



QUALITÉ DES DONNÉES

Définition, importance, solutions et mise en œuvre

Sommaire

Qu'est-ce que la qualité des données ?

Pourquoi la qualité des données est-elle nécessaire ?

Comment assurer la qualité des données en entreprise ?

Votre partenaire pour la qualification des données

Qu'est-ce que la qualité des données ?

La qualité des données, ou data quality, désigne les processus déployés pour maintenir la fiabilité et la cohérence des données détenues par une entreprise.

Elle regroupe l'ensemble des pratiques permettant l'identification, la correction et la gestion des erreurs dans les données. La data quality englobe également les protocoles de contrôle de qualité et la gouvernance des données.

La gouvernance des données est la capacité d'une organisation à assurer leur mise à jour et à maintenir leur qualité dans le temps par la mise en place de procédures.

L'objectif de la qualité des données est de produire des informations fiables à forte valeur ajoutée pour le développement de l'entreprise.

La qualité touche plusieurs dimensions



La complétude



La validité



La précision



La cohérence



La disponibilité



La mise à jour

Pourquoi la qualification des données est-elle nécessaire ?

La qualité est un enjeu crucial dans le processus de gestion des données d'entreprise. Elle permet de garantir un système d'information opérationnel pour les métiers et d'améliorer les performances économiques de la société.

Les données fiables évitent les erreurs stratégiques dommageables et améliorent l'image de l'entreprise. Les organisations qui ont le contrôle sur la qualité de leurs données sont plus performantes.

La non-qualité des données représente un coût important pour l'organisation. Elle a un impact direct sur la productivité des collaborateurs et la réputation de la société.

3 grands avantages à utiliser des données qualifiées en entreprises

Prendre de meilleures décisions et agir



L'objectif de la donnée est d'alimenter les différentes activités de l'entreprise et donc de développer du business. Sans data quality, les métiers sont amenés à agir en se basant sur des informations inexactes.

La qualité des données joue un rôle primordial dans la capacité des organisations à prendre les bonnes décisions et agir de manière éclairée. Une stratégie d'entreprise efficace ne peut donc pas être déployée à partir de données non qualifiées.

Le processus de qualification permet également de personnaliser les actions à destination des prospects et clients. La data quality impacte ainsi directement les performances des activités marketing et commerciales et bien d'autres.

Travailler en toute confiance et plus efficacement



Les données non qualifiées engendrent des coûts opérationnels importants et impactent la productivité globale des entreprises. Si une société utilise des données imprécises ou inexactes, les utilisateurs métier perdront du temps à trouver la bonne information.

La qualification et l'accessibilité des données augmentent l'efficacité et la productivité des métiers. Ces derniers peuvent travailler en toute confiance sans perdre de temps avec des valeurs erronées ou incomplètes. Ils peuvent se recentrer sur des tâches à forte valeur ajoutées.

Des données de qualité permettent également à l'entreprise d'anticiper les risques et d'être plus réactive en cas de changement important.



Limiter les risques et éviter les pertes

La non-qualité des données englobe l'inexactitude, l'incohérence, la non complétion, la non actualisation ou encore l'indisponibilité. Les causes les plus courantes de non-qualité sont liées à l'erreur humaine.

Les données non qualifiées représentent un risque pouvant parfois être très dommageable. Elles mènent à de mauvaises décisions, au déploiement de stratégies bancales et entraînent ainsi des coûts financiers pour l'entreprise.

La non-qualité peut donc porter directement atteinte à la réputation d'une société.

Le coût des données de mauvaise qualité

10x

Il revient 10 fois plus cher d'exécuter une tâche à partir de données erronées.

83%

83% des entreprises se disent affectées par la non-qualité des données.

15%
20%

Le coût des données de mauvaise qualité peut atteindre de 15% à 25% du CA total.

15M

La non-qualité des données à coûté 15 millions de dollars aux entreprises en 2017.

Comment assurer la qualité des données en entreprise ?

Les entreprises doivent maîtriser la qualité de leurs données pour pouvoir les convertir en connaissances exploitables, les exploiter efficacement et en tirer parti.

Il est impossible de réussir une stratégie de data quality sans outils appropriés face aux volumes de données produits quotidiennement par les entreprises.

Il existe aujourd'hui de nombreuses solutions sur le marché pour répondre aux problématiques de management de la qualité des données. Plusieurs d'entre elles sont disponibles en version open source et accessibles à tous les niveaux d'expertise.



A propos de Talend Data Quality

Talend Data Quality est une solution open source dédiée à l'évaluation de la précision et l'amélioration de la qualité des données d'entreprise.

Cet outil propose des fonctionnalités avancées telles que le nettoyage, la déduplication, la standardisation et l'enrichissement des données. Il peut être utilisé pour traiter n'importe quel format et volume de données.

Talend Data Quality permet aux entreprises de produire rapidement des données propres et fiables pour les utilisateurs métier sans utiliser le code.

Les fonctionnalités de Talend Data Quality



Nettoyage des données

Talend Data Quality propose des fonctionnalités de nettoyage intégrant la déduplication, la validation et l'élimination des incohérences.

Ce module recherche les valeurs manquantes de manière intelligente dans les bases et les met en évidence afin que l'entreprise puisse renseigner les données incomplètes rapidement.

Talend permet aux sociétés d'automatiser le nettoyage des données en appliquant des règles personnalisées basées sur du machine learning.

Enrichissement des données



Talend Data Quality propose la détection automatique de modèles prédéfinis et la mise en correspondance des données.

Les données de plusieurs systèmes sont fusionnées et consolidées avant leur intégration. Les valeurs manquantes sont complétées automatiquement au sein des bases de l'entreprise.

Cette solution permet également d'enrichir ses données en utilisant des bases internes ou sources externes (références publiques, adresses postales, numéro de téléphone etc...)

Profilage des données



Le profilage est un processus qui consiste à examiner les data pour en générer une vue d'ensemble exploitable. Cette analyse graphique facilite l'élimination des erreurs dans les bases.

L'établissement de profils permet d'évaluer la qualité des données par rapport à des seuils personnalisés ou des normes légales.



Masquage des données

Talend inclut une fonctionnalité de masquage qui répond à certaines exigences du Règlement Général sur la Protection des Données (RGPD).

Cette technique permet de masquer des informations personnelles sensibles détenues par l'entreprise lorsqu'elle doit effectuer des analyses ou des tests. L'identité et la vie privées des individus sont protégées, mais les données restent toujours accessibles et utilisables par les différents services.

Monitoring des données

Talend Data Quality inclut une fonctionnalité de monitoring des données permettant aux entreprises de surveiller la qualité en temps réel.

L'administration de l'outil propose des rapports et des tableaux de bord de suivi de la qualité des data. Elle permet d'automatiser le processus de nettoyage et de configurer des alertes pour une meilleure surveillance.



Gouvernance des données

La gouvernance regroupe les procédures permettant d'encadrer la collecte des données et leur utilisation au sein d'une organisation. Elle ne concerne pas seulement le service informatique, mais l'ensemble des métiers de l'entreprise.

La gouvernance des données a plusieurs dimensions : la disponibilité, l'utilisabilité, l'intégrité et la sécurité. Talend Data Quality permet aux entreprises d'impliquer les utilisateurs dans les processus de gouvernance au sein d'un environnement de travail intuitif.



4 Votre partenaire pour la qualification des données

Axysweb accompagne les entreprises dans leur stratégie de qualification des données. Notre équipe d'experts techniques et consultants est spécialisée dans l'intégration de la solution Talend Data Quality.

Nous vous conseillons sur les meilleures techniques de nettoyage et mettons en œuvre pour renforcer l'intégrité et la pertinence de vos données sur le long terme.

Axysweb assure aussi bien les missions de conseil, gestion et mise en œuvre de vos projets de qualification des données. Notre objectif est de permettre à nos clients de prendre des décisions avisées et de définir des stratégies efficaces.



Qui sont nos clients ?

La qualification des données est une problématique qui concerne les entreprises de toutes tailles et de tous les secteurs d'activité.

Axysweb travaille principalement avec la Direction des Systèmes d'Information et le service informatique au sein des PME et ETI (Entreprises de Taille Intermédiaire) en France.

Le plus souvent, ces entreprises n'ont pas encore développé de service dédié exclusivement à la gestion des données au contraire des grands groupes qui ont des postes de data scientists, chief data officer etc...

Merci !



www.axysweb.com



05 24 57 11 02



PLACE DES GRAND HOMMES
1 ER ETAGE
CS 22029
33001 BORDEAUX CEDEX

