

On ne résout pas un problème en le mettant de côté.
Winston Churchill

La novation est le secret de jeunesse des industries,
récompense de la technique intelligente et nécessaire.
Fernand Braudel

Management du système d'information décisionnel

Le système d'information décisionnel est un sujet très important pour les directions générales car c'est un moyen incontournable pour gérer la performance de l'entreprise, de la définition de la stratégie au pilotage opérationnel des activités. Si la contribution des systèmes décisionnels à la création de valeur ne fait pas débat et est largement reconnue, le management de ces systèmes n'est la plupart du temps pas considéré comme optimum, leur contribution étant généralement reconnue comme inférieure aux attentes, les DSI n'en assumant pas une maîtrise technique totalement satisfaisante et les directions métiers ne s'impliquant pas suffisamment.

Dans les entreprises le décisionnel est encore souvent un domaine moins mature du système d'information, du fait d'une offre technologique globalement récente qui a continué à fortement évoluer ces dernières années, et aussi car les DSI ne l'ont pendant longtemps pas mis en tête de leurs priorités. Pour que les systèmes décisionnels créent de la valeur et constituent une source d'avantage concurrentiel, il est nécessaire d'assurer leur adéquation avec les besoins métiers et pour cela il convient de définir un système d'information décisionnel cible, des priorités, des étapes et mettre en place les moyens adéquats.

Que l'on parle d'urbanisation ou de schéma directeur, peu importe le vocabulaire, la planification et la création d'un système décisionnel doit intégrer les mêmes contraintes que pour la réalisation des autres grands composants du système d'information. Concrètement il s'agit de se préoccuper sérieusement de certains sujets clés, par exemple de la gouvernance, des méthodes de développement, du choix des solutions technologiques et de créer des

conditions favorables au ROI. Cependant la mise en place de systèmes décisionnels présentent des caractéristiques spécifiques qu'il convient de ne pas sous estimer, notamment pour l'élément fondamental de ces moyens à savoir la gestion des données historiques, en particulier en matière de modèle de données, de données de référence, de qualité de données et de protection des données.

Modélisation des entrepôts de données

Un bon système d'information décisionnel permet de développer la capacité de réflexion et d'action de l'entreprise (décider avant le concurrent, décider en temps réel, tirer parti d'un avantage concurrentiel, réaliser des économies grâce à l'identification des bonnes méthodes). Mais l'information n'existe pas en soit, c'est un « produit » élaboré par un processus de transformation itératif, interactif, à partir de données élémentaires relatives aux différents acteurs et aux événements qu'ils vivent. Par exemple les données clients et les événements gérés par de nombreux services (marketing, commercial, facturation, comptabilité clients, service après vente, réclamations, contentieux).

Historiquement les premiers systèmes décisionnels ont été développés pour répondre aux besoins en information d'une organisation fonctionnelle particulière, ignorant les besoins des autres entités et s'adressant à un nombre limité d'utilisateurs. Ces systèmes qui géraient des données dupliquées et redondantes, offraient de multiples vues de la réalité, de multiples perspectives fonctionnelles. Faute d'une intégration des données dans un modèle unique d'entreprise, chaque organisation élaborait l'information selon sa perception, il n'y avait pas de vue globale, pas de vérité d'entreprise. Cela cloisonnait la connaissance, contrariait la mise en place de processus transversaux et limitait la prise de décision au niveau global.

Les entrepôts de données qui sont aujourd'hui mis en place sont notamment conçus pour pallier aux inconvénients décrits ci-dessus. Ils gèrent des données normalisées, partagées entre toutes les organisations, orientées par sujet ou par thème, qui ne sont plus cantonnées à une seule sphère fonctionnelle mais valent pour l'entreprise toute entière. Ils offrent une vue unique des données, une information plus complète, plus précise qui facilite d'autant la prise de décision.

Il s'agit donc d'organiser les données intéressantes dans des structures cohérentes, en éliminant les redondances de données inutiles, et en définissant des relations entre les structures de données qui ont des rapports entre elles. Dans une première phase du projet de développement de l'entrepôt de données, il convient de réaliser un modèle conceptuel de données (MCD), puis celui-ci est affiné en un modèle logique de données (MLD) qui est dépendant du système de gestion de base de données utilisés. Par exemple dans un SGBDR (Système de Gestion de Bases de Données Relationnelles), les données sont définies dans des tableaux à deux dimensions (lignes et colonnes). Enfin le MLD est traduit en un modèle physique de données (MPD), c'est-à-dire écrit dans le langage de définition de données spécifique de la base choisie, généralement le SQL.

Comment modéliser un entrepôt de données pour satisfaire les utilisateurs ?

En matière de modélisation de systèmes décisionnels deux grands types de modèles sont mis en œuvre : le modèle en troisième forme normale et le modèle du schéma en étoile.

Le modèle en troisième forme normale dans lequel chaque attribut (colonne) doit être un fait relatif à la clé primaire - toute la clé et rien que la clé. Chaque donnée est alors implantée dans une table sans autre duplication, ni donnée dérivée. Ce modèle permet aux utilisateurs

de poser à tout moment n'importe quel type de questions à propos de la totalité des données disponibles.

Le modèle du schéma en étoile qui se compose d'une table de faits et d'un certain nombre de tables dimensionnelles. La table de faits est associée à une clé multi-composants. Chaque élément de la clé est lui même une clé autonome liée à une table dimensionnelle spécifique. Les autres champs de la table factuelle sont des faits qui peuvent être considérés comme autant de mesures prises à l'intersection de toutes les dimensions retenues. Ces modèles en étoile permettent d'améliorer les performances lorsqu'il s'agit de questions ou de requêtes qui ont pu être anticipées au moment de la conception.

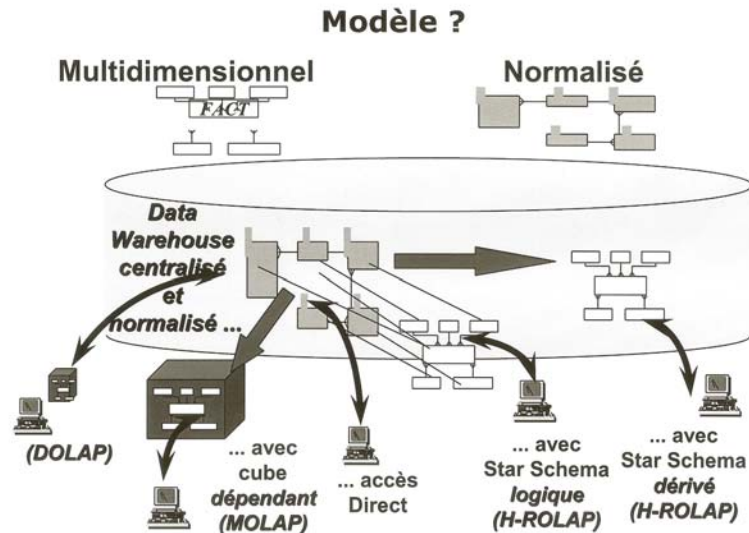


Illustration - 24

Ceux qui sont chargés de la modélisation des bases de données décisionnelles connaissent bien l'intérêt d'un modèle logique en troisième forme normale, mais parce que de nombreux moteurs de bases de données sont sujets à des contraintes physiques, ils sont contraints de faire des compromis. Les quatre opérations les plus difficiles à réaliser pour une base de données sont les suivantes : 1) rapprocher les tables, 2) cumuler les données, 3) trier les données et 4) analyser des volumes importants de données. Pour surmonter les limitations des bases de données, certains développeurs proposent des modèles qui évitent les rapprochements, ils prévoient alors d'utiliser des données déjà récapitulées pour éviter les contraintes liées au cumul, ou encore de stocker des données déjà triées et de faire un usage intensif d'index pour éviter l'analyse de données volumineuses.

En fait il existe peu de bases de données qui disposent d'une puissance et d'une maturité suffisantes pour faire usage, dans un système décisionnel, d'un modèle physique type troisième forme normale lorsque le volume de données est important. En raison de leurs limitations physiques, d'autres bases de données font appel à un modèle en étoile pour améliorer leurs performances, ce faisant, elles se privent de la possibilité de traiter des requêtes ad-hoc et perdent le bénéfice du data mining.

Après quelques années d'expérience en vraie grandeur, il est apparu aux experts que la troisième forme normale correspondait bien au modèle à utiliser pour les entrepôts de données d'entreprise. En effet, ce modèle permet aux utilisateurs de formuler à tout moment

n'importe quelle question sur les données disponibles. Telle est bien la philosophie fondamentale de l'entrepôt de données d'entreprise qui conduit également à tirer parti de la puissance des requêtes ad-hoc et du data mining. Au moyen de semblables techniques, des outils évolués peuvent faire apparaître des relations qu'il ne serait pas facile de détecter et qui peuvent être peu prévisibles.

Le modèle du schéma en étoile améliore les performances lorsqu'il s'agit de requêtes connues, mais il sera limité par les suppositions qui auront guidés sa conception. S'il peut arriver que ces suppositions soient exactes pour la première application, elles peuvent ne pas le rester pour d'autres. Par conséquent, pour construire un entrepôt de données de classe mondiale, il faut commencer par élaborer l'entrepôt de données d'entreprise autour de données détaillées en recourant au modèle de la troisième forme normale. Puis, à mesure que des requêtes sont réalisées et que les performances de certaines requêtes répétitives deviennent prioritaires, des data marts en étoile peuvent être créés au sein de l'entrepôt de données d'entreprise ou en aval. L'utilisateur a alors accès à la fois aux data marts - pour certaines requêtes répétitives - et à l'entrepôt d'entreprise pour toutes les autres requêtes.

Un entrepôt de données d'entreprise devrait toujours être conçu de manière à ce que tous les intervenants de l'entreprise puissent se procurer en temps opportun l'information dont ils ont besoin pour leurs activités. Il existe deux types de demandes différentes : des demandes répétitives dont le cadre général est connu et des demandes itératives imprévisibles. Quarante pour cent des requêtes sont répétitives, mais 80 pour cent du retour sur investissement provient généralement des 20 pour cent de requêtes qui portent sur les données détaillées dans une approche itérative. Par conséquent, en utilisant la troisième forme normale pour le modèle de l'entrepôt de données centralisé et l'indexage en étoile pour les data marts, toute entreprise devrait pouvoir maximiser son retour sur investissement, tout en bénéficiant de performances très optimisées sur 80 pour cent de ses requêtes. Voilà dans quelles conditions l'entrepôt de données peut donner satisfaction à tous les intervenants d'une organisation.

Modèle logique de données : ou comment fonder ses connaissances sans jargon

Les analyses n'ont jamais tenu un rôle aussi vital dans les entreprises, les pressions sur les coûts et les revenus, ainsi que le risques accrus lié aux affaires, rendent incontournable une vision intégrée de la performance des activités. De plus, les fusions et les acquisitions, l'essor de nouvelles technologies, et plus important encore les changements au niveau des besoins et des attentes de la clientèle, donnent naissance à un marché où le savoir, et principalement les connaissances sur les clients, procurent un avantage incontestable sur la concurrence.

La connaissance des clients est au cœur même de toute activité. Il s'agit de l'élément fondamental permettant d'identifier, d'acquérir, de fidéliser et de développer des clients à haute valeur qui soutiendront la croissance et la rentabilité. Il s'agit par exemple de savoir qui sont les clients les plus ou les moins rentables, qui doit être ciblé pour un nouveau produit ou un nouveau service, quelle est la tendance au niveau du nombre de comptes ayant des factures impayées ou contestées, comment mesurer et analyser les activités à haut risque potentiel.

Pour répondre à ces questions clés que les sociétés qui réussissent se posent chaque jour, les entreprises bâtissent des entrepôts de données pour supporter leurs analyses. Mais la

construction de n'importe quelle structure, qu'il s'agisse d'une maison ou d'une activité commerciale, requiert un plan. Il en va de même pour la construction d'un entrepôt de données : il faut débiter par un plan pour poser les fondations d'analyses pertinentes. Le plan en question ici, c'est le modèle logique de données, qui définit l'organisation des données au sein de l'entrepôt de données. Il fournit une structure de stockage des données représentatives des activités de l'entreprise et il permet des analyses pour répondre aux questions clés évoquées plus haut, pour gérer les clients avec plus de précision, pour optimiser les opérations et pour anticiper les risques.

Le concept d'infocentre, ancêtre de l'entrepôt de données d'aujourd'hui, ayant maintenant plus de 30 ans, différentes sociétés ont tiré parti de leurs expériences pour proposer des modèles de données qui contiennent des structures souples, pour aborder l'ensemble des domaines de gestion d'une entreprise. Ces modèles logiques de données apportent bien plus qu'une feuille de route, ce sont des modèles de planification d'entreprise qui définissent les données à capturer dans l'entrepôt, pour répondre aux questions sur les activités et produire les indicateurs en fonction des exigences du pilotage des opérations. Ces modèles prêts à l'emploi, peuvent faire gagner des semaines ou des mois dans le développement d'un entrepôt de données et épargner des ressources humaines et financières.

Un modèle logique de données doit regrouper par exemple, les informations capitales relatives au client, depuis tous les événements et les concepts relatifs aux comptes jusqu'à l'utilisation de produits. Il doit permettre de générer une compréhension globale de chaque client pour toutes les fonctions de l'entreprise. Il peut être conçu comme le dictionnaire de l'entreprise, une collection de termes et de références clairement définis pour chaque membre de l'organisation. Ainsi est obtenu un vocabulaire commun pour toute l'entreprise, ainsi est évité l'impasse des jargons métiers ou départementaux.

Grâce à un modèle logique de données, il est possible de construire plus vite et plus intelligemment, d'économiser des heures de développement, sans parler des euros, tout en réduisant le risque associé à la mise en œuvre d'un entrepôt de données. Et dans la mesure où le modèle adopte la troisième forme normale (3NF), il ne faut saisir les données qu'une seule fois. Cela élimine la redondance des données et signifie qu'il n'est pas nécessaire de réorganiser tout le modèle chaque fois qu'une nouvelle question importante pour l'activité doit être abordée.

Il convient d'utiliser le modèle logique de données comme un cadre pour la business intelligence, comme une première étape dans la pose des fondations d'un entrepôt de données robuste et flexible, fondations qui aideront à : fournir des données et des relations entre les données indispensables pour prendre de meilleures décisions, plus vite ; éviter de coûteuses modifications ; apporter discipline et structure à la tâche complexe que représente la mise en œuvre d'un entrepôt de données ; faciliter la communication entre les utilisateurs et les spécialistes informatiques.

Un modèle logique de données offre toute une série d'avantages parmi lesquels : une source unique d'informations qui repose sur des données détaillées pour des analyses approfondies ; la possibilité de décisions plus rapides et plus judicieuses sur la base d'analyses détaillées des performances à tous les niveaux de l'entreprise ; une représentation holistique des données de gestion qui contribuent à l'amélioration des pratiques et qui dégagent des réserves pour les opportunités d'optimisation des activités ; une amélioration des capacités d'analyse et de la conformité aux réglementations ; une amélioration de la modélisation des données en

garantissant d'avoir envisagé tous les domaines nécessaires avant de déployer l'entrepôt de données ; une perspective sur les activités dans plusieurs canaux.

Un modèle logique de données montre comment les mêmes données peuvent être utilisées pour exploiter différentes opportunités d'amélioration de l'activité dans divers secteurs de l'entreprise. Grâce au modèle logique de données il est possible d'identifier rapidement les données relatives à une opportunité d'amélioration particulière, et de voir également dans quelle mesure l'entrepôt de données apporte de la valeur à l'activité, via des analyses de données, la fixation d'objectifs spécifiques et des actions adaptées.

Construire un entrepôt de données pour identifier et saisir des opportunités

Lorsque qu'une entreprise examine sa situation une des questions fondamentales qu'elle se pose au-delà du bilan de ses actions passées et de ses résultats présents, est comment identifier des opportunités d'améliorer la conduite de ses affaires (vendre plus, réduire ses coûts, réduire ses risques, améliorer son cash flow, faire progresser son image). Bien sûr l'examen des résultats fait parfois ressortir des voies claires d'amélioration dues à des choix antérieurs erronés ou à des constats explicites de nouvelles tendances par exemple, mais la plupart du temps l'entreprise ne peut faire l'économie d'une recherche spécifique d'idées et donc elle est conduite à réaliser de nombreuses analyses, à raisonner de nombreuses métriques, à envisager des actions marché, des changements internes, des remises en cause de ses processus.

Dans ce jeu de piste aux bonnes idées, la première manche se joue sur la disponibilité et la qualité des données historiques qui vont être utilisées pour identifier et évaluer les opportunités. A ce niveau il convient de disposer d'un gisement de données détaillées, fondé sur un modèle qui autorise une vision d'ensemble des activités de l'entreprise, et permet ainsi d'éviter les pièges inhérents aux données agrégées, qui moyennent les résultats et masquent certaines réalités. La deuxième manche se joue sur la façon de structurer ses questions métier, et enfin la troisième manche se joue dans la définition des analyses et des métriques que l'on va élaborer pour répondre aux questions métier compte tenu des données disponibles.

Pour illustrer mon propos je vais prendre des exemples dans le monde de la Distribution. Les principales questions métiers peuvent être regroupées en une dizaine de domaines de gestion (assortiment, promotions, opérations magasins, ...). Chaque domaine se décompose en moyenne en une demi-douzaine de thèmes clés (les ventes perdues, rétention client, ...). Chaque thème clé peut faire l'objet de plus d'une douzaine d'analyses (quantité, marge, ...). Chaque analyse est fondée sur une ou plusieurs métriques qui nécessitent des données émanant de différents systèmes front ou back office qu'il faut rapprocher et suivre dans le temps. En particulier certaines analyses (préférences clients par exemple) nécessitent de disposer du détail du ticket client. Enfin chaque analyse débouche sur des scénarios d'actions qui peuvent aussi nécessiter des analyses.

Dans le cas du secteur de la distribution il faut donc envisager plus de cinq cents analyses différentes et la situation est la même pour d'autres secteurs (Banque, Assurances, Télécommunications, Transport). Au niveau des données à gérer, les entités se comptent en centaines et les attributs en milliers. Il est clair que les entreprises ne peuvent mettre en place que progressivement de telles approches analytiques qui nécessitent beaucoup de moyens

humains et techniques. D'autant plus qu'en cours de route elles ont généralement de mauvaises surprises au niveau de la qualité de leurs données, et de plus leurs gestionnaires doivent apprendre à mener de nouveaux raisonnements quantitatifs et qualitatifs. De façon générale, les clés du succès sont l'itération, le lotissement, l'évolutivité de l'infrastructure, une vision d'entreprise, la granularité des données du gisement de base et la prédominance des besoins métiers pour tous les choix.

Il existe sur le marché de nombreux fournisseurs de solutions d'infrastructure (serveur et base de données), de modèles de données, d'applicatifs (BI, CRM, Gestion de la Demande) et de prestations de conseil. C'est dans le cadre de ce dernier type d'activité qu'ont été mis au point des modèles sectoriels « d'opportunités décisionnelles », dont j'ai esquissé la mécanique dans les lignes ci-dessus. Ces modèles autorisent une meilleure vision des liens entre les questions métier, les analyses, les métriques et les données ; ils guident les maîtres d'ouvrage dans la définition de leurs besoins et dans l'ordonnancement de leurs projets, et l'expérience a montré leur efficacité, en particulier leur impact sur la qualité des dialogues (entre utilisateurs et entre les utilisateurs & la maîtrise d'œuvre).

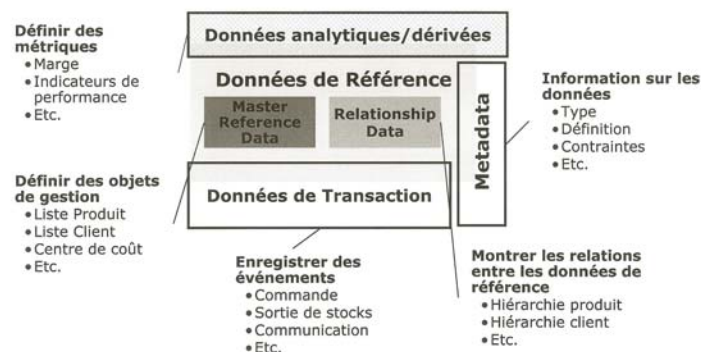
Les modèles « d'opportunités décisionnelles » permettent aux entreprises de mieux conduire leur projet d'entrepôt de données, en facilitent la définition de la première itération et la vision des questions métier clés susceptibles au-delà de l'objectif initial, de profiter de la disponibilité de données historiques déjà dans l'entrepôt ou qui pourraient venir le compléter. Ainsi ces modèles peuvent aider une entreprise à mieux utiliser et capitaliser sur ses données en identifiant les façons de les exploiter, ce qui permet de maximiser la rentabilité des investissements de l'ensemble du domaine de la Business Intelligence d'une entreprise.

Données de référence

Les entreprises s'appuient sur de nombreux moyens informatiques pour fonctionner et réaliser leurs activités physiques (production, logistique, ...) ou de gestion (comptable, ressources humaines, ...). Ces moyens gèrent des données communes, par exemple les données d'identification des clients ou des produits, mais le plus souvent les définissent de façon hétérogènes. L'hétérogénéité des définitions a deux causes, d'une part les besoins des différents acteurs de l'entreprise ne sont pas exactement les mêmes, et d'autre part les solutions informatiques sont différentes. Même les entreprises qui utilisent un progiciel de gestion intégré n'échappent pas à cette problématique, car leur PGI ne couvre généralement pas tous leurs besoins informatiques, et dans les grandes entreprises il est rare qu'elles en aient fait une mise en œuvre homogène.

La gestion des données de référence, en Anglais Master Data Management (MDM), a pour objet de répondre à cette problématique, elle consiste à gérer la qualité et la cohérence des données du système d'information de l'entreprise. Il s'agit de regrouper, l'ensemble des données permanentes ou dont la durée de vie est longue, dans un système qui a pour fonction de fournir la version standardisée d'une donnée, qui doit être systématiquement utilisée pour mettre à jour les différentes applications informatiques de l'entreprise. Concrètement ce référentiel contient l'ensemble des objets essentiels à la vie de l'entreprise, et décrit les liens qu'ils entretiennent entre eux, par exemple : numéro de référence clients, adresse de livraison, code produit, hiérarchie produit, plan de comptes, classification des postes, etc. Grâce à un référentiel et à des processus de contrôle et de validation, les données sont gérées de façon cohérente et les doublons évités.

Données métier de référence (MDM)



Les données métier de référence concernent toutes les données de l'entreprise

Illustration - 25

Une bonne gestion des données de référence présente un grand intérêt dans de nombreux domaines. En matière de relation client par exemple, il est essentiel que tous les acteurs de l'entreprise puissent avoir les meilleures données possibles pour réaliser leurs activités. Des données incohérentes ou incomplètes peuvent conduire à une dégradation du service rendu ou à rater des opportunités de vente. Il y a d'autant plus besoin de bien gérer l'homogénéité, qu'il y a un grand nombre d'acteurs ou de systèmes en relation avec le client, c'est particulièrement le cas pour les industries qui vendent à travers de multiples canaux

(internet, centre d'appel, borne automatique, point de vente, ...). Au-delà de la dimension opérationnelle et quelque soit le domaine fonctionnel, il est aussi intéressant d'avoir une bonne gestion des données de référence pour le système d'information décisionnel, par exemple pour les statistiques, les tableaux de bord et les prévisions.

Gérer l'ADN du système décisionnel

En matière de système décisionnel, la gestion des données de référence est l'ADN de la génération de rapports de qualité pour la prise de décision. Les données de référence sont le fondement des exigences en matière d'analyse, de reporting et d'audit de conformité, pour la transparence des données et leur niveau de détail pour un certain nombre de réglementations. C'est ce qui permet aux responsables de produire des rapports qui croisent avec confiance des données relatives à l'ensemble des activités d'une entreprise, données issues de nombreux systèmes opérationnels le plus souvent hétérogènes.

Les entreprises doivent relever de sérieux défis lorsqu'elles tentent de mettre en œuvre une solution efficace de gestion des données de référence. Ces défis comprennent dans les grandes entreprises les problèmes liés à la complexité des processus et la propension à créer des silos de données. Les silos de données sont créés par les secteurs d'activité qui cherchent à optimiser de façon isolée leur fonctionnement et qui considèrent le partage des données comme un risque pour l'intégrité de "leurs" données (et par conséquent, de leurs résultats). Ces approches en silos contribuent à l'incohérence des données et crée un fardeau pour le rapprochement des données, car chaque secteur d'activité, bien qu'il utilise les mêmes éléments, les adapte sans penser aux besoins de gouvernance et de génération de rapports de l'ensemble de l'entreprise.

Cet amasement spécifique de données et ce manque de confiance à l'égard de la centralisation des données s'expliquent notamment par des cas où l'on a vu les demandes des utilisateurs devancer la capacité du personnel informatique central à créer, capturer et distribuer de nouvelles données dans l'entreprise. Un exemple classique nous est offert par la vitesse à laquelle de nouveaux clients peuvent être créés. Les données relatives aux clients constituent une source stratégique de données de référence pour les utilisateurs. Souvent, les utilisateurs au niveau des secteurs d'activité n'ont pas le temps de demander au service informatique de créer un nouveau compte ou de vérifier l'existence d'un compte. Ils veulent des systèmes qui permettent de "conclure l'affaire" et n'acceptent aucun retard. La création de silos de données est la conséquence des architectures qui ne facilitent pas la synchronisation des données et en même temps la création de ces silos aggrave la situation.

Jusqu'à récemment, il n'existait aucune solution logicielle fiable pour centraliser et synchroniser les données de référence transversalement dans le cadre de la gouvernance d'entreprise. Mais aujourd'hui il existe des solutions spécifiques de gestion des données de référence, elles permettent de prendre en charge la configuration des processus de l'entreprise et de synchroniser les données tout en exploitant les plates-formes existantes d'intégration des données et les applications opérationnelles. Ces solutions spécifiques sont conçues pour être utilisées dans les processus clés des principales fonctions tels que la production, la commercialisation, la finance, les ressources humaines, la recherche et le développement. Elles permettent de fournir des données précises et actualisées aux systèmes d'information, de donner accès à des données couvrant divers systèmes, de maintenir un

thésaurus unique et fiable des données dans toute l'entreprise, de réduire le coût de revient total des systèmes informatiques.

D'autres part les silos de données historiques peuvent être regroupés dans un ensemble homogène, et il est possible d'atteindre une gouvernance efficace de la gestion des données de référence si l'entreprise fait confiance à une infrastructure de données groupe. La combinaison d'un entrepôt de données central et d'une solution spécifique de gestion des données de référence, permet aux entreprises de répondre aux défis de qualité et de transparence évoqués plus haut. Ensemble, ces éléments fournissent une solution de gestion des données viable et automatisée pour toute l'entreprise. La réussite se concrétise par des utilisateurs qui peuvent accéder directement aux données et réaliser des analyses en fonction des besoins des processus.

Une application de gestion des données de référence associée à un entrepôt de données central est une solution de services de données indispensable, elle permet aux entreprises de maîtriser leurs données et de les utiliser pour créer des informations clés qui leur apporteront un avantage concurrentiel. Cette approche de la gestion des données décisionnelles garantit l'exactitude, la cohérence et l'audit des données de toute l'entreprise. Elle crée des environnements dynamiques et évolutifs pour répondre aux besoins d'analyse approfondie de milliers d'utilisateurs simultanés.

De nombreuses sociétés de service œuvrent depuis longtemps dans la gestion des données de référence dans le cadre de la mise en place de grand entrepôt de données centraux. Depuis plus récemment elles mettent en place des applications de gestion des données de références pour l'ensemble des systèmes d'information (opérationnels et décisionnels) des entreprises. Ces approches permettent d'assurer des niveaux sans précédent en matière de qualité des données, d'intégrité, de synchronisation avec des coûts de déploiement et de maintenance très avantageux.

Gérer le référentiel métier dans l'entrepôt de données

Malgré tous les efforts déployés dans les grands groupes ces dernières années pour diminuer la diversité des applications analogues, les systèmes d'information continuent à se développer et à se complexifier, et une bonne gestion des données de référence apparaît de plus en plus essentielle.

La gestion des référentiels métiers (master data management en anglais, MDM) a pour but de référencer, contrôler, synchroniser toutes les données à travers l'entreprise (par exemple les données clients & produits), de façon à ce qu'elles soient précises, conformes, exploitables par les systèmes et leurs utilisateurs. Il s'agit par une bonne gestion des référentiels, de simplifier les processus métiers, d'éliminer les redondances, d'assurer la qualité et l'utilisation des données.

Les enjeux pour l'organisation se déclinent au niveau métier en termes de vitesse, d'agilité, de facilité d'intégration de ses processus internes ou externes avec ceux de ses partenaires, et au niveau système d'information en termes de facilité et de rapidité d'évolution. Seule une vue factuelle des activités de l'entreprise fondée sur des données de qualité, permet de créer un contexte d'information favorable aux affaires. Par exemple en matière de gestion des

risques, une bonne maîtrise de la conformité des données et de leur accès, permet de réduire les informations contradictoires à l'origine de biens des problèmes.

Les responsables d'entrepôts de données s'intéressent en particulier à la gestion des référentiels métiers, car les données principales des entreprises résident déjà dans leurs systèmes. Dès lors il est pertinent de profiter des ressources et des savoirs faire mis en œuvre dans le cadre de l'entrepôt de données, pour en faire profiter toute l'entreprise et globalement simplifier l'architecture de gestion des données, réduisant en conséquence le coût total du système d'information. L'objectif est donc d'éliminer la multiplication des référentiels métiers et des charges de gestion afférentes. Ces dernières années les entreprises ont utilisé différentes solutions de gestion de référentiels métiers du marché. Elles ont de ce fait évalué les solutions disponibles et peuvent aujourd'hui choisir les technologies les meilleures. La solution de gérer le référentiel métier dans l'entrepôt de donnée, est de nature à permettre aux entreprises d'apporter une valeur complémentaire significative à leurs systèmes décisionnels, tout en facilitant l'intégration du système d'information.

Par exemple pour une banque, il s'agit par une bonne gestion des référentiels métiers de pouvoir répondre aux exigences de BÂLE II en matière de gestion des risques. En établissant un gisement de données unique, centralement administré et maintenu, il est plus facile de gérer les processus du cycle de vie des données des référentiels métiers. Une solution de MDM dans l'entrepôt de données permet de répondre à ce besoin, et en particulier facilite les processus d'audit ou de garantie de la qualité des données. En particulier il est ainsi possible de corriger les erreurs plus rapidement, de faciliter la prise en charge de la problématique de la qualité des données par les utilisateurs eux même, et de produire des reportings ou des analyses dont la qualité fonde la crédibilité des décisions.

Toutes les entreprises peuvent tirer bénéfice d'une solution de MDM dans l'entrepôt de données, car elle leur permet d'augmenter leur efficacité globale par une meilleure utilisation de leur système informatique en général, et de leur système analytique en particulier. La gestion des référentiels métiers dans l'entrepôt de données facilite la simplification des processus. Quand les entreprises doivent regarder dans de nombreux systèmes pour trouver leurs données de référence, puis doivent réconcilier ces données, les processus (par exemple, création d'un nouveau client ou introduction d'un nouveau produit) peuvent devenir complexes, erratiques, et sources d'erreurs. Non seulement la mise en place d'une solution MDM dans l'entrepôt de données améliore de tels processus, autorisant une gestion plus rapide et plus simple des nouveaux produits ou clients, mais elle permet aussi aux utilisateurs d'éliminer des données et des processus superflus, de rendre les données réutilisables et d'améliorer la qualité globale des données.

De nombreux règlements gouvernementaux exigent maintenant de stocker et de pouvoir récupérer des données précises. Ces exigences concernent par exemple les données clients (dans la finance, la santé, les télécommunications), ainsi que les informations financières des entreprises. Si les données de référence d'une entreprise sont imprécises, elle va produire de mauvaises informations suite à des requêtes légales et peut encourir des sanctions. Si les règlements exigent d'être capable de récupérer et d'auditer rapidement des données sur un point particulier, les entreprises peuvent largement réduire leurs risques de produire des données fausses, grâce à une solution de MDM dans l'entrepôt de données qui peut garantir la cohérence des référentiels métiers, faciliter la transparence de l'information dans toute l'entreprise.

Gestion des données métier de référence (MDM) : où en sommes-nous ?

Force est de constater que les systèmes d'information des grandes entreprises sont systématiquement hétérogènes y compris ceux qui utilisent largement des progiciels, en effet sauf cas particulier, les entreprises ont généralement des applications spécifiques pour gérer leur cœur de métier. Dans le meilleur des cas on a un mille feuilles d'applications avec une intégration des données réduite aux principaux objets de gestion (client, produit, ...) pour répondre aux besoins minimum de suivre financièrement l'activité. Mais dans le contexte de globalisation d'aujourd'hui ceci se révèle insuffisant, et ne permet pas aux différents acteurs de l'entreprise de parler le même langage et de manœuvrer ensemble de façon totalement cohérente.

Idéalement les systèmes d'information devraient être intégrés et capable de répondre à deux besoins différents mais complémentaires, besoin concernant le système opérationnel (cohérence de la mise en œuvre de la stratégie) et besoin concernant le système décisionnel (vision globale de l'ensemble des activités pour définir la stratégie et piloter les opérations). Cependant, l'intégration de l'ensemble des systèmes d'information dans une grande entreprise globale est un défi particulièrement important, il peut être plus facilement relevé dans les industries qui par nature ont une gestion très centralisée (c'est le cas du transport Aérien par exemple), mais il semble plus difficile à atteindre pour certaines grandes entreprises qui se sont constituées suite à de nombreux rachats (c'est souvent le cas dans le secteur de l'Assurance par exemple).

Concrètement, intégrer suppose de gérer de façon unifiée les données de référence pour l'ensemble de l'entreprise, il s'agit d'échanger des informations entre applications et consolider la vision des activités pour optimiser la gestion des processus, et par exemple mieux servir les clients. Cette gestion unifiée des données de référence, c'est l'objet du Master Data Management (MDM) qui permet donc de consolider et de partager la description des objets de gestion à travers toute l'entreprise.

Globalement ce domaine du MDM n'est pas totalement mature et fait surtout l'objet d'opérations partielles, car au-delà de la technique les principales difficultés sont politiques et organisationnelles. Mettre d'accord plusieurs divisions ou fonctions sur un même concept métier et en définir les processus de gestion (création, validation, ...) peut dans certains cas être très difficile. Tous les analystes recommandent donc une approche MDM progressive, le mieux étant de partir d'une problématique métier du type traitement homogène du client, ou consolidation de certaines données d'un domaine de gestion à des fins analytiques. Il convient donc dans un premier temps de développer différents projets partiels en parallèle pour mieux intégrer des données, et obtenir des effets à la fois dans les systèmes opérationnels et décisionnels.

Jusqu'à récemment, il n'existait aucune solution logicielle fiable pour centraliser et synchroniser les données de référence transversalement dans le cadre d'une approche globale d'entreprise. D'autre part dans de nombreuses entreprises on a assisté à une multiplication de silos spécialisés de données historiques (Data Marts). Dans ces conditions les données de référence n'ont pas été bien gérées, mais aujourd'hui il est possible d'atteindre une gestion efficace si l'entreprise fait confiance à une infrastructure de données groupe. La combinaison d'un entrepôt de données central et d'une solution spécifique de gestion des données de référence permet de répondre aux défis évoqués plus haut, et la réussite d'un tel projet se

concrétise par des utilisateurs qui peuvent comprendre leurs données et réaliser des analyses en fonction de leurs besoins.

D'un point de vue architectural, différentes solutions sont envisageables, comme par exemple des applications dédiées par objets de gestion et reliées par un hub ou une application dédiée pour l'ensemble des systèmes d'informations de l'entreprise. Une application dédiée peut être développée de façon spécifique ou intégrée à un ERP ou à un entrepôt de données. Dans tous les cas elle est fondée sur une base de données (Oracle, DB2, Teradata) qui gère un modèle global de toutes les entités métier. Cette base est alimentée via des outils de consolidation (Informatica, Ab Initio, ODI), de fédération (IBI, ...) ou de propagation de données (MQ Series, Tibco). Cette base supporte des applications MDM (SAP, Oracle) qui gèrent un historique des différentes versions des données de référence (client, produit, finance, ...) et par exemple fournit à l'ensemble des systèmes d'information les adresses successives d'un même client. Cependant il convient de noter que pour les grandes entreprises qui ont des activités multiples et des systèmes d'information complexes, la cible la plus réaliste est probablement de mixer au mieux différentes solutions d'architecture.

Enfin à noter que Nationwide Insurance a eu en 2006 un « Best Practice Awards » de l'Institut du Data Warehouse (TDWI) pour son projet de MDM hébergé dans son entrepôt de données.

Qualité des données du système d'information décisionnel

Pour toutes les sociétés, la qualité des données est une question importante. A mesure que de nouveaux domaines thématiques sont ajoutés dans un entrepôt de données, la valeur des informations se multiplie. Malheureusement, l'impact des erreurs se multiplie également. Le but d'une équipe de gestion de la qualité des données, est de développer des règles communes et une terminologie cohérente, qui seront utilisées par les diverses unités de l'entreprise. Un programme de gestion de la qualité des données permet de promouvoir l'instauration d'une information exploitable dans l'ensemble de l'entreprise, afin que celle-ci mette le cap sur la rentabilité.

Selon le Data Warehousing Institute, la médiocre qualité des données coûte chaque année des milliards d'euros aux entreprises. Dans une récente enquête, près de la moitié des personnes interrogées considère que les données de leur organisation sont «pires que ce qu'ils pouvaient imaginer». Dans ces conditions l'amélioration de la qualité des données, même à petite échelle, peut considérablement influencer la prise de décision dans les entreprises, favorisant ainsi une augmentation des bénéfices. Malheureusement, de nombreuses initiatives en matière de qualité des données ne voient jamais le jour, car la tâche est toujours perçue comme difficile. Ainsi la plupart des entreprises se focalisent principalement sur la rentabilisation à court terme de leur système d'information décisionnel, et n'investissent pas dans la qualité des données.

Cependant, une procédure simple consiste à pointer les principaux problèmes de qualité des données, à créer des règles pour isoler, réparer les erreurs et utiliser ensuite un tableau de bord, pour contrôler et évaluer les zones de qualité des données de manière constante.

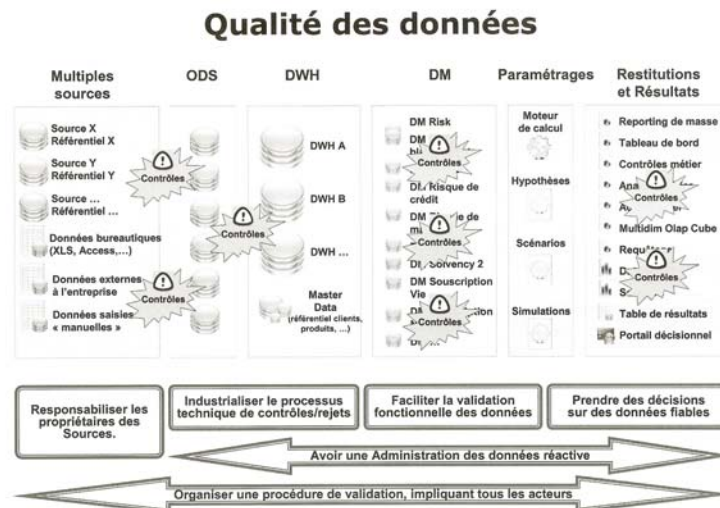


Illustration - 26 Capgemini

Un tableau de bord de la qualité des données bien conçu, aide les entreprises à mieux comprendre les problèmes de qualité des données, à évaluer les possibilités d'amélioration et à mesurer le progrès au fil du temps. Il est possible d'élaborer un tableau de bord de la qualité des données, et d'autres rapports de gestion des données en seulement trois semaines.

Le tableau de bord peut être implémenté à l'aide d'un portail ou à l'aide de l'outil de reporting qu'utilise déjà l'entreprise.

Bien qu'un plan d'action de qualité des données soit impératif pour toutes les entreprises, cette approche est plus facile à mettre en œuvre dans les entreprises qui commencent à mettre en place leur entrepôt de données. En abordant très tôt les problèmes de qualité des données, les actions importantes nécessaires au succès sont plus faciles à organiser que lorsqu'il faut revenir sur des réalisations déjà mises en production. Il convient aussi dans le cadre du système d'information décisionnel, d'industrialiser le processus technique des contrôles et des rejets, et donc de mettre en place un outil générique pour les contrôles et les rapprochements.

Enfin, la qualité des données n'étant pas qu'une problématique technique, et comme elle touche avant tout les utilisateurs métiers de l'entreprise, il faut impérativement responsabiliser les propriétaires des sources et faciliter le processus de validation fonctionnelle des données. Pour cela il convient d'organiser une structure de gouvernance viable, en particulier de définir des gestionnaires de données.

Cap sur la gestion de la qualité des données

Toutes les entreprises ont des problèmes de qualité de données. Les mesures standard de la qualité des données sont la validité, la précision et l'exhaustivité. Mais la qualité des données comprend également toutes les caractéristiques permettant aux données de répondre à l'usage que l'on a prévu d'en faire. Et plus particulièrement leur accessibilité, leur fraîcheur et leur cohérence.

Grâce au développement de la Business Intelligence et à la mise en œuvre d'entrepôts de données pour aider à la prise de décisions, les sociétés accèdent à une vision détaillée de l'information sur l'ensemble de l'entreprise. Par exemple dans le secteur des Assurances, les données clients sont reliées aux données sur les produits, les souscriptions, les primes et les sinistres. La valeur de l'information se multiplie à mesure que s'ajoutent de nouvelles dimensions. Malheureusement, l'impact des erreurs se multiplie également. Il est fréquent que ceux qui utilisent les données ne soient pas vraiment sûrs de leur signification, ou qu'ils les utilisent à des fins autres que celles prévues.

Les entreprises installent un entrepôt de données pour avoir accès à toutes les informations dont elles ont besoin pour prendre de meilleures décisions. L'accès s'est amélioré, mais nombre d'entre elles sont déçues de constater qu'elles ont maintenant un meilleur accès à des volumes plus importants de données de mauvaise qualité.

La qualité des données n'est pas simplement une question de précision, elle doit également fournir une meilleure représentation et un historique de la façon dont chaque donnée est obtenue ou calculée. Ces « données sur les données » s'appellent les métadonnées. Malheureusement, la plupart des entreprises stockent leurs méta données un peu partout, sur des documents, des tableaux, des post-it ou dans la mémoire d'un employé. Savoir où se trouve l'information et la mettre à la disposition de tous ceux qui en ont besoin n'est pas aisé. Un bon répertoire de métadonnées permet aux informaticiens de gérer efficacement l'entrepôt de données. Il permet d'identifier les doublons (les données sont-elles déjà présentes ailleurs ?), les données interdépendantes (cette donnée sert-elle à en obtenir

d'autres ; que se passera-t-il si elle est modifiée ?) et le taux d'utilisation des données (est-elle souvent utilisée ?). Il permet en outre aux utilisateurs de comprendre les règles afférentes aux données. Ceci inclue la validité des codes et des conditions d'utilisation, ainsi que toute information complémentaire à propos des données présentes dans l'entrepôt, comme par exemple à qui appartient cette donnée et quand est elle mise à jour.

Pour résoudre les problèmes de qualité des données et mieux profiter d'une vision d'ensemble des données de l'entreprise, de nombreuses sociétés mettent en place une équipe de gestion de la qualité des données. L'équipe de gestion de la qualité des données a pour objectif de permettre une bonne exploitation des informations-clés d'une société. Cette équipe multifonctionnelle réunit des représentants de chaque ligne d'activités et de chaque fonction de l'entreprise. Toutes ces personnes sont chargées de gérer les données et les informations pour l'ensemble de la société. Des représentants des services informatiques fournissent support et conseils à l'organisation sur de nombreux points, et plus particulièrement ceux liés à l'utilisation des outils d'infrastructure.

L'objectif de l'équipe de gestion de la qualité des données est de développer des règles communes et une terminologie cohérente, elles seront utilisées dans les différents secteurs d'activité et dans l'ensemble de la société. Une équipe de gestion de la qualité permet de surpasser les « demandes » conflictuelles de chaque unité et de promouvoir la mise en place d'une version unique de la vérité. Cette équipe de gestion de la qualité des données est composée de plusieurs profils.

Le sponsor est un cadre supérieur à qui a été délégué la gestion de la qualité des données. À mesure que les travaux de l'équipe de la gestion de la qualité des données avancent, il est nécessaire d'initier des changements de processus et de politique. Cette organisation, par le biais de son sponsor, doit disposer de l'autorité indispensable à la mise en œuvre de ces changements qui visent à améliorer la qualité et l'utilité des données en fonction des besoins.

Les propriétaires de données sont les unités fonctionnelles qui ont demandé la constitution de gisements de données historiques. Le propriétaire prépare les exigences métier et définit les règles qui seront associées aux données.

Les gestionnaires de données gèrent les besoins d'information sur un ensemble d'unités, en gardant à l'esprit les objectifs globaux de l'entreprise. Les informations créées par différents départements concernent divers domaines. Les données sont reliées les unes aux autres, et chaque problématique commune est résolue par décision consensuelle.

Le gardien des données est posté dans le secteur informatique, il doit s'assurer qu'elles sont collectées et traitées sur la chaîne d'information selon les spécifications développées par le propriétaire. L'informatique est responsable du nettoyage, de l'accessibilité et de la protection des données. Ceci comprend la sauvegarde et la récupération ainsi que l'accès sécurisé aux données.

De nombreuses entreprises ont pu expérimenter que quand les pilotes et les outils appropriés sont en place, l'information peut être exploitée dans l'ensemble de l'entreprise, et alors cette dernière met le cap sur la rentabilité.

Maîtriser la qualité des données : une tâche pour Sisyphe ?

La qualité des données est un des problèmes fondamentaux des systèmes d'information des entreprises. La multiplication des moyens technologiques mis en œuvre ces dix dernières années, n'a fait que complexifier la situation, et régulièrement les analystes évaluent en milliards de dollars, les pertes annuelles dues à une insuffisance de qualité des données.

Deux exemples pour concrétiser cet état de fait. Le premier concerne le secteur des télécommunications, où depuis quelques années nous avons vu se multiplier les programmes de « Revenu Assurance ». Il s'agit pour un opérateur d'être en mesure de bien facturer les prestations qu'il a effectuées pour le compte de ses clients. Basique penseront certains, pas du tout, les experts estiment que de nombreux opérateurs perdent jusqu'à cinq pour cent de leur chiffre d'affaires, du fait d'une qualité insuffisante des données qui ne leur permet pas de toujours aller au bout de leur processus de facturation. Le deuxième concerne la Nasa, qui a perdu un satellite d'une valeur de plusieurs centaines de millions de dollars, du fait que l'équipe internationale qui avait travaillé sur ce projet, n'a 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.

Un système d'information décisionnel a pour objet d'aider les décideurs à effectuer des choix pertinents en utilisant des données historiques. Ceci implique généralement de prendre en charge de grandes quantités de données disparates, afin de les calibrer en informations suffisamment précises et sûres, pour qu'en confiance des actions puissent être définies et lancées. Ce défi classique pour les architectes de systèmes décisionnels, est amplifié ces derniers temps par l'augmentation des exigences en matière de variété, de fraîcheur et de profondeur des données historiques à gérer. Dans ce contexte, la vitesse avec laquelle des données peuvent être nettoyées, transformées et intégrées dans un entrepôt devient essentielle pour la compétitivité des entreprises.

La qualité d'une donnée dépend d'abord du contexte dans lequel elle a été initialement saisie, mais d'un point de vue décisionnel, son intérêt dépend de l'usage qu'un utilisateur peut en faire. La qualité d'une donnée ne doit donc pas être appréciée dans l'absolu, mais de façon relative à son intérêt métier, et les programmes d'amélioration de la qualité des données doivent être ciblés en fonction des intérêts métier. Mais les besoins métiers évoluant, l'appréciation de la qualité d'une donnée ne peut donc jamais être fixée de façon définitive, et les programmes d'amélioration de la qualité des données, représentent souvent une charge de travail digne de Sisyphe.

Tous les systèmes d'information opérationnels génèrent des données qui présentent quelques défauts, et la meilleure place pour essayer d'améliorer leur qualité est bien entendu leur application d'origine, mais comme cela n'est pas toujours possible, il convient de le faire dans le cadre du processus ETL (Extract, Transform, Load), dans différentes zones tampons ou dans l'entrepôt de données lui-même. Dans cette opération un des moments clé est celui du « profilage » des données, c'est-à-dire lorsque sont mis en lumière certaines caractéristiques des données (exactitude, complétude, unicité, consistance, soit concrètement des valeurs manquantes, fausses, corrompues, désynchronisées).

Il existe maintenant des outils qui peuvent alléger les charges de travail des opérations d'amélioration de la qualité des données, comme ceux d'Informatica, de Business Objects, de Trillium par exemple, mais par-dessus tout il convient d'organiser la gouvernance des données. Pour cela il faut mettre en place un comité de direction spécifique incluant des utilisateurs, des informaticiens et désigner un utilisateur métier pour prendre la responsabilité à plein temps de ce programme. A ce niveau ont définis des principes de gouvernance des données et une organisation comprenant :

- des responsables de grands ensembles de données, qui s'appuient sur des utilisateurs métiers issus de toutes les directions de l'entreprise, et des informaticiens issus de la DSI.
- un programme de gestion proactive des données, incluant des processus de réponse aux besoins des utilisateurs et de gestion des projets de développement, des réunions régulières, des revues mensuelles, des indicateurs de suivi, etc.

Cette organisation doit permettre de traquer pro activement des opportunités d'amélioration de la qualité des données, de la qualité des modèles, des technologies mis en œuvre, des architectures et des projets de développement. En termes de résultat, au niveau de l'utilisation des données cela doit se traduire par une meilleure compréhension des informations produites par les systèmes décisionnels, et le développement de nouveaux usages métier des gisements de données.

Impact de la qualité des données sur les résultats des compagnies d'assurances

Lorsqu'on lit dans un rapport que les entreprises américaines perdent 600 milliards de dollars par an à cause de la qualité des données, il n'est pas difficile d'imaginer que les assureurs, qui jonglent avec des montagnes de données issues de diverses sources, représentent un pourcentage non négligeable de ces pertes. Il faut compter non seulement les coûts directs liés à l'analyse et à la correction des erreurs, mais également les coûts indirects. Par exemple, lorsque des erreurs sont révélées aux clients et aux organismes régulateurs, des amendes peuvent être imposées et des demandes d'évolution de la manière dont fonctionne l'entreprise peuvent entraîner une avalanche de modifications coûteuses. Ces coûts peuvent faire oublier l'intérêt même d'un entrepôt de données, même si la liste des usages de tels systèmes est longue. Par exemple, les responsables du marketing utilisent des données client pour développer de nouveaux produits ciblés et décrocher de nouvelles parts de marché. Grâce aux données, les actuaires peuvent chiffrer le risque avec plus de précision et évaluer les provisions pour sinistres en suspens. Les agents exploitent les données pour développer et préserver les relations avec les clients.

L'environnement actuel donne à ce problème une dimension encore plus aigue. Les nouveaux canaux de distribution et la concurrence accrue sont autant de nouveaux défis que les assureurs doivent relever. Ils doivent mieux comprendre l'efficacité, la rentabilité et les interactions des canaux de distribution, pour pouvoir agir afin de les optimiser. Ils doivent également réduire la durée du cycle de développement des produits et des tarifs, des ventes, des services à la clientèle et des communications. Toutes ces opérations requièrent un accès facile à des données fiables. Le nettoyage des données doit donc être une priorité pour les assureurs.

Avant de nettoyer les données, il faut tout d'abord comprendre les sources de données de l'entreprise. Dans le secteur des assurances, ces sources ont proliféré. Les sources externes

principales sont les organisations gouvernementales, les agences d'évaluation du crédit et de déclarations des sinistres, les courtiers, les tiers fournissant des données marketing ou des informations démographiques et enfin directement les consommateurs via le web. Au niveau interne, dès que les compagnies d'assurances ont reconnu l'importance de ne plus limiter l'accès aux données qu'aux cadres supérieurs, les forces de vente et de service, les agents exclusifs et les acteurs de back-office ont commencé à multiplier la saisie, le stockage et l'exploitation de toutes sortes de données. Malheureusement, la majorité de toutes ces sources présentent un grand risque de données inexactes.

De plus en plus de compagnies d'assurances ont commencé à consolider leurs données, issues de leurs systèmes opérationnels ou de bases diverses, au sein d'entrepôts de données centralisés. L'objectif poursuivi est d'obtenir une vision des clients et de l'activité pour l'ensemble de l'entreprise. Ces opérations représentent des occasions uniques pour nettoyer les données. Au début d'un projet de consolidation des données, la majorité des entreprises mettent en place un processus d'extraction, de transformation et de chargement. La partie transformation permet de rechercher les erreurs et les problèmes dans les données entrantes et de les corriger. Le processus de nettoyage a la réputation de souvent dépasser les budgets et les plannings, mais grâce aux nouveaux outils, la maîtrise de ce type de projet a été considérablement améliorée. Par exemple, une grande compagnie a profité des progrès technologiques en matière d'entrepôt de données, pour inverser les deux dernières étapes du processus, devenu Extraction, Chargement et Transformation. Dans le cadre de ce projet de consolidation de données, cette compagnie d'assurances estime que le passage au processus d'extraction, de chargement et de transformation lui a permis de gagner trois à six mois sur le projet, tout en ramenant le pourcentage de données sales de 20 à 0 pour cent.

Les économies de temps et d'argent associées à la confiance dans la fiabilité des informations sont de solides arguments en faveur d'un projet de consolidation des données. Vu le très grand nombre de sources de données, et leur croissance constante, il est primordial de posséder un programme adéquat de gestion des données pour garantir un flux constant de données propres et fiables. Un programme de gestion des données efficace est piloté par la direction et est caractérisé par des équipes pluridisciplinaires qui regroupent des représentants des principaux secteurs techniques et commerciaux, y compris le marketing, la souscription, les ventes, les déclarations de sinistre, les services financiers, juridiques *et* l'informatique. Ces équipes garantissent non seulement que les collaborateurs comprennent leur rôle dans le maintien de la qualité des données, mais renforcent également la conscience au niveau de l'entreprise de l'importance des données historiques de façon à ce que leur utilisation soit optimisée. Les informaticiens, quant à eux, bénéficient d'une meilleure vision d'ensemble qui leur permet d'aligner entièrement leurs priorités sur la stratégie globale de l'entreprise.

Malheureusement, peu d'entreprises ont déjà mis en place un programme de gestion des données tel qu'esquissé ci-dessus.

Protection de la vie privée

Les approches marketing et commerciales bénéficient des progrès technologiques qui apportent de nouvelles possibilités de connaissance des clients, de ciblage, de personnalisation et de mesure des opérations. En même temps les entreprises cherchent à développer leurs approches à moindre coût et pour cela automatisent de plus en plus leurs relations clients. Elles mettent en place des systèmes qui fournissent aux clients des informations sur mesure, leur permettant de faire seuls des opérations de base. Ces systèmes peuvent se substituer aux personnels de vente traitant les cas ordinaires, les employés n'étant sollicités que pour les cas exceptionnels, ils peuvent aussi assister les vendeurs à exécuter des tâches complexes pour le client (par exemple visualisation 3D d'agencement de cuisine).

Pour mettre en place ce type de support, les entreprises sont obligées de gérer de plus en plus de données nominatives plus ou moins consciemment fournies par les clients. Certes les entreprises responsables demandent à leurs clients l'autorisation de les mettre en base, mais est-ce que les porteurs d'une carte de fidélité, d'un accès à un site internet, ont-ils toujours conscience que dès qu'ils utilisent leur carte ou se connectent, leurs faits et gestes peuvent être enregistrés dans une base de données nominative ? Bien sûr, chaque entreprise avec qui une personne est en relation ne connaît qu'une petite partie de ses comportements, de ses habitudes dans les hôtels, les transports, les restaurants, du profil de sa santé, de ses finances, de ses loisirs numériques, etc. Cette vue partielle la plupart du temps ne semble pas bien gênante et la loi interdit le rapprochement de fichiers entre entreprises.

Théoriquement tout est bien dans le meilleur des mondes, mais les enquêtes montrent une inquiétude latente des consommateurs. Beaucoup de personnes vivent très mal certaines pratiques agressives de communication, de recrutement sur le net, les appels trop fréquents ou les dialogues inutilement intrusifs menés par des téléacteurs à qui on a demandé de poser d'innombrables questions, qui ne semblent pas toujours en relation avec le prétexte de l'appel, ou qui vous rappellent des informations collectées lors de contacts précédents. Au final beaucoup de consommateurs craignent que leurs données finissent dans les mains d'un big brother qui pourrait porter atteinte à leur vie privée.

Le sujet doit être pris très au sérieux sur le plan de l'organisation des approches marketing ou commerciales et de la gestion du système d'information.

Gestion des données nominatives et protection de la vie privée

Toute personne a droit au respect de son intimité (domicile, image dans un lieu privé, correspondances, télécommunications, paroles prononcées à titre privé ou confidentiel, fichiers informatiques, images de vidéosurveillance si elles sont constituées en fichier nominatif). Cependant les consommateurs ont le sentiment que leur vie privée est menacée, du fait de l'évolution des technologies de l'informatique & des communications qui facilitent de nouvelles approches de gestion de la relation clients, qui exploitent de plus en plus de données pour traquer le client en temps réel (segmentation interactive, proposition en temps réel).

Des informations sur les clients peuvent être collectées via différentes sources, au delà des données liées aux transactions commerciales (distribution, banque, transport, hôtel), il est possible de suivre les activités web de quelqu'un, de le localiser via son téléphone portable, de savoir les programmes de télévision qu'il regarde ou de disposer de photos ou de films enregistrés à son insu. Au final par collecte et analyse de données, une entreprise peut savoir d'une personne ses nom, prénom, adresses, âge, CSP ; ce qu'il a : son patrimoine mobilier et immobilier ; ce qu'il voudrait : ses projets (court, moyen, long terme) ; ses relations (qui il connaît et dans quel contexte ? études, loisir, travail) ; ses charges (nombre de personnes dans le ménage, mensualités) ; ses revenus (salaires, primes, héritages) ; ses valeurs (ce qui compte pour lui, ses critères de satisfaction, ses goûts, ses loisirs) ; son évolution par rapport à un historique ; sa rentabilité passée et présente ; sa rentabilité potentielle à court, moyen et long terme.

Au-delà des indiscretions, des fuites d'information, du piratage, des données fausses, des pannes, des malveillances, des erreurs, des usurpateurs, le danger vient surtout du rapprochement de données issues de sphères de la vie privée (finance, santé, loisirs, ...) qui n'ont pas à être connues globalement par des tiers. Des processus de prêt ou d'embauche par exemple peuvent être perturbés par des rapprochements d'informations. Il est fondamental pour la vie en société que l'on puisse avoir confiance dans les systèmes de gestion des données qui doivent garantir la sécurité et la confidentialité des données nominatives.

Une organisation est responsable des données qu'elle gère et elle doit respecter les principes suivants : la collecte de données est autorisée pour des finalités définies et légitimes ; les données demandées doivent être pertinentes et proportionnées aux finalités ; dans tous les cas le client doit être informé de sa mise en fichier ; le client peut (sauf exceptions) s'y opposer ; les données ne peuvent être conservées au-delà de la durée nécessaire aux finalités déclarées ; le client a un droit de contrôle de la qualité, de l'utilisation et un droit de correction ; aucune entreprise ne peut transmettre des données nominatives concernant un citoyen de l'union dans un pays ne respectant pas la Directive Européenne.

Trois points clés sont à bien maîtriser. Le consentement du client pour la gestion des données qui le concerne (consentement implicite ou explicite). L'interdiction de gérer certaines données sensibles (race/origine ethnique, opinions politiques, religion, affiliation à un syndicat, santé physique, mentale, vie sexuelle, condamnation). Les limites à l'utilisation des données : mailing, faxing, télémarketing incessants, cross-mailing à l'intérieur d'un groupe, spam, scoring, diffusion interne d'informations nominatives, ventes de données. A noter qu'il existe des réglementations spécifiques pour certains secteurs d'activités, c'est le cas par exemple pour le secteur des Télécommunications.

Concrètement pour une entreprise il convient de : définir un responsable, publier une charte d'engagement vis à vis de ses clients, mettre en place des moyens pour notifier l'enregistrement des données et les utilisations prévues, gérer les choix restrictifs individuels, permettre de consulter, de corriger, de rendre anonyme des enregistrements, de comptabiliser les utilisations, enfin de restreindre l'accès et l'utilisation des données nominatives, notamment en matière d'analyse de données, de décision automatisée, de marketing direct et de transmission à des tiers.

Les entreprises sont fortement incitées à respecter les règles de protection de la vie privée par les dispositions pénales suivantes :

- Utiliser sans autorisation le Répertoire national d'identification des personnes physiques : 5 ans et 300K€
- Entraver l'action de la CNIL : 1 an et 15K€
- Ne pas respecter l'autorisation préalable pour les traitements d'informations nominatives : 3 ans et 45K€
- Manquer de précaution pour le traitement de données nominatives : 5 ans et 300K€
- Utiliser un moyen frauduleux, déloyal, ou illicite pour la collecte de données nominatives ou utilisation malgré l'opposition de la personne : 5 ans et 300K€
- Conserver certaines données (origines raciales, opinions politiques, ...) : 5 ans et 300K€
- Dépasser la durée de conservation prévue lors la demande de création du système : 3 ans et 45K€
- Détourner les informations de leur finalité : 5 ans et 300K€
- Porter atteinte à la considération de quelqu'un ou à l'intimité de sa vie privée : 1 an et 15 K€
- etc..

Pour faire cesser l'atteinte à sa vie privée (en cas d'urgence) il convient de saisir le juge des référés. Pour accéder aux fichiers, il faut demander directement à l'organisme en question, sauf dans quelques cas (via la CNIL pour les RG, via la CNCIS pour les écoutes téléphoniques, via un médecin pour le dossier médical). Pour faire réparer un préjudice, il faut demander des dommages et intérêts auprès du tribunal de grande instance (assistance d'un avocat obligatoire). Enfin vous pouvez toujours demander à la CNIL de vous aider en cas de doute sur la démarche.

Il existe des solutions pour intégrer les informations nécessaires à la gestion de la confidentialité dans les modèles de données des systèmes décisionnels. Les moyens à mettre en œuvre doivent permettre aux différents acteurs de l'entreprise de n'accéder qu'aux données clients qui leurs sont nécessaires pour leurs activités, et de ne disposer que de vues restrictives, le plus souvent anonymes.

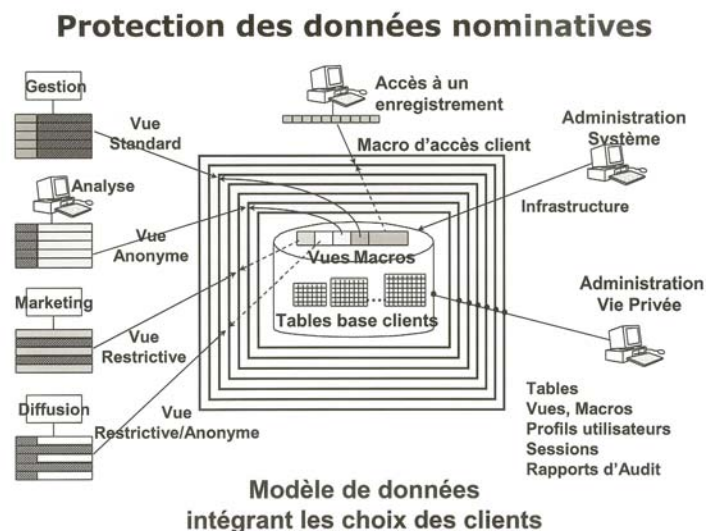


Illustration - 27

Les entreprises doivent investir pour gérer les données nominatives et protéger la vie privée de leurs clients, cependant il ne faut pas voir cela comme une contrainte, mais plutôt comme

une opportunité de développer, de renforcer la relation avec le client en améliorant les services et la confiance, et d'affiner la connaissance du profil client avec son accord, dans le cadre de programme de fidélisation.

Concilier marketing personnalisé et protection de la vie privée : de big brother à big mother

Un environnement de confiance, de confidentialité et de sécurité est indispensable pour les affaires. Or depuis quelques années, les nouvelles technologies de l'information favorisent la collecte, le stockage, l'exploitation, la diffusion d'une masse sans précédent d'informations nominatives, et le développement d'approches marketing individualisées impliquant la mobilisation de toutes les ressources des entreprises, pour optimiser les relations clients au sein d'un éventail toujours plus riche de canaux de distribution (internet, force de vente, centre d'appels, publipostage, etc.). Tout ceci met en question la protection de la vie privée des consommateurs, et ce point est devenu particulièrement crucial aujourd'hui avec les nouvelles approches de marketing événementiel, qui visent à interagir en temps réel avec le client où qu'il soit, et amènent les entreprises à largement partager en leur sein des informations clients.

Le Marketing se renouvelle constamment et au-delà d'une créativité certaine en matière de vocabulaire, de nouvelles approches sont réellement expérimentées comme le Permission Marketing, le Knowledge Marketing, le Community Marketing, le Customer Empowerment ou le Marketing Collaboratif pour n'en citer que quelques unes. Toutes ces approches se fondent sur de nouveaux usages des technologies de l'information et des communications qui permettent entre autre de connaître ses clients avec plus de précision en facilitant l'obtention et la gestion d'informations clés qui offrent des opportunités d'analyses et d'actions pour les professionnels du marketing !

Cependant, il faut savoir résister au syndrome de Big Brother. D'abord parce que la législation actuelle ne permet pas tout, et que les risques en cas de dérive sont réels pour l'entreprise : amendes et peines de prison sont prévues par la loi. La rigueur voulue dans ce domaine par la Communauté européenne, ainsi que les projets de loi en cours d'élaboration obligeront à terme à encore mieux protéger les citoyens, et donc tout cela nous invite à prendre le sujet de la gestion des données clients avec sérieux. Surtout que malgré toutes les précautions prises, les entreprises ne sont jamais à l'abri de problèmes. Mais, au-delà des obligations légales et des risques, il faut bien avoir conscience des attentes des clients. En effet dans les approches de connaissance approfondie des clients, il est facile de se retrouver plus dans une logique d'enquête brutale, que de coopération intelligente, et il est toujours très difficile, une fois le client brusqué, de recréer une relation de confiance.

L'intérêt de l'entreprise va bien au-delà de ses responsabilités légales ou de la gestion des risques, et il doit être possible de transformer en opportunité relationnelle, le souhait légitime des consommateurs de protéger leur vie privée, de maîtriser les communications qui leur sont adressées, tout en étant des acteurs engagés du marché, en participant par exemple à l'innovation produit. Il s'agit alors de mettre en place une forme d'approche que l'on pourrait qualifier de "big mother". Elle permet d'être à l'écoute du client, d'affiner la connaissance de son profil avec sa participation, de ne pas se limiter à mettre en œuvre les protections prévues, mais de prendre l'initiative d'aller plus loin en fonction des spécificités du métier de l'entreprise, et de garantir ainsi pour longtemps la gestion de relations intimes

avec les clients. Ce n'est que dans ces conditions qu'une relation commerciale intelligente, constructive, peut être tissée et gérée grâce aux possibilités des outils technologiques actuels.

Pour pouvoir développer de nouvelles approches Marketing telles qu'évoquées ci-dessus, il convient pour une grande entreprise de disposer de moyens technologiques appropriés qui doivent :

- Fournir une vue complète du client, une vue unique et consolidée de toutes les activités clients dans toute l'entreprise.
- Mettre à disposition un ensemble d'outils d'analyse (du tableau de bord à la fouille de données) pour les différents acteurs concernés (des statisticiens, aux responsables d'activités en passant par les analystes métier).
- Gérer des campagnes coordonnées au niveau de toute l'entreprise, des plus simples au plus complexes en plusieurs étapes, par de multiples canaux (des plus traditionnels au web 2.0 en passant par les réseaux sociaux ou les sms).
- Gérer les offres, en tenant compte des comportements et des cycles de vie.
- Gérer les interactions, en fournissant des renseignements au point d'interaction avec la clientèle, en facilitant des prises de décision en temps réel, en automatisant des réactions.
- Optimiser les communications, en ordonnant et adaptant toutes les communications client à travers tous les canaux.
- Gérer les ressources marketing, en supportant tous les processus de gestion des actifs et des moyens du marketing.

Nouvelles données gérées par les systèmes décisionnels

L'emploi de systèmes décisionnels ne cesse de se développer ces dernières années. Non seulement ils sont mis en œuvre par toute les fonctions de l'entreprise, mais au-delà de leur vocation à supporter les décisions stratégiques, depuis quelques temps ils sont aussi utilisés pour supporter des décisions opérationnelles. Aujourd'hui alors que la plupart des systèmes décisionnels traitent des données correspondant aux objets traditionnels de gestion de l'entreprise du type, commande, facture, ordre de fabrication, bon livraison, etc., certains prennent déjà en compte de nouvelles données comme par exemple des données IP, des données GPS et des données RFID.

Données IP

Une directive européenne récente a imposée de nouvelles exigences en matière de rétention des données de communication. La directive stipule que les opérateurs doivent conserver toutes les données liées aux communications électroniques qu'ils génèrent ou traitent, et être en mesure de les mettre à disposition des organismes définis dans le cadre d'enquêtes et d'investigations concernant des crimes graves. Il s'agit en particulier de pouvoir retrouver les appels passés sur et vers des téléphones fixes ou mobiles, SMS, MMS, WAP, les appels VoIP, les courriels, les messageries instantanées, les connexions internet. Ces données capturées et archivées par les opérateurs doivent pouvoir servir de fil d'Ariane et être exploitées pour identifier des activités illégales.

Ces données sont à conserver de six à vingt quatre mois ce qui représente des téraoctets de données à gérer. Concrètement il s'agit de mettre en place un entrepôt de données capables de capturer et restituer les données sans délai pour analyse. Au-delà des volumes, tout ceci pose des défis techniques tant en termes de qualité des données, de sécurité, de célérité des processus, de traçabilité, de confidentialité et de complexité des analyses à réaliser qui doivent permettre de relier des données de télécommunications commutées à des communications sur IP du même abonné, de la même adresse ou du même destinataire.

Données GPS

Par exemple un assureur britannique a lancé une police d'assurance automobile « Pay As You Drive » basée sur l'utilisation du véhicule. A partir des données recueillies sur les comportements de conduite, cet assureur britannique propose une police qui facture le coût de l'assurance à partir de 1 penny seulement par mile (1,6 km).

Des systèmes GPS embarqués permettent à l'assureur de relever le comportement de chaque automobiliste. Les clients reçoivent des factures mensuelles basées sur l'utilisation de leur véhicule, notamment en fonction des plages horaires, du type de route et du kilométrage, ce qui constitue une innovation dans le domaine de l'assurance automobile. Les factures ressemblent donc à celles des opérateurs de téléphonie mobile, les primes d'assurance étant calculées et totalisées pour chaque déplacement. Cette approche transparente en matière d'assurance automobile permet aux clients de contrôler le coût de leur assurance d'une façon qui était jusqu'ici impossible.

Données RFID

La RFID apporte un nouveau standard d'identification des objets de gestion et automatise l'interface avec eux. Pour les produits par exemple on peut mettre en place une identification individuelle de chaque article, et enregistrer automatiquement tous les événements de leur vie de leur naissance à leur destruction. A noter que l'on peut mettre des puces RFID en bien d'autres endroits que sur des produits, sur des conditionnements (palettes, containers), sur des cartes (paiement, fidélité, badge). Les applications sont très nombreuses en particulier dans les domaines du paiement, du contrôle d'accès, de la sécurité et de la logistique, mais aussi pour toutes sortes d'opérations commerciales, industrielles ou relatives aux personnes.

Concrètement par exemple pour le secteur de l'industrie et en restant au niveau du produit, l'utilisation de la RFID est de nature à impacter fortement les domaines fonctionnels suivants : gestion des stocks, suivi des procédés de production, contrôle qualité, suivi des transports et de la logistique, carte d'identité produit, traçabilité du produit, dossier de la vie du produit & maintenance, lutte contre le vol et la contrefaçon, gestion des retours et des garanties, etc. En effet l'accélération de la saisie et de la circulation de l'information apportée par la RFID associée à un entrepôt de données actif permet de réinventer, d'optimiser les processus de gestion des produits et de réduire la charge de travail nécessaire.

Progiciel, entrepôt de données et RFID pour gérer les processus de demain

Aujourd'hui les grandes entreprises sont occupées à rationaliser leurs systèmes d'information dans toutes leurs entités, elles sont donc engagées depuis quelques temps dans la réduction de la variété des applications qu'elles exploitent et dans la généralisation de l'utilisation de progiciels. L'objectif est de réduire les coûts et d'harmoniser les processus. En mettant en œuvre les mêmes progiciels elles finiront par neutraliser l'apport des NTIC en matière d'avantages concurrentiels. Pour s'aider avec son SI à prendre la tête de la compétition, à court terme il y a le domaine du décisionnel où il est possible de mettre en place des solutions qui permettent de faire la différence, d'autant plus que de nombreuses sociétés persistent dans l'impasse des data marts. A long terme il y a la technologie RFID qui va permettre de faire évoluer les processus, et ouvrir de nouvelles possibilités d'analyses et donc d'optimisation.

Dans un premier temps la RFID peut être simplement utilisé pour automatiser la saisie des événements sans changer les processus, mais très rapidement les processus sont remis en cause et les systèmes gèrent en continu des flux à un niveau beaucoup plus détaillés qu'aujourd'hui car il n'y a plus les limites liées à la saisie de l'information. Dans ces conditions nous passons d'un mode « photo » par exemple l'état du stock, à un mode « film » le suivi de l'objet de gestion aux différents stades de sa vie. Les quantités d'informations sont importantes, les décisions peuvent être prises en quasi temps réel et le concept d'entrepôt de données actif prend tout son sens.

L'utilisation de la RFID a déjà commencé, par exemple la plupart des entreprises de la Distribution qui font plus de 3 Milliards de Dollars de chiffre d'affaires ont des projets de RFID en cours ou prévus qui impliquent certains de leurs fournisseurs. Le principal avantage de cette technologie pour les fabricants et la grande distribution actuellement envisagé est l'amélioration de la visibilité de la chaîne d'approvisionnement. Mais les experts s'accordent à penser que sauf exception, dans les deux ans à venir 80% des expériences dans la distribution consisteront à mettre en place des moyens au niveau des conditionnements et dans des circuits fermés impliquant quelques partenaires sélectionnés. La limitation de ces

expériences est due aux principaux défis de la RFID que sont le coût et le manque de normalisation des éléments entrant dans ce genre de solutions ? Ce sont aujourd'hui les grands distributeurs (Wal Mart, Metro, Tesco) qui sont les plus moteur sur le sujet et qui poussent à la mise en place de standard universel.

Possibilités du web analytics

Ces dernières années de nombreuses entreprises se sont mises à collecter et analyser des données sur l'utilisation de leurs sites web, ce type d'analyse est connu sous l'acronyme de Web Analytics. Les premières mesures visaient surtout à optimiser les sites et leur utilisation par les visiteurs, mais rapidement ces données ont été rapprochées d'autres sources d'information et intégrées aux différentes approches analytiques qu'elles pouvaient enrichir (par exemple : analyse des ventes, des promotions, de la logistique, de la dynamique multi-canal, de la fidélité, de l'attrition). En particulier ces données permettent de renforcer l'ambitieuse approche d'une vision à « 360° » des clients.

Les technologies utilisées sont diverses, analyse des logs, pose de marqueurs ou analyse de trames réseau. Il s'agit de générer des indications variées comme le nombre de visiteurs, de visites, de pages & de bannières vues, les enchainements ou les recherches effectués, etc. Mais au-delà de données de navigation anonyme, on peut aussi suivant les cas collecter des données d'identification (adresse IP, cookies, e-mail), et des données relatives à des opérations, par exemple d'enregistrement ou d'achat. Au final selon le niveau de détail des données gérées, on peut réaliser des analyses plus ou moins fines de l'impact d'un site en matière de communication (audience, contenu, durée, ...), ou d'opération de recrutement, de vente (directe, indirecte via d'autres canaux), etc. On peut aussi gérer des actions interactives comme afficher en temps réel à chaque visiteur des contenus adaptés (communication, promotion, tarification, ...), ou des actions en temps différé via le canal internet (e-mail, bannière et offre pour la prochaine visite), ou classiques (centre d'appels, agence ou magasin).

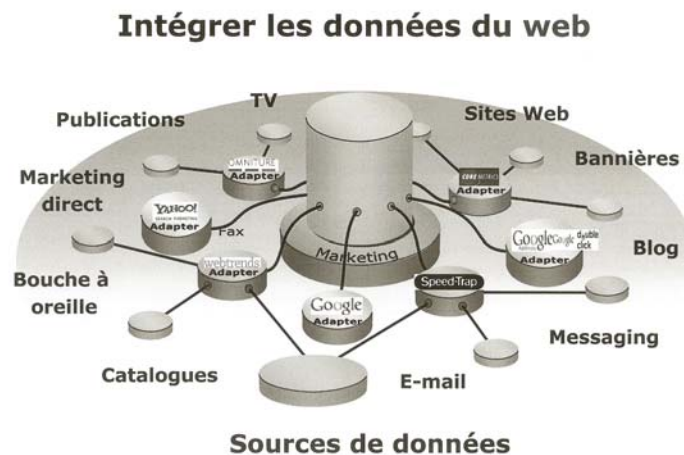


Illustration - 28

Au niveau des moyens, au-delà des indispensables entrepôts de données et outils de BI ou de data mining habituels à tout système d'information décisionnel, il convient de mettre en œuvre des moyens spécifiques de web analytics, comme ceux fournis par WebTrends,

SpeedTrap, Coremetrics, Omniture, Double Click ou Google dont la solution « Analytics » est en fort développement en concurrence ou en complément des autres offreurs cités. A noter que ce domaine est en pleine effervescence et que nombreuses nouvelles solutions apparaissent régulièrement.

De nombreuses entreprises ont déjà mis en place des moyens de Web Analytics, mais certains sont cependant plus en pointe sur ce sujet comme e-Bay, Travelocity, Overstock.com, Wells Fargo Bank, ABN Amro ou Paypal. Sans entrer dans le détail, ces organisations développent tout ou partie des actions ci-dessous :

- Des e-mailings de promotion sont créés qui tiennent compte de toutes les informations disponibles sur les clients, y compris tous les éléments de leur activité web (segmentation, optimisation des campagnes, ...).
- 2 secondes après qu'un client se soit identifié sur un site, la meilleure offre parmi des dizaines peut être sélectionnée et présentée pendant la session en cours (merchandising, bannières, ...).
- Des e-mailings sont générés avec des propositions spéciales en fonction des navigations web des clients qui la veille ne sont pas allés au bout de leur transaction et ont abandonné leur panier (campagne personnalisée, ...).
- Les dernières informations de navigation web sont mises à disposition du centre d'appel, ou du point de vente pour compléter la connaissance du client, déclencher ou orienter les dialogues ou propositions (alerte, coordination multicanal, support à la décision, ...).

Les activités de web analytics (métriques web, optimisation de campagnes, e-marketing, intégration multicanal, management de la performance) permettent d'obtenir d'excellents résultats. Comme, une multiplication par 3 des taux de réponse aux campagnes marketing intégrant ce type de données ; une augmentation de 90% des taux de succès sur appels entrants ; une multiplication par vingt des ouvertures de pages dans le cas de pages personnalisées en temps réel ; une augmentation de 100% de la valeur des paniers lors des « campagnes de masse personnalisées ». Mais attention à ne pas être uniquement obnubilé par les résultats et à mettre en œuvre ces approches avec finesse, car les clients n'aiment pas être suivis, espionnés, voir trop explicitement compris. D'autre part il convient de bien tenir compte des règles de protection de la vie privée, lors de la mise en place d'une base de données nominatives très exhaustives.

Compléter votre entrepôt de données d'entreprise avec des données géo spatiales

Dans leur système d'information décisionnel les entreprises stockent des données relatives à leurs activités, par exemple marketing, vente, production, achat ou logistique. Les données disponibles permettent d'analyser diverses dimensions de ces activités comme le qui, le quoi, le quand, mais généralement le où n'est appréhendé au mieux que via une donnée de type code postal. Or fonder ces analyses sur des données regroupées via un code postal limite fortement les possibilités, et notamment ne permet pas de faire par exemple des calculs de distance précis.

Pour aller au-delà du code postal, de nombreuses entreprises tant privées que publiques utilisent des données spatiales pour enrichir d'une réelle dimension géographique l'approche analytique de leur métier. L'utilisation la plus répandue concerne la réflexion sur les implantations, par exemple de points de vente, de distributeurs de billets dans le secteur

bancaire ou d'antennes de télécommunications. Il s'agit d'intégrer des informations sur des zones géographiques pour calculer un potentiel localisé qui tient compte de l'éloignement de différentes zones, et d'adapter la présence en fonction des ambitions et des moyens envisagés.

Pour ce faire les entreprises doivent gérer des données numériques relatives à des territoires, ces données doivent être ancrées dans un système de coordonnées géographiques ou de projection bien défini, de sorte qu'on puisse les situer sur une portion du globe terrestre. Il s'agit de données qui concernent des entités physiques, des activités ou des événements. Mais savoir quelles informations sont disponibles et pouvoir y accéder, sont deux des plus gros obstacles à l'utilisation de données géo spatiales existantes. En effet il est souvent difficile de trouver des données géo spatiales de qualité, c'est-à-dire adaptées aux besoins de l'entreprise, en termes d'accessibilité, de précision ou de fréquence de mise à jour. Les niveaux de granularité des découpages administratifs et des données statistiques officielles qui y sont associées ne répondent parfois pas à certains types de besoins d'analyse du territoire (analyse de marché, micro marketing).

Si nous prenons le secteur de la distribution, des données géo spatiales peuvent être utilisées pour conserver les clients fidèles et acquérir de nouveaux clients, en améliorant la commercialisation, l'assortiment et la tarification sur la base d'une meilleure compréhension de la localisation des clients, la proximité des concurrents et d'autres données sur la zone concernée (démographie, urbanisme, ...). Concrètement il s'agit de réaliser divers types d'analyses come : identifier les clients localisés dans une zone à risque par rapport à un concurrent actuel ou futur ; analyser la distance, le temps de déplacement pour se rendre dans un magasin et calculer en fonction des données démographiques de la zone la part de marché potentielle ; analyser l'impact des actions des concurrents positionner dans la zone, par exemple évaluer l'impact des variations des horaires d'ouverture sur les différentes zones d'attraction. Globalement toutes ces analyses doivent faciliter des prises de décisions fondées sur la connaissance de la situation compétitive de la zone, des possibilités de 1 : 1 marketing, et permettre de mieux cibler l'assortiment à la lumière du lieu et du positionnement des concurrents, de travailler de manière proactive les meilleurs clients, d'augmenter la marge grâce à un marketing plus ciblé et une tarification ajustée de façon plus efficace, ou enfin d'identifier les magasins qu'il convient de remodeler compte tenu du potentiel, voire de fermer.

Tous les secteurs industriels et commerciaux sont concernés par l'utilisation de données géo spatiales pour répondre à des questions clés qui les concernent, comme dans la Banque (Où implanter une nouvelle agence ?), Assurance (Quels engagements avons-nous dans telle zone sinistrée ?) Transport (Comment puis-je répondre rapidement aux demandes d'expédition des clients ?), ou Télécommunications (Quels sont les clients qui pourront bénéficier de ce nouvel équipement réseaux ?).

Récemment ont été développées des solutions qui permettent de convertir les données de localisation basiques d'un entrepôt de données, en coordonnées communes de latitude et de longitude, un processus connu sous le nom de géocodage. Une fois converties, les données géo spatiales peuvent être intégrées et analysées conjointement avec les autres données de l'entrepôt et ainsi offrir de nouvelles perspectives d'analyse. La possibilité de réaliser des analyses au sein de l'entrepôt de données permet d'éliminer le recours à des data marts spécifiquement dédiés aux données géo spatiales, évitant ainsi des mouvements de données,

des redondances, des intégrations de systèmes, et la production d'information commerciales imprécises.

Il convient donc d'encourager les entreprises à intégrer leurs données géo spatiales dans leur entrepôt d'entreprise. Il s'agit de tirer profit des capacités géo spatiales de leur base de données et de solutions spécialisées comme FME de Safe Software. FME permet aux utilisateurs de traduire, transformer, intégrer et distribuer rapidement des données spatiales, à partir de plus de 225 formats différents, de manipuler ces données avec le modèle de données requis et de partager ces données spatiales via le Web.

Rapprocher les mondes de la technique et des affaires

Depuis quelques temps, nous voyons de plus en plus d'opérateurs de Télécommunications intégrer des données techniques et des données commerciales pour mieux connaître leurs clients. Le marché des Télécommunications se sature et pour augmenter leur chiffre d'affaires, les opérateurs sont créatifs et lancent de nouvelles offres combinant de façon innovante des matériels, des applications et des réseaux. D'autre part les opérateurs souhaitent prioritairement nouer des relations privilégiées avec les bons clients, ils cherchent donc à les identifier et à leur proposer des services de qualité. Dans cette perspective la connaissance détaillée des expériences des clients est fondamentale. Concrètement pour les opérateurs il s'agit de définir les matériels ou les services les plus appréciés, de savoir ce que chaque client fait, les services qu'il utilise le plus, le chiffre d'affaires engendré, la rentabilité procurée, et de bien situer l'impact des problèmes de qualité de réseau (perte de revenus, voir attrition).

Pour avoir une vue d'ensemble de la performance et de la qualité de service d'un réseau, il convient de se référer aux données de signal qui permettent de recenser toutes les activités d'un réseau, et pour obtenir ses données il faut mettre en place des moyens de mesures spécifiques tels que ceux proposés par la société Agilent qui est le leader mondial de ce domaine. Pour obtenir la vision métier et technique des expériences client (communications réussies ou ratées : rupture de ligne, appels infructueux, ...), il faut intégrer des données de signal dans un entrepôt de données d'entreprise.

Agilent a signé des partenariats stratégiques afin de proposer aux opérateurs de Télécommunications une vue unique sur leurs propres clients. Une feuille de route pour renforcer et approfondir l'intégration technique a été fixée, et une équipe a été mise en place pour promouvoir les meilleures pratiques d'utilisation des solutions des partenaires. Dans ces partenariats, Agilent apporte sa connaissance approfondie des réseaux, et ses solutions qui produisent quotidiennement des quantités très importantes de mesures ; les partenaires fournisseurs de plateformes décisionnelles offrent la possibilité d'obtenir une vue unique du réseau et des clients, en permettant l'intégration de toutes les données nécessaires, habituellement réparties.

Au final ce type de proposition conjointe permet d'un point de vue financier d'obtenir une plate-forme hautement évolutive pour un TCO très réduit, du fait de la rationalisation apportée par la consolidation des Data Marts, et d'un point de vue métier réalise l'intégration des données clients avec des données du réseau, ce qui augmente les possibilités d'analyser et de connaître les clients, de développer des programmes d'amélioration de la satisfaction des clientèles et de réduction des désabonnements.

Propos sur le développement d'un système décisionnel

Beaucoup d'entreprises se laissent submerger par les données internes et externes qu'elles utilisent. Elles ne savent pas comment créer une infrastructure analytique adéquate qui leur permet de convertir des données en information, des informations en opportunités et des opportunités en actions. Pourtant le domaine du décisionnel est mature et un grand nombre de spécialistes savent très bien ce qu'il faut faire pour aligner les collaborateurs, les processus et les moyens informatiques avec la stratégie, pour construire une plate-forme matérielle et logicielle rentable, évolutive, pouvant répondre aux besoins présents et futurs, pour identifier et concevoir des applications adaptées au secteur d'activité, pour tirer profit de la croissance explosive des données dans l'entreprise, pour proposer une solution globale adaptée à un budget donné, pour mettre en place une prise en charge technique permanente, enfin pour assurer que cette solution satisfait les attentes en matière de ROI.

Par exemple pour développer de bons outils d'aide à la décision il convient de bien maîtriser toute la chaîne de valeur du projet, depuis la définition des enjeux pour l'entreprise jusqu'à la traduction opérationnelle sur le terrain de la mise en œuvre des solutions, et la réalisation du retour sur investissements. Les facteurs clés de succès d'une telle démarche sont surtout la pluri-compétence des équipes, en termes de connaissance des métiers ou des technologies associées, et la capacité à gérer un projet, en respectant le périmètre, le budget ou les contraintes de délais. Il s'agit en particulier de maîtriser les risques, d'obtenir rapidement des résultats "métier", de fonder son système sur des solutions techniques permettant une expansion, de limiter l'investissement initial et enfin de chercher à maximiser le ROI

Au-delà des aspects techniques il convient aussi de savoir conduire les changements induits par l'introduction de nouveaux moyens d'aide à la décision. Pour cela il faut comprendre les changements et les enjeux associés, anticiper les impacts, les moteurs, les freins, les leviers et compléter les traditionnels plan d'administration du changement par des actions dynamiques d'implication et d'incitation.

Comme toujours en matière de nouvelles technologies, les entreprises qui débutent n'ont pas intérêt à trop déléguer à des tiers les études et la réalisation d'un nouveau système, mais elles doivent s'impliquer fortement, et pour beaucoup faire avec leurs propres équipes quitte à les renforcer et en particulier à intégrer des conseillers dans le groupe de pilotage. Ceci est vrai pour la réalisation d'un grand projet comme pour l'introduction de nouveautés en matière d'architecture comme la SOA, ou en matière de méthodes comme les approches RAD par exemple.

Entrepôt de données & SOA

Depuis quelques temps les éditeurs de progiciels ont trouvé la solution miracle pour pallier à leur incapacité à répondre à l'ensemble des besoins des entreprises : la SOA, l'architecture orientée services. Une des promesses est de pouvoir assembler entre eux des composants issus d'applications différentes pour supporter au mieux les processus métier de l'entreprise. Actuellement les réalisations sont peu nombreuses, et les possibilités offertes par certains éditeurs se limitent la plupart du temps à assembler des éléments de leurs propres offres et de

quelques partenaires. Cependant des réalisations montrent que la SOA peut être facilement mise en œuvre pour faire dialoguer en temps réel un système décisionnel et des systèmes opérationnels hétérogènes, auxquels sont apportés des services d'aide à la décision de différentes natures, par exemple un simple accès à un historique, un scoring à la volée ou un ensemble de services pour gérer un type d'événements.

Les experts en systèmes décisionnels sont particulièrement intéressés par ces approches et font déjà depuis plusieurs années la promotion du concept d'entrepôt de données actif (ADW en Anglais). Un tel système doit offrir les caractéristiques suivantes : une alimentation active de certaines données (temps quasi réel & batch) ; une gestion active des événements (analyse de l'activité opérationnelle, déclenchement automatisé d'actions grâce à des services applicatifs) ; une intégration active au système d'information (interopérabilité avec le reste du SI) ; des possibilités d'accès actifs opérationnels (interactif ou automatisé) ; une gestion active de la charge (gestion dynamique des ressources, charges mixtes) ; et enfin une disponibilité active (continuité de service opérationnel jusqu'au 24x7x52).

Un ADW, c'est-à-dire un entrepôt de données dans une architecture orientée services, permet de développer avec un retour sur investissement rapide et important, des applications différenciatrices spécifiques au cœur du métier des entreprises. Par exemple de façon non limitative, dans la Banque de détail des services interactifs sur GAB ; dans le Transport de colis du Track & Trace, de l'optimisation des trajets ou des chargements ; dans la Distribution le déclenchement de promotions en temps réel selon les volumes de ventes, de stocks ou l'analyse des tickets ; dans le Manufacturing avec l'utilisation de RFID (Radio Frequency Identification) la gestion de nouvelles interactions bidirectionnelles avec les produits (pour gérer la logistique, la qualité, ...) ; dans l'Assurance le développement des nouvelles offres Pay As You Drive qui implique une facturation à l'usage à partir de GPS installés dans les véhicules des clients.

Un autre usage de la SOA et d'un entrepôt de données est de mettre en œuvre des services de gestion des données de référence (MDM en Anglais), au bénéfice des systèmes opérationnels et décisionnels. Il s'agit dans ce cas de mettre en place une application à même d'assumer les fonctions suivantes : gestion des métadonnées (modèles, règles, gouvernance) ; workflow des données MDM (publication / extraction, temps réel & batch, événementiel) ; gestion de la qualité et référencements croisés ; réconciliation de formats, tables de correspondance. Toutes les entreprises peuvent tirer bénéfice d'une solution de MDM dans l'entrepôt de données, en effet quand les entreprises doivent regarder dans de nombreux systèmes pour trouver leurs données de référence, puis doivent réconcilier ces données, les processus (par exemple de création d'un nouveau client ou d'introduction d'un nouveau produit) peuvent devenir complexes, erratiques et sources d'erreurs. Sans parler des exigences de qualité et de traçabilité des données dans le cadre d'analyse, de reporting et d'audit de conformité pour certaines réglementations.

Pour en revenir à la SOA en général, il est clair qu'elle est une voie d'avenir pour aligner les processus, les collaborateurs et les moyens informatiques avec la stratégie d'une entreprise. Elle peut permettre de construire une plate-forme matérielle et logicielle, rentable, évolutive, pouvant répondre aux besoins organisationnels présents et futurs. Elle facilite la mise en place d'applications très adaptées aux secteurs d'activité, et permet de tirer profit de la croissance explosive des données dans l'entreprise en valorisant notamment les gisements de données historiques.

De grandes entreprises comme 3M, Continental Airlines, Wal*Mart, Ping, Safeway, AT&T, London Stock Exchange peuvent représenter autant de sources d'inspiration pour mettre en œuvre des approches SOA dans le cadre d'un entrepôt de données.

Entrepôt de données & développement rapide d'applications (RAD)

Il ne suffit pas de disposer d'une base de données pour garantir la mise en œuvre efficace d'un entrepôt de données ou pour activer l'intelligence d'une entreprise. D'autres moyens sont nécessaires et en particulier pour une première approche du décisionnel il est recommandé de s'appuyer sur un partenaire expérimenté. Aujourd'hui les méthodes de planification, conception, implémentation, support, optimisation d'un entrepôt de données sont bien rodées, elles permettent un travail commun efficace des maîtres d'ouvrage et des maîtres d'œuvre.

Cependant il convient d'avoir conscience qu'un projet d'entrepôt de données d'entreprise doit assumer les contraintes inhérentes à toute démarche qui mobilise sur un objectif commun des personnes venant de différents horizons. La lourdeur, la lenteur, la production tardive de résultats - le fameux effet tunnel - ne sont malheureusement que trop régulièrement constatés. Il faut en outre compter avec les difficultés spécifiques de l'étape primordiale d'évaluation des besoins. En effet en décisionnel il est toujours difficile de définir précisément les besoins, qui sont dans certains domaines souvent peu matures, au mieux estimés, évolutifs, voire éphémères. Dans ces conditions la planification d'étapes de développement successives s'appuyant sur un cahier des charges détaillé s'avère parfois aléatoire.

Si pour la création des fondations d'un entrepôt de données d'entreprise une démarche très rigoureuse (méthodologie, plan, équipes structurées, normes industrielles, contrôles) s'impose, pour son développement ultérieur des démarches RAD (Rapid Application Development) peuvent être très intéressantes. La méthode de développement rapide d'applications permet en effet de sortir du carcan des méthodes classiques dite « en cascade » centrées sur la réalisation d'un plan global, pour proposer de nouvelles approches fortement participatives, itératives, incrémentales, adaptatives et pilotées par la valeur.

Dans une approche classique le projet est découpé en briques élémentaires et chaque élément fait l'objet d'une définition, d'une estimation des ressources et du temps nécessaires par des maîtres de l'art, le tout étant fixé dans un planning d'ensemble intégrant les contraintes d'ordonnancement. Il s'agit de trouver la meilleure organisation possible qui garantisse la qualité, permette de terminer le projet dans les meilleurs délais et de maîtriser les tâches critiques, c'est-à-dire les tâches qui ne doivent souffrir d'aucun retard sous peine de retarder l'ensemble du projet.

Dans le cas d'une méthode RAD il s'agit de développer par itérations successives, planifiées au fur et à mesure de l'avancement, une solution dont le périmètre final sera fonction des moyens, des savoirs faire et du temps accordés à chaque itération. L'un des intérêts de la méthode est d'impliquer très fortement les utilisateurs lors de toutes les phases et pas seulement celles de conception et de recette. Il convient cependant de noter que les méthodes RAD sont plus utilisées pour le développement d'un prototype, d'une application pilote ou pour la réalisation d'un test que pour la mise en place d'une application de production.

Dans le cas des entrepôts de données les méthodes RAD sont très utiles pour répondre à la demande d'agilité, demande qui croît avec la maturité du domaine décisionnel. Les méthodes classiques et les bonnes pratiques durcies par des années d'expérience ne permettent pas toujours d'être suffisamment réactif, de piloter de façon flexible les projets d'évolution. Le cadre souple des méthodes RAD est plus adapté pour des projets innovants par exemple, il offre plus de possibilités de se rendre compte rapidement de la validité d'une idée. Il vaut mieux se rendre compte qu'un projet est voué à l'échec après un mois, plutôt que quinze mois plus tard.

Lorsqu'un entrepôt de données a atteint une certaine maturité, il est fréquent de créer des espaces privatifs ou bacs à sable d'essais taillés sur mesure pour des utilisateurs finaux ou des informaticiens, pour tester des données, des outils de chargement ou de restitution, réaliser des prototypes, des applications pilotes ou des services bureau temporaires. Dans le cadre de ces bacs à sable décisionnels, les méthodes RAD sont utilisées de façon quasi systématique. Ainsi par exemple il est possible d'intégrer de nouvelles données en complément de celles gérées par le système décisionnel existant, de réaliser toutes les approches analytiques imaginables des plus simples aux plus complexes sur toutes les données des systèmes décisionnels de l'entreprise permanents ou temporaires, et de faciliter le test de certains choix de conception ou de moyens. A noter cependant que les démonstrations étant faites, l'expérience recommande de développer de façon classique et industrielle toute solution devant être mise en production durablement.

Au-delà de la flexibilité, l'agilité apportée par un centre d'expertise

« La seule chose qui ne changera jamais est que tout change toujours tout le temps ». Cette pensée extraite du Yi Jing, un texte Chinois vieux de 35 siècles, explique l'intérêt que les entreprises portent aujourd'hui à l'analyse de données, vue comme un des moyens d'appréhender les changements. Les entreprises doivent continuellement intégrer les évolutions de leurs clients, partenaires, concurrents et autres parties prenantes. Les clients sont de plus en plus exigeants, les concurrents innover, les financiers veulent plus de sécurité, les actionnaires plus de rendement, et la société dans son ensemble a des attentes plus grandes en matière d'environnement propre et hygiénique, de pratiques éthiques ou de respect de valeurs morales.

Pour comprendre et répondre à l'évolution des demandes, l'entreprise doit être au minimum ouverte et souple. Sa performance dépend dans ce contexte de sa capacité à adapter sa stratégie, ses structures, ses processus d'affaires et ses systèmes d'information. Il ne s'agit pas simplement d'optimiser la productivité à court terme, en se projetant, en planifiant, mais d'être réactif, voir d'être pro-actif. En fait au-delà de la flexibilité, l'entreprise doit être agile, c'est-à-dire capable de saisir des opportunités et d'innover. L'agilité est un gage de pérennité face aux évolutions. Pour le long terme, seule la capacité d'innovation de l'entreprise à tous les niveaux peut lui assurer une croissance durable. Thomas Edison a dit « il faut faire une innovation mineure tous les dix jours et une majeure tous les six mois ». L'exigence est élevée, mais l'entreprise ne peut se contenter d'une innovation ponctuelle, il faut au-delà de rares évolutions radicales, accumuler constamment les nouveautés pour transformer profondément une entreprise et la garder en phase avec son environnement.

Flexibilité réactive ou agilité créative, concernent tout le monde dans l'entreprise. Elles se mesurent en termes de vitesse et de pertinence des réponses apportées individuellement et

collectivement face aux changements. Par exemple, le fameux « Time to Market » mesure la capacité d'une entreprise à réagir à des fluctuations du marché, par la création ou tout au moins l'adaptation de nouveaux produits ou services. Les experts en management nous disent que l'agilité et la flexibilité sont largement améliorées par la promotion au sein de l'entreprise de certains comportements facilitateurs : partage, confiance, responsabilité, recherche des faits et du progrès. Or la grande entreprise dans sa forme traditionnelle ne facilite pas ses comportements car elle est : élitiste, procédurale, disciplinaire et adepte de la langue de bois.

Le système d'information se révèle généralement être un frein aux changements qui impliquent souvent le développement de nouvelles applications ou la modification des applications existantes. Sans revenir sur la lourdeur légendaire des PGI, les systèmes décisionnels eux-mêmes ne sont pas exempt de reproche. En effet, du fait de la place cruciale qu'ils ont pris dans la conduite des affaires, il a été nécessaire pour assurer le niveau de qualité requis et le rythme de fonctionnement (24h/24, 365 jours par an), de leur imposer des normes industrielles tant pour leur développement que pour leur exploitation. Au final s'ils ont gagné en qualité, ils ont perdu en réactivité métier, en liberté, en créativité, en capacité d'expérimentation, en innovation, en support à l'agilité.

Pour qu'un système décisionnel aide l'entreprise à être agile, il faut que classiquement il permette une vision et une analyse de la situation de l'entreprise par département, fonction et processus. Il faut aussi qu'il offre des facilités pour exploiter de façon expérimentale, toutes sortes de nouvelles informations internes ou externes, qui peuvent en complément des données déjà utilisées, aider l'entreprise à décrypter les jeux dans lesquels elle est engagée et trouver des solutions d'avenir. Cependant attention, il ne s'agit pas seulement de moyens technologiques mais du travail collaboratif de spécialistes métier, d'experts en analyses avancées et d'informaticiens. Ainsi un des indicateurs clé de la potentielle agilité du système décisionnel d'une entreprise est l'existence d'un centre d'expertise dédié, appelé BICC par les analystes anglo-saxons (Business Intelligence Competency Center).

Entrepôt de données : quelques recommandations pour échouer

Régulièrement des entreprises font part des déboires qu'elles ont rencontrés lors de la mise en œuvre de leur entrepôt de données, et il ressort systématiquement de leurs discours qu'elles ont faits certains choix dans la conduite de leur projet. Sans entrer dans la problématique du choix des fournisseurs, les lignes ci-dessous recensent sous forme de recommandations, les principales idées qu'elles ont suivies pour assurer leur échec.

Pour réaliser votre entrepôt de données, ne consultez aucune personne extérieure à votre organisme. Votre cas est unique et seules des personnes de chez vous peuvent trouver la solution qui convient. Ne cherchez pas à savoir ce que font vos concurrents, ne regardez pas les références de fournisseurs de solutions décisionnelles.

Fixez-vous comme premier objectif de spécifier de façon exhaustive tous les besoins décisionnels de votre organisme, en particulier soignez particulièrement bien la définition détaillées des restitutions attendues.

Cherchez à satisfaire les utilisateurs en leur réalisant des systèmes personnels sur mesure. Concrètement il s'agit de multiplier les data marts dans une logique de construire pour

chaque division, département voire service, des moyens totalement adaptés au périmètre de leur problématique, en particulier à leur vocabulaire, à leur façon de voir les données sans se soucier d'autres choses que de leur vision spécifique des affaires. A chacun sa vérité.

Planifiez de multiples projets décisionnels en parallèle, sans vous souciez de fixer des priorités, ni de produire rapidement des résultats.

Evitez le plus possible de stocker des données détaillées et fondez les systèmes sur des données agrégées, des compteurs du type RFM, et faites une large place aux données d'enquêtes, d'études ponctuelles et d'informations qualitatives que les utilisateurs maîtrisent bien, voire produisent eux-mêmes.

Au niveau de la modélisation, bannissez la 3eme forme normale, fondez le plus possible les data marts sur des modèles en étoile ou en flocon, de façon à optimiser au maximum les performances des requêtes qui sont envisagées au moment de la conception du système.

Pour les restitutions, simplifiez le plus possible la vie des utilisateurs en fournissant des résultats facilement compréhensibles comme des moyennes, des médianes ou des métriques synthétiques, qui permettent de suivre simplement des sujets complexes.

Pour faciliter l'appréciation des performances fondez exclusivement les suivis sur des comparaisons entre données de prévision et de réalisation, en veillant à ne pas trop multiplier les chiffres et à rester très synthétique, en particulier il convient de limiter la profondeur de l'historique des données.

Au niveau des moyens technologiques, privilégiez des solutions prêtes à l'emploi spécifiques, quitte à multiplier les fournisseurs et les moyens.

Défendez bec et ongles l'idée que la vérité est dans les statistiques.

Investissez un minimum dans votre entrepôt de données, garder le plus possible votre argent au fond d'un coffre, il y sera en sécurité.

Exemple de mise en place de moyens décisionnels

Les entreprises sont toujours à l'écoute de leurs clients, de leur consommation, de leur satisfaction, de leurs nouvelles attentes et grâce à une adaptation inventive de leur offre, elles cherchent à répondre à leurs exigences. Pour cela elles comptent depuis de nombreuses années sur leur système d'information en général et leur système décisionnel en particulier, mais si cette voie est clairement identifiée et reconnue, son parcours est long et nécessite persévérance et ténacité.

La remarque ci-dessus peut être illustrée par l'exemple d'une entreprise du monde de la Grande Distribution. Cette enseigne a près de 300 magasins (un magasin peut proposer jusqu'à 60 000 références), 20 000 collaborateurs et un chiffre d'affaires consolidé (normes IFRS) de 3 575 millions d'Euros HT. Fin 2002, elle se dote d'un Entrepôt de Données dans le but de pouvoir analyser et piloter ses activités. Pour cela il lui a fallu dans un premier temps consolider les systèmes décisionnels existants en trois étapes centrées sur l'analyse des comportements clients, l'analyse de l'offre (produits, promotions, assortiments), et la

constitution d'un référentiel unique incluant les indicateurs de pilotage. Puis dans un deuxième temps développer et mettre en place successivement plusieurs applications métier.

Technologiquement, l'enseigne a recherché en priorité une valeur ajoutée pour les Métiers, et pour cela a retenue une solution qui comprenait un modèle de données et une suite analytique spécifiques à la distribution, une base de données et une plateforme massivement parallèle dotée d'une surface disques de 2,4 TO pour gérer l'ensemble des détails tickets de l'enseigne. Elle a en outre choisi un ETL, pour interfacer ses 300 magasins avec son système central, et consolider des données permettant de réaliser des analyses en rapprochant les informations relatives aux achats, à l'approvisionnement, au suivi des stocks, au traitement des inventaires, aux systèmes d'encaissement et aux opérations promotionnelles.

A ce stade le nouveau système a permis de répondre aux besoins de différentes directions. D'abord à ceux des Directions du Marketing et des Ventes du Siège, en exploitant les données ventes issues des remontées des tickets de caisse des magasins, complétées peu après des données des porteurs de la carte fidélité (~1 million de clients encartés), afin de : faciliter l'analyse des ventes et du comportement des clients en magasins, pour mieux cibler l'offre produits par type de magasins et améliorer la performance opérationnelle ; diminuer les temps de réalisation des tableaux de bord pour se concentrer d'avantage sur l'analyse ; permettre l'homogénéité ; favoriser la transversalité ; et enfin globalement faciliter un langage commun.

L'Entrepôt de Données a été ensuite enrichi d'autres données, pour répondre aux besoins de pilotage d'autres directions du siège, comme par exemple : pour le Bureau d'Achat analyser la vente par produit ; pour les Magasins et le Siège publier sur l'intranet l'analyse de la fréquentation et du CA de la veille ; pour la Direction de la Supply Chain détecter les « non ventes », mesurer le chiffre d'affaire perdu magasin par magasin en raison des produits identifiés en rupture de stock la veille (par jour, par tranches de journée), centraliser l'information sur les entrepôts dont dépendent les magasins et créer des alertes pour les approvisionneurs afin de corriger ces ruptures par des réassorts plus réactifs ; pour la Direction Marchandises des analyses des ventes, de la performance des produits, et l'envoi des ventes vers le système opérationnel pour le calcul d'une partie du réassort automatique ; et enfin pour la Direction Financière / Contrôle de Gestion la publication d'indicateurs de pilotage du CA des magasins. Au final le système publie quotidiennement différents tableaux de bord et indicateurs pour suivre l'activité des enseignes, et permet de répondre aux demandes d'indicateurs de toutes les directions fonctionnelles.

L'enseigne a aussi placé le client au centre de ses préoccupations, grâce à la mise en place d'une solution de data mining. La principale motivation était de mieux connaître les clients et de partager ce savoir dans toute l'entreprise depuis la direction générale, les responsables marketing, le service commercial, jusqu'aux chefs de rayon, vendeuses ou hôtesses de caisse dans les magasins. Plus précisément, il s'agissait de mieux cerner les besoins et motivations d'achat des consommateurs, et comprendre leurs comportements (et leurs évolutions). Objectif : adapter l'offre produits afin d'optimiser et maximiser les ventes. Les applications décisionnelles permettent ainsi de confirmer ou d'affiner des "ressentis" ou des "intuitions".

En 2007 le système comprenait 13 TO de disques et était utilisé par environ 1500 personnes en magasin et 300 personnes au siège, qui exploitaient des données, des indicateurs ou des tableaux de bord (statiques ou dynamiques), issus de l'Entrepôt de Données d'Entreprise et publiés sur intranet, ou reçus directement dans les messageries. Mais attention, à ce stade de

nombreux besoins restaient encore à traiter comme le suivi des marges, ou en matière de performance celles relatives à la logistique, aux fournisseurs, et à l'animation commerciale. Il a été lancé alors un schéma directeur décisionnel, qui a permis de confirmer et compléter les choix technologiques, fixer des évolutions de la gouvernance pour un meilleur support à la stratégie du groupe, définir un nouveau plan d'actions sur cinq ans pour prendre en charge de nouveaux besoins fonctionnels, et enfin lancer différents chantiers d'urbanisation.

Sujets d'analyse dans la grande distribution

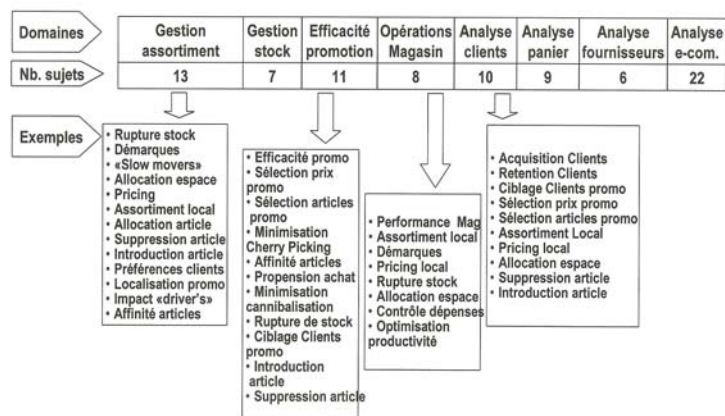


Illustration - 29

Aujourd'hui, grâce à tous les moyens décisionnels mis en place, les collaborateurs de cette enseigne disposent et partagent des informations de qualité pour analyser, décider et conduire les actions qui lui permettent d'être un leader emblématique, apportant régulièrement des idées nouvelles à la profession, réalisant des performances financières remarquables, notamment en portant le nombre de magasins à 330 et en maintenant une rentabilité élevée, sans parler des résultats des enquêtes de satisfaction clients, ni de la qualité de l'image de marque de cette enseigne dans le grand public.

Solutions technologiques spécifiques pour le système décisionnel

Il n'y a pas de jugement possible, pas d'appréciation saine de la réalité sans repère. Si l'on veut avoir une réflexion stratégique par exemple, il est préférable de l'inscrire dans une période longue et de se fonder sur des références solides pour projeter les différents avènements possibles. Les entreprises gèrent des volumes considérables de données opérationnelles, mais sans un système d'information décisionnel correctement agencé, les responsables ou les opérationnels n'auront pas facilement accès aux données nécessaires pour réagir rapidement notamment aux situations non prévues.

Si nous regardons les entreprises qui ont un système d'information décisionnel performant, nous pouvons constater le grand nombre de solutions technologiques spécifiques au domaine du décisionnel qu'elles emploient (matériels, bases de données, modèles, outils et applications). En effet de très nombreuses solutions propres au monde du décisionnel ont été développées en parallèle du monde opérationnel pour couvrir les besoins particuliers qui ne pouvaient pas être pris en charge par des solutions standards. Il est ici amusant de se rappeler, que pendant longtemps certains fournisseurs de matériels, de base de données ou d'ERP, ont cherché à faire croire que leurs solutions standards permettaient de mettre en place des entrepôts de données ou des applications BI de qualité, mais depuis ils ont renoncé et les plus grands ont même rachetés des sociétés spécialisées pour se constituer une offre à la hauteur de leurs ambitions. Le domaine du décisionnel requiert des solutions technologiques spécifiques et pas seulement pour faire de la fouille de données.

Il est certain qu'un système décisionnel ne peut pas être trouvé tout fait prêt à l'emploi sur l'étagère d'un fournisseur, il faut obligatoirement le construire, l'agencer en fonction des besoins particuliers de l'entreprise. Si l'on excepte quelques rares domaines particuliers d'analyse plus matures que les autres pouvant faire l'objet de data marts, il n'existe pas encore de « progiciel décisionnel d'entreprise ». En matière de construction d'un système décisionnel, même si globalement la démarche obéit aux règles générales de conduite des projets informatiques, il convient cependant de ne pas sous-estimer l'importance des caractéristiques spécifiques au domaine du décisionnel, et donc de privilégier le recours à des personnes expérimentées. Par exemple déterminer les modalités de transformation de données en information pour répondre à certains besoins métier d'une entreprise, n'est absolument pas trivial et nécessite de l'expérience.

Enfin il convient aussi de noter que si certaines solutions spécifiques au décisionnel sont clairement reconnues et largement mises en œuvre c'est le cas de l'olap, des modèles en étoiles ou en flocon par exemple, d'autres restent très confidentielles, comme les serveurs spécialisés massivement parallèles qui ont été jusqu'à très récemment avec l'arrivée des appliances pratiquement ignorés, malgré leur domination dans les très grandes entreprises.

Mettre les appliances à leur juste place

Depuis quelques temps le monde du décisionnel est agité par de nouvelles offres « tout en un » en anglais Appliance. L'idée générale est d'offrir une solution spécifiquement décisionnelle, complète (matériel + base de données), prête à l'emploi, à même de supporter

des applications de BI. Le phénomène est réel et j'ai pu recenser 20 compétiteurs se réclamant de cette approche : Alterian, Calpont, Celequestbox , Clareos, Datallegro, Dataupia, Enterprisedb, Exasol, Exegy, Extenddb, Greenplum, Hyperroll, Infocyclone, Netezza, Nexql, Panca, Paracel, Qlicktech, RightOrder, Sandtechnolo. Cependant il convient d'être prudent car il n'existe pas une définition précise de ce qu'est un appliance décisionnel et il est en fait difficile de les distinguer des plateformes préconfigurées dédiées au décisionnel. Théoriquement l'appliance est censé être totalement prêt à l'emploi, mais sauf pour des solutions décisionnelles très limitées ceci n'est en fait pas le cas.

L'apparition des appliances rappelle l'importance et les besoins spécifiques de la partie matérielle pour les systèmes décisionnels, elle révèle aussi la valeur de ce domaine pour les entreprises et l'existence d'un marché qui attire les convoitises. Les offreurs d'appliance cherchent à transformer le décisionnel en marché de commodité, mais cela est très largement prématuré selon le Gartner qui positionne ces solutions dans le stade 1 sur 5 de sa courbe d'évolution des solutions, avec une perspective de maturité d'ici 5 à 10 ans. A court terme les offreurs d'appliance étant nombreux, ils sont condamnés à une guerre des prix qui en verra survivre très peu. Actuellement le nombre d'appliance en production est très limité et les analystes cherchent toujours le premier site pouvant servir de « best practice » de mis en place d'un entrepôt de données d'entreprise avec un appliance.

Dans sa forme actuelle l'appliance se focalise sur certains aspects du problème qu'une plateforme décisionnelle doit assumer. Les appliances cherchent à offrir de la haute disponibilité, de la facilité d'utilisation, des faibles coûts d'acquisition, une exploitation économique (électricité et refroidissement), ils ne répondent pas ou mal à d'autres aspects de la demande en matière d'entrepôt de données : gestion d'un grand nombre d'utilisateurs et surtout intégration dans le système d'information de l'entreprise en quasi temps réel pour certains processus, avec des activités parallèles de mises à jour et d'analyses plus ou moins complexes avec pour certaines des exigences de niveaux de services élevés sur des données très fraîches. Il ressort de tout cela que l'emploi optimal d'un appliance semble être de réaliser un data mart ou une application analytique isolée ayant peu d'intégration avec le reste du système d'information de l'entreprise.

Spot for data warehouse appliances

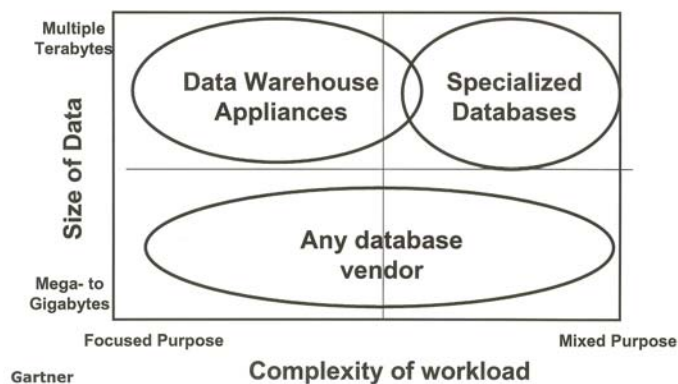


Illustration - 30

Si nous prenons un peu de recul, nous constatons donc que les appliances cherchent avant tout à profiter du marché des infrastructures décisionnelles qui s'élèvent mondialement à 17,4 Milliards de dollars et devrait progresser de 6,5% par an dans les 3 ans à venir. Dans ce

marché ils visent à court terme à offrir la possibilité de mettre en place des moyens décisionnels basiques à faible coût, limités le plus souvent à une application spécifique d'un département. Nous sommes donc loin de ce que souhaitent les grandes entreprises, c'est-à-dire de faire passer les moyens décisionnels d'un rôle « passif » à un rôle « actif ». Dans cette optique, il ne s'agit plus de seulement supporter les décisions stratégiques (détermination des buts, des politiques, définition des objectifs des organisations) au niveau de l'entreprise, d'une fonction ou d'un département. Il s'agit de supporter aussi des décisions tactiques en dotant de moyens décisionnels des opérationnels clés, par exemple ceux qui sont en relation quotidienne avec les clients ou les fournisseurs. Il s'agit aussi d'ouvrir les systèmes décisionnels aux partenaires. Mais pour faire tout cela il faut une capacité d'intégration avec les systèmes d'information des entreprises qui est aujourd'hui hors de portée de tous ces nouveaux appliances.

Virtualisation et entrepôts de données

La virtualisation, de serveurs, de stockage et d'applications, devient de plus en plus une technique essentielle pour une utilisation plus efficace des ressources des centres de données, et depuis de nombreuses années certains fournisseurs de plateformes ont pris en compte ce concept dans la conception de leurs produits, autorisant par exemple dans une même configuration une hétérogénéité de serveurs. Avec la virtualisation, il est possible de remplacer un grand nombre de serveurs peu utilisés, par un plus petit nombre pleinement utilisés, ainsi un seul entrepôt de données d'entreprise peut remplacer de multiples data marts, coûteux et à l'efficacité limitée.

La virtualisation n'est pas une fin en soi, mais un moyen de construire une infrastructure flexible, efficace, permettant d'augmenter la disponibilité et de faire des économies, notamment d'énergie, ce qui en fait aussi une approche verte politiquement correcte. La virtualisation élimine la nécessité de comprendre les détails physiques d'un entrepôt de données avant de pouvoir l'exploiter et son utilisation en est donc simplifiée. Pour les serveurs, la virtualisation élimine la contrainte d'exécuter une application, utilisant un système d'exploitation (OS) sur une plate-forme matérielle dédiée. Un environnement virtuel permet à de multiples applications et à leur OS hôte de coexister et de fonctionner sur une seule plate-forme matérielle.

La virtualisation des serveurs s'obtient grâce à une couche de logiciel qui interagit avec le matériel. Cette couche de virtualisation permet au serveur d'héberger plusieurs environnements de système d'exploitation, alors que dans un environnement non virtualisé, chaque OS a besoin d'un serveur distinct. Pour la virtualisation des serveurs, le premier bénéfice vient bien sûr de la réduction du nombre de serveurs. Cela permet des économies en matière de coût des équipements, de maintenance, d'alimentation, de refroidissement et d'espace nécessaire. En effet il est estimé que la plupart des serveurs non virtualisés sont utilisés à moins de 20% de leur capacité. Mais le plus grand avantage pour un système bien configuré, est la capacité de fournir une plus grande puissance de traitement à chaque application quand elle tourne.

La virtualisation de stockage est mise en œuvre en utilisant une plate-forme matérielle de contrôle spécialement conçue pour un ou plusieurs systèmes de stockage. À partir d'un serveur, cette plate-forme spécialisée a pour objet de fournir les capacités de stockage nécessaires aux applications, en gérant les multiples systèmes de stockage et en facilitant la

gestion de l'augmentation des besoins de stockage. Dans le cas de la virtualisation du stockage, on supprime totalement la nécessité de connaître les vitesses, les spécifications des moyens d'interface (SCSI, ATA, Fibre Channel, etc) et les paramètres des disques du type « cylindre / tête / secteur ». Grâce à la virtualisation, le stockage devient tout simplement un ensemble d'octets sur un disque (gigaoctets, téraoctets ou pétaoctets), un espace réservé pour les applications et l'on peut définir des caractéristiques de stockage spécifiques, par exemple de performance ou de protection des données. Cependant il faut avoir conscience qu'un des grands problèmes des environnements virtualisés est que, comme on ne connaît pas à priori la ressource exacte qui va rendre le service demandé, on a du mal à en garantir à priori les performances (les temps de réponses). Mais grâce à certains logiciels spécifiques tous les détails techniques sont cachés, la virtualisation est facile à configurer, à gérer, à faire évoluer, et les niveaux de services requis peuvent être garantis.

La virtualisation est une technologie qui permet donc d'accroître l'efficacité des centres de données. La virtualisation du stockage apporte une dimension supplémentaire, en particulier pour les environnements à évolution rapide et aux exigences de protection des données élevées.

Sécurité du système d'information et des entrepôts de données

Plus du tiers des grandes entreprises ont déjà fait l'objet d'une intrusion dans leur système d'information ou d'une tentative sérieuse, ceci fait que la sécurité est devenue un domaine clé d'investissement et que le marché des solutions de sécurisation se porte bien. Cependant la confiance ne peut pas être obtenue que par l'achat de technologie, et au de la technique il convient de bien agencer différents moyens organisationnels, juridiques ou humains, pour garantir la sécurité des informations. Concrètement il s'agit d'assurer l'intégrité, la confidentialité et la disponibilité de l'information, c'est-à-dire de garantir que l'information ne puisse être altérée, révélée à des tiers non autorisés ou indisponible au moment voulu.

Les entreprises sont confrontées à de nombreuses nouvelles menaces & contraintes, dues au renouvellement de leurs approches commerciales, à l'évolution de leurs organisations, à l'attaque de nouveaux marchés, ou aux nouvelles exigences en matière de normes ou de réglementations. De plus l'évolution des systèmes d'information apporte aussi son lot de fragilités, du fait de l'intégration de plus en plus poussée des systèmes de l'entreprise avec ceux de partenaires, du recours accru à internet pour mener des opérations, des nouvelles technologies de réseaux sans fil ou des réseaux VPN d'accès distant aux intranets. La sécurité généralement a du mal à suivre le rythme des évolutions.

Les informations à protéger sont en fait nombreuses. En ce qui concerne les clients, leur liste, leurs achats, leur carte de paiement par exemple. Les données sur les employés, sur leur rémunération en particulier, ainsi que les informations financières de la société, ou celles relatives à la planification de la production, aux achats, sans parler des bases de connaissance, des brevets.

Les menaces viennent à la fois de l'intérieur et de l'extérieur des entreprises. La grande majorité des problèmes de sécurité sont générés par l'utilisateur agréé du système d'information (employés, partenaires d'affaires), généralement insouciant et inconséquent, mais aussi parfois créateur d'erreurs volontaires. Il faut aussi compter sur de la malveillance

de pirates externes, sur des virus largement diffusés ou des attaques très ciblées (déni de service, prolifération de points de présence internet, ...).

Sans être totalement paranoïaque, il faut bien convenir qu'il est particulièrement difficile de mettre en place des moyens de parer toutes les attaques. Un expert du domaine disait d'une façon provocatrice que : « Le seul système vraiment sûr, est celui qui est coulé dans un bloc de béton et placé dans une chambre forte, surveillée par des gardes armés. Mais même ainsi, j'ai des doutes ».

Pour protéger un système d'information, il faut commencer par évaluer les risques, et pour cela il existe des approches spécifiques comme la méthode OCTAVE (Operationally Critical Threat, Asset and Vulnerability Evaluation) ou MEHARI (Méthode Harmonisée d'Analyse des Risques). Il s'agit de déterminer avec rigueur quelles sont les informations sensibles, qui peuvent être parfois de simples données. Il faut évaluer la sensibilité des différents éléments recensés, caractériser les menaces et déterminer les vulnérabilités. Par exemple dans la méthode Méhari, le produit de l'impact de chaque menace et de la potentialité d'occurrence, permet de définir un degré de gravité. Ces derniers peuvent servir à déterminer des objectifs de sécurité, qui sont l'expression de la volonté de contrer des risques identifiés, et être déclinés en fonctions de sécurité, à mettre en œuvre sur le système d'information.

La conception de la sécurité se doit de prendre le problème de façon globale, et il est recommandé de mettre en place une défense en profondeur, où chaque sous système est sécurisé où chaque acteur est mobilisé. La sécurité doit être l'affaire de tous, même s'il est essentiel de nommer un responsable de la sécurité du système d'information. Enfin, la criticité des systèmes est devenue telle, qu'il est indispensable de prévoir un plan de sécurisation organisant soit le redémarrage rapide des systèmes (PRA : plan de reprise d'activité), soit de pouvoir maintenir l'activité grâce à une redondance d'infrastructure (PCA : Plan de Continuité d'Activité). Avec le développement d'entrepôts de données actifs, sollicités par les systèmes opérationnels dans le cadre de processus clés, il y a de plus en plus d'entreprises qui mettent en place des PRA ou des PCA.

Protéger votre entreprise contre le coût des temps d'arrêt informatique

Selon le Gartner, les temps d'arrêt des systèmes informatiques peuvent coûter un million de dollars par heure aux grandes entreprises américaines. Ce chiffre peut paraître étonnant mais si l'on tient compte de l'importance des systèmes informatiques dans les entreprises d'aujourd'hui, il n'est pas si surprenant d'apprendre qu'elles courent d'énormes risques financiers lorsque leurs systèmes informatiques ne sont plus disponibles. En effet, des pertes de chiffres d'affaires, la réduction de la productivité, des pénalités ou des amendes peuvent entraîner des coûts directs considérables et nuire à la réputation de l'entreprise. Les systèmes d'aujourd'hui tirent la majeure partie de leur valeur des montagnes d'informations stockées dans les entrepôts de données, et ceci d'autant plus que les sociétés qui utilisaient leurs entrepôts uniquement pour supporter des analyses stratégiques recourent désormais activement à ces entrepôts pour appuyer la majorité des fonctions clés de leurs activités. Dans ce contexte, même si le chiffre d'un million de dollars par heure est plus ou moins exact, ce niveau de montant interpelle et souligne l'importance pour les entreprises de protéger leurs systèmes stratégiques en général et leur entrepôt de données en particulier.

La meilleure solution face à ce problème des temps d'arrêt est d'utiliser deux systèmes de production opérationnels. Cette configuration permet au second système d'entrer en jeu lorsque le premier système n'est plus disponible ou lorsque ses performances sont amoindries. Ce type de configuration est communément appelé "système redondant". Pour assurer la continuité de l'activité il convient de relever trois défis principaux. Le premier défi est celui de la disponibilité des systèmes et des données. Vu l'ampleur des enjeux, les sociétés doivent être certaines qu'en cas d'arrêt intempestif d'une partie du système, des moyens restent opérationnels et les données accessibles. La deuxième préoccupation se situe au niveau de la récupération du système. Les clients peuvent-ils se permettre d'attendre plusieurs semaines pour que l'activité soit à nouveau opérationnelle ? Attendront-ils ? Ensuite vient la continuité des performances. Les entreprises doivent maintenir un niveau de service à la clientèle acceptable pendant ces défaillances. A noter que dans de nombreuses entreprises, lorsqu'une partie du système tombe en panne, les performances du reste du système se détériorent.

Diverses solutions peuvent être mises en œuvre. L'attention accordée à chacune des préoccupations citées ci-dessus dépendra des sociétés. Par exemple, si une société peut supporter une défaillance de quelques jours, alors un système partagé dans un centre de récupération distant est la meilleure solution. La plupart des entreprises qui ont opéré une migration vers un entrepôt de données unique doivent en général tenir compte des trois préoccupations majeures déjà évoquées. Dans ce cas, les solutions partielles peuvent être combinées dans une solution unique et globale : deux systèmes actifs avec des fonctions de sauvegarde.

Les systèmes redondants actifs (Dual Active) se composent de deux systèmes qui permettent de synchroniser des données et de répartir le travail entre eux, alors qu'ils sont habituellement situés dans des emplacements géographiques différents. Ces systèmes offrent un certain nombre d'avantages, tout d'abord ils réduisent sensiblement les temps d'arrêt, car ils fournissent des fonctions d'entreposage de données qui permutent automatiquement en cas de pannes, et offrent un service continu pendant que l'un des systèmes est arrêté pour des tâches de maintenance ou de mise à jour de routine. Ensuite, les systèmes redondants actifs permettent de restaurer les applications stratégiques en quelques minutes, afin de répondre aux besoins essentiels des sociétés. Enfin, en faisant du second système un système de production constamment synchronisé, les sociétés peuvent traverser les périodes de pointe ou de défaillances partielles sans problèmes.

Tout est à conditionner en fonction de l'évaluation des besoins qui consiste à comprendre ce que les temps d'arrêt et les performances réduites coûtent à l'organisation. Les principales questions à se poser sont : Combien de temps l'entreprise peut-elle supporter une panne avant que le chiffre d'affaires ou des activités stratégiques ne soient affectées ? Toutes les applications touchent-elles les fonctions clés de la même manière ? Existrent-ils des priorités claires qui peuvent être identifiées et faire l'objet d'une solution spécifique ? Quelle est la complexité de l'environnement informatique ? Quelles sont les contraintes financières ? Il est recommandé aux entreprises de travailler ces questions de continuité des activités avec un consultant. Il convient de sélectionner des consultants qui ont déjà consacré un certain temps à évaluer ce type de besoin, et qui disposent d'une expérience concrète pour relever ce type de défi relativement récent pour les systèmes décisionnels.

Gouvernance des investissements informatiques

Les Directions générales voient habituellement assez bien les dépenses liées à l'informatique, mais beaucoup moins clairement les bénéfices apportés. Dans la majorité des entreprises, seuls les grands projets font l'objet d'une évaluation des bénéfices. Mais, même dans ce cas, seul un projet sur deux bénéficie d'un suivi ou d'une analyse de rentabilité.

En fait, les maîtrises d'ouvrages ont du mal à justifier leurs projets, alors que les directions générales attendent une implication forte des directions métier, celles-ci ne répondent pas vraiment à leur demande. C'est sans doute là que se situe le principal problème des directions informatiques, qui n'ont pas assez le soutien des directions métier, et qui peinent à démontrer la valeur ajoutée des dépenses informatiques, les 3/4 sont convaincues d'amener à leur organisation une valeur ajoutée supérieure aux dépenses engendrées, mais près de 90 % d'entre-elles avouent être dans l'incapacité de le démontrer à leur direction générale.

Il n'y a pas de fatalité, une gouvernance des investissements informatiques est possible et pour être effective, elle exige que les directions informatiques maîtrisent les engagements de dépenses, établissent la valeur des investissements et des résultats livrés par les projets informatiques. Ceci peut se faire à travers des processus d'alignement stratégique et de planification (schéma directeur, étude d'opportunité, urbanisation, plan d'évolution), d'une gestion de portefeuille de projets, d'une gestion budgétaire, d'une gestion des projets informatiques focalisée sur la prise en main par les utilisateurs et la réalisation des bénéfices.

Gouvernance : acteurs et processus

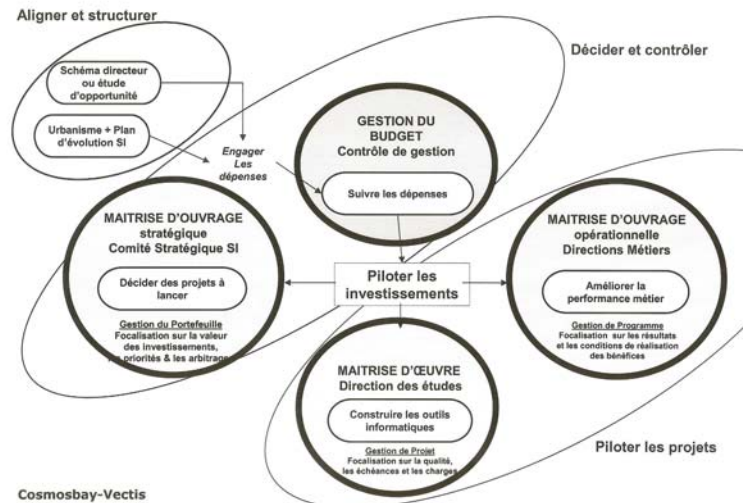


Illustration - 31 Cosmosbay-Vectis

Normalement la mise en place d'un nouvel applicatif doit permettre à une maîtrise d'ouvrage de faire évoluer ses résultats métier, mais pour cela il convient de comprendre la logique de production des résultats, de définir une stratégie de réalisation des bénéfices, afin de les piloter et de livrer les meilleurs résultats possibles (résultats recherchés, à éviter, indicateurs utilisés, photo avant / photo après). La conduite du changement doit être fondée sur une démarche proactive de production des résultats.

Chaque projet doit faire l'objet d'une évaluation selon quatre critères : alignement stratégique (clarté et précision de la contribution), contribution métier (crédibilité des résultats annoncés), risques et enfin valeur financière calculable (clarté et crédibilité du ROI).

Faut-il et si oui, comment, se préoccuper du retour sur investissement des projets NTIC ?

Existe-t-il une corrélation entre le niveau des investissements dans les nouvelles technologies de l'information et des communications (NTIC) et la performance d'une entreprise ?

Voilà une question à laquelle de nombreuses personnes aimeraient répondre depuis de nombreuses années. Régulièrement des consultants publient des études intéressantes qui montrent que les entreprises qui ont une gestion « avancée » de l'information, sont en moyenne plus rentable que les entreprises qui sont « à la traîne », et que les avancées dépassent largement les « à la traîne » aux niveaux du chiffre d'affaires et des parts de marché. Ce que l'on ne sait jamais très précisément dans ces études, c'est si les investissements dans les NTIC ont générés les bons résultats ou s'ils ont été effectués du fait des bons résultats.

Renonçant à une démonstration au niveau global de l'entreprise, certains consultants ont mis au point des approches pour apprécier projet par projet les bénéfices apportés. Par exemple il propose de segmenter les bénéfices en grands types : productivité, effets de levier financiers, qualité de service, conditions de travail. Pour chaque type, des critères d'évaluation sont définis qui permettent de cerner des dimensions quantitatives et qualitatives. Dans la pratique ces approches sont difficiles à mettre en œuvre. La difficulté majeure réside dans le fait d'attribuer à un investissement technique donné, la paternité d'un résultat métier observé. D'autre part de nombreux avantages ne sont pas financiers, il faut alors relever le défi de fixer des métriques pour les apprécier, et il reste dans ce cas à synthétiser une batterie d'éléments divers qui vont plus ou moins dans le même sens et qui peuvent difficilement s'additionner arithmétiquement. En fait la plupart du temps les entreprises simplifient leur problème et privilégient une approche financière en termes de recettes ou d'économies, et négligent tout ce qui ne peut pas être facilement financièrement valorisable, comme par exemple le fait de faire mieux que le marché, de diversifier sa clientèle ou ses canaux, d'améliorer sa notoriété, de motiver ses collaborateurs, etc...

Quels modèles financiers utiliser ?

Trois approches financières sont plus fréquemment utilisées.

- La période de recouvrement de l'investissement : délai pour atteindre le moment à partir duquel l'investissement va faire gagner de l'argent.
- La valeur actuelle nette (VAN) : la valeur actuelle (escomptée) des futurs gains diminuée de la valeur actuelle des investissements et de toutes les futures dépenses associées.
- Le taux de rendement interne (TRI) : taux de rendement minimal attendu pour justifier la décision d'investissement.

Ces différentes méthodes sont abondamment documentées, et généralement bien maîtrisées par les directions financières des entreprises qu'il convient d'associer systématiquement à ce genre d'exercice.

Quelle démarche pour garantir le résultat ?

En complément de ces approches prévisionnelles de la rentabilité des projets, de nombreux cabinets de conseil ont défini des règles pour assurer la rentabilité des investissements informatiques. Par exemple ils recommandent de : lancer des projets qui ajoutent de la valeur aux activités plutôt que de réfléchir en termes de systèmes informatiques ; retarder le moment d'arrêter le choix d'une technologie et tester au moins deux solutions ; quantifier le plus précisément possible les coûts, les revenus et se servir de ces évaluations pour sélectionner un ensemble cohérent de projets ; adopter un plan d'investissement large mais l'engager étape par étape ; responsabiliser les directions fonctionnelles et informatiques sur des résultats métier précis qui concernent leurs domaines ; imputer tous les coûts informatiques (investissements et exploitation) aux différentes directions fonctionnelles.

En matière de NTIC, les spécialistes des entrepôt de données ont aussi dégagé les caractéristiques générales des projets qui réussissent : ils sont fondés sur une vision future de l'activité, une "réflexion métier" à long terme ; ils visent à obtenir des avantages concurrentiels à court terme, comme accroître les ventes (1^o objectif) ou augmenter la focalisation sur l'extérieur (clients, revendeurs, concurrents) ; ils consistent à réaliser des applications stratégiques, s'appuyant sur une infrastructure pérenne ; ils sont mis en œuvre de façon itérative.

Toujours en suivant l'expérience des spécialistes, nous voyons que les projets réussis mettent en évidence les points suivants : la culture d'entreprise est le principal facteur clé de succès ; le premier frein est la résistance au changement ; définir les informations stratégiques est un travail qui n'est pas trivial ; les utilisateurs s'impliquent surtout au début du projet ; les domaines les plus favorables pour commencer sont l'analyse des ventes, l'aide à la vente, l'analyse des profils clients, la gestion des promotions et la gestion des stocks ; les gains de productivité dans l'analyse des données sont spectaculaires et les utilisateurs deviennent indépendants dans ce domaine.

Conclusion

Les projets de mise en œuvre de Nouvelles Technologies de l'Information et des Communications (NTIC), sont soumis aux mêmes exigences de rentabilité que les autres investissements des entreprises et ont affaire à une forte concurrence, notamment dans le cadre actuel d'une plus grande rigueur de gestion. Les investissements effectués dans les NTIC sont handicapés face à leurs concurrents, car ils présentent des caractéristiques défavorables bien connues en matière de risque et de rentabilité. En effet le monde du système d'information n'a pas encore atteint un bon niveau de maîtrise de la qualité des délais, des budgets, de la couverture des fonctionnalités prévues, et aussi de l'utilisation effective des moyens mis en place. D'autre part, les investissements en NTIC n'apportent que des potentialités, et nécessitent la plupart du temps de modifier les processus pour donner tous leurs avantages, ce qui n'est pas toujours effectué et pénalise donc les solutions techniques réussies.

Le ROI n'est qu'un instrument imparfait d'aide au pilotage, le plus important est de savoir où aller et pour cela, s'il est souhaitable d'avoir une conception d'ensemble, il est rare de pouvoir faire une démonstration détaillée du bien fondé de la vision. Mais même s'il est toujours difficile de calculer un retour sur investissement notamment pour les moyens d'infrastructure, se préoccuper d'anticiper la rentabilité d'un projet est clairement un facteur clé de succès, d'autant plus si vous ne vous contentez pas de faire des calculs, mais si vous multipliez les actions pour assurer la réussite des différentes étapes. Le processus de planification, d'obtention et de mesure de la rentabilité des investissements dans les NTIC, exige au delà des méthodes, de la discipline, de la cohérence et de la persévérance. Amat victoria curam (la victoire aime qu'on s'occupe d'elle).

Gouvernance du système d'information décisionnel

Dans l'entreprise la notion de gouvernance fait souvent l'objet de commentaires critiques et de scepticisme. Au-delà du fait que les modes de direction évoluent plus ou moins vite selon les entreprises, le souhait d'une régulation fondée sur une plus grande implication des parties prenantes est lui par contre très partagé. Le système d'information est l'un des domaines où une plus grande participation aux décisions est recherchée. L'informatique a pris une telle importance pour toutes les activités d'une organisation, que de nombreux choix qui la concerne ont un impact majeur voir stratégique pour les grandes fonctions, ils peuvent déterminer le devenir de l'entreprise. Pour les utilisateurs il ne s'agit donc pas simplement de veiller à ce que la direction informatique « aligne les technologies sur la stratégie d'entreprise », mais bien de travailler ensemble pour gérer la maison commune.

Partie importante du SI, les systèmes décisionnels font plus que d'autres l'objet d'une demande d'évolution de la gouvernance, et compte tenu de la complexité du domaine, les directions informatiques sont souvent plus facilement ouvertes à de nouvelles pratiques. Cependant, rares sont les entreprises qui ont radicalement transformé leur façon de diriger, d'administrer et de contrôler cette partie du SI, les directions informatiques très jalouse de leurs prérogatives, mettent toujours en avant les dimensions techniques des choix pour conserver tout leur pouvoir. Malgré tout il est fréquent de voir des avancées en matière de gouvernance de données et le cas de British Airways est en ce sens représentatif.

British Airways a depuis longtemps mis en place son premier entrepôt de données et l'a depuis fait évoluer pour répondre aux besoins de l'ensemble de ses activités, cherchant à supporter aussi bien des décisions stratégiques, qu'opérationnelles, avec quinze domaines fonctionnels aujourd'hui couverts. Dans le cadre de ses développements successifs, British Airways a toujours eu en ligne de mire d'améliorer les relations avec ses clients en s'appuyant sur une bonne gestion de l'information, en particulier une bonne qualité des données. Pour cela la compagnie aérienne a engagé un projet conséquent de gouvernance des données.

La première phase de ce projet a consisté à identifier les différentes parties prenantes de cette problématique et à trouver des supports pour cette initiative de gouvernance, au plus haut niveau de la compagnie (direction générale, financière, commerciale). Cela s'est concrétisé par la mise en place d'un comité de direction spécifique incluant des utilisateurs, des informaticiens, et la désignation de quelqu'un du métier du transport aérien pour prendre la responsabilité à plein temps de ce programme de changement. A ce niveau ont été définis des principes de gouvernance des données et une organisation comprenant :

- 18 responsables de grands ensembles de données (données clients, coupons de vol, contrôle de gestion, trafic, maintenance), s'appuyant sur 30 responsables métiers issus de toutes les directions métiers de l'entreprise et 12 responsables techniques issus de la DSI.
- Un programme de gestion proactive des données, incluant des processus de réponse aux besoins des utilisateurs et de gestion des projets de développement, des réunions régulières, des revues mensuelles, des indicateurs de suivi, etc.

Au-delà de l'implication, de la responsabilisation et de l'organisation de la prise en charge des problèmes quotidiens de gouvernance des données, ce projet a permis de traquer pro-activement des opportunités d'amélioration de la qualité des données, de la qualité des modèles, des technologies mis en œuvre, des architectures et des projets de développement. En termes de résultat, au niveau de l'utilisation des données cela s'est traduit par une meilleure compréhension des informations produites par les systèmes décisionnels, et le développement de nouveaux usages métier des gisements de données.

Sans rentrer dans tous les détails du fonctionnement de la gouvernance des données chez British Airways, le responsable de ce programme donne les conseils suivants aux personnes qui auraient la même mission que lui :

- Assurez-vous que vous avez la bonne infrastructure de gestion des données en termes de modèle de données et de plateforme.
- Faites en sorte que le comité de direction vous soutienne et vous fournisse le financement nécessaire au projet.
- Mettez en place une solide structure de gouvernance.
- Cherchez passionnément le succès.
- Soyez résilient.

Gouvernance des systèmes décisionnels

En matière de décisionnel il n'existe pas de solutions toutes prêtes sur étagère qu'il suffirait d'acquiescer et de déployer. La mise en place de moyens au niveau de l'ensemble de l'entreprise ne peut se faire que progressivement, aussi bien pour des raisons techniques, d'usages et d'apprentissages organisationnels. Par exemple un distributeur s'est doté en 2002 d'un entrepôt de données destiné dans un premier temps à répondre aux besoins d'analyses de la Direction Marketing et de la Direction des Ventes du Siège. Dans un deuxième temps il a progressivement répondu aux besoins de plus en plus de directions fonctionnelles. Aux termes de ces développements de nombreux besoins étant encore à traiter ce distributeur a donc lancé en 2007 un schéma directeur décisionnel qui a permis en outre de fixer des évolutions de la gouvernance pour un meilleur support à la stratégie du groupe.

Cet exemple illustre l'importance de la qualité de la gouvernance si l'on veut que les systèmes décisionnels apportent une aide déterminante à la définition et la mise en œuvre de la stratégie de l'entreprise. Il s'agit donc de piloter l'étude, la mise en place et l'utilisation de moyens décisionnels. Les experts de la gouvernance nous disent que ce pilotage doit être anticipatif, collaboratif, concerté, rétro-correctif, transparent, équitable et efficient. Concrètement il convient de :

- organiser une structure spécifique (mise en place d'un organe de direction incluant des représentants des parties prenantes et nomination d'un responsable de la gouvernance),
- définir des processus (une dizaine de processus clés, en particulier : budget/refacturation des services, gestion des demandes et du portefeuille de projets, gestion des projets, des services (SLA), des ressources humaines, et des actifs du système décisionnel - urbanisation du SID, applications décisionnelles, management de l'information -),
- fixer des objectifs de résultats et la façon de les mesurer notamment sur les points suivants : création de valeur, conformité, risques, efficacité des processus et aspects financiers.

Une bonne gouvernance doit assurer la qualité des dialogues entre les métiers et la direction du système d'information. Elle doit en outre permettre de mettre en place des solutions techniques appropriées. Dans le cadre des systèmes décisionnels un point technique est particulièrement à surveiller, c'est la capacité à évoluer des solutions retenues (en termes de volumétrie, de puissance, de nombre d'utilisateurs, de variété des usages, de complexité des services rendus). Par exemple le distributeur évoqué ci-dessus a vu son système passé au fil des développements de 2,4 T0 à 13 T0 en cinq ans. Cependant, assumer une plus grande volumétrie est une chose, et être en mesure d'intégrer les données de l'ensemble de l'entreprise en est une autre, et là ce n'est pas un problème de puissance mais de qualité de la conception des modèles de données. Or s'il n'y a pas de solution toute prête en matière décisionnel, il y a des solutions plus susceptibles d'évoluer que d'autres.

Conseil pour une bonne gouvernance des systèmes décisionnels

L'objectif de la gouvernance du système d'information est de maximiser l'apport de valeur du SI à la stratégie de l'entreprise, tant au niveau de sa définition que de sa mise en œuvre. La gouvernance des systèmes décisionnels doit tenir compte des spécificités de l'étude, de la mise en place et de l'utilisation des systèmes décisionnels. Par exemple les utilisateurs de systèmes décisionnels ont généralement beaucoup de difficultés à fournir des spécifications détaillées de leurs demandes, ils doivent passer par des phases d'apprentissage qui impliquent que les systèmes soient construits progressivement de façon itérative. D'autre part les projets décisionnels sont dans un premier temps, très centrés sur l'intégration de données en provenance de multiples applications, et ce moment est toujours à l'origine de la révélation de nombreux problèmes de qualité des données des systèmes opérationnels en place. Pour répondre à cette situation, il faut alors un bon système de gouvernance pour organiser les dialogues nécessaires impliquant les différentes entités métiers de l'entreprise et la direction du système d'information.

Les spécialistes de la gouvernance de systèmes décisionnels recommandent de centrer les efforts sur les points suivants :

- La structure de gestion de la gouvernance : avoir un représentant du comité de direction, choisir avec soin les participants, définir une chartre de gouvernance.
- La définition de la gouvernance souhaitable : prendre en compte la stratégie de l'entreprise, son organisation, sa culture et définir des principes, une politique et les missions des intervenants.
- La gestion du portefeuille des demandes : organiser l'identification des opportunités métiers, définir des critères de hiérarchisation des projets et des mesures de résultats, auditer systématiquement les systèmes trois ans après leur mise en production.
- La gouvernance des données : créer un comité de surveillance, définir les « propriétaires » des données, définir des politiques et des processus de gestion des données, établir des normes et des procédures pour l'utilisation des données, mettre en œuvre les vérifications et les contrôles.
- La gestion des niveaux de services (SLA) : impliquer toutes les parties prenantes dans le développement des SLA, établir des règles pour travailler ensemble, définir et mettre en œuvre les SLA.
- La gestion et l'animation des parties prenantes : impliquer toutes les parties prenantes le plus tôt possible, définir des plans de communication et de formation.

Mais attention, la gouvernance c'est avant tout de la communication. La gouvernance doit permettre la mise en place de solutions techniques appropriées et dans le cadre des systèmes décisionnels un point est particulièrement à surveiller, c'est la capacité à évoluer (en termes de volumétrie, de puissance, de nombre d'utilisateurs, de variété des usages, de complexité des services à rendus). Mais une bonne gouvernance doit par-dessus tout assurer la qualité des dialogues entre les métiers et la direction du système d'information. La gouvernance c'est avant tout de la communication et de la gestion des ressources humaines. L'organisation de nombreux va-et-vient réguliers entre les métiers et l'informatique améliorent à la fois la motivation et les compétences des hommes. Lorsque l'on met en place et anime un programme de gouvernance, il faut toujours avoir en tête la phrase de Lao-Tseu : « Aucun de nous ne sait ce que nous savons tous, ensemble ».

Développeurs fantômes et centre de compétence BI : l'expérience de Belgacom Group

Belgacom Group est le huitième opérateur mondial de Télécommunications, il est numéro un en Belgique pour les services de ligne fixe et de communications mobile avec par exemple 4,6 Millions de clients pour le mobile. Le groupe emploie 17000 personnes, est possédé à 53,5% par l'Etat Belge et développe des services de téléphonie fixe sous les marques Belgacom, Telindus, Skynet, des services de téléphonie mobile sous la marque Proximus et des services de téléphonie internationale sous la marque Belgacom International Carriers Services. D'un point de vue positionnement stratégique, Belgacom a développé une approche spécifique en jouant de façon équilibrée de différents leviers : connaissance client, leadership produit et concurrence par les prix.

Les moyens décisionnels de Belgacom s'appuient sur plusieurs systèmes qui représentent globalement un peu plus de 90T0, et ont connu depuis 1997 quatre grandes phases d'évolution :

- Intégration progressive de données issues des systèmes opérationnels : facturation, trafic, détention produit, comptabilité, gestion des réseaux, ...
- Intégration de données externes : profil client, données géographiques, informations sur les sociétés, ...
- Développement de nombreuses applications d'analyse : marchés, campagnes marketing, scores, canaux de distribution, types de contacts préférés, ...
- Développement de l'aide à la gestion opérationnelle des événements clients : centre d'appels, réclamations, actions de promotion, installations, réparations, gestion des canaux de distribution, ...

En parallèle du développement des moyens décisionnels, les relations entre les utilisateurs & l'informatique ont elles aussi évoluées, passant par des hauts et des bas dans la collaboration. Lors de la mise en place du premier entrepôt de données, l'équipe dédiée a entretenu d'excellentes relations avec les utilisateurs et l'introduction de nouveaux moyens notamment en 1998 (facturation, base de données clients, SAS Reporting) s'est passé de façon nominale. Les utilisateurs ont pu obtenir le soutien qu'ils souhaitaient pour répondre aux besoins de l'époque : forte compétition sur le marché du fixe ou à l'international et début de l'internet.

Les années passant, la direction informatique s'est laissée accaparer par la gestion quotidienne des moyens décisionnels déjà en place, a manqué de ressources pour répondre à toutes les demandes émanant des utilisateurs métiers, elle s'est en particulier heurtée à la

difficulté de développer des requêtes complexes. Les utilisateurs dans ce contexte, surtout ceux du monde Marketing/Vente, ont cherché à être plus indépendants et ont formé des équipes de développeurs fantômes au sein de leur propre division ou département, qui les ont aidé à réaliser un nombre toujours plus grand de rapports, d'analyses et surtout de requêtes ad hoc. Cette approche a été créative et a permis de couvrir des besoins à court terme, mais elle s'est accompagnée d'une explosion des coûts et d'une dégradation de la qualité des résultats produits. En effet à cette époque sur un même sujet on pouvait avoir plusieurs réponses différentes, il n'y avait plus de « vérité d'entreprise ».

Dans un troisième temps, la direction informatique s'est organisée pour consacrer plus de ressources pour développer des moyens pour les autres directions, elle a en particulier utilisé Business Objects pour créer des « univers décisionnels » et développer de nouveaux tableaux de bord. Cependant les utilisateurs ont conservé une large autonomie par rapport à l'entrepôt de données. Ils ont continué de façon indépendante à charger leurs propres données, à développer des activités de data mining, de modélisation et ont même de façon autonome commencé à mettre en place du support d'aide aux décisions opérationnelles. Les problèmes de qualité de données ont continué à se multiplier, de nombreux rapports contradictoires ont été produits, des conflits entre directions se sont créés, et beaucoup de départements ont finalement été très frustrés n'ayant pas réussi à développer tout ce qu'ils souhaitaient par manque de ressources.

Récemment une nouvelle approche a été proposée par la direction informatique, elle comprend la création d'un Business Intelligence Competency Center (BICC) et des investissements sur de nouvelles technologies (Ab Initio, Informatica, Microstrategy, le modèle logique de données). Le BICC se compose de 3 Centres d'Expertise (CoE) :

- Le Data Integration CoE a pour objectif d'intégrer les données, c'est-à-dire concrètement qu'il prend en charge les problématiques de sources de données, les processus ETL, les ODS, l'alimentation et la cohérence de l'entrepôt de données d'entreprise, de façon à garantir une « vérité d'entreprise ».
- Le Data Delivery CoE a pour objectif de mettre à disposition des données à des fins d'utilisation décisionnelle, c'est-à-dire concrètement qu'il prend en charge la création de vues métiers, le calcul de données dérivées et la mise à disposition d'objets BI.
- Le Reporting and Analytics CoE est lui très orienté métiers il prend en charge le développement de moyens pour répondre aux besoins des différentes directions de l'entreprise en matière de suivi d'activité (tableaux de bord, dash board, scorecard), d'analyse métier (requête pré formatée, olap, ad hoc), de services décisionnels pour aide à la décision opérationnelle (alerte, score à la volée), de data mining.

Le BICC regroupe donc les compétences métiers, informatiques et d'analyses avancées qui sont nécessaires pour constituer et bien exploiter un entrepôt de données d'entreprise actif. Il a permis de dissoudre et d'intégrer les groupes de développeurs fantômes, d'élever le niveau de compétence de tous les participants, et en particulier d'améliorer les connaissances métiers des informaticiens, notamment des architectes. A noter que si la mise en place du BICC s'est globalement bien passée, il a fallu cependant traiter de façon plus ou moins diplomatique quelques résistants extrémistes. Au final :

- Toutes les équipes BI sont en ligne et contrôlées par le BICC.
- Le coût des activités BI est sous contrôle et jugé optimisé.
- Les équipes BI ont conservé de la flexibilité et peuvent répondre à des besoins de court terme.

- La fourniture d'une « vérité d'entreprise » est garantie grâce à l'attention particulière accordée aux processus d'intégration des données.
- Le BICC apparaît comme une opération pilote en matière de nouvelles façons de faire travailler ensemble des utilisateurs et des informaticiens, et le top management le voit comme un moyen de maximiser l'apport de valeur du décisionnel à la stratégie de Belgacom Group.

Système décisionnel : comment concilier industrialisation et capacité d'innovation

Pour réussir la mise en place d'un système décisionnel, il faut comme pour tout autre projet de système d'information procéder avec professionnalisme et rigueur. Dès que l'opération prend une certaine ampleur, ce qui est systématiquement le cas dans les grandes entreprises, il faut s'appuyer sur une organisation industrielle par exemple aux niveaux des structures, des méthodes et des outils. Ceci devient d'autant plus nécessaire que les systèmes décisionnels sont aujourd'hui de plus en plus souvent essentiels voir critiques, pour le bon déroulement quotidien des opérations, et qu'ils doivent assurer des exigences de qualité de services de plus en plus élevées.

En mettant en œuvre une organisation industrielle, on obtiendra au final la qualité recherchée mais on en subira aussi les contraintes par exemple en termes de lourdeur, de délai d'obtention de résultats (le fameux effet tunnel), de sentiment de perte de contrôle de la part des maîtres d'ouvrage, et de délitement du support des sponsors. Outre ces difficultés, il est très difficile de mettre en œuvre une approche industrielle pour tous les lots d'un système décisionnel, c'est particulièrement le cas pour des gisements de données qui n'ont pas encore fait l'objet d'une utilisation décisionnelle, dont la qualité réelle est mal connue, ou dont les possibilités d'analyse, d'apport de connaissance et d'action restent à être déterminées.

Le besoin de flexibilité dans la construction d'un système décisionnel est d'autant plus nécessaire, qu'il s'agit d'un projet sans fin car les activités des entreprises évoluent de façon organique ou par rachat, car de nouvelles idées métiers voient le jour régulièrement, parce que la technologie offrent constamment de nouvelles possibilités, enfin parce les entreprises sont toujours à la recherche de nouveaux avantages concurrentiels et que les systèmes décisionnels sont absolument clé pour cela. L'entreprise a besoin d'une grande flexibilité pour s'adapter au marché, pour supporter l'innovation et concrètement comprimer au maximum le délai de transformation d'une idée en action.

La construction et la vie d'un système décisionnel impliquent donc d'assumer des besoins d'évolution rapides voir inopinés, de prendre en charge de nouvelles données, de nouveaux rythmes d'alimentation, de nouvelles intégration ou production de données, d'indicateurs, d'analyses, de restitutions diverses, pour des populations d'utilisateurs qui s'élargissent sans cesse. Dans ce contexte pour pouvoir concilier rigueur et flexibilité, il faut mettre en place en parallèle deux approches complémentaires, l'une classiquement industrielle et l'autre spécifiquement dédiée à l'apprentissage et l'innovation.

Pratiquement il s'agit en parallèle, du développement, de l'exploitation classique et industrielle de systèmes décisionnels, d'offrir un environnement spécifique d'apprentissage et d'innovation, ayant son fonctionnement propre en termes de gouvernance, d'organisation, de méthodologie, d'architecture, de processus, de technologie et d'exploitation. Un tel

dispositif, appelé DataLab ou SandBox doit offrir une flexibilité qui permette de rapidement & facilement tester la faisabilité, et évaluer l'intérêt métier ou technique de façons de gérer ou d'utiliser des données historiques.

Concrètement, il s'agit d'offrir à des utilisateurs métier et des informaticiens, des espaces privatifs offrant un cadre de fonctionnement adapté à leur problématique spécifique. Par exemple il convient pour un espace offrant des facilités d'apprentissage à des utilisateurs finaux, de minimiser les exigences de gestion, alors que pour un espace dédié à de la recherche et de l'innovation en matière d'analyses avancées de données, il convient de les élever. Ces espaces peuvent être utilisés pour réaliser tout ou partie des opérations ci-dessous,

- Gestion de nouvelles données : possibilité de charger, d'analyser et d'intégrer des données en complément de celles gérées par le système décisionnel existant.
- Bi & Analyses de pointe : possibilité de réaliser toutes les approches analytiques imaginables des plus simples aux plus complexes sur toutes les données des systèmes décisionnels de l'entreprise permanents ou temporaires
- Prototypage : faciliter la réalisation de prototypes d'applications décisionnelles afin de tester certains choix de conception ou de moyens.
- Services bureau : possibilité d'offrir temporairement aux utilisateurs métier des services réguliers de production de supports d'aide à la décision (indicateurs, tableaux de bord, scores).
- Dépannage production (patch run) : possibilité pour l'informatique d'épauler provisoirement son système décisionnel de production en attendant la réalisation d'une mise à niveau des moyens existants.

D'un point de vue infrastructure, différentes possibilités sont envisageables mais comme pour un entrepôt de données et pour globalement les mêmes raisons, l'approche intégrée du DataLab/SandBox est préférable à toute autre. Enfin pour rester court, il convient de noter que la deuxième clé majeure de la réussite d'un DataLab/SandBox réside dans la coopération entre utilisateurs métier & informaticiens, et dans le cas de la réalisation d'approches analytiques avancées il faut aussi enrôler des experts de l'analyse de données dans le centre de compétence spécifique qu'il convient de créer.

Les gains à attendre d'un dispositif DataLab/SandBox sont une plus grande capacité à innover et concrètement une plus grande capacité à générer, tester, industrialiser et déployer de nouvelles idées métiers utiles à la conduite des affaires de façon beaucoup plus rapide et économe. A noter que de nombreuses entreprises mettent déjà en œuvre ce genre d'approche, comme e-Bay, Bank Austria, Cabelas et en France, SFR, Leroy Merlin, FT Orange, Banque Populaire, Bouygues Telecom.