


# *L'écueil caché qui fait échouer les projets de Big Data.*

*Comment éviter la plus grosse erreur que vous puissiez  
commettre en matière de Big Data.*



*Plus elles sont  
grandes, plus  
dure est  
la chute.*



*De plus en plus de grandes entreprises prennent conscience des possibilités extraordinaires que représentent les Big Data.*

# La possibilité de :



Comprendre les clients, ainsi que les relations qu'ils entretiennent avec l'entreprise.



Leur proposer de meilleures recommandations sur les produits.



Savoir comment ils interagissent sur différents canaux.



Fournir aux gestionnaires des risques des tableaux de bord d'analyse en temps réel.



Optimiser la gestion des stocks.



Gérer plus efficacement les relations avec les fournisseurs.

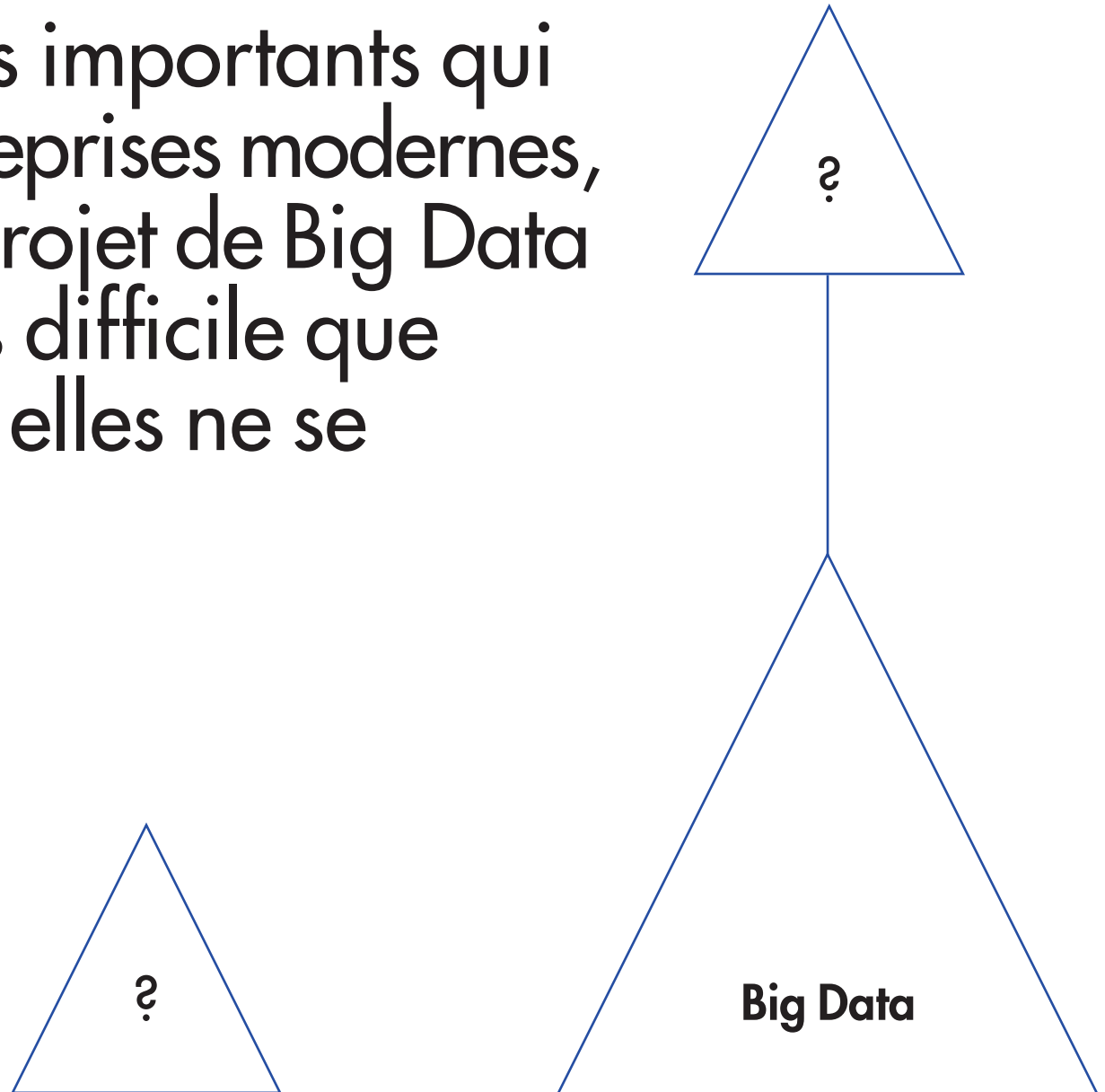


Distribuer les produits en fonction de la demande en temps réel.



La liste des avantages tangibles offerts par les Big Data est pratiquement illimitée.

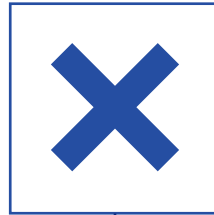
Pourtant, même si les Big Data recèlent la solution à certains des problèmes les plus importants qui se posent aux entreprises modernes, mener à bien un projet de Big Data est beaucoup plus difficile que la plupart d'entre elles ne se l'imaginent.



Plus de la

*moitié*

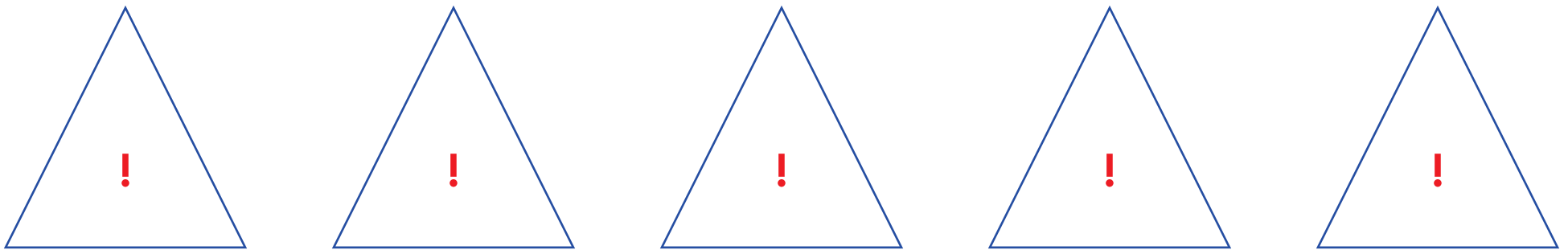
des projets de Big Data  
n'aboutissent pas.<sup>1</sup>



Les raisons sont nombreuses :  
dépassements de budget, délais  
excessifs, objectifs trop vagues,  
pénurie de spécialistes  
en données, etc.<sup>2</sup>



Il arrive souvent, cependant,  
que sous toutes ces raisons  
se cache la menace la plus grave  
qui pèse sur les projets de Big Data :



*Les données de  
mauvaise qualité.*



# Des données en double, incohérentes, incomplètes, bourrées d'erreurs, peu fiables.



En double car provenant  
de plusieurs systèmes sources.



Incohérentes car stockées  
dans différents formats.



Incomplètes car saisies  
sans être vérifiées.

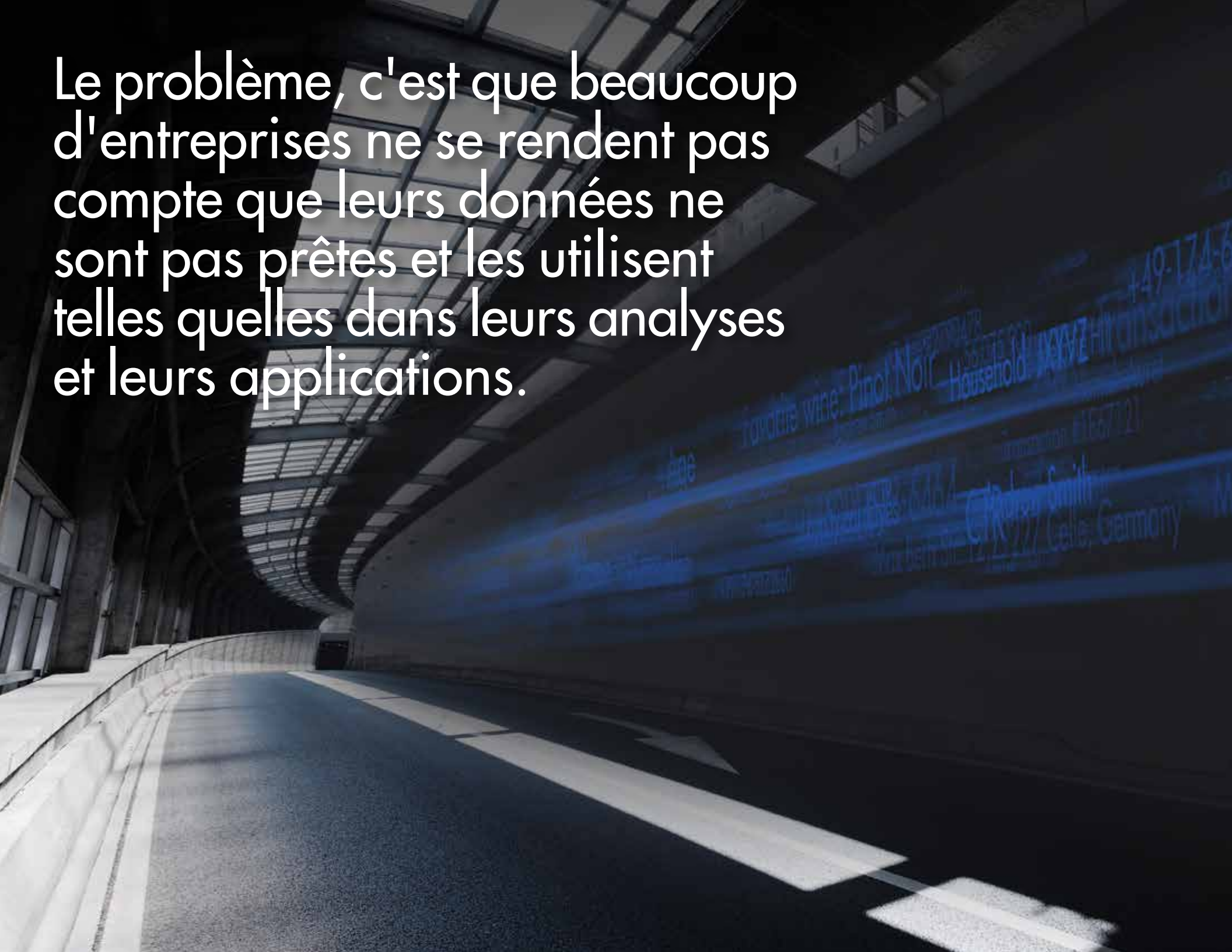


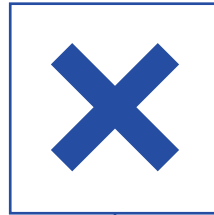
Bourrées d'erreurs  
car jamais nettoyées.



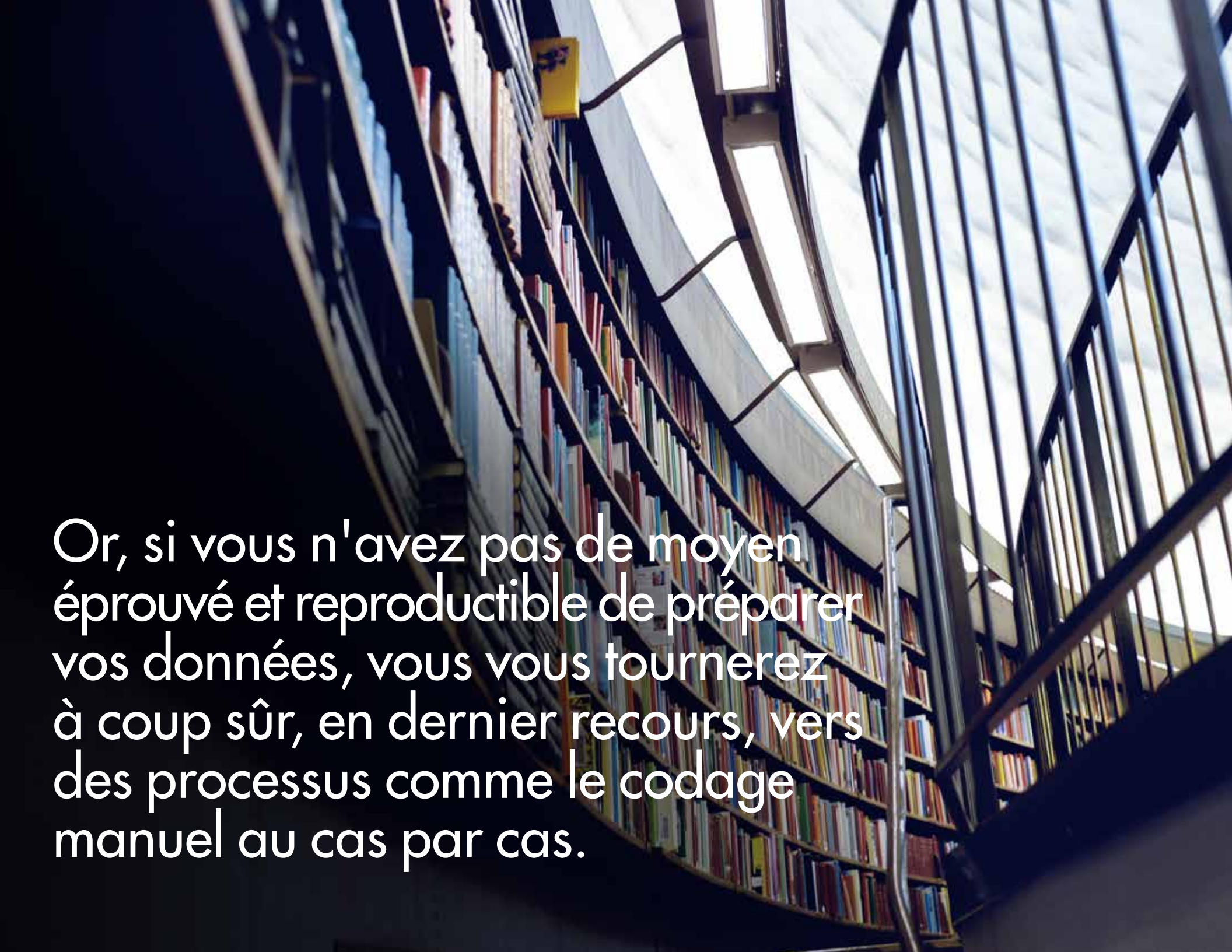
Peu fiables car vous travaillez  
sur des versions contradictoires  
de la vérité qui ne sont  
probablement pas à jour.

Le problème, c'est que beaucoup d'entreprises ne se rendent pas compte que leurs données ne sont pas prêtes et les utilisent telles quelles dans leurs analyses et leurs applications.





Elles ont beau disposer de technologies d'analyse hypersophistiquées, elles courent droit à l'échec si les données sous-jacentes ne sont pas prêtes.



Or, si vous n'avez pas de moyen éprouvé et reproductible de préparer vos données, vous vous tournerez à coup sûr, en dernier recours, vers des processus comme le codage manuel au cas par cas.

Si cette stratégie peut fonctionner à petite échelle, elle devient ingérable lorsqu'il s'agit des Big Data, étant donné leur complexité et le volume de travail nécessaire pour en venir à bout.

Les résultats ne sont jamais très beaux à voir.



# 1

Vos spécialistes en données, tout brillants, onéreux et difficiles à trouver qu'ils soient, consacrent le plus clair de leur temps aux tâches ingrates de préparation des données...<sup>3</sup>

et une infime partie de leurs journées à remplir la mission pour laquelle ils ont été embauchés : tirer des informations exploitables des données.



# 2

Vos dirigeants doivent prendre des décisions fondées sur les données peu fiables qui leur sont transmises. Mais, s'ils ne peuvent pas se fier aux données, ils ne devraient pas non plus se fier aux analyses ni, à plus forte raison, aux « informations exploitables » qui en sont l'aboutissement.

Au final, la prise de décision axée sur les données dans laquelle ils devaient investir prend des allures de fiasco.

# 3

Les tâches manuelles de préparation des Big Data ne se contentent pas de ralentir les projets et de coûter trop cher : elles sont aussi ingérables sur le long terme. Incapables de s'adapter aux évolutions, elles génèrent de nombreuses erreurs et sont si chronophages qu'il est très difficile de les reproduire.

Voilà que votre programme de Big Data tient de plus en plus de la grosse bourde : trop cher, trop long, trop approximatif, trop lourd pour l'entreprise.

*Nous avons  
toutefois une bonne  
nouvelle pour vous :*



Il existe un moyen plus efficace  
de gérer les Big Data.

C'est un moyen qui se résume  
en trois lettres : **M, D, M**  
(ou **gestion des données  
de référence en français**).

MDM

Si vous savez de quoi il s'agit, vous voudrez  
peut-être savoir quel rôle joue le MDM dans  
la prise de décision fondée sur les Big Data.

**Lire le livre  
blanc.**

Le MDM est une discipline qui apporte à vos projets de Big Data plusieurs avantages importants :



Elle vous procure un panorama complet des données les plus importantes concernant vos clients, vos produits, vos fournisseurs et vos sites.

Le MDM est une discipline qui apporte à vos projets de Big Data plusieurs avantages importants :



Elle permet de gérer toutes les données au moyen de profils de référence exhaustifs qui résolvent automatiquement les doublons, de sorte que, si vous avez un enregistrement au nom de « J. Robert » et un autre au nom de « Julie Robert », ils seront fusionnés pour produire une seule version de la vérité.

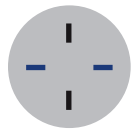
Le MDM est une discipline qui apporte à vos projets de Big Data plusieurs avantages importants :



Elle garantit l'exactitude et l'actualité de toutes vos données ; elle va jusqu'à réinjecter les données nettoyées dans vos systèmes d'analyse et vos applications, d'où une amélioration des processus et de la prise de décision dans l'ensemble de l'entreprise.



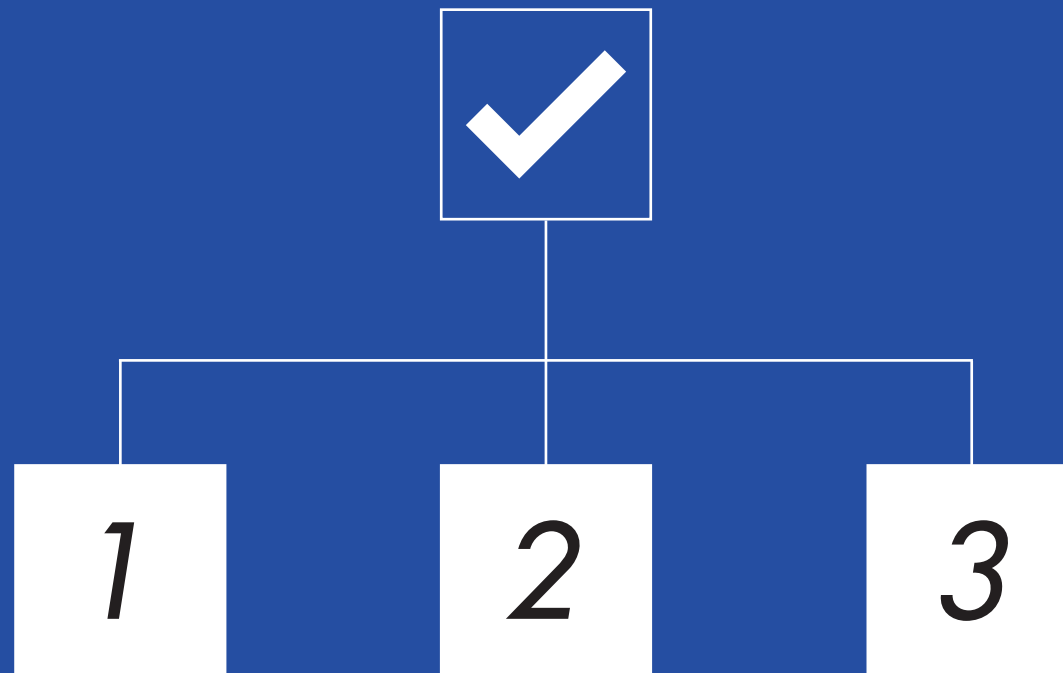
Le MDM est une discipline qui apporte à vos projets de Big Data plusieurs avantages importants :



C'est parce qu'elle gère ces profils de référence qu'elle peut vous montrer les liens qui existent entre eux. Vous pouvez donc visualiser toutes les interactions entre Julie et, par exemple, l'un de vos produits. Ou découvrir que Julie est mariée à un autre de vos clients, Georges.

*C'est tout bonnement  
le meilleur moyen  
de préparer vos  
Big Data à la prise  
de décision.*

C'est une bonne chose,  
et ce pour 3 raisons :



1

Vos brillants experts en données (qui sont aussi difficiles à trouver et qui vous coûtent cher) peuvent se concentrer sur la mission que vous leur avez confiée, à savoir la production d'informations exploitables, au lieu de perdre leur temps à nettoyer les données.

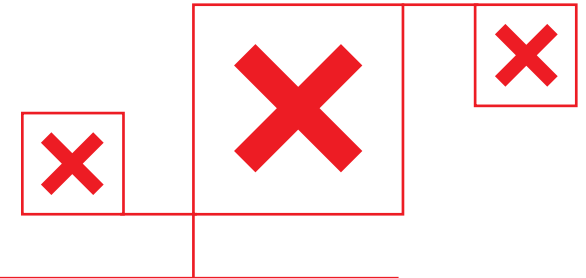
2

Vos dirigeants peuvent prendre des décisions fondées sur des données fiables, au lieu de prier pour que les graphiques qu'on leur a transmis soient exacts.

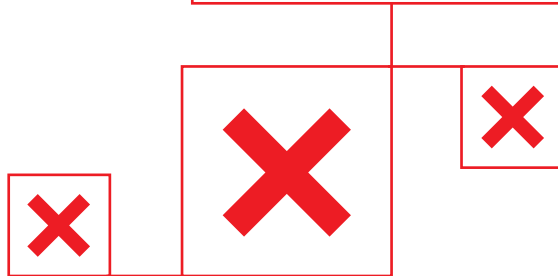
3

Votre projet de Big Data bénéficie de la fiabilité et de l'évolutivité nécessaires pour produire les informations dont votre entreprise a besoin, au lieu de vous coûter une fortune et un temps précieux à chaque fois que vous devez le mettre à l'échelle.

Bref, le MDM fait disparaître  
le plus gros écueil qui menace  
vos projets de Big Data :



*Les données de  
mauvaise qualité.*





Il vous procure également une vue exacte de toutes les données et toutes les interconnexions importantes.



*Voilà comment on  
maîtrise les Big Data  
et comment on réussit.*

Voilà comment  
on devient

*orienté décision.*

# À lire pour en savoir plus.

Pour découvrir les avantages du MDM pour la gestion des Big Data, lisez :

Master Data Management in a Big Data World.

Making your enterprise decision ready. (La gestion de données de référence à l'ère des Big Data.

Préparer votre entreprise à la prise de décision.)



**Lire**  
**maintenant.**

# À propos *d'Informatica.*

Informatica aide les entreprises de toute taille à gérer leurs Big Data pour se préparer à la prise de décision. Notre solution de MDM procure à nos clients un panorama complet de leurs clients, leurs produits, leurs fournisseurs et leurs sites ; elle rend aussi les utilisateurs métiers plus autonomes.

**Discutons-en.**

# Sources

1. <http://www.analytics-magazine.org/july-august-2014/1074-the-data-economy-why-do-so-many-analytics-projects-fail>
2. <http://blogs.gartner.com/svetlana-sicular?s=0lbid>
3. [http://www.nytimes.com/2014/08/18/technology/for-big-data-scientists-hurdle-to-insights-is-janitor-work.html?\\_r=0lbid](http://www.nytimes.com/2014/08/18/technology/for-big-data-scientists-hurdle-to-insights-is-janitor-work.html?_r=0lbid)