



# LES ENJEUX DE LA DATA ET DU CLOUD POUR VOTRE ENTREPRISE EN 2019



T'SOFTWARE  
SOLUTIONS  
<http://www.tse-online.fr>



# À PROPOS DE L'AUTEUR



Jean-Christophe THOMAS

Bénéficiant de plus de 10 ans d'expérience autour de la conception et du développement d'applications de gestion web pilotées au niveau serveur de données, Jean-Christophe est avant tout un passionné de data.

Certifié Expert Microsoft en management et analyse de données, et après avoir mené de nombreux projets au profit de PME et grandes entreprises, il recentre son expertise et ses services autour de la Data.

Convaincu de l'importance grandissante de la gestion de la qualité des données dans un référentiel maîtrisé, il accompagne aujourd'hui ses clients dans des projets visant à pérenniser et maîtriser leurs coûts d'exploitation autour de leurs données à l'aide de méthodes à la fois structurantes et innovantes pour les accompagner dans leur développement.

# SOMMAIRE



- 1 INTRODUCTION : LE CONSTAT ..... p3 à p4
- 2 PRESENTATION DE LA « BI » OU BUSINESS INTELLIGENCE ..... p5
- 3 TPE/PME : LES PRINCIPALES PROBLEMATIQUES  
DES SOLUTIONS E-COMMERCE OPEN SOURCE AU NIVEAU DATA ..... p6
- 4 LES APPLICATIONS D'EXPLOITATION ET D'E-COMMERCE SUR MESURE ET LEUR ENJEUX AU NIVEAU DATA ..... p7 à p8
- 5 POSSÉDER ET MAÎTRISER SON RÉFÉRENTIEL DE DONNÉES ..... p9
- 6 LE CLOUD ..... p10
- 7 LES IDÉES PRÉCONÇUES SUR LE CLOUD LES PLUS FRÉQUEMMENT RENCONTRÉES ..... p11
- 8 LES PRINCIPAUX AVANTAGES INCONTOURNABLES DU CLOUD ..... p12

## INTRODUCTION : LE CONSTAT

Toutes les entreprises ou presque débutent leur activité avec les mêmes contraintes et objectifs :  
Mettre en place un investissement minimal avec mise en production rapide pour un ROI à court terme voir moyen terme.

Pour atteindre cet objectif la majorité des TPE et certaines PME démarrent avec une solution e-commerce Open Source (Wordpress, Prestashop, etc...) facile et rapide à mettre en œuvre, tandis que d'autres dotés de plus de moyens, vont démarrer avec une solution ERP plus robuste pour l'avenir de leur données.

Quelle que soit la solution initiale retenue, le résultat est que le modèle initial de données de chaque entreprise est établi dans plus de 95% des cas par le modèle de l'application web commerciale retenue.

Il en résulte dans les mêmes proportions un référentiel initial de données dégradé dont les principales causes sont (et/ou) :

- Rareté de développement applicatif web en bases de données épaisses pour les applications développées sur mesure.
- Une culture insuffisante ou absence de compétences en bases de données des développeurs même dit « full stack » de façon générale.
- Des modèles de données open source dénormalisés à fortement dénormalisés et surtout inadaptés aux processus spécifiques de chaque entreprise qui sont obligés de s'adapter à l'outil utilisé et de travailler en faisant des compromis dans leur gestion interne.



Aujourd'hui et encore plus demain on assiste à la multiplication du besoin d'interconnexion avec plusieurs référentiels de données dont les deux principaux facteurs sont :

- Le besoin d'automatiser son processus de vente : connexion avec différents catalogues fournisseurs et/ou transporteurs principalement (sites e-commerce, ...).
- L'utilisation d'outils spécialisés dans chaque département de l'entreprise et/ou sites de l'entreprise, qui possèdent leur propre référentiel(s) de données afin d'être autonomes (SAGE, CRM, etc...)

D'autre part, dans le développement sur mesure d'applications/outils spécifiques d'exploitation, les délais imposés sont de plus en plus tirés vers le bas, et les budgets de plus en plus ajustés et/ou réduits.

La qualité de ces mêmes applications suit par conséquent la même tendance et délèguent le plus souvent le contrôle de qualité de saisie des données à l'opérateur de saisie (réduction coûts et délais), quitte à alourdir considérablement le besoin et les coûts de formation et d'exploitation pour ces applications pour l'entreprise en prime.

En conclusion, le constat est sans appel :

- 20 à 35% d'erreurs de saisie sont généralement constatés dans les données d'entreprise.
- Présence de doublons (client, etc...) dans le système d'information rendant difficile voire impossible l'identification des données à jour parmi celles qui sont obsolètes.
- Le référentiel de données n'est pas représentatif des besoins ni des spécificités d'exploitation de l'entreprise
- Le référentiel de données n'est pas maîtrisé par l'entreprise car le plus souvent externalisé voire non accessible.



## PRESENTATION DE LA BI

### — — — — — OU BUSINESS INTELLIGENCE

Aussi appelée « **informatique décisionnelle** » ou plus communément « BI », la business intelligence désigne l'ensemble des outils et services visant à fournir une aide à la décision aux entreprises en se servant d'un élément précieux : **leurs datas**. En 2016, contrairement aux métiers du Big Data ou de l'analyse prédictive, la business intelligence se centre plus particulièrement sur les **données des systèmes d'entreprise**.



C'est l'extraction de ces données, puis leur **transformation intelligente** en tableaux de bords et indicateurs de performance (dit aussi « KPI », « Key Performance Indicator ») qui permettra d'opérer des **choix stratégiques** grâce à la grande visibilité que ces solutions offrent aux utilisateurs et aux décideurs.

Reporting, état de l'art, simulations, projections, etc. Ces informations pèsent lourd quand il s'agit d'orienter la feuille de route de l'entreprise. Ainsi, aider à mieux voir, comprendre et anticiper en dévoilant les trésors qui se cachent derrière les datas : telle est la vocation de la business intelligence.

Longtemps cantonnée aux départements finance et vente, la BI connaît aujourd'hui une vraie transformation, en devenant une **expertise transverse** avec une **véritable portée stratégique**.



## TPE/PME : LES PRINCIPALES PROBLEMATIQUES DES SOLUTIONS E-COMMERCE OPEN SOURCE AU NIVEAU DATA



- ☞ Ces solutions ne **sont rentables** dès le moyen terme **que si elles couvrent au moins 80% des besoins fonctionnels nativement**
- ☞ Elles **imposent de fait un modèle de données peu adapté à l'entreprise** et surtout **dénormalisé** à fortement dénormalisé (présence de redondance, de duplication d'information, et de façon générale ne répondant pas aux formes normales de CODD)
- ☞ Elles utilisent MySQL comme moteur de données et par conséquent :
  - ★ Ne possèdent pas une capacité de montée en charge acceptable pour un usage professionnel sur le long terme
  - ★ Ne possèdent pas de capacités d'indexation suffisante au niveau des données limitant fortement leur performance
  - ★ Ne permettent pas de gérer la cohérence des données ni un référentiel de données de façon qualitative
  - ★ Ne permettent pas de faire du reporting dans un contexte transactionnel de façon acceptable
- ☞ Elles nécessitent un usage abondant de plugins / modules qui gèrent leurs propres données sans logique globale qui vous soit forcément pertinente
- ☞ Elles obligent à surdimensionner l'infrastructure au niveau data très rapidement pour maintenir un service « acceptable »



## LES APPLICATIONS D'EXPLOITATION ET D'E-COMMERCE SUR MESURE ET LEUR ENJEUX AU NIVEAU DATA

Développer une application sur mesure est la plupart du temps synonyme d'une application conçue et réalisée dans les règles de l'art, avec un modèle de données associé conforme aux besoins, aux contraintes et aux règles internes de gestion et de validation de l'entreprise.

De même, le chef d'entreprise valide les choix et les compromis devant être faits en connaissance de cause durant les développements, car il est accompagné par un ou des professionnels tout au long du processus pour ce faire.

La plupart du temps les compromis qui sont fait (budget, délais, etc...), le sont **toujours fait à tort au détriment de la data**. Pourquoi ? Les trois principaux facteurs sont :

- ★ Absence de conception du modèle de données, voire absence de dictionnaire de données et/ou d'implémentation des règles de gestion et de validation dans le modèle de données.
- ★ Le contrôle de validité de saisie est supprimé et seul un contrôle de type et/ou de longueur de donnée est parfois mis en place. Par exemple l'absence de caractères non numériques ne sera pas contrôlé pour des codes postaux Français avant validation en base de données.
- ★ Absence de contrôle de validité des règles de gestion de données avant validation dans le modèle de données. Ce contrôle est reporté au niveau de l'utilisateur / opérateur de saisie.

**☞ L'utilisateur pouvant faire, saisir, valider des choses qu'il n'est pas censé pouvoir faire, la qualité, la cohérence voire l'intégrité des données est presque systématiquement impactée à fortement impactée.**  
**Ces défauts sont généralement également présents dans les solutions e-commerce Open Source**



 **NE PAS OUBLIER**

Ce genre de choix **fait économiser une fois** de l'argent de manière importante  
**MAIS**  
**Génère des coûts d'exploitation conséquents de manière récurrente et crée de nouveaux besoins pour corriger les problèmes générés systématiquement**

« Ainsi j'ai pu constater des dépenses de quelques centaines d'euro par mois pour des TPE à plusieurs centaines de milliers d'Euros par mois pour des PME/Grandes Entreprises de dépenses récurrentes en développeurs et techniciens pour résoudre principalement des problèmes de qualité de données au quotidien au sein de leur entreprise. Cela génère également de nouveaux projets non prioritaires en gestion de qualité de données conséquents et qui ne règlent environ qu'entre 50 et 75% des problèmes rencontrés au quotidien au niveau production. »

### **Le Conseil de l'expert**

Dans le cas où les prérequis de votre projet sont flous, ne faites pas « du provisoire » en pensant refaire votre modèle de données ou votre application une fois que le périmètre fonctionnel sera clair et validé. L'expérience montre, qu'une fois en production il vous sera délicat de la remplacer sans rupture de service et sans conséquences sur vos employés et/ou clients que cela soit au niveau productivité, prise en main, formation, expérience utilisateur, habitudes pour ne citer que ces cas.

De même ne faites jamais de compromis sur la gestion de qualité de vos données ou il pourra vous en coûter cher sur le long terme.



# POSSÉDER ET MAÎTRISER SON RÉFÉRENTIEL DE DONNÉES



Quelle que soit l'entreprise, tout projet data, comme un projet BI, commence par un état des lieux dans lequel le contrôle de la qualité des données est fondamental.

Les données doivent être en effet préalablement propres, uniques et fiables pour inspirer confiance aux décideurs.

Aujourd'hui posséder un tel référentiel de données est un avantage concurrentiel certain pour toute entreprise sur son marché.



**N'hésitez pas à faire auditer vos données par un consultant BI si vous avez des doutes quant à leur fiabilité.**

Certains signes ne mentent pas : redondance des informations, latence de mise à disposition des données, lenteur inhabituelle des processus de reporting ou dérive des projets... Ne cherchez plus : vos données sont certainement éparpillées, dupliquées ou mal nommées!



**Le bon réflexe**

**Mettez en place une stratégie visant à fiabiliser vos données dans un référentiel maîtrisé au plus tôt !**

*Parmi les nombreux avantages ....*

- ★ Vous gardez la maîtrise sur votre entreprise et son développement
- ★ Coûts d'exploitation optimisés, plus grande compétitivité sur votre marché
- ★ Investissements pérennisés sur le long terme : vous ne changez pas de référentiel lors d'un changement applicatif
- ★ Les coûts généralement élevés d'une telle démarche sont réduits et étalés dans le temps
- ★ Vous n'attendez pas de pouvoir vous lancer dans un projet BI pour bénéficier de bons indicateurs pour piloter votre entreprise

## LE « CLOUD »



Le cloud est le terme pour représenter une infrastructure informatique managée distante qui offre plusieurs types de services qui lui sont spécifiques dont les principaux sont :

- Délivrer un service web au plus proche de vos clients et prospects partout dans le monde en quelques secondes de déploiement
- Mise à disposition de stockage, bande passante réseau et de ressources CPU en « illimité » dans un modèle où l'on paye que ce que l'on consomme.
- Gestion du « scaling out » et « scaling up » nativement en temps réel ou presque.

L'approche "scale-out" consiste à ajouter des unités de manière "horizontale" à un ensemble de ressources. Par exemple, d'ajouter un serveur à un ensemble de serveurs dans un centre de données ou à un cluster. Il peut s'agir également d'augmenter le nombre de serveurs web frontaux à un site web pour faire face à un pic de connexions simultanées dépassant les capacités réseau des serveurs déjà en place.

L'approche "scale-up" consiste à ajouter des ressources à un serveur existant. Exemple : plus de mémoire, disques, CPU, etc...

### *Ce qu'il faut savoir*

La plupart des solutions cloud (Google, Amazon AWS, cloudflare, ...) sauf Windows Azure notamment ne garantissent pas la géolocalisation de vos données en cas de perte ou vol. Faites attention si vous avez des contraintes de juridiction en cas de recours ou de conformité légale vis-à-vis de vos données.

## Les idées préconçues sur le cloud les plus fréquemment rencontrées



1) « Une application web développée pour le cloud ne peut pas fonctionner en dehors du cloud »

 **Faux!**

En général le cloud est un espace plus restrictif et sécuritaire par défaut que peut l'être un serveur dédié ou un VPS. Une application optimisée pour le cloud est en général une application capable de prendre en charge le « scaling-out » et des serveurs de gestion de cache comme Redis. Donc, qui peut le plus peut le moins.

L'inverse par contre n'est pas toujours vraie :

- une application web hébergée sur un VPS/serveur dédié/serveur mutualisé nécessitera une mise à jour si elle est migrée dans le cloud et qu'elle doit supporter le « scaling-out ».
- L'envoi de mails à partir d'une application web, ne pourra plus fonctionner si elle est migrée dans le cloud sans faire appel à un service de mail externe via API comme par exemple mailjet.

2) « Une solution e-commerce hébergée dans le cloud revient toujours moins chère qu'une solution classique d'hébergement »

 **Faux!**

À ressources utilisées identiques entre le cloud et un hébergement standard, le cloud est un peu plus cher. Le cloud a également des coûts qui lui sont propres que l'on ne retrouve pas on-premises (sur site en local). Pour connaître la rentabilité du cloud vs autre solution, de nombreux facteurs sont à prendre en compte comme les besoins de scaling-out et le scaling up à moyen terme, les besoins d'administrations et d'audit de la solution concernée et de SLA.

Dans tous les cas, faites étudier votre besoin réel par un professionnel pour rationaliser vos coûts d'exploitation.

## Les principaux avantages incontournables du cloud



- ★ Les outils intégrés d'audit de performance, de disponibilité et d'administration au niveau infrastructure et applicatif
- ★ Le modèle universel dans lequel on paye uniquement pour les ressources consommées
- ★ Ressources illimitées ou presque, disponibles à la demande en quelques clics
  
- ★ Centralisation des ressources et des datas avec un contrôle et une maîtrise incontestés.
- ★ Une SLA allant de 99,99% à 99,999%
- ★ Mettre à disposition l'information au plus proche des internautes partout dans le monde en quelques clics.
- ★ Outils ETL (Extraction Transform & Load de données) intégrés
- ★ Environnement idéal pour stocker et gérer des data lakes, entrepôts de données et de la big data.



## Pourquoi faire appel à TSE ?

Parce que TSE possède une équipe d'experts en data et BI autour de Microsoft SQL Server que cela soit pour du conseil ou en technique.

Créée en 2009, TSE est une société indépendante de **services sur mesure** en data, business intelligence et applications de gestion web pilotées au niveau data, qui met l'innovation technique au service de ses clients

### TSE-T'Software Engineering

2a rue Jeanne d'Arc, 68720 Zillisheim

+ 33 6 59 78 66 52

+ 33 3 89 37 37 06

admin@tse-online.fr

