



Cinq idées reçues sur les lacs de données

Au fur et à mesure que les entreprises prennent conscience de la valeur de leurs données et de leur capacité à offrir de nouvelles opportunités et de nouveaux modèles de revenus, elles sont inondées de technologies qui offrent la « meilleure » approche pour gérer toutes ces données et générer des informations exploitables.

Ce n'est pas un hasard si le lac de données est l'une de ces architectures de référence qui attire le plus d'attention. Cependant, il s'accompagne de diverses idées reçues qui font perdre un temps précieux aux entreprises et aux responsables informatiques lorsqu'ils évaluent les solutions.

1.

Un lac de données ne peut être déployé que dans un seul cloud → Faux

Le lac de données n'est pas limité à un seul emplacement ou cloud. Il ne se limite pas non plus à un déploiement sur site. Vous pouvez créer un lac de données sur plusieurs clouds avec des options hybrides. Le lac de données est une architecture de référence indépendante de la technologie. C'est une approche qu'utilisent les entreprises pour placer les données au cœur de leurs opérations afin de faciliter l'accès à différents types de données à grande échelle et de mettre les analyses à la disposition des utilisateurs en mode libre-service.

2.

Hadoop est le seul lac de données → Faux

Même si le terme « lac de données » est souvent associé à Hadoop, ou au stockage d'objets orienté Hadoop, un lac de données peut être développé et utilisé efficacement sans intégrer Hadoop. Par exemple, un lac de données efficace peut être basé sur différents systèmes de gestion de bases de données relationnelles. Le lac de données regroupe une variété de technologies afin de créer des systèmes d'analyse permettant aux scientifiques des données d'effectuer une exploration agile en vue de répondre aux besoins métier.

3.

Vous pouvez placer n'importe quelles données dans un lac de données sans avoir besoin de gouvernance → Faux

Même si les logiciels et le matériel sont des composants clés d'une solution de lac de données, le catalogage, la qualité et les processus de gouvernance et de gestion des données sont tout aussi importants.

Tout comme certains entrepôts de données sont devenus de gigantesques trous noirs d'où de grandes quantités de données ne s'échappent jamais, un lac de données peut se transformer en un véritable « marécage » si des règles de gouvernance appropriées ne sont pas appliquées.

4.

La fourniture d'accès est le paramètre qui détermine la réussite d'un lac de données → Faux

Déverser des données dans un emplacement central ne constitue pas en soi une solution d'analyse. L'objectif est d'effectuer des analyses des données qui produisent des informations commerciales exploitables, mais aussi de découvrir de nouvelles sources de revenus, de nouveaux modèles de fidélisation client ou de nouvelles extensions de produits.

Ces données doivent cependant être fiables, pertinentes et accessibles à tous les consommateurs de données. Un lac de données nécessite un catalogue de métadonnées intelligent et capable de faire le lien avec la terminologie métier, permettant de déplacer des données codées cryptiques et de les rendre plus compréhensibles grâce à la contextualisation. De même, il affecte des actifs informationnels structurés et non structurés, ainsi qu'une matrice de gouvernance, aux données en fonction de leur source et de leur qualité afin de garantir la protection, la standardisation, la gestion efficace et la fiabilité des informations.

5.

Le lac de données remplace l'entrepôt de données → Faux

Le lac de données peut regrouper plusieurs entrepôts de données d'entreprise (EDW), ainsi que d'autres sources de données telles que les réseaux sociaux ou l'IdO. Toutes ces sources sont regroupées dans le lac de données où la gouvernance peut être intégrée, simplifiant ainsi la découverte fiable des données pour les utilisateurs de toute l'entreprise.

Par conséquent, le lac de données vient enrichir les environnements EDW et permet aux scientifiques et analystes des données d'explorer facilement leurs données et de découvrir de nouvelles perspectives, idées et opportunités afin d'accélérer l'innovation et la croissance de l'entreprise.



Avantages du lac de données multicloud

Ne vous laissez pas bernier par la désinformation ! Un lac de données gouverné peut fournir, à l'échelle de toute l'entreprise, un accès à un large éventail d'actifs informationnels structurés et non structurés, tout en garantissant leur fiabilité et leur sécurité, où qu'ils se trouvent.

Lorsqu'il est optimisé selon les besoins de l'entreprise, le lac de données gouverné peut accélérer l'analytique et améliorer la précision des informations car :



Les données résident sur une infrastructure sécurisée et fiable.



Des flux de données contrôlés remplissent le lac de données d'informations fiables, puis documentent les ressources d'information, leurs métadonnées et leur contexte métier. Ils constituent un flux de données en temps réel vers le lac de données.



La qualité, l'origine et la lignée des données sont bien comprises.



Les données sont présentées en langage commercial, ce qui permet aux scientifiques des données de se mettre immédiatement au travail plutôt que se débattre avec des termes mystérieux.



Les données sont correctement classées, protégées et gouvernées.

La vérité éclate au grand jour

Le [portefeuille de solutions de gestion de données d'IBM](#) a permis aux clients d'éviter les pièges courants et les idées reçues sur les lacs de données. Il les a guidés tout au long des étapes critiques vers la mise en œuvre réussie d'un lac de données gouverné.

Découvrez les solutions de lac de données d'IBM

Prenez des décisions avisées en exploitant différents types de données provenant de sources plus nombreuses.



Lire l'eBook

Découvrez comment les lacs de données gouvernés permettent d'obtenir des informations métier essentielles.

