

Le BIG DATA et les Progiciels ERP : Enjeux et Défis Pour l'Entreprise

BIG DATA and ERP software packages: issues and challenges for the company

Aziz SATHAOUI

*Enseignant-chercheur à l'université O.E.B
Laboratoire Economie et Développement
(L.E.D), université de Bejaia
sathaoui_a@hotmail.com*

Matouk BELATTAF

*Professeur à l'université de Bejaia,
Laboratoire Economie et
Développement (L.E.D), université de
Bejaia
matoukb@yahoo.fr*

Received date: 14/05/2017 **Revised Paper:** 10/06/2017 **Accepted paper:** 15/06/2017

JEL: M1

Résumé:

Actuellement, grandes quantités de données structurées ou semi-structurées sont accessibles dans les BD de l'entreprises ainsi, que sur la Toile. Aussi ces entreprises ont-elles besoin de méthodes et d'outils capables d'extraire les connaissances pertinentes pour les décideurs, de les stocker, analyser et de les diffuser aux différents acteurs de l'organisation.

Devant ce phénomène de croissance exponentielle des volumes de données générées « Big Data » qui traduit usuellement par « données massives », toutes les entreprises seront concernées à terme, le défi qui attend ces entreprises sur ce sujet est : comment capturer, stocker, partager, analyser et visualiser les données qui doivent être redéfinis ? Par ce que les perspectives du traitement de grande quantité de données sont énormes.

L'objectif de cet article est de montrer comment gérées cette grosse quantité de données par l'entreprise qu'ils en deviennent difficiles à travailler avec des outils classiques de gestion de base de données. Comment l'entreprise va supporter ce challenge devant ce mélange entre "le stockage" pur et "le traitement"? Comment profitera de la disponibilité immédiate et au temps réel de l'information ?

Cet article de recherche met en perspective les défis et les avantages de Big Data pour l'entreprise afin de booster sa compétitivité, et l'amélioration de sa relation clientèle devant ce boom de volume des données de nature informatique.

Mots-clés : Big Data, données massives, l'entreprise numérique, information, bases de données

ملخص:

حاليا هناك حجم كبير من البيانات والمعطيات المهيكلة أو شبه المهيكلة متاحة في قواعد بيانات المؤسسة وكذا على الشبكة العنكبوتية، فالمؤسسة أمام هذا الكم الهائل من المعلومات والبيانات لأجل تخزينها، تحليلها وتوزيعها إلى مختلف الفاعلين من طرف صانعي القرار في المؤسسة.

أمام هذه الظاهرة من النمو الهائل في حجم البيانات «Big data» التي عادة ما تترجم «données massives»، فإن معظم المؤسسات ستكون معنية في نهاية المطاف، والتحدي الذي ينتظرها في هذا السياق هو: كيف تلتقط، تخزن، تحلل، توزع وتعرض هذه البيانات التي يجب إعادة تقييمها؟

إن الهدف من هذه الورقة البحثية هو عرض كيفية التحكم في هذه الكمية الضخمة من البيانات والمعلومات وكيفية إستغلالها من طرف المؤسسة، حيث أصبحت هناك صعوبة العمل بالأدوات الكلاسيكية لتسيير قاعدة البيانات. لذلك فهي مجبرة على مواكبة هذا التحدي والعمل للاستفادة من توفير المعلومة في الوقت المناسب من أجل تعزيز مكانتها التنافسية وتحسين علاقتها مع الزبون.

الكلمات المفتاحية: البيانات الضخمة، المؤسسة الرقمية، المعلومات، قاعدة البيانات.

Introduction

L'essor des technologies de l'information et de la communication, dont celles du Web, a multiplié le volume des informations et données numériques disponibles. Une partie de ces ressources est stockée dans des bases de données, plus ou moins structurées, développées en vue de les gérer, les retrouver et les exploiter. Cette richesse est presque toujours conservée sous deux formes : le système opérationnel de stockage des données et l'entrepôt de données. Disons pour simplifier que les données entrent par les applications opérationnelles et que nous les retirons par l'entrepôt de données (Datawarehouse).

Les utilisateurs d'une application opérationnelle font tourner les rouages de l'organisation. Ils prennent des commandes, signent des contrats avec de nouveaux clients et enregistrent des réclamations. Ils ont presque toujours affaire à un enregistrement à la fois. Ils répètent les mêmes tâches opérationnelles un grand nombre de fois.

Les utilisateurs d'un entrepôt de données, quant à eux, regardent tourner les rouages de l'organisation. Ils comptent les nouvelles commandes et les comparent avec les commandes de la semaine dernière, demandent pourquoi les nouveaux clients ont signé et sur quoi portent les réclamations des clients. Les utilisateurs d'un entrepôt de données n'ont presque jamais affaire à une seule d'information à la fois. Au contraire leurs questions exigent souvent de chercher des réponses parmi des centaines ou des milliers de lignes et de condenser dans un unique jeu de réponse. Pour compliquer davantage les choses, les utilisateurs d'un entrepôt de données modifient tout le temps le type de questions qu'ils posent.¹

La maîtrise de l'information est devenue indispensable et le système d'information qui accompagne l'intégration et la fluidité des processus joue un rôle essentiel. Toutes les fonctions de l'entreprise sont désormais couvertes par le SI. Tout le monde est concerné par des données de qualité, puisque celles-ci supportent la plupart des flux de l'entreprise. Les seules personnes qui n'ont pas besoin de s'en préoccuper sont celle qui n'en crée pas et n'en utilisent jamais. Or dans un monde irrigué d'information et exigeant de plus en plus de réactivité, qui peut prétendre vivre à l'écart des données.

1- L'émergence de l'entreprise numérique

Tout les changements associés à une redéfinition considérable des modes opératoires dans les entreprises, ont créé les conditions propices à une entreprise fortement « numérique ». une entreprise fortement numérique peut se définir selon différent aspects. Presque toutes le relations d'affaires significatives avec les clients, les fournisseurs et les employés sont établis et facilitées de manière numérique. Les procédures liées au cœur de métier s'établissent à l'aide de réseaux numérique présent dans toute l'entreprises ou reliant plusieurs entreprises entre elles.²

le "numérique" est un mot qui tend à se généraliser dans notre langage, ce terme désigne tout les usages liés à l'informatique et aux technologies de communication dont notamment Internet. Il est au cœur de tous nos usages quotidien. Il y a quelques années encore, il était envisageable de débiter une activité professionnelle avec seulement quelques stylos, un agenda papier et un bloc de factures pré-imprimées.

Mais cette époque est aujourd'hui révolue. Tout fondateur d'entreprise est d'emblée confronté aux questions du choix d'un outil informatique afin de communiquer par e-mail, établir ses premières factures ou faire connaître son activité, etc., pour ne citer que les principales tâches.

1-1 La numérisation de l'entreprise

La numérisation concourt à la virtualisation croissante des activités économiques. La virtualisation de l'entreprise peut se décomposer en trois niveaux d'analyse: ³

- *Virtualisation du travail*: Il s'agit de la dématérialisation du poste de travail. Les technologies les plus récentes le rendent de fait totalement mobile.

- *Virtualisation des processus opérationnels* : Les processus opérationnels sont de plus en plus numériques dans la mesure où les clients entrent en contact avec l'entreprise par une interface numérique : le Web. Ainsi, la gestion de la relation client (GRC), la chaîne logistique, la production et les processus internes sont de plus en plus pris en charge directement par le système d'information, comme la facturation, la gestion des stocks.

- *Virtualisation de l'organisation*: On considère désormais l'entreprise comme un ensemble de ressources matérielles (outils de production) et immatérielles (brevets, savoir-faire). L'entreprise virtuelle consiste à articuler des ressources internes et externes grâce au système d'information afin de créer des avantages concurrentiels, reconfigurables de façon instantanée pour répondre aux évolutions de la demande. Le développement des technologies de l'information rend

désormais possible cette interconnexion des ressources disséminées dans l'espace au travers de réseaux Extranet ou de place de marchés. L'organisation de cette offre est virtuelle dans la mesure où elle ne répond qu'à une demande à un moment donné pour un client donné.

Plusieurs technologies rendent possible cette virtualisation. Le tableau ci-dessous rassemble les outils électroniques et les principales applications qui correspondent au trois niveaux de virtualisation.

Tableau 1 : Outils électroniques et numérisation

Degré de numérisation	Outils électroniques	Applications
Bureau virtuel	Ordinateur de bureau Ordinateur portable Téléphone mobile Assistant personnel électronique Visioconférence	Courrier électronique Intranet Forum interne Agenda partagé Workflow Groupware
Processus virtuels	Progiciel de gestion intégré (ERP) Applications métiers CRM	Gestion d'une commande Gestion de stocks Facturation Gestion de notes de frais Gestion de la relation client
Organisation virtuelle	EDI, Extranet Place de marché	Gestion de la chaîne logistique intégrée

source : Michel KALIKA, et al., op. cit., p.47.

2- Définition & généralités sur le Big Data

2-1 Définition du Big Data

« Une expression qui circule depuis quelque temps dans la niche hi-tech de l'informatique dématérialisée (computer in the cloud) et qui fait référence aux outils, processus et procédures permettant à une entreprise de créer, manipuler et gérer de très larges quantités de données ». ⁴

Le terme anglo-saxon "Big Data" est difficile à traduire. En effet, le terme "grosse donnée", il n'existe pas de définition officielle et difficile de définir le Big data en quelques mots. Le mouvement Big data a été amorcé il y a une dizaine d'années avec l'émergence de données trop volumineuses pour être manipulées à l'aide de techniques traditionnelles telles que les bases de données relationnelles et les requêtes SQL. Les entreprises de moteurs de recherche ont été les premières à être confrontées à d'importants volumes de données pas forcément structurées. Face à ces difficultés, Google créa en 2001 une base de données compressées intitulée *Big Table* ainsi que l'algorithme MapReduce. Google publia alors ces éléments en 2004.⁵ « **Big Data = data x V3** : formule marketo-mathématique qui dégage les trois critères essentiels (Volume, Vitesse, Variété) qui permettent de distinguer ce qui est de l'ordre du Big Data de ce qui ne l'est pas ». ⁶ « Big Data est un concept à quatre dimensions (Volume, Vitesse, Variété et Véracité) ce qui me fait dire, en faisant l'analogie avec la formule marketo-mathématique, que Big **data = data x V4** ». ⁷

2-1 Information et données :

Même si l'on parle encore d'informatique, l'évolution de terminologie traduit une préoccupation fondamentale : optimiser la gestion de l'entreprise par la maîtrise de l'information. A l'inverse du système informatique qui est un ensemble de matériels et de logiciels, le système d'information n'est pas dépendant d'une plate-forme matérielle.

Il est très difficile de définir le mot *information* car il revêt un sens très large et se retrouve dans différents domaines : presse, télévision, droit, économie, gestion, ... ce qui en fait une notion

abstraire. Les dictionnaires indiquent : « action d'informer ou de s'informer ; renseignement quelconque » et confirme ainsi le caractère très étendu de ce mot.

Dans le contexte de la gestion, l'information est un ensemble d'éléments reflétant une réalité économique ou physique, susceptibles d'apporter de la connaissance utile à l'exercice de l'activité de l'entreprise. Tous les acteurs de l'entreprise, de l'ouvrier au dirigeant, ont besoin d'information pour exercer leur métier dans les meilleures conditions. C'est clairement une matière première vitale nécessaire au fonctionnement de l'entreprise.

Pratiquement, *l'information est déduite d'un ensemble de données. Les données sont donc des éléments porteurs d'information*, les deux mots ne sont pas synonymes, bien qu'ils soient souvent utilisés l'un pour l'autre. Les données peuvent aussi bien être définies par des individus que générées par le système d'information. Ainsi, la date de création d'une commande, le chiffre d'affaire d'une entité, la quantité en stock d'un article sont des exemples de données.⁸

2-3 L'origine des données du Big Data

Les données traitées par le Big Data proviennent notamment :⁹

- Du **Web**: journaux d'accès, réseaux sociaux, e-commerce, indexation, stockage de documents, de photos, de vidéos, linked data, etc. (ex: Google traitait 24 petaoctets de données par jour avec MapReduce en 2009) ;
- Plus généralement, de **l'internet** et des **objets communicants**: RFID, réseaux de capteurs, journaux des appels en téléphonie;

- Des **sciences**: génomique, astronomie, physique subatomique (ex: le CERN annonce produire 15 petaoctets de données par an avec le LHC), climatologie (ex: le centre de recherche allemand sur le climat gère une base de données de 60 petaoctets), etc.;
- Données **commerciales** (ex: historique des transactions dans une chaîne d'hypermarchés);
- Données **personnelles** (ex: dossiers médicaux);
- Données **publiques** (open data).

2-4 Les différents types de données

Si l'univers de la gestion d'entreprise utilise majoritairement les données sous forme texte, l'arrivée du multimédia et d'internet permettent d'exploiter d'autres formats de données, ce type de données est manipulées par l'ordinateur :¹⁰

- Le texte, les dessins vectoriels, l'image matricielle, la vidéo numérique et le son.

2-4-1 Les données purement transactionnelles :

On peut affirmer, sans exagérer, que ce type de données est intrinsèque au développement du commerce mondial. Rien qu'aux États-Unis, les établissements de services financiers transfèrent des trillions de dollars chaque jour pour le compte de leurs clients, ce qui représente une très grande valeur dans de très petites quantités de données.¹¹

2-4-2 Les données de reporting pour les clients :

Le reporting des données financières, notamment pour les clients commerciaux et institutionnels, présente autant d'intérêt – sinon plus –

que le traitement des transactions. Les clients savent qu'ils ne peuvent pas gérer de manière efficace leurs fonds s'ils ne savent pas où ceux-ci se trouvent. Un reporting fiable et en temps voulu, fourni par les établissements financiers leur permet d'optimiser leurs opérations financières quotidiennes.¹²

2-4-3 Les données de reporting réglementaire

La multiplication des réglementations complexifie la gestion des données d'entreprise. La rapidité, la qualité des données et la possibilité de les intégrer, bien qu'elles soient issues de sources différentes, constituent sans conteste des leviers de performance. Bien que la multiplication de sources ne signifie pas nécessairement que nous parlons des Big data, les exigences réglementaires de ces données, peu importe leur volume et leur nombre, indiquent bien que nous parlons des big data stratégiques.

Tableau 2 : Principales caractéristiques de qualité des données

Qualité du contenu	<ul style="list-style-type: none">• Justesse• Adéquation / besoins• Compréhension
Accessibilité	<ul style="list-style-type: none">• Disponibilité• Facilité d'accès
Flexibilité	<ul style="list-style-type: none">• Evolutivité• Cohérence avec d'autres sources• Traduction
Sécurité	<ul style="list-style-type: none">• Respect de la confidentialité• Fiabilité• Traçabilité• Intégrité

Source : Christophe Brasseur, Op cit, p 71.

3- Les données : pour faire quoi et pour qui ?

La concurrence de plus en plus vive entre société impose désormais d'être très vigilant quant à la qualité de l'information, car c'est un

gage de compétitivité.

L'information dans l'entreprise est variée et dépend des besoins des différentes fonctions. Tous les métiers de l'entreprise ne peuvent être exercés avec efficacité que si l'information adéquate est disponible au bon moment, avec le niveau de précision adapté. Certaines informations doivent par ailleurs être disponibles pour des raisons légales et fiscales. On peut distinguer :¹³

- Les contraintes légales ;
- Les besoins opérationnels ;
- Les besoins d'analyse et l'aide à la décision.

Les personnes qui sont concernées par l'information :¹⁴

-Les dirigeants : ils sont concernés au plus haut point, puisque les tâches de définition de la stratégie, de prise de décisions, de pilotage de l'entreprise nécessite plus que jamais d'information de grande qualité. Les dirigeants sont conscients qu'avec des données de qualité, leur entreprise est plus forte, plus réactives, plus flexible donc plus compétitive. Aussi, mieux connaître ses clients pour les satisfaire davantage est l'une de leurs préoccupations majeurs. La qualité des informations contenues dans la base clients est à cet égard fondamental.

-Les responsables opérationnels : ces managers opérationnels gèrent au quotidien une grande quantité de tâches et de problèmes variés. Ils sont utilisateurs d'un grand nombre d'informations et, en conséquence, la qualité des données impacte directement leur travail. Ils doivent en effet s'assurer à la fois de la bonne marche des opérations.

-Les collaborateurs opérationnels : au quotidien, le travail de grande

majorité des employés est largement supporté par des données. Ainsi, un approvisionneur doit connaître pour chaque article un niveau de stocks fiable, sous peine de subir des ruptures de stock intempestive ou au contraire de commander à tort de la marchandise. Cet exemple pourra être multiplié à l'infini et il n'est pas exagéré d'affirmer que tous les métiers opérationnels de l'entreprise nécessitent à un moment ou autre la création ou l'utilisation de données correctes.

-Les responsables informatiques : ils se préoccupent en permanence de la qualité de l'information qui circule dans le système d'information. Si un problème de données survient, c'est naturellement vers leurs équipes que l'on se tourne. Le rôle le plus important des responsables informatiques et de leurs équipes à l'égard des données est sans aucun doute l'animation et la promotion du rôle-clé de la qualité de l'information. Faire prévention, former, conseiller, aider, accompagner les utilisateurs font partie intégrant de leur job.

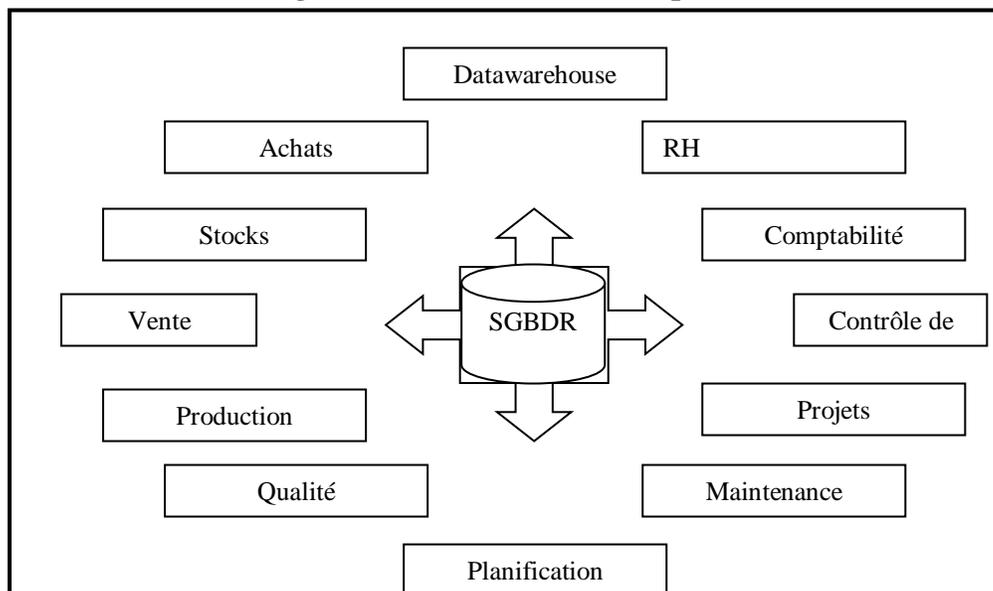
4- Le Big Data & SI de type ERP

L'accès des entreprises aux nouvelles technologies, à Internet en particulier, tend à modifier la communication entre les différents acteurs, notamment entre l'entreprise et ses clients, le fonctionnement interne de l'entreprise et la relation de l'entreprise avec ses différents partenaires et fournisseurs. L'agrégation au sein de l'entreprise d'outils basés sur les technologies de l'information et la communication, en l'occurrence les *Enterprise Resource Planning* (ERP).

Un ERP est un progiciel qui permet de gérer l'ensemble des processus d'une entreprise intégrant l'ensemble de ses fonctions comme la gestion des ressources humaines, la gestion financière et

Le BIG DATA et les progiciels ERP... Aziz SATHAOUI & Matouk BELATTAF
comptable, l'aide à la décision, la vente, la distribution, l'approvisionnement et la production ou encore du e-commerce.

Figure 1 : Domaines couverts par un ERP



Source : Christophe Brasseur, Op cit, p 48

Le principe fondateur d'un ERP est de construire des applications informatiques correspondant aux diverses fonctions citées précédemment de manière modulaire sachant que ces modules sont indépendants entre eux, tout en partageant une base de données unique et commune au sens logique. L'autre principe qui caractérise un ERP est l'usage de ce qu'on appelle un moteur de workflow * ¹⁶ et qui permet, lorsqu'une donnée est enregistrée dans le système d'information, de la propager dans les modules qui en ont l'utilité, selon une programmation prédéfinie.

4-1 Les raisons de l'essor des ERP

Le succès des ERP s'explique par plusieurs facteurs :¹⁷

- Recentrage des entreprises sur leur métier d'origine : dans

environnement économique difficile, les entreprises ont tendance à se recentrer sur leur métier d'origine, toutes les directions et services sont concernés par ce mouvement de recentrage, les directions informatiques ont largement été impliquées dans cette démarche et les développements d'applications spécifiques souvent longs, coûteux et incertains ont laissé peu à peu place à l'implémentation de solutions packagées de type progiciels métier ou ERP.

- Optimisation des processus : dans chaque domaine fonctionnel, l'ERP s'appuie sur les meilleures pratiques de gestion, aussi, la mise en place d'un ERP a l'avantage de réduire voire supprimer les coûts liés au développement du logiciel conçu et développé, l'entreprise peut alors se consacrer au paramétrage du progiciel (pour que celui-ci colle à son organisation et sa stratégie), à la conduite du changement, etc., la mise en place d'un ERP est souvent l'occasion de repenser et améliorer les processus de l'entreprise.
- Intégration et cohérence transversale des données de l'entreprise : c'est incontestablement un point fort des ERP, s'appuyant sur un SGBD relationnel, les données sont partagées par toutes les fonctions et ne sont saisies qu'une seule fois à la source, les échanges entre différents services ou départements d'une entreprise sont fluidifiés, l'ERP joue donc un rôle moteur dans le décloisonnement des structures de l'entreprise, qui devient ainsi plus agile.
- Maîtrise des coûts informatiques : les coûts liés à la technique informatique sont diminués au profit d'un investissement dans l'optimisation des processus, la mise en place d'un ERP se fait généralement avec l'aide de prestataires extérieurs, cette sous-

traitance permet clairement une meilleure maîtrise des coûts, tout en s'appuyant sur de solides compétences produit et métier.

Les raisons de ce succès ne doivent cependant pas masquer les obstacles parfois rencontrés lors de la mise en place d'un ERP. L'impact sur l'organisation et les processus nécessite généralement un effort d'adaptation non négligeable pour l'entreprise.

4-2 Les facteurs-clés de succès d'un projet ERP

La réussite d'un projet ERP repose globalement sur : ¹⁸

-L'implication du top management de l'entreprise : l'ERP touche de près à l'organisation, aux structures et aux processus, il est donc essentiel que le top management soit moteur tout au long du projet et que la ligne directrice soit clairement tracée par les dirigeants, la motivation et l'appropriation de l'ERP sont à la hauteur de l'implication des managers ;

-Une équipe projet de qualité : les équipes projet doivent être de qualité car une forte implication est requise, la maîtrise du produit, la compréhension du métier et des processus, le respect des délais, des facultés d'adaptation et de communication sont nécessaires, aussi, il convient de ne pas négliger la dimension technique avec la mise en place de compétences spécifiques dans ce domaine, enfin, l'équipe de management du projet joue un rôle-clé dans la réussite de projet (planification, budget, suivi, arbitrage, gestion des compétences, ...), cet accompagnement dans l'expertise produit, métier, technique et management est généralement réalisé par ou avec l'aide de consultants expérimentés dans ce type de produit et de projet.

Le marché des ERP est largement dominé par SAP. De nombreux

autres éditeurs se partagent le reste du marché : Oracle, Peoplesoft (racheté par Oracle), Intentia, JD Edwards, Adonix, ... et bien d'autres.

4-3 Les bénéfices fonctionnels attendus d'un ERP

Le tableau suivant souligne seulement quelques exemples qui, tout en étant limités, montrent bien les dimensions d'intégration, de réduction de temps de cycle et de réduction des erreurs ; bref, une vraie rupture et un meilleur service pour le client final de l'entreprise.

Tableau 3 - Exemple des bénéfices attendus d'un ERP

Fonctionnalité	Bénéfices
Prix en temps réel sur les commandes clients	Réduction des erreurs de prix et des efforts manuels
Identification physique automatique des produits à livrer	Réduction des erreurs, élimination de l'identification manuelle des produits
Possibilité d'annuler ou d'inverser une expédition avant facturation	Gain de temps et d'effort pour procéder aux multiples opérations nécessaires
Disponibilité d'un suivi de commande client, de la cotation à la facturation	Possibilité multiples de recherche et de suivi à n'importe quel moment
Visibilité sur inventaire et fabrication pour planifier les commandes clients	Réduction de temps et d'effort pour s'engager avec un client
Définition de critères client spécifiques pour expédier une révision de produit	Assurance du traitement intégral de la demande spécifique d'un client

Source : J-L. TOMAS et Y. GAL, *ERP et conduite des changements*, 6^e édition, DUNOD, Paris, 2011, p 15

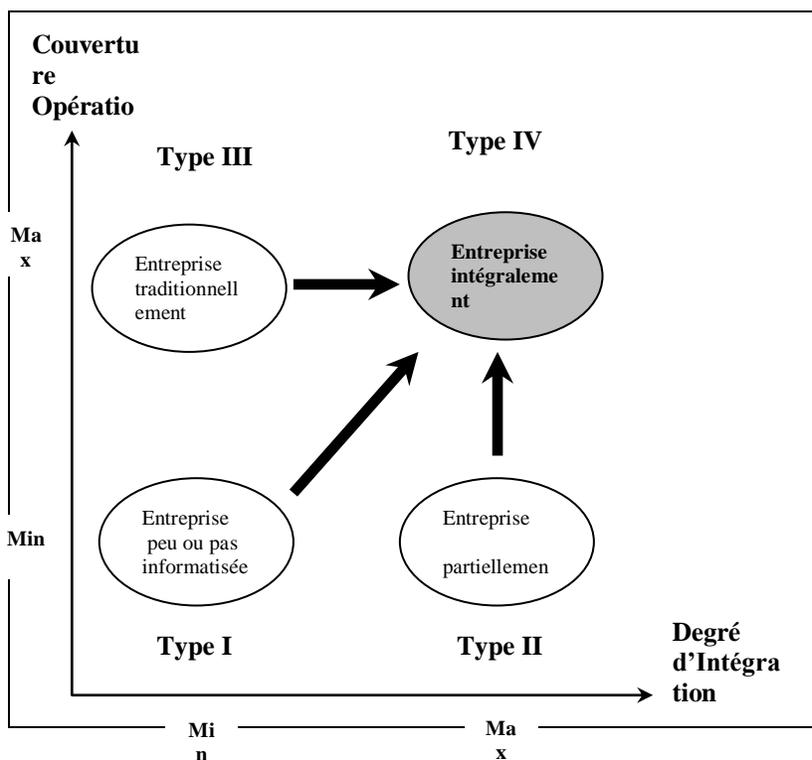
5- Le degré d'informatisation des entreprises d'aujourd'hui

Toutes les appellations d'ERP ne recouvrent pas bien sûr les mêmes concepts, et nous aurons l'occasion de souligner les différences. Du fait de la place prépondérante qu'ils utilisent sur les marchés actuel et futur, de leur problématique et de la complexité de leur mise en œuvre, cette diversité prend en compte deux dimensions capitales dans le concept d'ERP : le **degré d'intégration** et la

couverture opérationnelle qui peuvent se définir de la façon suivante :¹⁹

Le **DI** définit la capacité de fournir à l'ensemble des acteurs de l'entreprise une image unique, intégrée, cohérente et homogène de l'ensemble de l'information dont ils ont besoin pour jouer pleinement leur rôle. La **CO** définit la capacité de fédérer l'ensemble des processus de l'entreprise dans chacun des domaines qui la constituent, et ce, dans une approche transversale qui optimise sa productivité. Nous pouvons dire que les entreprises d'aujourd'hui sont, selon leur degré d'informatisation, de quatre grands types.

Figure 3 - Les quatre niveaux d'informatisation d'une entreprise



Source: Jean-Louis TOMAS, *ERP et progiciels intégrés*, 2^e édition, InterEditions, Paris, 1999, p 18

- L'entreprise de type (Min DI, Min CO) est ***peu ou pas informatisée*** : tout reste à faire. Aucun des processus de l'entreprise n'est couvert. Les fonctions de bureautique, lorsqu'elles existent, représentent les seuls îlots informatisés.
- L'entreprise de type II (Max DI, Min CO) est ***partialement informatisée*** : l'effort s'est concentré que quelques processus clés de l'entreprise qui relativement bien intégrés entre eux. En revanche, une grande partie des processus clés et des processus secondaires doivent encore être couverts.
- L'entreprise de type III (Min DI, Max CO) est ***traditionnellement informatisée*** : c'est l'entreprise type. Elle représente la grande majorité des entreprises aujourd'hui. L'ensemble des processus est plus ou moins bien couvert avec des solutions qui y ont été progressivement implantées. L'intégration des solutions est minimale, et seules des interfaces assurent le lien entre elles.
 - L'entreprise de type (Max DI, Max CO) est ***intégralement informatisée***. C'est le niveau idéal vers lequel toute entreprise doit tendre. L'ensemble des processus est couvert au sein d'une solution globale et intégrée qui satisfait pleinement les acteurs de l'entreprise. De toutes les solutions existant sur le marché aujourd'hui, seul un ERP permet à une entreprise de se rapprocher du type IV. Mais comme nous le verrons plus loin, les ERP sont aussi les solutions qui demandent le plus de temps et le plus de ressources pour être implantés »²⁰

6- L'organisation de la gestion des données

Après la bonne compréhension des besoins des utilisateurs qui est

le point de départ de toute activité de conception ou de maintenance de système d'information, il vient l'étape de la codification des données (définir le format, les valeurs possibles, s'appuyer sur les normes,...), et puis la documentation et l'administration des données, le nettoyage des bases de données (clean-up), la surveillance des données, la sécurité des données et l'auditabilité des données.²¹

6-1 La sous-traitance de l'entrepôt de données (Datawarehouse)

Au plus fort du tapage auteur d'internet, un espoir est né de voir les fournisseurs de service d'application (*en Anglais ASP, Application Service Providers*) décharger les services informatiques d'une bonne partie du travail de développement et de support des nouveaux types d'entrepôt de données qu'exige le commerce électronique. Certains d'entre eux étaient affiliés à des fournisseurs de solutions. L'intérêt éventuel d'une telle solution demeure très réel, mais ici encore, profitant de pause qui nous est octroyée, nous nous devons d'estimer les risques du modèle ASP, ainsi que ses avantages. Si nous ne sommes plus désespérément pressés d'implémenter notre e-business et notre e-Datawarehouse, pourquoi confierions-nous une responsabilité stratégique clé à un outsider ?

6-2 Les avantages du modèle ASP pour l'entrepôt de données

- L'ASP a déjà des compétences que le service informatique n'a pas et ne peut peut-être pas acquérir facilement ;
- L'ASP a configuré un ensemble complet de matériels et de logiciels dont on sait qu'ils fonctionnent bien ensemble ;
- L'ASP dispose de capacités disponibles en matériels pour faire face à une croissance explosive de besoins sur le Web ou au

rétablissement de l'exploitation ;

- L'ASP bénéficie d'économies d'échelle par la mise en commun des opérations de sauvegarde et de restauration ;
- Les coûts de ce modèle peuvent être isolés et gérés de façon plus claire que ceux d'un service interne ;
- L'ASP assure sa propre gestion de personnel.

Ces avantages potentiels considérables sont à mettre en regard des risques.

6-3 Les risques du modèle ASP pour l'entrepôt de données

Parmi les risques du modèle ASP sont les suivants :²²

- La définition d'un accord sur un niveau de service pour l'entrepôt de données est une tâche sophistiquée et l'industrie de l'entreprise parfois n'a pas beaucoup d'expérience en la matière. En tout cas, cet accord doit être élaboré par l'entreprise, non pas par le service de l'ASP ;
- Un ASP peut faire faillite. Un accord relatif à la fourniture des codes sources n'est pas une grande consolation dans cette hypothèse ;
- Un ASP peut avoir son propre programme d'évolution du logiciel, et il ne voudra probablement pas faire des modifications au logiciel standard qu'il propose s'il a de nombreux clients pour de nombreuses applications ;
- Un ASP peut avoir les concurrents de l'entreprise comme des clients.

7- Sur quelles technologies reposent les solutions Big Data ?

Le phénomène Big Data suscite des investissements majeurs donnant naissance au développement d'une variété d'applications ou d'outils adaptés :²³ *Graph Data bases*, de *Framework* comme *MapReduce* ou *Hadoop* et de systèmes de gestion de bases de données comme *BigTable*.

Pour analyser les hautes et disparates volumétries de données, les solutions peuvent s'appuyer sur trois types de technologies distinctes : une accélération matérielle à l'aide de mémoires dynamiques *DRAM* ou *Flash*, le recours à des bases de données massivement parallèles (*Massively Parallel Processing*) ou encore les solutions utilisant des formats de bases de données non relationnelles basées sur *NoSQL*.

Selon le cabinet d'études *IDC*, l'un des principaux moteurs de ce marché est la disponibilité croissante de données interactionnelles, d'attitude et de comportement à partir de sources telles que les médias sociaux (*Facebook*, *Google+*, *twitter*, ...) et d'autres applications web, et une volonté de creuser ces données au niveau de leur valeur.

Outre la rareté des outils et du personnel qualifié, un facteur qui pourrait inhiber la croissance du marché est la concurrence entre les éditeurs *Open Source* (*Source Ouverte*) et leurs homologues proposant des logiciels propriétaires, ce qui pourrait conduire à une pression sur les coûts de licences de ces derniers. Cela devrait résulter d'une croissance plus faible que si les projets *Open Source* ne représentaient pas une si grande proportion de ce marché.

7-1 Les différents types d'analyse du Big Data

L'analyse des données disponibles peut se classer en 2 catégories majeures, l'analyse des données externes à l'entreprise et l'analyse des données internes :²⁴

- L'analyse des données externes : les données externes sont les données qui ne sont pas produites par l'entreprise, comme les conditions climatiques, les informations fournies par les Etats ou les études de marché. L'analyse de ce type de données permet d'anticiper et de prévoir des procédures en cas de problèmes. C'est ce que l'on appelle l'analyse des chocs externes qui pourraient soit bénéficier, soit fragiliser la position d'une entreprise.
- L'analyse des données internes : l'analyse des données internes est en résumé ce que l'on appelle le Business Intelligence, la capacité à utiliser les données générées par les actions de l'entreprise pour identifier des tendances et des informations clé, comme la segmentation des clients, leur valeur dans le temps ou encore leur coût d'acquisition. Les entreprises comme Oracle et SAP sont des exemples d'acteurs de ce marché.

Conclusion

Les entreprises progressent dans l'organisation et l'examen de leurs données internes, telles que les interactions des centres d'appels, les transactions commerciales et les niveaux des produits. En revanche, elles ne sont pas préparées aux informations externes. Or, au final, ce qui constitue l'essence même des grandes quantités de

données, ce sont les informations externes sur le web, notamment les médias sociaux. Cependant, la plupart des entreprises se méfient de la validité ou de la véracité de ces données.

C'est ce que l'on appelle les « BIG DATA ». Avec la montée en puissance de la numérisation, les entreprises bien établies dans des secteurs tels que la télécommunication et les médias, la santé ou encore la finance ont accumulé des terabytes de données sur leurs légions de clients. Cette richesse de données numériques, déjà prisee comme un excellent moyen de répondre aux besoins évolutifs des clients et de repérer les tendances, peut aider les entreprises à créer de nouveaux produits et services et peut même les aider à initier de nouvelles affaires. Un grand nombre d'organisations ayant une stratégie orientée vers les clients ont une avance naturelle dans ce domaine ; pour elles, utiliser les BIG DATA représente une opportunité particulièrement lucrative.

Bibliographie :

-
- ¹ - Ralph Kimball et Margy Ross, *Entrepôts de données*, traduit de l'anglais par Claude Raimond, 2^{ème} édition, Vuibert Informatique, Paris, 2003, p2
 - ² - Kenneth LAUDON, et al., *management des systèmes d'information*, 11^{ème} éd., PEARSON, Paris, 2010, p. 11.
 - ³ - Michel KALIKA, et al., op. cit., p.46.
 - ⁴ - Ayoub Figui, *big data : la sécurité sera-t-elle au rendez-vous*, < <http://blogs.orange-business.com/securite/2012/12/big-data-la-securite-sera-t-elle-au-rendez-vous.html> >, 28/01/2013.
 - ⁵ - Christophe Brasseur, *Enjeux et usages du Big Data*, collection management et informatique, Lavoisier, Paris, 2013, p.33.
 - ⁶ - Ayoub Figui, Op cit.
 - ⁷ - Idem.

⁸ - Christophe Brasseur, *Data Management qualité des données et compétitivité*, Lavoisier, Paris 2005, p 23

⁹- Stefane Fermigier, *BIG DATA & OPEN SOURCE : UNE CONVERGENCE INÉVITABLE?*, VERSION 1.0 - MARS 2012, P 5.
[<http://www.fermigier.com/assets/pdf/bigdata-opensource.pdf>], 29/12/2012.

¹⁰ - Christophe Brasseur, Idem, p 29

¹¹ - Susan Feinberg, *Les Big data vraiment stratégiques de l'entreprise...*, Axway, [<http://www.decision-achats.fr/Breves/Les-big-data-vraiment-strategiques-de-l-entreprise-49282.htm>], 23/10/2012.

¹²- Susan Feinberg, Idem.

¹³ - Christophe Brasseur, Op cit, p 24

¹⁴ - Idem, p 15, 16

*- Un **workflow** est un flux d'informations au sein d'une organisation, la modélisation et la gestion informatique de l'ensemble des tâches à accomplir et des différents acteurs impliqués dans la réalisation d'un processus métier.

¹⁶ - [<http://www.techno-science.net>], 02/01/2013.

¹⁷ - Christophe Brasseur, Op cit, p 48, 49

¹⁸ - Idem, p 49, 50

¹⁹ - Jean-Louis TOMAS, *ERP et PGI*, 4^e édition, Dunod, Paris, 2005, p 10, 11

²⁰ - Jean-Louis TOMAS, *ERP et progiciels intégrés*, 2^e édition, InterEditions, Paris, 1999, p 17, 18

²¹ - Christophe Brasseur, Op cit, p 17

²² - Ralph Kimball et Margy Ross, Op cit, p 395, 396

²³[http://www.telecomreview.info/index.php?option=com_content&view=article&id=185:le-big-data--une-solution-pour-absorber-la-masse-dinformations-proteiforme&catid=27:new-to-joomla&Itemid=58],23/04/2017.

²⁴ - [<http://www.romainpechard.com/analyse-outils/big-data-methodes-outils-et-benefices-pour-lentreprise/5076/>], 23/04/2017.