IMPORTANCE DE LA QUALITE DES DONNEES : ETUDE AUPRES DES MANAGERS MAROCAINS

Said MARBI , Mehdi ELMEJDOUB
l'EST Casablanca, Morocco

Abstract

À mesure que la technologie prend de plus en plus d'importance dans la vie quotidienne, elle cesse d'être un support pour l'entreprise pour devenir un pilier de son fonctionnement.
On peut dans un sens considérer que la dernière crise des Sub-Primes est due à une défaillance du contrôle interne des divers établissements concernés, à cause d'une mauvaise qualité de données car incomplète.
Cette situation n'est pas sans poser des problématiques au niveau de la gouvernance de manière générale.
Afin d'apporter des solutions, des normes ont été développées telles que ISO 9001 pour encadrer les processus de traitement de données à l'intérieur de l'entreprise. Le projet ISO 8000 a pour ambition de couvrir les risques d'information incomplète, partielle ou non validée.
Je me réfère dans cet article aux raisons qui ont conduit le comité ISO 8000, sous la direction de M.

IJSER

REMERCIEMENTS

Enormes remerciements au comité d'organisation de cette conférence Internationale , sans qui ce grand événement n'aurait pu avoir lieu. Plus particulièrement au Dr Salman Bourekkadi Salmane qui n'a pas ménagé ses efforts pour la réussite de l'événement. Que ce soit à travers le support qu'il nous a offert, le temps qu'il nous a dédié ainsi que ses précieuses orientations.

INTRODUCTION

Le développement naturel des échanges mondiaux, couplé à une élévation continue du niveau technologique des opérations (mise en place de suivi informatique de la production, déclaration généralisée de manière informatique aux impôts, réseaux sociaux..) causent une croissance exponentielle du volume des données traitées annuellement. IDC (une firme Américaine spécialisée dans les études et analyses de marché) estime que le volume d'informations produites chaque année est passé de près de 70 exocets (milliard de gigaoctet) en 2006 à plus de 600exaoctets en 2012. Cette tendance ne va qu'en s'accentuant car IDC estime que ce même chiffre passera à plus de 30 ZettaOctets (Trillion de Giga-octets) en 2020. De plus, les technologies qui étaient auparavant des supports efficaces, est en train progressivement de se transformer et devenir un pilier incontournable dans plusieurs domaines (militaire, santé, logistique …). On imagine alors qu'une fraction aussi faible soit- elle, de « mauvaises » données peuvent avoir des conséquences massives sur la productivité mondiale.

LE COUT DE LA NON QUALITE DES DONNEES

Une donnée incomplète, légèrement ou carrément erronée … peut avoir des conséquences lourdes sur une organisation. Selon une étude de AT&T Lans & Research, la piètre qualité des données cause plus de 600 Milliards de Dollars chaque année à l'économie Américaine (temps de correction des données, efforts correction des rapports, nouvelles décisions
rectifiant les anciennes …). Ces coûts sont :

- Les coûts directs : représentés par les dépenses directes en monnaie, matières ou temps dû à l'erreur.. Mais surtout le ralentissement de tout le processus de production
- Les coûts indirects : les frais de constatation de l'erreur, réorganisation pour palier à l'erreur, et les coûts

d'opportunité
- Les coûts cachés : difficiles à quantifier tels que des visites de commerciaux supplémentaires, les retards de livraison causant une image négative auprès de la clientèle, …

Voici quelques cause à l'origine de la mauvaise qualité des données :

- L'erreur humaine évidement est à l'origine d'une grande partie des erreurs (en 2005, le Tokyo Stock Exchange a perdu en une journée 286 Millions d'Euros à cause d'un opérateur qui devait placer 610000 titres à 1 yen au lieu de 1 titre à 610000 yens lors de l'introduction en bourse)
- Une donnée mal récoltée à cause d'un besoin mal exprimé de la pat du client de la donnée
- Une mauvaise synchronisation entre les différents clients de l'information (le parachute sur Mars d'une sonde spatiale n'a pas fonctionné correctement car une partie des paramètres de vol était en Km et l'autre en Miles)
- Une information mal produite à cause des biais cognitifs (une situation imprévue peut faire prendre à un haut manager des décisions irrationnelles)
- Détérioration des données lors des migrations d'une application à une autre (ce qui a causé d'économes pénalités à Airbus lors du développement de l'A380 dont la conception du câblage électrique a été menée par des partenaires différents sur des applications de versions différentes)
- Non application de référentiels réglementant les données : règles de validation, procédures précises de transfert… (il y a des ébauches nationales de normes relevant de la qualité de la donnée informatique, mais ce domaine est encore en cours de développement tel que l'ISO 8000)

UNE DEMARCHE QUALITEE DE PLUS EN PLUS MOTIVEE

Beaucoup de structures ont adopté récemment une démarche qualité relatives aux données pour plusieurs raisons :

- Recherche d'amélioration de la productivité (dans un de ses rapports, l'armée de Terre de Grande Bretagne estime de 10% l'écart existant entre ses divers stocks référencés et la réalité)
- Conformité réglementaire : les référentiels (tels que ISO 9001) et les procédures d'audit interviennent pour normer ce domaine
- Une fusion entre 2 structures (Lafarge et Holcim étaient deux entreprises importantes dans le domaine de la cimenterie, avec chacune un système d'information à part et un portefeuille de clients à part, ont dû revoir leur système d'information)
- La mise en place d'un nouvel ERP (la verrerie Walterpserger a mis en place un nouvel ERP et *introduit le concept de qualité de données en plus de celui classique de la production). Cette migration a décelé en effet beaucoup de pratiques informelles effectuées par les managers pour assurer le fonctionnement de la structure. Ces actions causaient malheureusement souvent des confusions, mauvaise coordination,…
- Profiter pleinement du potentiel de collaboration intra-groupe : ou bien Schneider Electric qui a décidé de remanier l'organiser des activités pour profiter d'effets de synergie entre les unités du groupe et se devaient donc d'harmoniser les données qui étaient stockées nationalement dans une cinquantaine d'ERP à travers le monde. Ce qui a mis en exergue les lacunes de chacun de ces systèmes de données.
- Améliorer la collaboration avec l'environnement de l'entreprise : Plusieurs constructeurs automobiles mettent en place des Extranets avec leurs fournisseurs. La nomenclature des données fait alors l'objet d'un cahier de charge précis.

PCG : LA PREMIERE DEMARCHE QUALITE DE DONNEE AU MAROC

Les principes fiscaux et comptables constituent bien le premier projet œuvrant pour la qualité des

IJSER

données comptables grâce à des principes tels que :

- Principe de permanence des méthodes
- Principe de clarté : l'information doit être mentionnée dans la rubrique adéquate, sous la dénomination adéquate
- Principe d'importance significative : doivent impérativement figurer toute information susceptible d'influencer les évaluations et les décisions

Depuis le 1er Janvier 2016, la télé déclaration des impôts est devenue obligatoire. Cependant, est maintenue l'obligation de conserver tous les documents et calculs menant aux résultats déclarés. A l'entreprise alors d'exercer la bonne gouvernance de ses données (traçabilité, fiabilité, mise à jour périodique …) .

LE PROJET ISO 8000

Un membre du comité ISO 8000 discutant avec un expert en sécurité :

- comme vous savez dans les passeports de quelques pays (comme celui de France) on trouve le concept de couleur des cheveux et couleur des yeux . N'est ce pas ?
- Effectivement
- Êtes vous sûr que c'est mentionné dans le dictionnaire : couleur des cheveux et couleur des yeux ?
- Mais c'est absurde … tout le monde sait ce que ça veut dire.. vous savez .. la couleur des yeux !
- Ben .. je n'en suis pas sûr. Parlez vous de couleur biologique ou couleur observée ?
- Wow !
- Vous savez, dans mon passeport Américain, la mention de couleur des yeux et des cheveux a été effacée. Elle est inadéquate. La réponse que vous obtenez dépend avec laquelle vous avez posé la question, et combien explicite vous avez été à demander une information.

L'organisme International de Normalisation (ISO) a nommé un comité de très haut niveau formé de docteurs et d'experts mondiaux tels que Gerard Radack (Concurrent Technologies Corporation), Dr. Timothy King (LSI Group), Mr. Howard Mason (BAE Systems) ..
l'ensemble est dirigé par Peter Benson(expert mondialement reconnu dans le domaine des technologies de l'information , et Project leader de l'ISO 22475 qui est un dictionnaire des termes techniques).
La demande pour la norme est venue d'experts, organisations, cabinets conseil … des entités qui travaillent quotidiennement avec des données, et qui ont besoin d'évaluer à juste titre ce qu'ils achètent des fournisseurs de données. Le but de la norme est de rassembler toute la communauté des praticiens autour d'un seul standard. Ce standard doit impérativement correspondre à toutes les applications et faciliter ainsi l'échange et l'exploitation des données car comme mentionne la norme ISO 9001 : la qualité est la correspondance du produit à la demande.
Ce projet est en cours de chantier, mais les grandes lignes ont déjà été définies. La partie 1 est une définition des lexiques. Les parties de 2 à 99 traiteront de questions générales telles que la gouvernance. La partie 100 traitera de sujets tels que l'ingénierie des données, ou les données transactionnelles, elle concernera surtout les fournisseurs de données (qu'ils soient internes à l'entreprise ou externes) car l'on se rend de plus en plus compte que l'intégration des données est un facteur clé de succès . La partie 110 a été publiée en 2008. Puisqu'elle concerne des spécificités techniques(TS), le nom précis de la norme est ISO/TS 8000- 110:2008.

Rappel : eOTD (Open Technical Dictionnary) est le dictionnaire des termes techniques de la norme ISO 2745 dirigée par le même Peter Benson.
ISO 8000 vient en complément de ISO 9001. On peut considérer qu'elle a pour objet : « je ne m'intéresse pas à comment fabriquez vous vos données, montrez moi juste comment vous les utilisez ». Cette approche couvre une grande partie des risques , mais pas la totalité. D'où l'importance de l'ISO 8000 qui couvrira les risques restant . On peut donner l'exemple de la dernière crise des SubPrimes qu'on peut considérer d'une certaine manière dûe à des mauvaises informations car elles n'avaient pas été demandées correctement.

OBJECTIFS ET METHODOLOGIE

Pour une telle étude exploratoire, le choix d'une approche qualitative est le plus adapté vu la nature de notre sujet : la compréhension d'un standard de qualité et les difficultés rencontrées dans son application. Ces questions ne peuvent pas être examinées à travers une enquête ou un questionnaire.
A l'occasion de l'étude, 16 entretiens approfondis de 30 minutes au moins ont été conduits. Ils avaient

pour objet l'étude de la nature du problème, de ses causes et de ses effets. Une discussion libre et ouverte permet à l'interviewé d'exprimer son point de vue mais aussi ses propres pratiques à l'égard du problème.
Un questionnaire semi-ouvert a été utilisé. Il couvrait les sujets suivants:
- Informations sur l'interviewé
- Structure à laquelle il appartient
- Causes de la défaillance des données informatiques
- Conséquences de ces problèmes
- Connaissance des autres normes adoptées par l'entreprise et de leur pratique
- Défis pour ISO 8000
- Effets de l'application de la norme sur l'entreprise
- Difficultés rencontrées par la future norme ISO 8000

Les entretiens ont été menés avec 16 managers, de diverses entreprises de taille moyenne, plus ou moins exposés aux dangers de la non-qualité des données informatiques. Ils sont tous responsables dans leurs organisations respectives et couvrent plusieurs secteurs:
- banque
- Assurance
- Analyse marketing
- Planification média
- Industrie
-Administration publique
-Tourisme

Le profil des individus varie selon le sexe, formation, caractéristiques démographiques, années d'expérience et domaine de compétence.
Une fois les entretiens terminés, ils ont été transcrits, codés puis analysés à des fins d'interprétation.

PROPOSITIONS DANS LE CADRE DE L'ISO 8000

1) Une étude comparative :

Les solutions proposées par les managers s'approchent beaucoup de celles que j'ai connu à l'occasion de la recherche bibliographique , réalisées divers cabinets internationaux de techniques d'informations , des rapports de grands producteurs d'ERP, chacun d'entre eux a développé un modèle propre tels que :

- Livre blanc de Micropole, 2011 sur la qualité des données
- Modèle d'intelligence économique de l'AFDIE
- « Normes et lignes directrices concernant la documentation sur la qualité des données et la méthodologie » publié par STATISTICS CANADA en 2000

Tous ces modèles se concentrent sur la qualité intrinsèque de la donnée . A notre avis, la donnée doit être traitée comme un objet systémique. Son interaction avec les autres données d'une part, et avec les organes de l'entreprise doit être envisagée.

Cette problématique s'avère très présente dans les règles prudentielles appliquées aux banques et aux assurances. Ces 2 secteurs stratégiques et très sensibles à la gestion des informations ont très tôt pris conscience de l'importance de la qualité des données . Ils ont adopté depuis 1998 une série d'exigences de plus en plus drastiques, ayant pour objectif de leur assurer une meilleure stabilité et résistance face aux crises futures.

Dans Bâle II et III, la Directive Solvabilité 2 insiste sur l'importance de la production d'une donnée de qualité.
En 2016, est entrée en vigueur l'exigence BCBS 239 qui a pour objet la surveillance des établissements bancaires d'importance systémiques. Elle apporte 14 principes destinés à cadrer le traitement des données sur les risques. Parmi elles, 11 concernent les établissements bancaires :

- mise en place d'une gouvernance des données conformément aux principes de Bâle
- mise en place d'une architecture de système d'information apte à supporter des stress et respecter toujours les autres principes
- Intégrité des données de base mais aussi celles agrégées
- Exhaustivité :Chaque donnée devrait être vérifiable par plusieurs départements de diverses natures
- Une donnée à jour
- Une donnée de nature et de format parfaitement adaptée au besoin du client interne
- Exactitude des données
- Représentativité fidèle de l'importance de la donnée
- Clarté : les reportings et agrégations doivent être utiles, clairs et concis
- Fréquence : la production de rapports, de vérifications et vérifications doit se faire de manière suffisante, et augmenter de rythme en période de stress
- Distribution : au bon moment au bon

IJSER

destinataire

On retrouve dans ces principes la prise en compte de l'information dans son environnement, s'intégrant parfaitement dans la chaîne de valeur de Porter. Je préconise fortement une approche de Bench Marking car la BCBS 329 est exhaustive par rapport aux autres modèles .

2) Donner les moyens suffisants

Pour animer ces principes, il faut mettre la qualité des données dans la stratégie de l'entreprise. La première conséquence est que, comme on a un membre du personnel (ou plusieurs) dédié à la qualité des données comptables, il faudra dédier une personne à la maitrise de l'ensemble des données de l'entreprise, un sorte de Directeur des données qui se chargera de la sécurité des données, leur exactitude, coordination…ce rôle peut être joué par la DSI, DAF.. n'importe qui ait le pouvoir suffisant pout imprégner cette démarche à l'ensemble de l'entreprise depuis la phase d'implémentation du système par les opérationnels (on passe alors d'une Only Exective Information System à un EveryBody Information System)

4) Une stratégie proactive

La majorité des entreprises ont une stratégie réactive par rapport à la mauvaise qualité des données . L'on ne s'en rend compte souvent que lors d'audits ou de sanctions du marché . Experian Marketing services (un grand cabinet international de 17.000 collaborateurs) a effectué une étude dans ce sens .

Ne détecter l'erreur que lorsqu'elle nuit à l'entreprise peut plomber fortement la productivité. Une profonde réflexion dans les processus , organisation de l'entreprise, répartition des tâches, mise en place d'une gouvernance des données… est nécessaire pour garantir la qualité des données. Des logiciels existent aujourd'hui pour assister ce processus et détecter pro- activement les erreurs avant leur circulation dans le système (détection de doublons, erreurs typographiques, alertes lors de dépassement de seuils logiques prédéfinis …)

5) L'harmonisation des outils

Beaucoup d'entreprises investissent aujourd'hui dans des solutions informatiques de pointe, mais malheureusement ces efforts restent éparpillés : chaque solution se base sur un lexique technique qui lui est propre, un langage informatique qui lui est unique … l'information se trouve verrouillée sur autant de solutions existantes. Les données doivent plutôt être normalisées et exploitables par tous les services. Dans ce cadre, la technologie utilisée doit être plus sophistiquée pour être capable de prendre en compte l'ensemble des natures de données, mais par la suite l'exploitation des données est facilitée et coûte beaucoup moins cher que dans le cas contraire. C'est d'ailleurs un pré requis pour la mise en place d'une Business Intelligence.

6) Ne pas tomber dans la sur qualité

L'exigence quant à la précision des données n'est pas la même pour des domaines différents ou des besoins différents . sur investir peut même être contre productif . L'exemple le plus frappant est l'entreprise qui investira énormément pour produire au final une information peu utilisée comme beaucoup d'entreprises Marocaines qui ont investi des centaines de Milliers de Dirhams dans des bases de données et des logiciels complexes , pour que ces derniers tombent dans l'oubli moins de 6 mois après leur implémentation.

Enfin de compte, l'information doit être au service de l'entreprise et non le contraire.

CONCLUSION

Les entretiens ont révélé que toute entreprise mettant en place une politique de qualité rencontre des résistances de la part des collaborateurs à cause de :

- Résistance à la diffusion de l'information qui est source de pouvoir
- Divulgation d'informations que quelques uns tiennent confidentielles pour garder leur emploi
- Mise en évidence des dysfonctionnements et des erreurs d'ingénierie organisationnelle
- L'application des principes de gouvernance remettra en cause la répartition des tâches, certaines fonctions et nombre de postes le seront aussi.
- Remise en cause de certains principes officieux à l'intérieure de l'organisation, des habitudes confortables auxquelles s'est tenu les collaborateurs se trouvent bousculées

IJSER

- …

Du bon sens, et beaucoup d'empathie doit être déployé pour faire adopter une démarche qualité des données à l'organisation.

Ainsi, comme toute autre stratégie, l'approche qualité est une action permanente qui doit être maintenue comme le lit d'une rivière et qui doit être entretenue quotidiennement. A l'instar de la roue Deming, les étapes suivantes doivent être constamment suivies :
- Planifier: les objectifs de notre système d'information, le niveau de granularité requis et les mesures prises pour contrer les différents risques.
- Développer: élaborer des manuels de qualité et concevoir de nouveaux processus
- Suivi de la mise en œuvre: adapter les outils aux besoins spécifiques
- Agir : soit déployer de nouvelles solutions

Le résultat est l'amélioration de tous les éléments de la chaîne de valeur du modèle de Porter et l'obtention d'un avantage concurrentiel indéniable.

REFERENCES BIBLIOGRAPHIQUES

- Bloch L, Le management de la sécurité de l'information, Sécurité informatique, N°□□, décembre 2006
- Brasseur C, Enjeux et usages du Big Data – technologies, méthodes et mise en oeuvre – 2013
- Directives ISO/CEI, Partie 2, Règles de structure et de rédaction des Normes internationales, cinquième édition, 2004
- EIFR, article du 15 Septembre 2015
- Equinox-Coghizant, rapport de 2014
- Livre Blanc Experian Marketing Services, Février 2015, MIT Information Quality Industry , Symposium, July 15-17, 2009
- Sadok Ben Yahia and Jean- Marc Petit, editors, Extraction et gestion des connaissances, EGC, Actes, au janvier 2010, Hammamet
- Rampacek S. Sémantique, interactions et langages de description des services, web complexes. PhD thesis, Université de Reims , Champagne-Ardenne,Novembre 2006.

IJSER