian Weller, chercheur senior dans le domaine à l'Université de Cambridge : « La transparence est généralement considérée comme essentielle au déploiement effectif de systèmes intelligents » tels que le machine learning. En effet, elle garantit notamment que les modèles fonctionnent comme prévu et permet aux utilisateurs de faire confiance aux prédictions du système au moment de prendre des décisions.

Ce besoin de transparence a donné naissance à une pratique qui consiste à rendre les modèles de machine learning compréhensibles et transparents. Face à un résultat, les décisionnaires ont davantage de questions : pourquoi le modèle a-t-il donné cette réponse ? Quel est son niveau de certitude ? Qu'aurait-il répondu si les données avaient été différentes ? Au final, ils poseraient les mêmes questions à un expert humain avant de prendre une décision stratégique. Selon Richard Tibbetts, chef de produit pour l'intelligence artificielle chez Tableau, « les décisionnaires ont raison de se méfier lorsque les réponses fournies par un système d'intelligence artificielle ou de machine learning sont inexplicables.

L'analytique et l'intelligence artificielle ont pour objectif d'assister et non pas de remplacer entièrement, l'expertise et la compréhension humaine ».

Dans les entreprises, particulièrement celles qui sont exposées à des risques comme les services financiers ou les laboratoires pharmaceutiques, les responsables métier exigent que les équipes de data science utilisent des modèles plus explicables dont la construction est consignée dans un historique et bien documentée. Pour expliquer ces modèles aux utilisateurs métier, les data scientists se servent des plates-formes BI. En effet, celles-ci permettent d'analyser des conclusions de manière interactive pour mieux les valider.

Les entreprises ont fini par reconnaître la valeur de l'intelligence artificielle et du machine learning. Mais, pour avoir un réel impact, ces systèmes doivent se montrer dignes de confiance. Ils doivent pouvoir justifier leurs conclusions de façon intelligible, aussi simplement que possible et répondre de façon dynamique aux questions supplémentaires, afin d'aider les humains à mieux comprendre leurs données.

“

L'analytique et l'intelligence artificielle ont pour objectif d'assister et non pas de remplacer entièrement, l'expertise et la compréhension humaine.

Richard Tibbetts, chef de produit pour l'intelligence artificielle, Tableau

02

Le langage naturel humanise vos données

Les avancées des systèmes de traitement du langage naturel permettent à chacun d'interroger naturellement les données.

Le traitement du langage naturel combine l'informatique et la linguistique pour aider les ordinateurs à comprendre le langage humain. Aujourd'hui, les fournisseurs de solutions BI offrent une interface de langage naturel dans leurs visualisations. Ainsi, les utilisateurs peuvent interagir naturellement avec leurs données, en posant des questions telles qu'elles leur viennent à l'esprit et sans nécessairement avoir une connaissance approfondie de l'outil BI.

Dans le contexte de la BI moderne, le langage naturel sert à assister la conversation

analytique. Il s'agit d'une discussion entre l'humain et le système au sujet des données. Le système puise dans le contexte de la conversation pour comprendre l'intention de l'utilisateur et faire progresser le dialogue. Ainsi, l'expérience conversationnelle paraît plus naturelle. Par exemple, lorsqu'une personne a une question supplémentaire au sujet de ses données, elle n'a pas à reformuler sa question pour l'approfondir ou clarifier une ambiguïté. Vous pouvez par exemple demander à un outil BI de « trouver les tremblements de terre importants en Californie », puis lui demander

“

Grâce au langage naturel, un public beaucoup plus large peut désormais accéder à tous types de technologies. Le langage naturel élimine l'obstacle technologique : vous n'avez pas besoin d'apprendre à utiliser le logiciel. Vous n'avez même pas besoin d'avoir des connaissances analytiques. Vous devez simplement connaître le contexte métier afin de poser la bonne question.

Stephanie Richardson, directrice senior du marketing produit, Tableau

« et au Texas ? » sans avoir à mentionner les tremblements de terre à nouveau. Le machine learning permet aux systèmes d'améliorer leur connaissance d'un domaine au fil du temps. Cet apprentissage est alimenté par les des données de l'entreprise et le type de questions posées par les utilisateurs.

Vidya Setlur, responsable du développement pour l'équipe en charge du langage naturel chez Tableau, explique : « Éviter les impasses est l'une des caractéristiques clés de la conversation analytique. On doit pouvoir poser une question, obtenir un résultat et rebondir sur cette première question. Tout le monde veut réussir à dégager des informations exploitables de ses données. Le langage naturel est un moyen important pour permettre une meilleure analyse ».

Le langage naturel permet également aux utilisateurs de poser des questions en s'appuyant sur une visualisation de données. Ryan Atallah, ingénieur logiciel chez Tableau, nous donne un exemple : « Imaginons que je demande à mon outil BI de me sortir des informations sur des épidémies et qu'il m'affiche une visualisation. Je peux alors lui

demander de me donner des détails sur un pic représenté en orange. C'est une question supplémentaire, mais pas au sujet de mes données. Elle relève des encodages des visualisations ». Et lorsqu'une visualisation n'est pas pertinente dans le contexte de la question suivante, le système offre une alternative.

Le langage naturel représente un changement radical de la façon dont vous pouvez interroger vos données. Lorsqu'interagir avec une visualisation devient aussi naturel qu'avec un interlocuteur, l'utilisateur moyen peut explorer des zones du pipeline analytique qui étaient autrefois accessibles uniquement aux data scientists et aux analystes spécialisés. Il n'est plus limité par ses compétences analytiques, mais par la portée de ses questions. Les utilisateurs expérimentés peuvent également répondre à des questions plus complexes plus rapidement et proposer des fonctionnalités de tableau de bord plus attrayantes à leurs interlocuteurs. En gagnant du terrain dans le secteur de la BI, le langage naturel fera tomber les barrières qui limitent l'adoption de l'analytique dans les entreprises et contribuera à placer les données au cœur de la culture du travail.

Le marché de la
génération en langage
naturel devrait
représenter

**825,3 millions
de dollars**

d'ici 2023
(Markets and Markets)

03

Les analyses exploitables contextualisent les données

Les plates-formes BI évoluent pour placer les données au cœur des actions des utilisateurs.

Dans le domaine de l'analytique, les utilisateurs ont besoin d'avoir leurs données sous la main lorsqu'ils souhaitent agir. Plutôt que d'avoir un fonctionnement cloisonné avec la réalisation des analyses d'un côté et la prise de décision de l'autre, les utilisateurs doivent pouvoir effectuer ces deux tâches dans le contexte de leurs processus métier et de leurs workflows. Les plates-formes BI répondent à ce besoin en s'intégrant aux opérations, workflows et processus métier fondamentaux. Elles y parviennent grâce à des fonctionnalités comme l'analyse mobile, les analyses intégrées, les extensions de tableau de bord (également connues sous le nom de « compléments »)

et les API. En conséquence, les analyses exploitables accélèrent le processus de prise de décision pour les employés spécialisés ou non.

Ces fonctionnalités permettent aux utilisateurs d'analyser les données et d'agir après avoir décelé des informations exploitables sans faire d'allers-retours. Les analyses intégrées illustrent parfaitement la convergence des informations exploitables et du passage à l'action. Avec les analyses intégrées, les utilisateurs accèdent à leurs données et à leurs informations exploitables à l'endroit même où ils travaillent. Ainsi, ils n'ont pas à accéder à une autre application

“

Pour que les analyses soient exploitables, nous devons nous assurer de transmettre le bon message, à la bonne personne, au bon moment et d'une manière compréhensible.

Peter Benson, responsable des alliances stratégiques, Automated Insights

ou à un serveur partagé. Les analyses peuvent être intégrées à des portails internes (comme SharePoint) ou au sein d'autres applications communes.

Par exemple, de nombreuses entreprises intègrent l'analytique à un outil de gestion de la relation client (CRM) comme Salesforce. Les commerciaux peuvent y consulter en contexte de précieuses données relatives à un compte, par exemple les préférences de produit ou les dépenses dans le temps. Ces informations leur permettent de mieux préparer leur discussion ou de déterminer les prochaines actions à entreprendre avec le client.

Les extensions de tableau de bord abordent cette convergence sous un autre angle. Destinées aux utilisateurs qui passent la majorité de leur temps sur la plate-forme analytique, elles leur permettent d'accéder à d'autres systèmes sans quitter le tableau de bord. Les employés peuvent désormais agir sans avoir à quitter leur workflow analytique.

Par exemple, les responsables IT qui s'appuient sur l'analytique pour surveiller la file d'attente des tickets peuvent utiliser une extension pour modifier les informations relatives à un dossier ou effectuer des

actions directement depuis leur tableau de bord. Ainsi, ils n'ont plus à repasser sur leur système de gestion des tickets. N'ayant pas à se disperser, ils restent mieux focalisés, ce qui leur permet d'agir plus rapidement.

Alors que les extensions de tableau de bord et les analyses intégrées rassemblent les actions et les informations exploitables au sein d'une même plate-forme ou d'un même outil, les analyses mobiles, quant à elles, mettent physiquement ces fonctionnalités à disposition des utilisateurs à tout moment. Les consultants peuvent exploiter les données clients lorsqu'ils sont sur site et les mécaniciens peuvent utiliser des données relatives à l'Internet des objets (IoT) pour réparer un appareil sur le terrain.

Contextualiser les analyses est d'autant plus important que cela permet de personnaliser leur application pour une unité métier ou un secteur d'activité particulier. La convergence des analyses et des actions réduit le temps et l'effort à fournir entre la découverte d'informations exploitables et la prise de décision. Elle démocratise également l'accès aux données au sein des workflows métier, ce qui encourage davantage d'employés à intégrer les données à leurs prises de décisions quotidiennes.

D'ici 2022, **50 %** des projets de plate-forme technologique dans le cadre d'une transformation numérique établiront une corrélation entre les événements et les résultats de l'entreprise. (Gartner)

04

La collaboration autour des données a un impact positif important sur le plan social

Les efforts concentrés des organismes publics et privés renforcent le mouvement Data for good.

Les données ont redéfini les modes de fonctionnement de toutes les structures, y compris des organismes non gouvernementaux (ONG) et les organismes à but non lucratif. À présent, le mouvement Data for good fait fureur aussi bien dans le secteur public que dans le privé. Une étude Gartner révèle que « les mentions du mouvement Data for good sur les réseaux sociaux ont augmenté de 68 % l'année dernière ». Cela indique que le grand public est de plus en plus

conscient de l'impact positif que les données peuvent avoir sur la société.

Des entreprises du secteur privé, par exemple l'opérateur multinational de télécommunications Orange, s'appuient sur les informations exploitables décelées dans les données pour mettre au point des projets destinés à avoir un impact positif sur la société. Dans le cadre du projet OPAL, Orange a fondé un comité de gouvernance en

partenariat avec les organismes d'État locaux, pour contrôler la façon dont les données sont collectées, rendues anonymes et protégées. Cela lui permet de partager en toute sécurité des journaux d'appels avec des organismes œuvrant pour l'intérêt général. Récemment, au Sénégal, les statistiques d'envoi de SMS ont permis d'évaluer le taux d'alphabétisation. Les organismes concernés ont pu s'appuyer sur ces données au moment de décider des ressources à allouer aux programmes d'alphabétisation.

Jusqu'ici, les ONG et les organismes à but non lucratif ne disposaient pas des ressources

nécessaires pour investir dans des infrastructures de données sophistiquées ou dans des équipes spécialisées dans l'analyse de données. À présent, grâce à la flexibilité et à la rentabilité du cloud computing, ces organismes peuvent développer des environnements de données sophistiqués sans investissement conséquent sur site, qui ouvrent la voie à davantage d'initiatives sociales guidées par les données.

En résultent notamment l'émergence et la croissance de plates-formes de mise en commun de données, dont l'objectif est le

partage et la collaboration des organismes pour atteindre un objectif commun. Le Hutch Data Commonwealth, par exemple, est « une équipe pluridisciplinaire au service des chercheurs du Fred Hutchinson Cancer Research Center. Sa mission consiste à fournir à ces derniers des outils innovants de data science, des capacités en matière d'infrastructure et de leur permettre de collaborer facilement pour accélérer leurs recherches. » Dans un article publié par GeekWire, Matthew Trunnell, directeur des systèmes d'information et directeur exécutif du Hutch Data Commonwealth, explique que le cloud sert « d'espace d'échange des données

“

Grâce à la mise en commun des données, les entreprises peuvent partager des données entre elles et avec le monde extérieur de manière fiable, sécurisée et en protégeant la vie privée des particuliers dont les données sont collectées.

Neal Myrick, responsable Monde de Tableau Foundation

Les mentions du mouvement Data for good sur les réseaux sociaux ont augmenté de **68 %** l'année dernière.

scientifiques. Le cloud devient un lieu de réunion et de collaboration, un lieu où tout le monde se rassemble pour le bien commun ». Cette plate-forme de mise en commun s'appuie sur le partenariat avec d'autres instituts de recherche et fournisseurs de technologie, et met les données au cœur de sa mission.

Qu'ils viennent du secteur privé, public ou de ces plates-formes de mise en commun des données, ces partenariats reposent sur la confiance. Les organismes doivent déterminer les éléments essentiels d'un partenariat réussi, y compris les implications juridiques et les normes de gouvernance requises en matière de partage des données. Il leur faut donc estimer les risques en matière de confidentialité et mettre en place des protections autour du partage de données à caractère personnel.

Le Governance Lab de l'Université polytechnique de New York vient de publier un rapport sur les difficultés liées au partage de données relatives aux réseaux sociaux entre les organismes du secteur public et privé¹. Les principes qu'il établit peuvent s'appliquer à bien d'autres types de partenariats en matière de partage. Le GovLab prédit qu'à l'avenir, davantage d'entreprises feront appel à des responsables des données pour gérer les collaborations de ce type. L'objectif sera d'assurer « l'établissement d'un processus clair pour répondre aux demandes relatives aux données, un système pour filtrer ou prioriser certains types d'informations et une méthode pour garantir que les données dévoilées répondent bien à la demande et aux besoins du public ».

L'accès à une mine de sources de données diverses, si, comme dans le cas de Fred Hutch, il est accompagné de mécanismes de contrôle adaptés, peut avoir un impact considérable. Même si ces projets collaboratifs à grande échelle présentent d'autres difficultés, le mouvement Data for good prouve le potentiel altruiste du partage de données. Avec les avancées technologiques, l'augmentation de la maîtrise des données et l'accent mis sur la collaboration, nous créons un environnement adapté pour résoudre certains des plus grands problèmes de la planète.

¹<http://datacollaboratives.org/static/files/social-media-data.pdf>

05

Les codes de déontologie se mettent à jour sur les questions relatives aux données

Les responsables prédisent l'avenir des pratiques déontologiques en matière de données à la lumière des réglementations telles que le RGPD.

La confidentialité des données est au cœur de tous les débats. Aujourd'hui, les utilisateurs sont plus soucieux que jamais des données personnelles qu'ils partagent. Cela a un impact notable sur la façon dont les entreprises abordent la monétisation, la collecte et le partage des données. À l'aune de nouvelles réglementations comme le RGPD, les entreprises doivent redéfinir leur approche en matière de déontologie et de confidentialité des données dans le contexte de leurs activités métier quotidiennes. Pour ce faire, ils s'appuient sur :

Les codes de déontologie : des codes de déontologie s'appliquent déjà pour de nombreuses professions, particulièrement dans les secteurs juridique, médical et financier. Face au rôle croissant des données dans tous les secteurs d'activité, les entreprises cherchent à appliquer des principes similaires aux pratiques liées à l'analytique des données. Comme l'explique Gartner, « à l'ère du numérique, les frontières entre technologies et activités métier sont floues ». Les données sont désormais au cœur de toutes les stratégies. De plus en plus d'entreprises s'appuient sur les

Le nombre de CDO pour qui la déontologie fait partie de leurs responsabilités a augmenté de **10** points (Gartner)

données pour prendre des décisions métier au niveau de chaque service et pour chaque rôle. Cela signifie que la façon dont les données sont utilisées et partagées représente un enjeu pour un public plus large.

En réaction, les dirigeants, et plus particulièrement les CDO (Chief Data Officer), montrent l'exemple en mettant au point des règlements intérieurs relatifs aux données pour toute l'entreprise, dans le cadre de sa transformation numérique. À titre d'exemple, l'enquête de 2017 de Gartner sur le rôle des CDO révèle que « le nombre de CDO pour qui la déontologie fait partie de leurs responsabilités a augmenté de 10 points entre 2016 et 2017 ». Ces codes de déontologie serviront de cadre pour les prochaines décisions prises en matière d'infrastructure, de gouvernance et de recrutement.

Les modifications apportées aux processus métier : les entreprises posent un regard critique sur le cycle de vie complet de leurs données, depuis la collecte jusqu'à l'analyse. Ainsi, elles peuvent évaluer leur stratégie de gestion des données dans son ensemble pour s'assurer qu'elle est conforme à leurs codes de déontologie internes et aux réglementations. Ce processus de vérification n'a rien de ponctuel. Comme l'a noté Accenture dans

son rapport Les principes universels de la déontologie des données (en anglais), « les pratiques de gouvernance doivent être solides, connues de tous et revues régulièrement ». Elles doivent également s'adapter aux évolutions de l'entreprise.

Les principes déontologiques ne s'appliquent pas uniquement à la collecte ou à la gouvernance des données. Il faut également les respecter au moment de l'interprétation des données et de la prise de décision. Les plates-formes BI modernes ont démocratisé l'analyse de données. En conséquence, davantage de rôles seront créés pour surveiller le respect du code déontologique des données. Selon Bridget Cogley, consultante senior chez Teknion Data Solutions, toute personne analysant des données ou communiquant des informations exploitables devrait d'abord « s'interroger sur les partis pris éventuels et s'assurer que les faits sont présentés clairement ». Il est également essentiel de vérifier que « les limites des données sont bien comprises et correspondent à la question posée ». De plus en plus d'utilisateurs vont travailler avec les données. Il est donc essentiel que les pratiques déontologiques deviennent un élément central dans l'apprentissage et la connaissance des données, aussi bien dans un contexte professionnel que personnel.

“

La mise en pratique d'un code déontologique permet de prendre du recul et d'évaluer une situation d'un point de vue éthique. Les pratiques déontologiques liées aux données ont été créées avant tout pour servir de garde-fou dans le cadre du travail, afin que nous sachions comment faire face à un dilemme d'un point de vue personnel, mais également professionnel.

Bridget Winds Cogley, consultante senior, Teknion Data Solutions

06

La gestion des données converge avec les plates-formes BI modernes

La curation gouvernée des données comble le fossé entre les données et les unités métier.

Les sources de données sont de plus en plus complexes, diverses et nombreuses. En conséquence, les déploiements de solution BI moderne reposent plus que jamais sur une bonne gestion des données. Alors que davantage d'employés prennent des décisions en fonction des données, les entreprises doivent s'assurer de la précision de celles-ci, ainsi que de leur bonne utilisation dans les analyses.

Un accès élargi aux données pose des difficultés en matière de gestion et de contrôle des données. Pour y faire face, les entreprises se tournent vers la curation des données. Celle-ci englobe les méthodes de collecte, de nettoyage, de définition et d'alignement des données disparates d'une entreprise. Ce processus permet de combler le fossé entre les données et leur utilisation concrète.

“

La curation des données est le processus visant à identifier les sources de données nécessaires, puis à placer ces données dans le contexte de l'entreprise pour que les utilisateurs métier puissent interagir avec, les comprendre et les utiliser dans leurs analyses.

Mike Hetrick, responsable senior du marketing produit, Tableau

Dans un souci de clarté aussi bien interne qu'externe, les entreprises consacrent déjà un budget important à des technologies intégrant la définition de données à leurs outils analytiques. En réaction, les outils et processus de curation des données (comme les catalogues de données et la gouvernance sémantique) se combinent aux plates-formes BI pour lier les données au contexte des unités métier.

Un catalogue de données fonctionne essentiellement comme un glossaire métier de sources de données et de définitions de données communes pour l'entreprise. Les spécialistes, comme les ingénieurs Data et les responsables des données, peuvent ajouter des descriptions et des définitions à des champs et des sources de données. Elles sont ainsi clairement identifiables, ce qui favorise la découverte d'informations. Les experts peuvent également ajouter de précieux indicateurs de qualité des données, notamment des notifications de certification de contenu fiable, de maintenance ou de suppression d'ensembles de données.

Les utilisateurs ordinaires n'ont pas besoin de savoir où se situent les données au sein de leur source, mais ils veulent pouvoir comprendre ce que celles-ci représentent dans le monde réel. Par exemple, les analystes et les consommateurs de contenu doivent souvent vérifier l'origine

d'une donnée (on parle également « d'analyser la traçabilité »). Et, si les ensembles de données évoluent, les ingénieurs Data et les responsables des données auront besoin d'en analyser l'impact en aval sur leurs tables et schémas. Combiner un catalogue de données et une plate-forme BI aide à rationaliser toutes ces tâches. En effet, cela permet d'obtenir des statistiques d'utilisation pour identifier rapidement les sources de données et les tableaux de bord les plus fréquemment utilisés.

Aussi nécessaires que soient les catalogues de données, la gouvernance des métadonnées offre des opportunités encore plus prometteuses dans le domaine de la gouvernance sémantique. La sémantique permet de connecter le contexte des données, mais également l'intention des actions analytiques. Cela consiste, par exemple, à cartographier des synonymes pour relier des commandes telles que « volume de la commande » et « quantité ». Elle offre aux utilisateurs d'autres moyens d'interagir avec les données et de déceler rapidement de nouvelles informations exploitables. Les interactions en langage naturel en sont un parfait exemple : une plate-forme BI est capable de comprendre des couches de plusieurs questions, du type « mettre en évidence la valeur la plus élevée, la plus basse et la moyenne ».

À mesure que ces technologies et processus continueront de converger, la curation des données et la sémantique offriront une base de plus en plus solide au reste de l'expérience analytique. Cela permettra d'unifier des composants plus disparates de l'écosystème des données, par exemple le nettoyage et l'analyse en aval. Par ailleurs, les recommandations automatiques en matière de tables, de jointures et de modèles de données seront plus pointues. Enfin, grâce aux avancées en matière de curation des données, les employés ne se contenteront plus d'interroger leurs données pendant l'analyse : à l'avenir, ils interrogeront leurs activités métier.

Le volume
des données
numériques
augmentera à un
taux de croissance
annuel composé de
42 %
jusqu'en 2020 (IDG)

07

La mise en récit des données est le nouveau langage des entreprises

La recherche et la communication d'informations exploitables constituent désormais un sport d'équipe.

Peu importe le niveau d'automatisation, la taille de l'ensemble de données et la finesse du calcul : si vous ne pouvez pas communiquer vos trouvailles aux autres, vos analyses n'auront aucun impact. C'est là que la visualisation de données entre en jeu. La visualisation de données est un langage que tous les analystes doivent apprendre pour transmettre aux décideurs des informations à la fois exploitables et faciles à comprendre. Cette compétence, combinée à la capacité des analystes à expliquer les

étapes qu'ils ont suivies pour découvrir les informations exploitables, est souvent qualifiée de « mise en récit des données ».

Cette mise en récit des données est un élément essentiel du processus analytique. Et, face à une culture du travail en pleine évolution où l'analytique est au centre de toutes les préoccupations, il convient de la redéfinir. Alors que les entreprises développent chacune leur culture analytique, les analystes cherchent davantage à alimenter une conversation autour

“

En tant que membres du public, nous devons être désireux d'apprendre. Pour pouvoir interpréter les résultats, nous devons avoir une certaine connaissance du sujet. Si ce n'est pas le cas, c'est au designer de nous en faire comprendre le sens général.

Andy Kirk, fondateur, [VisualisingData.com](https://visualisingdata.com)

D'après une étude de marché réalisée par Dresner en 2018, **75 %** des personnes interrogées estiment que la mise en récit des données est soit essentielle, soit importante pour leurs initiatives en matière d'aide à la décision.

des données qu'à justifier une seule conclusion. Ces cultures analytiques favorisent également les efforts déployés pour apprendre à chaque utilisateur à véritablement comprendre les données et à participer à toutes les étapes de la conversation analytique, depuis la découverte jusqu'à la prise de décision métier.

Andy Kirk, spécialiste de la visualisation de données et fondateur de VisualisingData.com partage les [sept casquettes liées à la visualisation de données](#) (en anglais). L'une d'elles est celle du communicant, « une personne fondamentalement intéressée par toutes les relations interpersonnelles impliquées dans chaque projet (les commanditaires, les intervenants et le public) ». D'après Andy Kirk, « du point de vue de la communication, toutes les visualisations doivent être conçues pour le public ». Les utilisateurs doivent comprendre comment leur public tire une conclusion à partir d'une visualisation. En parallèle, le public doit maîtriser suffisamment le sujet pour interpréter les données et être « désireux d'apprendre ».

Cette évolution de la mise en récit des données se manifeste également dans les tendances de la visualisation de données. Les longs formats de mise en récit, que l'on peut

faire défiler ou qui se déclinent sur plusieurs pages d'un tableau de bord, sont de plus en plus répandus. Ils permettent aux analystes d'illustrer chaque étape de leur cheminement jusqu'à leur conclusion. Ces méthodes permettent aux analystes de démontrer la progression de leur analyse, en mettant en évidence les informations exploitables qu'ils ont dégagées des données et les conclusions qu'ils en ont tirées. L'étape suivante consiste à entamer une conversation ouverte au sujet de ces informations. Ainsi, les employés d'autres services ou ayant des postes différents peuvent partager le contexte de leur unité métier et apporter une perspective différente avant qu'une décision métier soit prise.

Face au nombre croissant d'entreprises qui créent des flux de travail et des équipes dédiés à la collaboration analytique, la mise en récit des données va continuer à s'imposer dans le monde du travail. Cette approche détermine la façon dont les entreprises utilisent les données pour capter l'intérêt, informer et tester des idées. Et, plus les utilisateurs seront nombreux à comprendre comment interpréter les données et expliquer leurs processus analytique, plus grand sera l'impact sur l'entreprise.

08

Les entreprises vont plus loin en matière d'adoption de l'analytique

Que se passe-t-il lorsque les leaders se focalisent moins sur l'adoption et davantage sur l'engagement ?

Les initiatives en matière d'aide à la décision ont souvent une date de début et de fin bien définies. Par ailleurs, il n'est pas rare qu'elles soient considérées comme « terminées » après avoir été déployées auprès des utilisateurs. Mais, pour que les utilisateurs adoptent véritablement les solutions d'aide à la décision, il ne faut pas se contenter de les rendre accessibles. Les CDO réévaluent l'importance de l'adoption de la BI dans les stratégies de modernisation. En effet, ce n'est pas la solution déployée qui compte, mais la façon dont vos employés vont l'utiliser pour prendre des décisions métier.

Croire que le simple fait d'avoir accès à une plate-forme BI permet à chacun d'en tirer des bénéfices peut au contraire freiner vos progressions en matière d'analytique. Comme l'explique Josh Parenteau, directeur de la veille stratégique chez Tableau, si vous évaluez votre retour sur investissement au nombre de licences achetées, « vous risquez de passer à côté d'opportunités d'apprentissage, de croissance et de réussite plus poussée ». ² Au lieu de s'intéresser uniquement à cet aspect primaire de l'adoption, les responsables regardent si les données et l'analytique influencent la prise de décision au sein de l'entreprise. Par exemple, si vous retirez la plate-forme BI à vos employés, comment réagiraient-ils ? Est-ce que cela aurait une influence quelconque sur leur prise de décision quotidienne ?

Vous n'utilisez pas forcément toutes les applications que vous téléchargez sur votre téléphone. De la même façon, ouvrir un rapport une fois par mois n'a pas nécessairement d'influence sur votre prise de décision. Les responsables se penchent donc sur les programmes qui encouragent l'engagement, comme la création de communautés internes et de groupes d'utilisateurs. Ces actions, qui étaient auparavant jugées comme de simples initiatives spontanées des utilisateurs, sont désormais considérées comme des éléments fondamentaux

d'une stratégie d'aide à la décision. Elles permettent aux utilisateurs de maîtriser l'outil plus vite, de devenir autonomes et d'obtenir des réponses plus rapidement. Résultat : la solution est adoptée, ce qui pousse les responsables à augmenter leur investissement et à aider les communautés à se développer.

Dirigé par l'IT, le centre d'excellence de JPMorgan Chase (JPMC) a utilisé ce modèle pour former des milliers d'analystes et accroître sa communauté d'utilisateurs. Il a organisé des sessions d'une journée entière (que Steven Hittle, Vice-président et responsable de l'innovation BI qualifiait de « sessions de thérapie par les données ») pour partager des meilleures pratiques en matière de gouvernance et de visualisation de données. Ces sessions n'étaient que l'une des nombreuses activités mises en place pour encourager à l'engagement et à l'échange entre les rôles et les services. Grâce à ses efforts, JPMC a pu déployer sa plate-forme BI auprès de plus de 30 000 employés.

Ces communautés internes forment elles-mêmes les employés à la plate-forme BI, ce qui permet aux entreprises de commencer à déléguer les responsabilités en matière d'analyses et à créer de nouveaux rôles de champions. À terme, la charge de travail en termes de maintenance et

de reporting, qui incombait traditionnellement à l'IT, s'en trouvera réduite. De plus en plus de champions internes vont émerger. Ils agiront en tant que spécialistes, transmettront les meilleures pratiques et feront en sorte que tout le monde ait la même définition des données. Toutes ces actions inciteront davantage d'employés à utiliser et à tirer parti des logiciels BI. Et, le plus important, vos équipes gagneront en efficacité et votre entreprise prendra une longueur d'avance sur la concurrence.

60 %

des DSI ont l'intention d'augmenter les dépenses en outils analytiques dans les 12 prochains mois. (IDG CIO Tech Poll: Tech Priorities 2018)

² <https://www.tableau.com/fr-fr/about/blog/2017/10/three-reasons-your-business-intelligence-adoption-has-stalled-77448>

“

Nous devons repenser notre manière de mesurer les avantages de la BI. Le nombre d'utilisateurs n'est pas l'indicateur le plus déterminant. Il faut également prendre en compte l'utilisation réelle des solutions analytiques pour prendre des décisions éclairées. C'est la véritable définition de l'adoption.

Josh Parenteau, directeur de la veille stratégique, Tableau

09

La démocratie des données élève le rôle du data scientist

Les data scientists développent des soft skills pour faire évoluer leur entreprise.

Les data scientists ont le vent en poupe. Dans son rapport de 2017, « U.S. Emerging Jobs Report » (Les métiers émergents aux États-Unis), LinkedIn révèle que « les postes de data scientist ont augmenté de 650 % depuis 2012 » et que « des centaines d'entreprises ouvrent ce type de poste » dans de nombreux secteurs. Et la demande commence à suivre : « les ingénieurs machine learning, les data scientists et les ingénieurs Big Data se classent parmi les premiers métiers émergents ».

Mais, étant donné que davantage de services et de rôles sont amenés à travailler avec les

données, les entreprises assistent à une augmentation générale de la connaissance des données et à l'apparition de davantage de data scientists ordinaires. [Gartner définit](#) le data scientist ordinaire comme étant « une personne qui crée ou génère des modèles utilisant des analyses diagnostiques avancées ou des fonctionnalités d'analyses prédictives et prescriptives, mais dont le cœur de métier n'appartient pas aux domaines des statistiques et de l'analytique ». Ces personnes ne remplacent pas les data scientists. Au contraire, elles deviennent leurs partenaires pour le développement et le test d'hypothèses.

“

Aujourd'hui, la connaissance des modèles statistiques et du machine learning sont des prérequis pour devenir data scientist. Ce qui fait la différence, c'est votre capacité à communiquer vos résultats de manière simple mais exploitable.

Sonic Prabhudesai, responsable des analyses statistiques, Charles Schwab

Les postes de data scientist ont augmenté de **650 %** depuis 2012 (LinkedIn)

Ce phénomène fait évoluer la définition de la data science et brouille les frontières entre les employés disposant d'une expertise classique en matière de données et ceux disposant d'une bonne connaissance des unités métier. Sonic Prabhudesai, responsable des analyses statistiques chez Charles Schwab explique : « De plus en plus d'employés des unités métier se familiarisent avec les données, tandis que les data scientists comprennent de mieux en mieux les mécaniques des unités métier ».

Aujourd'hui, on attend des data scientists qu'ils aient des compétences poussées en matière de statistiques et de machine learning, mais également un bon esprit stratégique pour prendre des décisions métier, et une connaissance approfondie de leur secteur d'activité. « La data science ne se résume pas à décortiquer des chiffres : c'est la mise en application de nombreuses compétences afin de résoudre un problème particulier dans un secteur spécifique », explique N. R. Srinivasa Raghavan, Chief Data Scientist Monde chez Infosys. « Les data scientists ont besoin de bien comprendre le domaine dans lequel leurs découvertes vont être appliquées ».

Les résultats produits par les algorithmes et

les modèles sont efficaces uniquement s'ils permettent de résoudre le bon problème dans le bon contexte. Il faut donc collaborer avec les intervenants pour identifier et redéfinir la problématique, ainsi que l'hypothèse, dès le début du processus et les impliquer tout au long du workflow. Puis, à la fin du workflow³, il faut savoir communiquer les résultats aux partenaires de façon pertinente et exploitable.

« Aujourd'hui, la connaissance des modèles statistiques et du machine learning sont des prérequis pour devenir data scientist », explique Sonic Prabhudesai. « Ce qui fait la différence, c'est votre capacité à communiquer vos résultats de manière simple mais exploitable ». Les data scientists ne se contentent plus de transmettre des résultats. Au contraire, ils ont un rôle central à jouer dans l'application de ces résultats dans les unités métier.

Grâce aux outils d'analytique en libre-service, les utilisateurs expérimentés peuvent, tout comme les data scientists, explorer et mieux comprendre leurs données. Cela permet de découvrir des informations exploitables qui peuvent orienter le reste de l'analyse et, au final, faire évoluer l'entreprise.

³ <https://medium.com/@sonicmsba/how-to-build-an-effective-business-context-for-data-analytical-problems-cb02906341cd>

10

L'accélération de la migration de données dans le cloud favorise l'adoption de la BI moderne

Les données passent au cloud plus vite que jamais, obligeant les entreprises à repenser leur stratégie de gestion des données.

Dans le cadre de la modernisation de votre stratégie des données, vous devez réfléchir à leur emplacement de stockage. De nombreuses entreprises migrent leurs données dans le cloud. Elles y trouvent de nombreux avantages, notamment une flexibilité et scalabilité accrues, ainsi qu'un coût total de possession (TCO) plus faible. Une étude Gartner révèle que « le marché des services de cloud public devrait augmenter de 21,4 %

en 2018 pour atteindre les 186,4 milliards de dollars ».⁴ Grâce au cloud, les entreprises peuvent capturer et intégrer plus facilement différents types de données. Mais pour ce faire, elles doivent passer d'un environnement rigide, où toutes les données étaient stockées dans des entrepôts sur site, à une infrastructure plus flexible et évolutive. Elles peuvent choisir de passer à une solution entièrement hébergée dans le cloud ou hybride.

⁴ <https://www.gartner.com/newsroom/id/3871416>

C'est là qu'intervient le concept de gravité des données, qui suggère que les applications et les services sont attirés vers le lieu de stockage des données⁵. De plus en plus d'entreprises migrent du contenu vers le cloud à une vitesse accrue. En conséquence, la gravité des données attire les processus analytiques vers le cloud. « Quand les entreprises passent à Google Cloud, leurs dirigeants repensent entièrement leur stratégie d'analytique des données et reconsidèrent la façon dont le cloud peut influencer leur activité et leurs résultats financiers, » explique Sudhir Hasbe, directeur de la gestion produit chez Google Cloud.

Ce changement gravitationnel s'explique par plusieurs facteurs : la latence, c'est-à-dire le temps requis pour effectuer une action, et le débit, c'est-à-dire le nombre de fois qu'une action peut être exécutée ou un résultat obtenu par unité de temps donnée. Que se passe-t-il lorsque les données, les applications et les services sont rapprochés ? La latence et le débit diminuent, et l'efficacité augmente. Naturellement, si les données résident dans le cloud, ces applications et services ne tarderont pas à les suivre.

Lorsqu'elles évaluent leur stratégie de gestion des données à grande échelle, les entreprises repensent également leur modèle analytique

et passent d'une approche classique de la BI à une approche moderne. McKinsey estime que la véritable valeur du cloud se révèle lorsque les entreprises abordent l'infrastructure et les systèmes dans le cloud « non pas comme une décision tactique ponctuelle, mais comme un élément d'une stratégie holistique de transformation numérique ».⁶

Les solutions BI classiques s'appuient sur le service IT pour répondre aux questions, ce qui crée des goulets d'étranglement et ne permet pas d'ancrer les analyses dans le contexte des unités métier. De la même façon, les déploiements de solution BI classique suivent souvent un modèle sur site rigide, conçu pour soutenir ce type de reporting.

Au contraire, l'analytique dans le cloud offre de nombreux avantages, par exemple la possibilité d'envisager de nouveaux modèles de déploiement. Et les responsables sont impatients de saisir ces opportunités.

Grâce à ces solutions, les employés de terrain peuvent, par exemple, consulter des tableaux de bord sur un appareil mobile pour accéder à des données sans avoir à désactiver de pare-feu. Le cloud permet également le partage sécurisé de

tableaux de bord avec des partenaires ou des clients. Ainsi, tout le monde dispose d'une source unique d'informations fiables qui dépasse le cadre des processus internes.

Même si toutes les entreprises ne sont pas prêtes à migrer toutes leurs données dans le cloud, elles sont nombreuses à tester des solutions hybrides pour exploiter diverses sources de données. Elles évaluent donc les plates-formes BI modernes en fonction de leur capacité à prendre en charge une future transition vers une analytique entièrement dans le cloud.

Le marché des services de cloud public devrait augmenter de 21,4 % en 2018 pour atteindre les **186,4 milliards de dollars** (Gartner)

⁵ <https://www.techrepublic.com/article/how-data-gravity-both-hurts-and-helps-cloud-adoption/>

⁶ <https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/cloud-adoption-to-accelerate-it-modernization>

“

Les dirigeants repensent entièrement leur stratégie d'analytique des données et reconsidèrent la façon dont le cloud peut influencer leur activité et leurs résultats financiers.

Sudhir Hasbe, directeur de la gestion produits chez Google Cloud

 + a b | e a u[®]