



Les données constituent la pierre angulaire des SI

Etes-vous sûrs de vos données ? La plupart des études sur le sujet montrent des résultats inquiétants sur la question. En effet, les conséquences de l'inadaption des données aux objectifs de chaque métier de l'entreprise sont souvent importantes et induisent des gaspillages qui pourraient facilement être évités avec un minimum de rigueur.

Les données et leur contrôle sont un sujet sensible qui fait l'objet de plus en plus d'attention. Preuve en est, en trois ans, la CNIL a enregistré une augmentation de 570% de son activité. Ce phénomène est tout aussi important au sein de l'entreprise, particulièrement en raison de l'utilisation toujours plus importante de l'outil informatique au travail. De fait, la masse de données entrantes dans l'entreprise explose. Elles y entrent sous des formats et des attributs différents. Difficile alors d'avoir des données uniformes et homogènes.

Et, il en est de même pour la sécurité et la validité de ces données. Qui vérifie leur exactitude ou leur conformité ? Les implications sont nombreuses : réputation de la marque, acquisition et fidélisation de la clientèle, réduction des coûts, respect de la réglementation et augmentation des revenus sont les points particulièrement impactés par la qualité des données.

Une prise de conscience sans suite

Cela semble une évidence et pourtant, peu d'entreprises mettent en place des outils ou des politiques adaptés. Différentes études récentes mettent ce laxisme en exergue. En particulier, celle réalisée il ya tout juste un an par Dynamic Markets pour Experian QAS, un acteur du secteur de la qualité des données notamment dans le domaine de la gestion de la relation client. Si la prise de conscience est là – 96% des entreprises interrogées reconnaissent l'impact négatif d'une mauvaise gestion des données sur les revenus – les solutions mises en œuvre ne sont pas nombreuses et sont loin de répondre à l'enjeu. Si outre-Atlantique les choses bougent déjà depuis quelques temps, les entreprises hexagonales semblent moins réactives « *Le marché US est plus mature, mais cela commence à arriver en France. Il reste encore de l'évangélisation à faire. Un projet où l'intégration de données est primordiale fournit une bonne occasion pour se mettre à niveau* » constate Cédric Carbone directeur technique chez Talend.

L'occasion qui fait le larron ? Bruno Labidoire, directeur technique France chez Informatica, semble partager le même avis: « *La migration sur un nouvel ERP est une opportunité pour s'attaquer au projet, lorsqu'il s'agit de choisir une normalisation des données sur le système cible à partir de différents logiciels existants pour arriver à quelque chose de plus fédérateur.* »

La relation client en péril

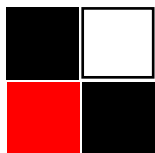
« *La prise de conscience est réelle. La plupart des entreprises n'ont cependant pas identifié qu'elles ont un problème de qualité de données, ni à quel point cela leur fait mal. Elles ne savent pas à quel point leurs données sont erronées, et ce qui est bon pour une entreprise peut être catastrophique pour une autre. L'aspect métier est donc très important* », ironise Marc Daverat, directeur général pour l'Europe du Sud chez Dataflux, un spin off spécialisé dans la gestion des données de l'éditeur de business intelligence SAS.

Jean-Marc Leclère, directeur commercial et des opérations pour la France chez [Uniserv](#), un acteur historique de la gestion de la qualité de la donnée, ajoute: « *L'émergence de l'intérêt pour la qualité des données survient du fait qu'elles ne proviennent pas de l'informatique mais des utilisateurs. Ils ont été les premiers à pointer les problèmes relatifs à la qualité des données alors que celles-ci sont en fait la matière première des informations sur lesquelles vous prenez des décisions. Or, les données restent en général incorrectes, incomplètes ou dupliquées. Elles sont aujourd'hui la cause la plus communément admise d'échec des projets de gestion de la relation client. Ce n'est pas le seul domaine atteint par ce phénomène. 60% des projets d'intégration ou d'agrégation de données, comme à la suite d'une fusion d'entreprise, échouent ou connaissent de graves dépassements en raison de ces problèmes.* »

Le marché de la qualité des données

En 2008, le marché de la qualité des données représentait, selon le Gartner Group, environ 300 millions de dollars. Il devrait doubler pour atteindre 677 millions de dollars en 2011 avec une croissance annuelle pondérée de 18%. Le marché est dominé par les éditeurs anglo-saxons et, en particulier, les fournisseurs d'infrastructures. Dans le domaine, seul Business Objects, racheté par SAP fait figure d'exception. L'objectif des outils actuels consiste à améliorer la qualité des données. De nombreuses solutions permettent de détecter des anomalies et de les corriger. Il s'agit donc, dans l'ensemble, d'un mode curatif, même si certains produits fonctionnent de manière préventive en détectant les anomalies à la source et en temps réel.

Pour toute information supplémentaire contactez



La moitié des entreprises sensibilisée

Seuls quelques projets aujourd'hui ont véritablement pour origine la qualité des données et ils sont destinés à résoudre des problèmes récurrents et graves. Ainsi, chez Reuters, l'outil de CRM mis en place peinait à apporter des retours sur investissement positifs. Plus que l'outil, le problème provenait de la qualité des données.

Seulement 46% des entreprises ont instauré une politique formelle de contrôle de la qualité des données. C'est déjà mieux que sur la vague précédente de l'enquête réalisée en 2005 (27 %). Il n'en reste pas moins que plus de la moitié des entreprises n'ont pas vraiment d'initiatives en la matière. L'impact est aussi significatif en termes d'argent perdu. A Singapour, près d'un tiers du budget des entreprises serait englouti par une mauvaise qualité des données. Pas plus de 8 % des entreprises entreprennent des vérifications exhaustives des données collectées, quand 34 % d'entre elles ne réalisent aucune vérification. Sur la conformité à la législation, les entreprises atteignent même des résultats inférieurs à ceux de la vague de 2005. Parmi les organisations interrogées, elles sont 27 % à se dire en conformité avec la législation pour les données présentes dans leurs bases, contre 37 % en 2005 ! Là encore le tableau se doit d'être modulé suivant les secteurs d'activité. Les bons élèves sont plutôt à rechercher dans la distribution, les organisations caritatives et les services financiers.

La conséquence principale est un énorme gaspillage financier. Les entreprises interrogées pensent qu'en moyenne 19 % des revenus de l'entreprise y sont ainsi consacrés. Elles n'étaient que 6 % penser cela en 2005. Les mauvais élèves se trouvent plutôt dans le secteur des transports... et dans le service public ! La France s'inscrit dans la moyenne des pays occidentaux avec 18 % de revenus perdus.

Des conséquences désastreuses pour l'entreprise

Fait sûrement le plus significatif et inquiétant, 17 % des entreprises interrogées ne savent même pas si une stratégie a été mise en place ou non. La taille de l'entreprise y est pour beaucoup.

53 % des grandes entreprises déclarent avoir une stratégie contre 38 % pour les PME. Il est à constater que la problématique concerne d'ailleurs plutôt les entreprises ayant différents systèmes sources avec des rafraîchissements d'informations fréquents.

Dans une tribune récente, Michel Bruley, directeur marketing chez Terradate pour l'Europe de l'Ouest, citait deux exemples parlant des conséquences d'une mauvaise qualité des données. Le premier provenait de l'industrie des télécommunications. La plupart des opérateurs perdent environ 5 % de leur chiffre d'affaires du fait d'une qualité insuffisante des données qui ne leur permet pas de toujours d'aller au bout de leur processus de facturation. Le second exemple est

la NASA qui a perdu un satellite d'une valeur de plusieurs centaines de millions de dollars, car l'équipe internationale qui avait travaillé sur ce projet n'avait pas utilisé la même unité de valeur pour l'un des paramètres de pilotage de l'engin. Armé de « fausses » indications sur sa position, l'ordinateur de bord a conduit, avec précision, le satellite à sa perte.

Une confiance aveugle dans les données collectées

Cela n'empêche pas les entreprises de faire confiance aux données qu'elles collectent, alors qu'elles ne réalisent globalement aucune vérification ni validation *a posteriori*. En tête de liste des données jugées exactes, on trouve le nom du client et ses coordonnées bancaires (91 %). Mais les adresses mails et la taille de l'entreprise descendent sous la barre des 75 % !

Les entreprises semblent se conduire comme un dégustateur de vin novice qui fait plus confiance à l'étiquette qu'à ce qui se trouve à l'intérieur de la bouteille. En ce sens, une étude menée par DataMonitor démontre que la qualité inadéquate des données a été identifiée comme le premier obstacle au succès des processus métiers, suivi par le manque de compréhension du [cycle de vie des données](#).

Cette incompréhension est présente dès le début de la démarche, et peu d'entreprises se posent même la question de savoir quels sont les critères pour disposer d'une donnée de qualité.

Les critères pour de bonnes données

Sans objectif, il n'existe pas de résultats. Il en va de même pour les données présentes dans le système d'information. Quels sont les critères retenus pour démontrer la qualité des données ? Ils ne sont pas légions ! Au plus ils sont au nombre de six.

Le premier est une évidence. La donnée doit être pertinente. C'est à dire qu'elle représente ce qu'elle doit être après avoir été comparée à une donnée de référence. Cette pertinence s'exprime de toute manière dans le contexte métier et peut donc évoluer d'une entreprise à l'autre, d'un service à un autre.

Le but est de trouver un critère de pertinence qui satisfasse tous les utilisateurs de cette donnée. Selon Jean Marc Leclère « *apporter des informations exactes est tout de même le but* » !

Cédric Carbone va dans le même sens : « *Les données doivent refléter la vérité.* »

Deuxième point les données doivent être complètes. Un exemple simple : pour le service commercial, le plus important est probablement l'adresse postale ou les coordonnées bancaires d'un client. Dans la même entreprise, le service marketing aura lui plutôt besoin du mail, et le service après-vente nécessite généralement de connaître les données issues de la direction technique pour

ses interventions. Si les données sur un client ne sont pas complètes, plusieurs services de l'entreprise ne pourront réaliser efficacement leur tâche. « *Si les données ne sont pas complètes elles peuvent devenir inexploitable du fait même que des renseignements sont manquants* », ajoute Cédric Carbone.

En plus de la complétude, il faut veiller à la cohérence. Ainsi, il n'est pas rare de voir dans les entreprises des calculs de chiffre d'affaires dont l'addition ne correspond pas à la somme des différentes données issues des commerciaux de l'entreprise, par exemple.

La validité est un autre aspect de la qualité de la donnée. Par validité, on entend qu'elle corresponde aux règles de l'entreprise ou de la législation en vigueur. Ainsi s'il peut sembler logique d'enregistrer l'adresse d'un client ou son téléphone portable, il n'est pas forcément nécessaire de connaître sa religion ou son régime alimentaire, tout du moins, si l'activité de l'entreprise ne nécessite pas de posséder de telles informations. Il convient aussi de vérifier systématiquement la structure de la donnée et voir si celle-ci correspond bien aux règles en vigueur dans l'entreprise.

Dernier critère mais non le moindre, la donnée doit être unique. Dans les problèmes relatifs à la qualité des données, c'est certainement le critère le plus souvent cité certainement du fait que beaucoup d'outils sur le marché peuvent le traiter par [dédoublonnage/déduplication](#) (opération qui consiste à identifier et éliminer les doublons).

Les entreprises ont souvent pratiqué ces opérations sur leurs données clients ou postales. Jean-Marc Leclère constate que celles-ci « *recouvrent le plus facile de la démarche de qualité de données et sont nécessaires, mais loin d'être suffisantes* ». Les secteurs du marketing direct ou de la gestion de la relation client et de l'e-business en sont de grands adeptes !

Malgré des efforts, seules les grandes entreprises ou celles directement concernées par la problématique de qualité des données se sont lancées dans des projets d'amélioration du panel d'informations dans leurs bases de données. Les autres sont le plus souvent inconscientes du phénomène, perdant au passage des sommes importantes du fait de nombreux manques à gagner. Il reste difficile d'expliquer à un dirigeant que les données collectées pendant des années sont médiocres, voire inexploitable d'un point de vue métier. Comme souvent l'occasion fait le larron, et un projet d'intégration de données, comme l'installation d'un ERP ou une fusion-acquisition, peut servir de déclencheur. ■

Les impacts d'une mauvaise qualité des données

La mauvaise qualité des données peut avoir des impacts divers sur l'entreprise. Le premier concerne la satisfaction du client. Les clients « victimes » de données incorrectes pardonnent difficilement. Une livraison tardive suite à une erreur sur la date de livraison par exemple. Avec Internet, la sentence est encore plus rapide. En effet un internaute s'apercevant que les données d'un site sont peu fiables ou incorrectes – comme des prix non actualisés – n'y reviendra sûrement jamais. Le second impact touche la publication de chiffres incorrects. Des erreurs dans les données comptables entraînent généralement de lourdes sanctions (amendes, condamnations). Et ce, même si cette situation ne résulte pas d'une manipulation volontaire. Suite aux récentes affaires de falsification des comptes, toute erreur dans la publication des chiffres paraît désormais suspecte. Au-delà de cet aspect réglementaire, de telles erreurs ont pour conséquence de détériorer l'image de l'entreprise et de freiner son expansion commerciale. De plus, l'impact boursier des sociétés cotées est immédiat, pouvant entraîner la chute rapide du cours de l'action.

Autres répercussion : des données incorrectes peuvent ralentir le fonctionnement opérationnel au sein de l'entreprise et rendre les décisions difficiles ou arbitraires ; cela génère aussi du stress et de la démotivation chez les collaborateurs, et le surcoût de travail engendré par des erreurs de données n'est jamais le bienvenu ; enfin la perte de confiance dans les informations manipulées n'est propice ni à l'initiative ni à un climat sain.

Définir une stratégie nécessite plus de temps si les informations utilisées sont de mauvaise qualité ou tout simplement erronées. Comment réfléchir à l'avenir si la situation présente n'est pas maîtrisée ? Les dirigeants peuvent évidemment se fier à leur intuition, mais ce type de démarche montre rapidement ses limites.

Les impacts énoncés précédemment engendrent évidemment des surcoûts. Les situations variant d'une entreprise à une autre, il est très difficile d'estimer le coût de la non-qualité. Certains experts estiment cependant que ce coût caché peut représenter plus de 10 % du chiffre d'affaires. ce qui ampute de façon conséquente la rentabilité.