

Informatica Data Quality et l'intégration de données

Comment assurer une qualité de données répondant aux besoins métiers, à l'échelle de l'entreprise

Cycles symbiotiques de profiling, d'intégration et de qualité de données

(poster)
TDWI, 2006

« Le profiling, la qualité et l'intégration de données sont trois pratiques indissociables, un peu comme le pain, le beurre et la confiture . . . Les professionnels de la gestion de données et leurs homologues métiers doivent coordonner leurs efforts et concevoir des projets intégrant efficacement ces trois pratiques. »

Malgré près de 40 ans d'avancées dans le domaine des technologies de l'information, ce n'est que maintenant que certaines entreprises comprennent la véritable valeur de leurs données. De plus en plus dépendantes de processus métiers automatisés, ces entreprises sont aussi confrontées à des réglementations de plus en plus strictes et à une concurrence toujours plus acharnée. Dans ce contexte, elles prennent conscience qu'elles doivent se concentrer sur les données qu'elles utilisent au quotidien pour améliorer leur service client, être en conformité avec les réglementations gouvernementales et en phase avec les attentes de leurs clients, et rationaliser leurs opérations internationales.

Les données doivent être facilement accessibles dans toute l'entreprise et certifiées exactes

Pour supporter les processus métiers actuels et atteindre les objectifs fixés, toutes les données de l'entreprise doivent être accessibles, flexibles, certifiées exactes et réutilisables. Les entreprises ont besoin d'en savoir davantage sur le contenu de leurs systèmes sources. Elles doivent pouvoir intégrer les données de plusieurs systèmes dans de nouvelles applications plus productives, à même de gérer un volume très important de données. Elles doivent pouvoir nettoyer et optimiser leurs données, et aussi de contrôler et gérer la qualité des données utilisées dans leurs différentes applications.

Les technologies de qualité et d'intégration de données ont évolué pour répondre aux besoins d'organisations devant traiter des données fragmentées et défectueuses. Ces deux technologies supportent indiscutablement des activités symbiotiques, mais où se termine l'une et où commence l'autre ? La réponse est que les deux fonctions doivent travailler ensemble de façon transparente.

BÉNÉFICES :

- Connectivité universelle pour profiler et nettoyer les données de tous les types de systèmes
- Performances et capacités de montée en charge permettant de déployer la qualité de données à l'échelle de l'entreprise
- Haute disponibilité et tolérance aux pannes garantissant la continuité des opérations de qualité de données portant sur de gros volumes
- Analyse partagée des métadonnées et réutilisation des règles métiers pour la qualité de données

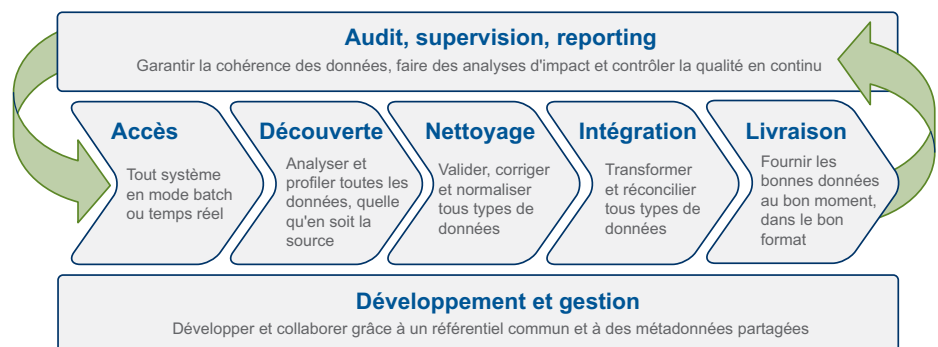


Figure 1 : La qualité et l'intégration de données supportent des activités intrinsèquement symbiotiques.

INTÉGRATION ET QUALITÉ DE DONNÉES : RENFORCER LA COLLABORATION ENTRE LES ÉQUIPES MÉTIERS ET INFORMATIQUES

La technologie d'intégration de données d'Informatica offre une connectivité avec presque tous les systèmes gérant des données structurées (bases de données relationnelles, applications d'entreprise, systèmes mainframe...), des données non-structurées (documents Word et Excel, e-mails, fichiers binaires...) et des données semi-structurées (XML et normes sectorielles telles qu'EDI, HIPAA, ACORD, FIX, SWIFT). La plate-forme d'Informatica automatise le cycle d'accès, de découverte et d'intégration de tous les types et formats de données historiques ou opérationnelles, dans la totalité de l'entreprise. En outre, cette plate-forme

réduit significativement les coûts et les risques d'implémentation en maximisant la réutilisation des composants.

Informatica, véritable référence dans le domaine de l'intégration de données, s'impose également comme un fournisseur leader de produits et de services de qualité de données. Dans le cadre de la qualité de données, l'analyse et le nettoyage sont devenus des activités orientées métiers, où la technologie de profilage sert à révéler les problèmes de qualité de données cachés, les écarts, les incohérences et les incompatibilités au sein des sources de données. Une fois les sources de données profilées, les résultats obtenus

sont utilisés pour concevoir et automatiser des stratégies de nettoyage visant à améliorer, enrichir et contrôler les contenus en fonction des règles métiers et des données de référence choisies. Les solutions de qualité et d'intégration de données d'Informatica permettent aux équipes métiers et informatiques de travailler efficacement ensemble pour valoriser au mieux le capital d'information de leur organisation. Pour l'entreprise, cela se traduit par de multiples avantages : prises de décision instantanées, visibilité à travers l'ensemble des divisions opérationnelles, synchronisation des opérations critiques et transparence en matière de conformité réglementaire.

Accès, découverte, nettoyage, intégration, livraison

Chez Informatica, nous pensons que la qualité de données est au cœur de toute initiative d'intégration de données. Avant que la plate-forme d'intégration de données puisse alimenter les data warehouses, les applications analytiques, les systèmes ERP ou CRM de l'entreprise, les données doivent être analysées, puis nettoyées. Avant toute intégration, l'analyse et le profilage des données sont des étapes clés du processus de planification et contribuent significativement à accélérer le développement des mappings et des workflows d'intégration de données. Le profilage initial aide l'entreprise à identifier et à comprendre ses données sources et, in fine, à réconcilier ces données avec les systèmes cibles.

Mais la qualité de données n'est pas une question qu'on traite une fois pour toute. Au fil du temps, la qualité des données se dégrade ou dégénère. C'est pourquoi, à l'instar de l'intégration de données, la qualité de données est un processus continu qui ne se limite pas à la recherche et à la correction des données manquantes ou inexacts. Ce processus a pour finalité de fournir à l'ensemble de l'entreprise des données complètes, cohérentes, pertinentes et actualisées, quelles que soient les applications, l'utilisation et la provenance des informations.

Qualité de données à l'échelle de l'entreprise rime avec intégration de données

La qualité de données s'étendant désormais à toute l'entreprise, la frontière entre qualité de données et intégration de données devient encore plus floue. De même que la plate-forme d'intégration de données profite de la qualité de données, la qualité de données à l'échelle de l'entreprise exploite et tire avantage des technologies d'intégration de données. Les initiatives de qualité de données sont de plus en plus étendues et plus profondément imbriquées dans les systèmes et les processus opérationnels. Elles nécessitent de ce fait une plate-forme d'intégration de données sous-jacente à même d'offrir les performances et la capacité de montée en charge permettant à l'entreprise de valoriser pleinement son capital d'informations.

Informatica au service de la gestion de qualité de données à l'échelle de l'entreprise

Informatica propose une plate-forme de qualité et d'intégration de données permettant à votre organisation de mettre en œuvre une approche complète pour gérer la qualité des données de bout en bout, tant en amont qu'en aval et à toutes les étapes de leur cycle de vie. Cette plate-forme d'entreprise comprend les solutions avancées de profilage, de qualité et d'intégration de données Informatica PowerExchange®, Informatica Data Explorer™, Informatica Data Quality™ et Informatica PowerCenter®.

Les fonctionnalités de la plate-forme d'intégration et de qualité de données d'Informatica vont bien au-delà de celles des produits autonomes disponibles sur le marché. La plate-forme Informatica offre notamment :

- la possibilité d'accéder aux données de pratiquement tous les systèmes, applications et bases de données et de livrer des données à ces mêmes environnements
- la performance, l'évolutivité, la haute disponibilité et la tolérance aux pannes nécessaires pour assurer le traitement continu de gros volumes de données
- un environnement de métadonnées partagé, facilitant l'analyse et la réutilisation des métadonnées

La plate-forme de qualité et d'intégration de données Informatica

Des entreprises de toutes tailles s'appuient sur les technologies d'intégration et de qualité de données d'Informatica pour exploiter toutes leurs données et renforcer la confiance des utilisateurs dans ces données. La suite d'intégration de données donne accès aux données de presque tous les systèmes métiers, quel qu'en soit le format, et permet de les intégrer puis de les distribuer, en temps voulu et selon la fréquence requise, à travers toute l'entreprise. Les solutions de qualité de données prennent en charge le profilage et le nettoyage des données permettant de certifier l'exactitude des données distribuées. Les solutions de la plate-forme Informatica travaillent de concert aux différents stades du processus d'intégration et de qualité de données pour fournir à toute l'entreprise des données robustes et de bonne qualité à partir d'un environnement unifié et intégré.

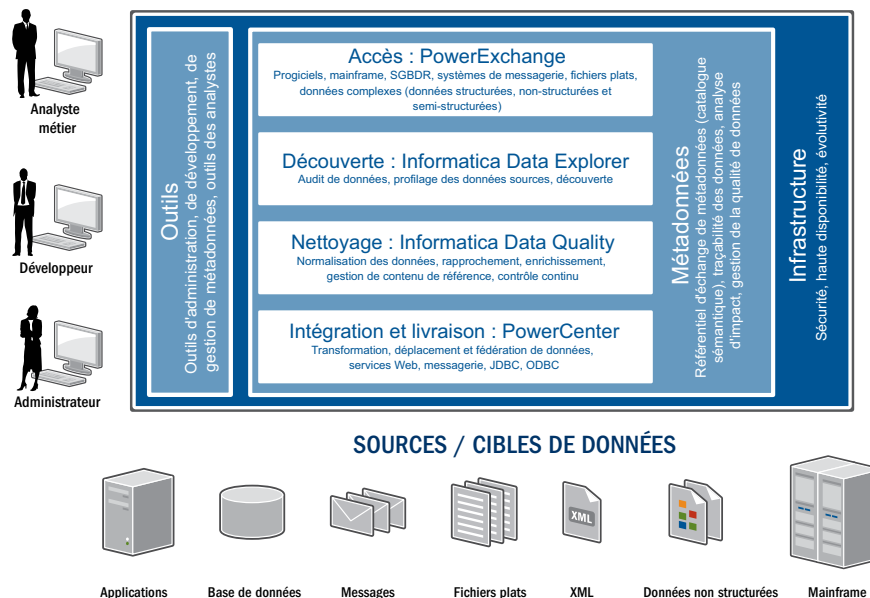


Figure 2 : La plate-forme Informatica est une solution de qualité et d'intégration de données d'entreprise permettant aux organisations de mettre en œuvre une approche complète pour gérer la qualité de données.

Accès

Informatica PowerExchange

Informatica PowerExchange automatise l'accès et la transformation de quasiment toutes les formes de données pour les rendre disponibles, pertinentes et utilisables dans toute l'entreprise. PowerExchange accède à tous les types de données électroniques : données structurées des systèmes mainframe, ERP et autres applications d'entreprise ; données basées sur des normes sectorielles spécifiques ; données non structurées comme celles des e-mails, des documents Word et autres applications bureautiques.

Découverte

Informatica Data Explorer

Informatica Data Explorer est une solution révolutionnaire qui fournit aux experts métiers et aux équipes informatiques des fonctionnalités de découverte et de profilage de données à la fois puissantes et simples à utiliser. Les analystes et les gestionnaires de données utilisent cette solution pour avoir une image complète et précise du contenu, de la qualité et de la structure des données de l'entreprise. Ces connaissances servent ensuite de base aux programmes de gestion de la qualité de données.

Nettoyage

Informatica Data Quality

Facile à utiliser, Informatica Data Quality est une solution spécifiquement conçue pour permettre aux équipes métiers de gérer et contrôler les processus de qualité de données. Elle regroupe de puissantes fonctionnalités de nettoyage, de rapprochement, et de contrôle de données et de reporting, donnant aux détenteurs d'informations métiers les moyens de mettre en œuvre et gérer des processus de qualité de données efficaces et durables, à l'échelle de l'entreprise. Informatica Data Quality comprend près de 50 composants facilement personnalisables pour analyser, normaliser et rapprocher les données. La souplesse de l'environnement de travail Informatica Data Quality Workbench permet aux analystes de données de construire des plans de qualité de données intégrant toutes les combinaisons de composants nécessaires pour atteindre une efficacité maximale.

Intégration et livraison

Informatica PowerCenter

Informatica PowerCenter est une plate-forme d'intégration de données unifiée permettant d'accéder, découvrir et intégrer les données quel qu'en soit le format, puis de les distribuer dans toute l'entreprise, à la vitesse souhaitée. PowerCenter aide les entreprises à tirer une valeur maximale de toutes leurs données, à rationaliser leurs activités et à développer leur chiffre d'affaires, tout en réduisant les coûts et la complexité de leurs systèmes informatiques.

Maximiser la valeur des données d'entreprise

Atteindre et maintenir un haut niveau de qualité de données est indispensable non seulement pour l'efficacité des opérations métiers et informatiques, mais aussi pour le succès des initiatives stratégiques et la compétitivité de votre entreprise à long terme. Longtemps, l'intégration et la qualité de données ont été associés aux projets de data warehouse, simplement parce que le déplacement des données des systèmes opérationnels vers le data warehouse était l'occasion d'affiner les données. La technologie d'intégration de données supportant ce processus est à présent utilisée pour d'autres types d'initiatives – migration, consolidation et synchronisation de données, gestion des données de référence, outsourcing... – qui sont autant d'opportunités d'amélioration continue de l'accès aux données et de leur qualité.

Les fonctionnalités de qualité de données d'Informatica permettent aux entreprises d'analyser, normaliser, valider, corriger, enrichir et superviser toutes les sources de données. Ces opérations renforcent l'intégrité et, par conséquent, la valeur du capital d'informations de l'entreprise et permettent de fournir aux utilisateurs métiers des données exactes et pertinentes pour eux. Grâce aux capacités d'intégration d'Informatica, ces fonctionnalités de qualité de données peuvent être déployées à l'échelle de l'entreprise en exploitant la connectivité, la gestion de métadonnées, la performance et l'évolutivité

linéaire qui caractérisent la suite de produits Informatica. Associés, ces produits fournissent l'infrastructure et les capacités de traitement nécessaires pour satisfaire les exigences de qualité de données des plus grandes entreprises.

La plate-forme Informatica protège également la qualité de ces processus en veillant à l'utilisation de définitions normalisées, au respect des normes de développement et à l'application des règles de sécurité d'accès. Ses fonctionnalités d'audit, de contrôle, de gestion de systèmes et de traitement des erreurs contribuent à la qualité globale des opérations.

Grâce à un ensemble complet de puissantes fonctionnalités de qualité de données, Informatica est à même de supporter de nombreux projets de qualité de données et d'en accélérer le cycle de vie, de façon à ce que toute l'entreprise ait confiance dans la cohérence, l'exactitude et l'actualité des données utilisées pour prendre des décisions métiers stratégiques. Ensemble, Informatica PowerExchange, Informatica Data Explorer, Informatica Data Quality et Informatica PowerCenter constituent la plate-forme idéale pour gérer la qualité de données à l'échelle de l'entreprise.

Pour plus d'informations à ce sujet, se reporter au livre blanc d'Informatica intitulé « Making Data Work: Addressing Data Quality at the Enterprise Level » (« Valorisation des données : comment parvenir à une véritable qualité de données à l'échelle de l'entreprise »).

Informatica Corporation propose des solutions et des services d'intégration de données destinés à résoudre un besoin auquel la plupart des grandes entreprises sont confrontées : la fragmentation des données dans de multiples systèmes hétérogènes. Informatica permet aux entreprises d'intégrer l'ensemble de leurs sources d'information afin d'en maximiser la valeur. Ouvertes et indépendantes des plates-formes d'exploitation, les solutions Informatica réduisent les coûts et accélèrent la réalisation de projets d'intégration de données, quelles qu'en soient la taille et la complexité. Depuis 1993, Informatica aide les entreprises et les organismes publics à maximiser la valeur de leur capital d'information en intégrant toutes leurs sources de données. C'est la raison pour laquelle la société Informatica est considérée comme un fournisseur de solutions d'intégration de données.



Siège social international, 100 Cardinal Way, Redwood City, CA 94063, USA

Téléphone : +1 650 .385 .5000 Télécopie : +1 650 .385 .5500 Appel gratuit depuis les États-Unis : +1 .800 .653 .3871 www.informatica.com/fr

Informatica dans le monde : Allemagne · Australie · Belgique · Canada · Chine · Corée · Espagne · États-Unis · France · Irlande · Italie · Japon · Pays-Bas · Portugal · Royaume-Uni · Singapour · Suisse

© 2008 Informatica Corporation. Tous droits réservés. Imprimé aux États-Unis. Informatica, le logo Informatica, PowerCenter, Informatica Data Explorer, Informatica Data Quality et PowerExchange sont des marques commerciales ou des marques ou des marques déposées d'Informatica Corporation aux États-Unis et dans d'autres pays. Tous les autres noms de sociétés et de produits cités sont la propriété de leurs détenteurs respectifs et peuvent avoir fait l'objet d'un dépôt de marque.

6787FR (26/08/2008)