



Points clés

- Interface utilisateur graphique conviviale dotée de connecteurs de données accessibles à l'aide d'un glisser-déposer.
 - Développement accéléré grâce à des fonctions de transformation intégrées.
 - Nettoyage des données complet et personnalisable
 - Reconnaissance et classification automatiques des données pour une gouvernance plus efficace.
-

IBM BigInsights BigIntegrate et BigInsights BigQuality

Le pouvoir de l'information pour votre écosystème Big Data

L'analyse du Big Data apparaît comme un facteur clé pour capitaliser sur la valeur métier et d'autres avantages organisationnels obtenus à partir des immenses volumes de données dont dispose l'entreprise. Si nombre d'entreprises expérimentées exploitent déjà les possibilités d'analyse et la rationalisation des coûts inégalées qu'offre Hadoop. Par nature, cet environnement n'est pas conçu pour l'intégration des données. En effet, la réussite de l'intégration des données nécessite de répondre à des exigences spécifiques, notamment concernant la gouvernance, la gestion des métadonnées, la qualité des données et les styles de production de données flexibles.

Cependant, les capacités essentielles d'intégration de données proposées dans l'offre Hadoop sont immatures, voire inexistantes, mais doivent néanmoins être prises en compte tout au long des projets. Plus important encore, la plupart des entreprises constatent que Hadoop n'est, à lui seul, pas assez robuste pour assurer une intégration des données efficace et que les compétences et les opérations nécessaires pour construire une solution complète ne sont pas viables.

Pour aider les entreprises à tirer pleinement parti du périmètre d'application et du potentiel de Hadoop, IBM® a conçu une solution puissante et rentable pour contribuer à la mise en œuvre de projets critiques et d'initiatives clés en matière d'analyse à l'échelle de l'ensemble d'un écosystème de données. Fondés sur la plateforme IBM InfoSphere Information Server, IBM BigInsights BigIntegrate et BigInsights BigQuality apportent les capacités d'intégration et de gouvernance de l'information de bout en bout dont les entreprises ont besoin pour :

- Comprendre les données ;
- Nettoyer, surveiller, transformer et diffuser les données ;
- Collaborer à l'échelle de l'entreprise, rapprocher les activités métier et la fonction Informatique.

En capitalisant sur les capacités d'intégration flexible qu'apportent IBM BigInsights BigIntegrate et BigInsights BigQuality, les entreprises peuvent s'assurer que les informations sur lesquelles reposent leurs initiatives métier et stratégiques – en passant de la notion de système d'enregistrement à celle de système de connaissances – sont fiables, cohérentes et gérées en temps réel.





IBM BigInsights : l'analyse appliquée aux grands ensembles de données

BigInsights apporte aux entreprises les moyens d'identifier des modèles en analysant l'ensemble des données, au lieu de se cantonner à quelques échantillons. Pour appliquer ces processus d'analyse à des ensembles de données extrêmement volumineux, BigInsights dispose d'un éventail complet d'algorithmes à apprentissage automatique massivement évolutifs (baptisés SystemML et créés par IBM Research), programmables en langage R, facile d'accès. Un ensemble de données suffisamment volumineux est essentiel pour construire un modèle statistique pertinent. En effet, les modèles les plus performants apportent des capacités prédictives plus précises et conduisent donc à de meilleurs résultats. Intégrés à la gamme de produits BigInsights, BigInsights BigIntegrate et BigInsights BigQuality apportent les moyens nécessaires pour assembler, transformer, nettoyer et mettre à disposition les données pour les analyses effectuées ensuite.

BigInsights BigIntegrate et BigInsights BigQuality : l'intégration des données en mode natif

BigInsights BigIntegrate et BigInsights BigQuality permettent aux entreprises d'intégrer et de transformer à l'aide de Hadoop toutes les données dont elles disposent, dans le cadre de leurs initiatives Big Data, en exploitant à la fois des sources de données existantes et nouvelles.

En outre, une interface graphique conviviale leur permet de transformer rapidement toutes les informations disponibles dans l'entreprise. La palette étendue de fonctionnalités proposées comporte des connecteurs vers un très large éventail de sources de données, en particulier vers les bases de données traditionnelles les plus courantes, toutes plateformes confondues (distribuées, IBM z/OS, etc.), les types de fichiers, les applications d'entreprise (notamment Oracle, Salesforce.com et SAP) et Hadoop. Les développeurs assurent l'accès à ces sources de données par de simples actions à la souris et peuvent adapter les paramètres de configuration pour maximiser les performances et la flexibilité.

En effet, BigInsights BigIntegrate et BigInsights BigQuality contribuent à l'amélioration des performances et offrent une extrême évolutivité. La connectivité native vers les sources de données les plus courantes est assurée par des interfaces spécifiques et plusieurs centaines de fonctions de transformation intégrées permettent d'accélérer le développement. Résultat : la plupart des opérations de codage sont effectuées infiniment plus rapidement qu'avec un codage manuel. Selon une étude, un développeur n'a consacré que deux jours à la concrétisation d'une tâche dont la réalisation en a nécessité 30, l'équivalent d'environ 2 000 lignes de code sans documentation appropriée, avec l'impossibilité de réutiliser un code existant et d'importantes difficultés de maintenance.

Les fonctionnalités plus récentes de Hadoop utilisent le même environnement de conception graphique simplifié qu'utilisent de longue date les clients d'IBM pour créer des millions d'applications d'intégration. Dans cet environnement, il est possible de créer des applications de traitement intensif de l'information utilisant Hadoop sans avoir à former de nouvelles équipes de développement à de nouveaux langages pouvant nécessiter un codage manuel et dépourvus de capacités de gouvernance.

Une collaboration étroite avec des outils de conception parfaitement intégrés

Les fonctionnalités de design intégré et le framework de métadonnées commun permettent d'assurer une collaboration mutuelle entre tous les groupes en s'appuyant sur des métadonnées identiques et cohérentes. Grâce à l'interface graphique, les utilisateurs peuvent automatiser le processus de spécification de la conception et enrichir les métadonnées en disposant d'une vue métier complète et contextualisée des actifs techniques. En outre, les outils de conception intégrés permettent aux développeurs de construire des jobs combinant des processus d'intégration et de qualité des données dans une tâche spécifique unique, ce qui permet d'assurer un niveau d'efficacité maximal. Grâce à une approche mutualisée de la conception, l'entreprise gagne également en flexibilité en intégrant l'assurance qualité dans le processus, là où elle est indispensable.



De plus, BigInsights BigIntegrate et BigInsights BigQuality apportent des fonctionnalités hautement performantes d'intégration et de gouvernance. Les jobs d'intégration et de transformation des données peuvent être directement exécutés dans un cluster Hadoop, permettant ainsi la disponibilité immédiate de puissantes fonctionnalités (connectivité, transformation, nettoyage, enrichissement et diffusion des données) dans la plateforme Hadoop. Grâce à cette approche, il est possible de résoudre certaines des problématiques les plus complexes en matière de données puisqu'elle permet de limiter au minimum les transferts de données, et donc d'accélérer la création de valeur métier.

BigInsights BigQuality : le nettoyage des données en mode natif pour Hadoop

BigInsights BigQuality apporte les capacités nécessaires pour nettoyer les données et en surveiller la qualité en vérifiant qu'elles sont bien transformées en informations fiables. En effet, la fiabilité des données est essentielle pour créer de la valeur à partir de ressources de données existantes, contribuant ainsi à l'amélioration du processus décisionnel et à la rationalisation des activités opérationnelles.

Pour ce faire, les fonctionnalités personnalisables et complètes de nettoyage des données, appliquées sous la forme de traitement par lots et en temps réel, permettent d'automatiser les recherches de données dans les sources et leur classification. Avec cette automatisation, l'équipe chargée de la coordination peut gérer de manière efficace les actifs de données et répondre aux objectifs métier grâce à des données fiables. Par ailleurs, l'automatisation facilite le traitement et la gestion des immenses volumes de données traités par la plateforme Hadoop, avec l'évolutivité requise pour y faire face.

De plus, l'augmentation rapide du volume des données conduit les communautés de fournisseurs et de consommateurs de données à se développer, entraînant des incertitudes quant à la confidentialité de ces informations et le respect des obligations légales qui en découlent. D'où l'intérêt de BigInsights BigQuality, qui permet de faciliter les investigations relatives aux différentes sources de données, y compris Hadoop, et de veiller à ce que la localisation et l'utilisation des données soient conformes aux politiques prédéfinies. En outre, BigInsights BigQuality permet de standardiser et de corréliser les enregistrements en fonction de règles métier personnalisables. En établissant des indicateurs de qualité des données en phase avec les objectifs métier, une entreprise peut rapidement détecter les problèmes de qualité et mettre en place un plan destiné à les résoudre. Ces capacités sont donc essentielles pour assurer une gouvernance efficace des données.

Pourquoi IBM ?

Les nouvelles capacités offertes par BigInsights BigIntegrate et BigInsights BigQuality pour Apache Hadoop viennent étendre la gamme IBM Analytics de solutions Big Data aussi bien en termes de logiciels et de services que de recherche et de matériel. Cette gamme associe les outils traditionnels d'entrepôts de données et les techniques et technologies actuelles de Big Data – notamment l'environnement Hadoop, l'informatique en flux continu, l'exploration des données, l'analytique évoluée, l'intégration dans l'entreprise et l'informatique cognitive avec IBM Watson – pour créer une solution intégrée permettant à n'importe quelle entreprise de capitaliser sur le Big Data et l'analytique.

Pour de plus amples informations

Pour en savoir plus sur BigInsights BigIntegrate, BigInsights BigQuality et d'autres plateformes d'intégration de données, contactez votre interlocuteur commercial IBM, ou cliquez ici : ibm.com/software/data/information-integration-governance



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

IBM, le logo IBM, ibm.com, BigInsights, IBM Watson, InfoSphere et z/OS 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 IBM est accessible sur le web sous la mention « Copyright and trademark information » à l'adresse 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 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.

Le client est responsable de sa conformité aux lois et aux réglementations qui lui sont applicables. IBM n'assure aucun conseil juridique, et ne déclare ou ne garantit en aucune manière que ses services ou ses produits assurent la conformité du client avec une loi ou une réglementation quelconques.

© Copyright IBM Corporation 2015



Veillez recycler
