

# Comment répondre aux enjeux du Big Data ?

Optimisation du retour sur les données Big Data avec Informatica 9.5



Le présent document contient des données confidentielles et exclusives, ainsi que des informations constituant des secrets commerciaux (« Informations confidentielles ») d'Informatica Corporation. Il ne peut être copié, distribué, dupliqué ni reproduit de quelque manière que ce soit, sans l'autorisation écrite préalable d'Informatica.

Même si tout a été mis en œuvre pour garantir que les informations contenues dans ce document sont exactes et exhaustives, il est possible qu'il contienne des erreurs typographiques ou des inexactitudes techniques. Informatica ne saurait être tenu responsable des pertes résultant de l'utilisation d'informations figurant dans ce document. Les informations contenues dans le présent document sont susceptibles d'être modifiées sans préavis.

L'intégration des attributs de produits étudiés dans ce document dans une quelconque version ou mise à jour d'un produit logiciel Informatica — ainsi que le calendrier de sortie de ces versions ou mises à jour — sont à la seule discrétion d'Informatica.

Protégé par les brevets américains suivants : 6,032,158 ; 5,794,246 ; 6,014,670 ; 6,339,775 ; 6,044,374 ; 6,208,990 ; 6,208,990 ; 6,850,947 et 6,895,471 ; ou par les brevets américains en instance suivants : 09/644,280 ; 10/966,046 ; 10/727,700.

Édition publiée en avril 2012

## Table des Matières

<b>Mutation informatique .....</b>	<b>2</b>
<b>Les 3 V : Volume, Variété, Vitesse .....</b>	<b>3</b>
<b>Optimisation de la valeur du Big Data .....</b>	<b>4</b>
<b>Big Data : éviter le cauchemar .....</b>	<b>4</b>
<b>Informatica 9.5 : optimisation du retour sur les données Big Data .....</b>	<b>5</b>
<b>Optimisation de la valeur des données .....</b>	<b>6</b>
<b>Réduction du coût des données .....</b>	<b>8</b>
<b>Concrétiser la promesse du Big Data. ....</b>	<b>9</b>
<b>Conclusion .....</b>	<b>11</b>

## Mutation informatique

Trois grandes tendances technologiques opèrent une mutation sur le monde de l'informatique. Le Cloud Computing, les réseaux sociaux et les mobiles apportent une valeur ajoutée incroyable. Mais, parallèlement, ils poussent les entreprises à adopter de nouvelles approches informatiques tout en adaptant les nouvelles réalités et exigences à leurs processus et infrastructure traditionnels.

Ces grandes tendances redéfinissent les questions liées à l'informatique, à savoir « où », « quoi » et « comment », et alimentent la croissance explosive des données Big Data — volumétrie, variété et vitesse accrues.

- Nous sommes passés des systèmes sur site aux systèmes dans le Cloud (« où »), ce qui a généré une nouvelle vague de fragmentation de données et modifié le modèle économique de l'informatique.
- Nous avons basculé des systèmes de traitement des transactions aux systèmes de « traitement des interactions » (« quoi ») — les personnes interagissent via des réseaux sociaux et les appareils électroniques via l'« Internet des objets » — déclenchant ainsi une avalanche sans précédent des données d'interaction.
- Quant au « comment », les appareils mobiles ont remplacé les postes de travail, ce qui a permis aux utilisateurs de bénéficier de services contextualisés basés sur la localisation et généré une augmentation des données d'interaction issues de capteurs et d'appareils électroniques.

La richesse des nouvelles données peut offrir des opportunités commerciales inouïes. En tirant parti du Big Data, les entreprises peuvent renforcer leur efficacité opérationnelle, réduire les coûts d'infrastructure informatique et de gestion des données et mieux gérer les marques et la relation client. Mais peu d'entre elles sont actuellement équipées pour capitaliser sur la valeur des données Big Data. En négligeant de s'adapter à l'augmentation des volumes, de la variété et de la vitesse des données, les entreprises s'exposent également à des coûts, à une complexité et à des risques imprévus.

## Les 3 V : Volume, Variété, Vitesse

En apparence, le concept de Big Data peut être assez trompeur. Il ne souligne que le volume. Le volume est, sans aucun doute, un aspect essentiel du Big Data qui affectera les entreprises, quel que soit le secteur d'activité, dans la mesure où, d'après une étude réalisée par l'institut de recherche IDC<sup>1</sup>, la volumétrie au sein des datacenters des entreprises devrait être multipliée par 50 au cours des dix prochaines années. Mais ce volume ne représente qu'un aspect du phénomène Big Data. L'augmentation parallèle de la variété et de la vitesse des données est tout aussi importante. Il faut tenir compte des éléments suivants :

- **Variété des données** : la croissance des réseaux sociaux — forums, LinkedIn, publications sur Facebook, tweets associés aux marques, etc. — a considérablement favorisé la multiplication des types de données accessibles aux entreprises, tout comme l'augmentation des interactions via l'informatique mobile et entre les appareils. Les données ne sont plus structurées et générées par des applications transactionnelles — elles s'apparentent à une concentration confuse d'informations imprévisibles et non structurées issues de sources très diverses.
- **Vitesse des données** : les données circulent à une vitesse incroyable dans un monde où la communication est omniprésente. Le déplacement des données s'effectuant en moins d'une seconde, la capacité des entreprises à identifier les événements en temps réel et à réagir en un temps record revêt une importance cruciale. Les transactions financières, la chaîne logistique, les télécommunications et le commerce électronique ne représentent qu'une partie des secteurs qui doivent s'adapter à la vitesse de circulation des données de plus en plus vertigineuse.

Cette croissance du volume des données, de leur variété et de leur vitesse — les trois V — est commune aux trois aspects du Big Data — à savoir, le Big Data de transaction, le Big Data d'interaction et le traitement des données Big Data.

- **Big Data de transaction** : les données transactionnelles classiques des bases de données et applications d'entreprise augmentent régulièrement, ce qui met en péril les performances des applications et génère une augmentation des coûts d'infrastructure. En outre, les entreprises sont contraintes d'agir de plus en plus rapidement face aux événements qui se déroulent en temps réel. Enfin, la fragmentation des données transactionnelles est encore plus importante dans la mesure où les entreprises migrent leurs processus métiers stratégiques et les données associées des systèmes sur site vers le Cloud, ce qui accentue la complexité de leurs environnements de données de transaction.
- **Big Data d'interaction** : les données sociales et des appareils issues des interactions générées par l'homme et la machine existent, tant au sein qu'au-delà du pare-feu. Outre les publications sur Facebook, elles comprennent les enregistrements détaillés des appels (CDR), les données géospatiales, l'enregistrement des navigations Internet et les informations issues des capteurs dans les secteurs scientifiques, médicaux, de la chaîne logistique et autres. Ces facteurs entraînent une augmentation sans précédent du volume et de la diversité des données.
- **Traitement des données Big Data** : face à ces bouleversements, les entreprises se tournent vers des technologies telles que Apache Hadoop et MapReduce, destinées au traitement et à l'analyse de vastes quantités de données. Ces plates-formes à faible coût ont modifié le calcul lié à la rentabilité des types d'application et d'analyse des données. Mais ces nouveaux environnements dédiés risquent de poser le problème des silos de données et nécessitent une stratégie d'intégration des données afin d'assurer l'interopérabilité avec les systèmes traditionnels.

<sup>1</sup> IDC, « The 2011 Digital Universe Study: Extracting Value from the Chaos », juin 2011.

## Optimisation de la valeur du Big Data

Le Big Data a été qualifié comme étant « le prochain défi pour l'innovation, la compétitivité et la productivité »<sup>2</sup>. Il offre à votre entreprise des opportunités inédites d'augmentation des revenus, des performances et des connaissances liées à votre activité, au marché et aux clients. Les données des médias sociaux en sont le parfait exemple. En s'appuyant sur la technologie pour capturer des informations issues de Facebook, LinkedIn, Twitter, Foursquare, Google+ et d'autres sources, votre entreprise peut affiner la gestion de la marque, découvrir de nouveaux canaux de distribution et renforcer la relation client.

Les données transactionnelles classiques permettent de savoir ce qu'il s'est passé avec votre client, mais de manière incomplète. Les données issues des médias sociaux vous offrent une vue plus pertinente et précieuse de l'avenir parce qu'elles peuvent inclure des précisions sur les préférences, opinions, relations, comportements ainsi que d'autres informations souvent disponibles uniquement sur ces réseaux. L'intégration de données est la technologie qui permet de révéler ce potentiel.

Les opportunités qu'offrent les données Big Data s'étendent également au-delà des réseaux sociaux et couvrent tous les secteurs. Dans le domaine de la fabrication et de la chaîne logistique, les flux de données en temps réel issus de capteurs peuvent aider les entreprises à optimiser la logistique, les inventaires et la production tout en leur permettant d'identifier rapidement les zones à risque. Dans le secteur des télécommunications, la maîtrise des enregistrements détaillés des appels (CDR), des messages textes et des données issues des mobiles réduit la perte de clients et optimise les réseaux. Les services financiers peuvent réagir en temps réel face aux événements du marché et gérer plus efficacement les risques financiers. Le secteur de la santé, entre autres, est en mesure d'affiner ses analyses en enrichissant les data warehouses traditionnels d'informations extraites de nouvelles sources.

## Big Data : éviter le cauchemar

Capitaliser sur la valeur du Big Data requiert une infrastructure de données évolutive, flexible et simple à gérer, ce qui fait défaut dans la plupart des entreprises. Ces sociétés pionnières qui mettent en œuvre les processus et systèmes adaptés tendent à prendre une longueur d'avance sur la concurrence qui ne profite pas de l'occasion. Sans stratégie cohérente pour le Big Data, votre entreprise s'expose aux risques suivants :

- Données fragmentées dans des environnements informatiques de plus en plus complexes
- Décisions inappropriées basées sur des données inexactes ou dupliquées
- Coûts élevés de la main d'œuvre, des logiciels et de l'infrastructure
- Opportunités manquées en termes de revenus et d'engagement des clients
- Faible rentabilité due à des processus redondants et à un défaut de réutilisation
- Risques inutiles en matière de conformité, d'efficacité opérationnelle et de performances financières

La plupart des entreprises n'atteindront pas la situation idéale, comme le prévoit l'institut de recherche Gartner. Selon Gartner, « Plus de 85 pour cent des entreprises figurant dans le classement Fortune 500 ne parviendront pas à exploiter le Big Data et à accroître leur avantage concurrentiel, et ce jusqu'en 2015 ». « La majorité des entreprises ne sont pas préparées à faire face aux problématiques techniques et de gestion que pose l'importante volumétrie des données ; peu seront donc en mesure de faire de cette tendance un avantage concurrentiel. »<sup>3</sup>

<sup>2</sup> McKinsey Global Institute, « Big Data: The Next Frontier for innovation, Competition, and Productivity », mai 2011.

<sup>3</sup> Gartner, « Gartner Reveals Top Predictions for IT Organizations and Users for 2012 and Beyond », communiqué de presse, 1er décembre 2011.

## Informatica 9.5 : optimisation du retour sur les données Big Data

Informatica® 9.5, la dernière version de la plate-forme Informatica, vous permet de maximiser le retour sur le Big Data — que vous élaboriez une stratégie sur la manière de gérer cette augmentation exponentielle du volume, de la variété et de la vitesse des données, de dégager une plus grande valeur des données plus traditionnelles ou de tirer profit de l'association de données nouvelles et conventionnelles.

Informatica 9.5 constitue un progrès majeur sur différents aspects de la valeur des données et offre des fonctionnalités nouvelles capables de renforcer leur caractère pertinent, fiable, opportun, irréfutable, exploitable, accessible, complet et sécurisé. Les services d'intégration dans le Cloud, par exemple, garantissent que les applications cloud disposent de toutes les données métiers pertinentes. L'historique des données en matière de gestion des données de référence (MDM) des réseaux sociaux crée, au fil du temps, une vue du client complète, irréfutable et en 4D pour améliorer le démarchage et la fidélisation.

Informatica 9.5 permet également de réduire le coût des données de bien des façons, en diminuant les dépenses liées à l'usage, à la main d'œuvre, aux logiciels, aux matériels et au stockage. Hadoop MapReduce, par exemple, accélère l'adoption de plates-formes informatiques à faible coût pour réduire les dépenses matérielles. Enfin, Informatica 9.5 vous permet de concrétiser la promesse des données Big Data en proposant de les traiter de manière innovante. L'environnement de développement visuel unifié sur Hadoop, par exemple, offre aux utilisateurs ne disposant d'aucune expertise Hadoop la possibilité de procéder à un déploiement Hadoop en réutilisant les compétences et ressources existantes pour en accélérer l'adoption.

Toutes les nouvelles fonctionnalités d'Informatica 9.5 sont disponibles dans la même plate-forme Informatica, ce qui vous permet de réutiliser vos compétences, procédures opérationnelles et processus de gestion existants et de vous consacrer à ces nouveaux défis en matière de gestion des données sans complications technologiques supplémentaires.

Fort de 19 années d'expérience en matière d'innovation et de leadership dans le domaine de l'intégration de données, Informatica est le seul à pouvoir vous aider à optimiser votre retour sur les données, même dans ce monde du Big Data où tout évolue rapidement. Grâce à Informatica 9.5, soyez acteur et non spectateur de cette nouvelle ère. Vous pouvez :

- Optimiser la valeur de vos données
- Réduire le coût des données
- Concrétiser la promesse du Big Data



« Plus de 85 pour cent des entreprises figurant dans le classement Fortune 500 ne parviendront pas à exploiter le Big Data et à accroître leur avantage concurrentiel, et ce jusqu'en 2015. La majorité des entreprises ne sont pas préparées à faire face aux problématiques techniques et de gestion que pose l'importante volumétrie des données ; peu seront donc en mesure de faire de cette tendance un avantage concurrentiel. »

— Gartner

### APERÇU CLIENT : T-MOBILE

Le fournisseur de télécommunications T-Mobile utilise la plate-forme Informatica pour intégrer des données issues d'une architecture fédérée disparate, notamment une implémentation Hadoop qui prend en charge l'analyse avancée des pertes de client d'après les enregistrements détaillés des appels, les logs Web, les données de facturation, les informations issues des réseaux sociaux, etc. En combinant les données Big Data de transaction et d'interaction, T-Mobile obtient une vue plus exhaustive de ses clients et des raisons motivant leur départ à la concurrence. L'entreprise a ainsi pu diviser par deux le nombre de ces désabonnements en un seul trimestre.

## Optimisation de la valeur des données

Informatica 9.5 a été conçue pour aider votre entreprise à renforcer la valeur des données sur huit aspects, allant de la pertinence à la fiabilité.

VALEUR	FONCTIONNALITÉ
<b>En temps voulu</b>	Le streaming des données traite les grands volumes de données de transaction en temps réel pour maintenir un avantage concurrentiel
<b>Exploitable</b>	Un contrôle proactif génère des alertes pour optimiser les opportunités et réduire les risques
<b>Accessibles</b>	Le système Hadoop met également à disposition de l'entreprise tous types de données complexes
<b>Pertinentes</b>	Les services d'intégration dans le Cloud garantissent que les applications disposent de toutes les données pertinentes
<b>Complètes</b>	La virtualisation des données intègre les informations et les grands volumes de données existants de l'entreprise pour offrir de nouvelles perspectives
<b>Sécurisées</b>	La confidentialité globale de vos données sensibles est assurée, quel que soit leur emplacement, afin de réduire les risques
<b>Fiables</b>	La gouvernance des données permet de disposer d'informations cohérentes et précises inspirant la confiance des utilisateurs
<b>Irréfutables</b>	L'historique des données permet un suivi des modifications apportées sur une certaine période afin de fournir une vue en 4D des clients, partenaires et produits

### Streaming des données | En temps voulu

Le nombre d'appareils en ligne ne cesse d'augmenter et les entreprises doivent faire face à une croissance exponentielle des événements en temps réel et du traitement des données. Les services financiers, le secteur industriel, le secteur de l'énergie, des transports et de la chaîne logistique sont autant de secteurs pouvant bénéficier du streaming de données issu d'appareils, d'applications, d'entrepôts de données et de capteurs, ainsi que des systèmes transactionnels avec vélocité. Grâce aux fonctionnalités de streaming de données que propose Informatica 9.5, les entreprises disposent d'une architecture de messagerie sans intermédiaire combinée à une réplique des données en temps réel avec capture différentielle permettant de recueillir et de distribuer les données en streaming.

### Surveillance proactive | Exploitable

Les volumes et la complexité des données sont en constante augmentation, tout comme les risques d'erreur qui peuvent impacter les opérations et les niveaux de service contractuels. La surveillance proactive rassemble des données en temps réel et historiques, identifie les centres d'intérêt et génère des notifications aux individus, groupes, tableaux de bord ou autres systèmes afin d'apporter rapidement une réponse. Grâce à ces fonctionnalités, les entreprises optimisent leurs revenus et limitent l'impact des problèmes dans de nombreux domaines allant de l'informatique au service client.

### Système Hadoop | Accessibles

Hadoop étant intégré dans leur environnement, de nombreuses entreprises doivent recourir à des scripts coûteux et fastidieux pour traiter les données et les intégrer à l'aide d'applications traditionnelles. Informatica 9.5 propose le premier environnement de transformation de données optimisé pour Hadoop, capable de convertir des données complexes ou non structurées dans des formats relationnels et inversement. Grâce à Hadoop, les développeurs exploitent la valeur des types de données complexes pour Hadoop — données d'évènement, enregistrements détaillés des appels (CDR), réseaux, applications, logs Web, etc.

### Cloud Computing | Pertinentes

L'expertise et les efforts des fournisseurs d'applications cloud doivent être axés sur l'application proprement dite, plutôt que sur la manière d'intégrer les données dans leur application. Informatica propose des services d'intégration de données dans le Cloud qui peuvent être utilisés de manière transparente dans les applications cloud. Ils gèrent les difficultés liées à l'intégration de données en s'appuyant sur une connectivité prête à l'emploi et des modèles prédéfinis et permettent ainsi aux fournisseurs d'applications cloud de se consacrer à leur cœur de métier tout en fournissant des informations pertinentes aux utilisateurs métiers.



## Virtualisation des données | Complètes

Malgré une demande croissante de données actualisées et complètes, les entreprises connaissent des retards de livraison des données en raison d'infrastructures complexes et de systèmes multiples. La virtualisation de données fournit une couche d'abstraction permettant de dissimuler la complexité des sources sous-jacentes ainsi qu'un environnement fédéré et global pour l'intégration des données. Grâce à des services de données réutilisables à travers les systèmes, la virtualisation de données accélère la Business Intelligence (BI), répond aux services commercial et marketing en temps voulu et comble l'écart entre Hadoop et les systèmes traditionnels.

## Confidentialité globale des données | Sécurisées

Avec la prolifération des données aussi bien dans les systèmes traditionnels que dans les nouveaux environnements offshore, Cloud et Hadoop, il est désormais primordial de protéger les informations sensibles. La confidentialité globale des données permet aux entreprises de garantir le respect des règles de confidentialité et de réduire les risques de détournement ou d'accès non autorisé aux données de production ou hors production. L'identification automatique des données sensibles, une console unifiée pour la définition des politiques de confidentialité ainsi que des rapports d'audits des données masquées pour les systèmes de production et hors production accélèrent les délais de déploiement des applications et garantissent la conformité.

## Gouvernance de données complètes | Fiables

Le Big Data peut compliquer la gouvernance et introduire de nouveaux éléments de complexité dans les processus de gestion des données qui peuvent s'avérer difficiles à contrôler et à gérer. La gouvernance de données complètes offre aux experts techniques la visibilité et le contrôle des données et métadonnées en leur permettant de bénéficier d'une qualité de données réutilisables, de confidentialité des informations et de règles de conservation ainsi que d'une solution unique capable de fournir des informations précises, cohérentes, sécurisées et pertinentes. Les workflows de prise en charge de la qualité des données, les tableaux de bord en matière de gouvernance de données, la traçabilité des données et l'analyse d'impact rationalisent la gouvernance alors que les métadonnées sémantiques favorisent la collaboration entre les équipes métiers et informatiques.

## Historique des données | Irréfutables

L'entreprise évolue, tout comme les données. Les principales entités métiers doivent impérativement pouvoir suivre et anticiper ces changements pour gérer efficacement les opérations et limiter les risques. Informatica 9.5 propose l'historique des données permettant d'effectuer un contrôle des versions pour consulter et gérer les versions des entités et des hiérarchies à un moment précis — passé, présent ou futur. Elle offre également une visibilité de toutes les modifications apportées chronologiquement, notamment des informations relatives à l'action exécutée, à la personne qui en est à l'origine et aux données impactées.

## APERÇU CLIENT : US XPRESS

Dans le cadre du programme « No Data Left Behind », pour ne négliger aucune donnée, la société de transport US Xpress a collecté 900 éléments de données provenant de dizaines de milliers de systèmes de transport — données des capteurs d'utilisation des pneus, de consommation du carburant et de fonctionnement des moteurs, données géospatiales pour bénéficier d'un suivi de la flotte et plaintes consignées sur les blogs des chauffeurs. À l'aide d'Hadoop et d'Informatica, US Xpress traite et analyse ces grands volumes de données pour optimiser l'utilisation de sa flotte, limiter les temps d'immobilisation et la consommation de carburant, et ainsi économiser plusieurs millions de dollars par an.

## Réduction du coût des données

L'éventail des technologies pour la gestion des données d'Informatica peut aider votre entreprise à réduire les coûts liés à l'usage, à la main d'œuvre, aux logiciels, aux matériels et au stockage.

AXE	FONCTIONNALITÉ
<b>Coûts d'usage</b>	La validation de données empêche de prendre de mauvaises décisions en évitant les données de mauvaise qualité
<b>Coûts de la main d'œuvre</b>	L'identification des données accroît la productivité grâce à l'automatisation de la détection des entités métiers et des données sensibles
<b>Coûts logiciels</b>	La plate-forme informatique hybride vous permet de bénéficier des avantages économiques du Cloud tout en gardant le contrôle de toutes vos données
<b>Coûts matériels</b>	Hadoop accélère l'adoption de plates-formes à faible coût pour réduire les dépenses matérielles
<b>Coûts de stockage</b>	Les partitions intelligentes diminuent les coûts d'infrastructure pour Hadoop, les appliances et les données relationnelles

### Validation de données | Coûts d'usage

Des données incohérentes et inexactes peuvent être à l'origine de mauvaises décisions et d'opportunités manquées pouvant impacter les résultats, sans oublier les coûts inutiles et la perte de revenus qui risquent d'en découler. La validation de données favorise la rentabilité et réduit les coûts d'usage en garantissant des données fiables et complètes pour une meilleure prise de décision. Les développeurs peuvent facilement créer des règles efficaces pour automatiser les tests et la validation de données. Informatica 9.5 offre des fonctionnalités optimisées pour rapprocher les mises à jour de données avant leur chargement dans les systèmes de production, éliminant ainsi les données de mauvaise qualité tout en garantissant que les informations répondent aux besoins métiers.

### Identification de données | Coûts de la main d'œuvre

Les sources et la complexité des données ne cessent de croître et les entreprises sont souvent contraintes de recourir à l'identification et au profilage des données, via un processus manuel et fastidieux, afin de comprendre les informations de l'entreprise et leur pertinence pour son activité. La technologie d'identification des données d'entreprise proposée par Informatica 9.5 automatise le processus pour les structures de données et métadonnées, les relations entre les données et les domaines à travers des milliers de tables, accélérant ainsi la valeur tout en réduisant la codification manuelle coûteuse. L'identification des données met en évidence les anomalies dans les informations structurées et non structurées et met en évidence les dépendances, réduisant de ce fait les risques.

### Plate-forme informatique hybride | Coûts logiciels

Les avantages économiques du Cloud Computing peuvent être réduits à néant si les entreprises doivent gérer séparément les données des systèmes sur site et celles dans le Cloud avec des logiciels et des ressources redondants — ou pire encore, si elles se risquent à créer de nouveaux silos de données incohérentes. Informatica offre une plate-forme informatique hybride unique permettant d'intégrer, de nettoyer et de gérer les informations en toute transparence à travers un ensemble d'environnements sur site et dans le Cloud. Les entreprises peuvent ainsi gouverner leurs données et capitaliser pleinement sur les avantages-coûts du modèle offert par le Cloud.

### Système Hadoop et MapReduce | Coûts matériels

L'infrastructure du réseau informatique traditionnelle est coûteuse et les grids peuvent accroître les coûts réseau dans la mesure où ils peuvent atteindre des téraoctets, voire même des pétaoctets de données. Informatica 9.5 ouvre la voie à l'adoption des systèmes à faible coût Hadoop et MapReduce grâce à sa capacité à profiler, traiter et transformer les données en mode natif dans Hadoop et à convertir automatiquement les règles et les mappings en langage MapReduce. Les entreprises peuvent supprimer les investissements matériels inutiles et optimiser la valeur de l'infrastructure distribuée Hadoop.

## Partitions intelligentes | Coûts de stockage

Une gestion efficace de la croissance des données est essentielle pour maîtriser les coûts de stockage et améliorer les performances dans la mesure où les données Big Data se multiplient dans les environnements traditionnels, Hadoop et Cloud. La fonctionnalité de partitions intelligentes, disponible dans Informatica 9.5, fait partie d'une solution complète de gestion du cycle de vie des données visant à automatiser les stratégies complexes de partitionnement, d'application et d'archivage de bases de données. Les administrateurs de bases de données et d'applications peuvent créer des règles cohérentes en matière de cycle de vie des données pour centraliser la gouvernance du partitionnement et de l'archivage, permettre de réduire les coûts de stockage, d'infrastructure et de maintenance et veiller au respect des contrats de niveau de service.

## Concrétiser la promesse du Big Data

En s'appuyant sur Informatica 9.5, votre entreprise peut concrétiser la promesse des données Big Data pour dynamiser votre activité de manière plus rapide et plus intelligente.

AXE	FONCTIONNALITÉ
<b>Big Data de transaction</b>	La réplication de données livre des données de transaction au traitement analytique Hadoop pour réduire les coûts de traitement
<b>Big Data d'interaction</b>	La gestion des données de référence (MDM) intègre les profils clients issus des réseaux sociaux aux données existantes afin d'améliorer les processus d'acquisition et de fidélisation des clients
<b>Traitement des données Big Data</b>	Hadoop exploite en natif le potentiel d'Hadoop pour accélérer la valeur métier
<b>Traitement des données Big Data</b>	L'environnement de développement visuel unifié pour Hadoop permet de créer une logique métier des données sans ressources Hadoop spécialisées

## Data Replication | Big Data de transaction

La réplication des données en temps réel, parmi les systèmes traditionnels et les nouveaux systèmes de grands volumes de données, constitue l'élément clé pour exploiter pleinement les informations de l'entreprise. Informatica 9.5 offre une plate-forme unique pour tous les modèles de réplication de données et des capacités de data warehouse et d'analyse en temps réel dans des environnements hybrides pouvant intégrer Hadoop, des plates-formes de Cloud et des appliances. Grâce à l'utilisation de la réplication universelle pour livrer des volumes importants de données de transaction, Hadoop peut fonctionner comme point d'agrégation des informations, réduisant la taille et les coûts de stockage des data warehouses.

## Gestion des données de référence et réseaux sociaux | Big Data d'interaction

Les interactions des consommateurs sur les réseaux sociaux ont donné une dimension supplémentaire à la recherche d'une vue unique du client pour accompagner les activités de vente et de marketing. Informatica 9.5 fournit des fonctionnalités de MDM visant à intégrer les données des profils clients issues des réseaux sociaux dans les données existantes. Les entreprises cernent ainsi les préférences et le comportement de leurs clients et bénéficient d'informations uniquement disponibles sur ces réseaux. Grâce à l'association avec l'historique des données, Informatica 9.5 est en mesure d'offrir une vue du client en 4D pour attirer de nouveaux clients, consolider la relation client et augmenter sa part de marché.

### APERÇU CLIENT : HEALTHNOW

HealthNow, un assureur et fournisseur de services de santé du nord-ouest de l'État de New York, utilise les fonctionnalités de visualisation des données d'Informatica pour rationaliser une architecture orientée services qui englobe 16 sources d'entreprise. Cela lui permet de réduire la complexité de son infrastructure de données et de bénéficier d'une vue unique et fiable. Sa réussite dans ce domaine lui a permis d'exploiter les grands volumes de données, notamment les données issues des réseaux sociaux et les données non structurées émanant des demandes d'indemnisation et des certificats médicaux, pour prendre en charge une analyse prédictive des problèmes de santé.



« Westpac commence à exploiter les données issues des réseaux sociaux pour renforcer l'efficacité des campagnes marketing en vue d'augmenter la fidélisation et l'acquisition des clients. Nous sommes ravis des nouvelles fonctionnalités qu'offre Informatica 9.5 grâce auxquelles nous gérons et nous traitons les données issues des réseaux sociaux avec Hadoop et MDM, afin de bénéficier d'une vue en 4D de nos clients. Nous sommes ainsi en mesure d'optimiser notre retour sur les données Big Data. »

– Torrance Mayberry, Responsable des systèmes Informations de gestion, Westpac Information Technology

### Transformations natives Hadoop | Traitement des données Big Data

Personnaliser et adapter la logique de transformation des données sur un cluster Hadoop est un processus long, coûteux et source de difficultés pour les développeurs Hadoop. Informatica 9.5 leur permet de bénéficier de transformations en mode natif prêtes à l'emploi pour Hadoop, notamment de capacités de résolution d'identité et d'analyse, de profilage, de rapprochement et de nettoyage des données. Grâce à ces innovations, les entreprises peuvent réduire les coûts de développement et libérer rapidement des données à des fins d'utilisation inter-entreprises et valoriser ainsi toutes les informations au maximum.

### Environnement de développement visuel unifié pour Hadoop | Traitement des données Big Data

La mise en œuvre du traitement des grands volumes de données sur Hadoop peut nécessiter des ressources de développement très spécifiques et très onéreuses. L'environnement de développement visuel unifié pour Hadoop proposé par Informatica 9.5 permet aux développeurs Informatica de mettre facilement à profit leurs compétences, sans formation particulière. Grâce à cette technologie, ils peuvent utiliser l'outil de développement intuitif Informatica qu'ils connaissent pour définir la logique métier, les transformations et les mappings, ce qui permet de réduire les coûts, d'optimiser l'utilisation des ressources et de libérer toute la puissance des données Big Data.

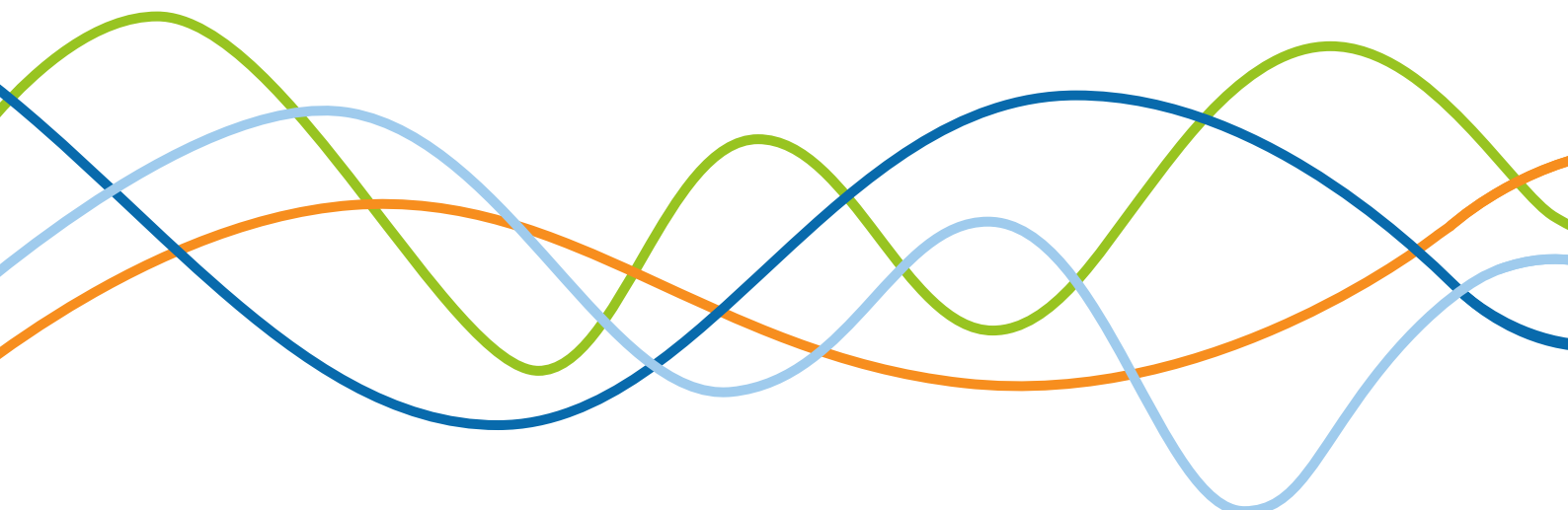
## Conclusion

Le changement est inévitable. Le Big Data, alimenté par les grandes tendances que sont le Cloud, les réseaux sociaux et la mobilité, révolutionne les processus métiers, l'infrastructure informatique et la manière dont nous recueillons et utilisons les informations générées par l'entreprise, les clients et les machines. Aucune certitude, par contre, quant à la capacité des entreprises à s'adapter aux nouvelles réalités du Big Data.

Pour réussir, votre entreprise doit s'appuyer sur une évaluation stratégique de vos défis et opportunités en matière de données ainsi que sur une transition vers une infrastructure de données de pointe, suffisamment souple et évolutive pour s'adapter à l'évolution des besoins métiers. Grâce à de nouvelles fonctionnalités et à une intégration plus poussée, Informatica 9.5 permet à votre entreprise de bénéficier de technologies innovantes pour apporter une plus grande valeur ajoutée, réduire les coûts et capitaliser sur la promesse du Big Data.

## À PROPOS D'INFORMATICA

Informatica Corporation (NASDAQ : INFA) est le leader des fournisseurs indépendants de solutions d'intégration de données. Les sociétés du monde entier font confiance à Informatica pour optimiser le retour sur les données et répondre à leurs principaux impératifs métiers. Plus de 4 630 entreprises dans le monde s'appuient sur Informatica pour tirer pleinement profit de leurs ressources en matière d'informations hébergées sur site, dans le Cloud et sur les réseaux sociaux.



**INFORMATICA®**

Siège mondial, 100 Cardinal Way, Redwood City, CA 94063, États-Unis  
France : Téléphone : 01 42 04 89 00 - [info\\_fr@informatica.com](mailto:info_fr@informatica.com) - [@InformaticaFr](#)  
[informatica.com](http://informatica.com) [linkedin.com/company/informatica](https://www.linkedin.com/company/informatica) [twitter.com/InformaticaCorp](https://twitter.com/InformaticaCorp)

© 2012 Informatica Corporation. Tous droits réservés. Imprimé aux États-Unis. Informatica, le logo Informatica et The Data Integration Company sont des marques commerciales ou déposées appartenant à Informatica Corporation aux États-Unis et dans d'autres pays. Tous les autres noms de sociétés et de produits sont la propriété de leurs détenteurs respectifs et peuvent avoir fait l'objet d'un dépôt de marque.

IN09\_0412\_01942FR