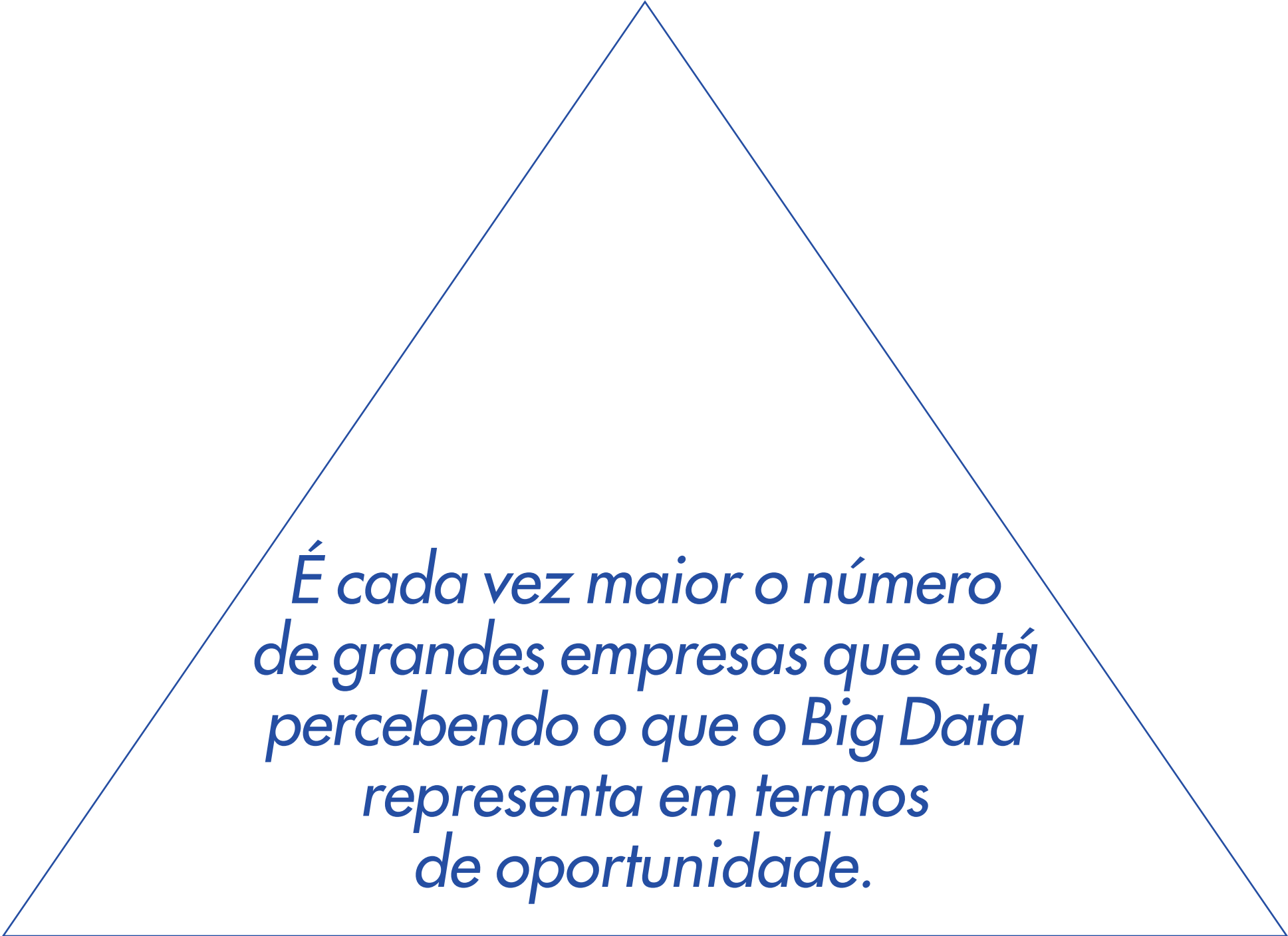


# O vilão silencioso de projetos de Big Data.

*Como evitar o maior erro de Big Data e que você não pode se dar ao luxo de cometer.*










*Quanto  
mais alto,  
maior o  
tombo.*

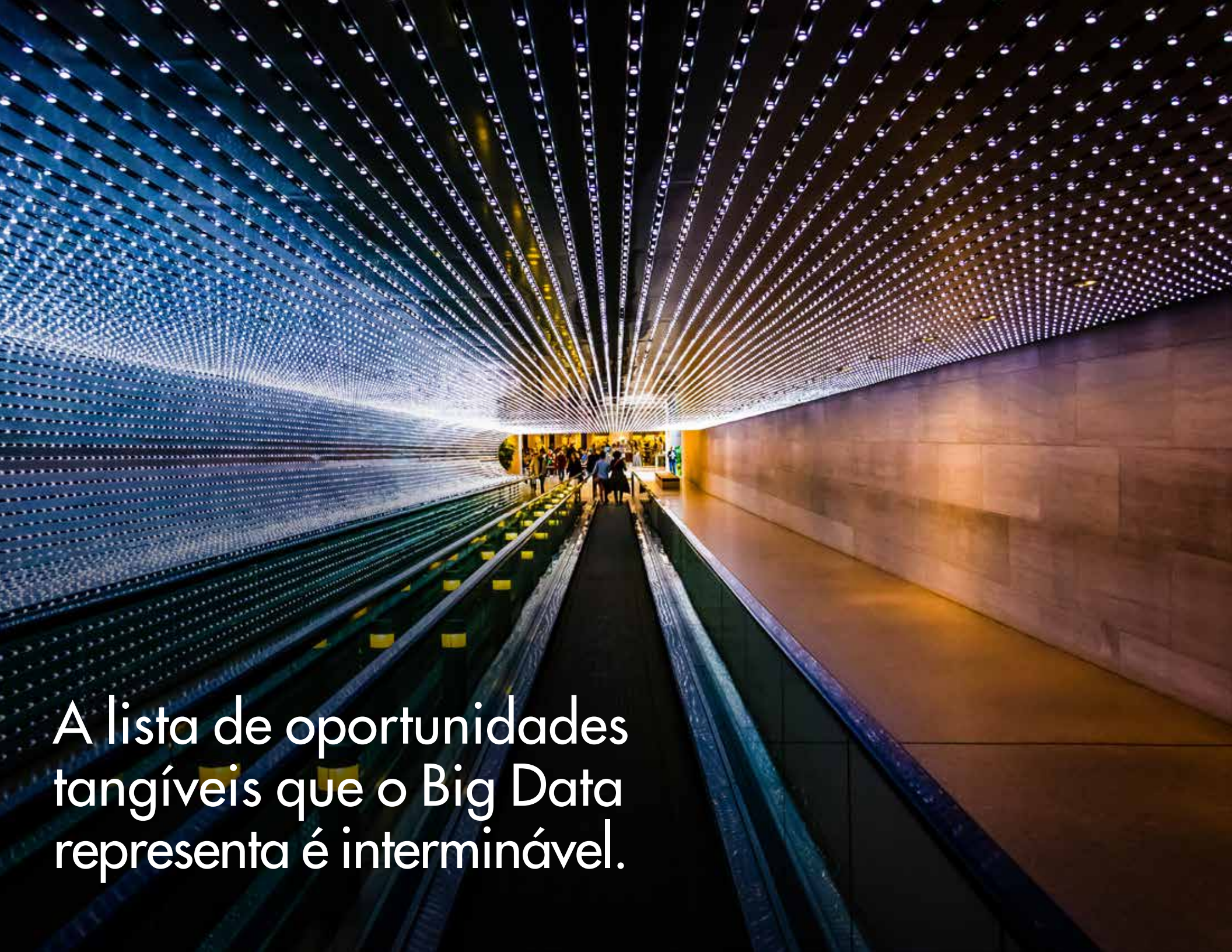


*É cada vez maior o número  
de grandes empresas que está  
percebendo o que o Big Data  
representa em termos  
de oportunidade.*

# Oportunidade de:

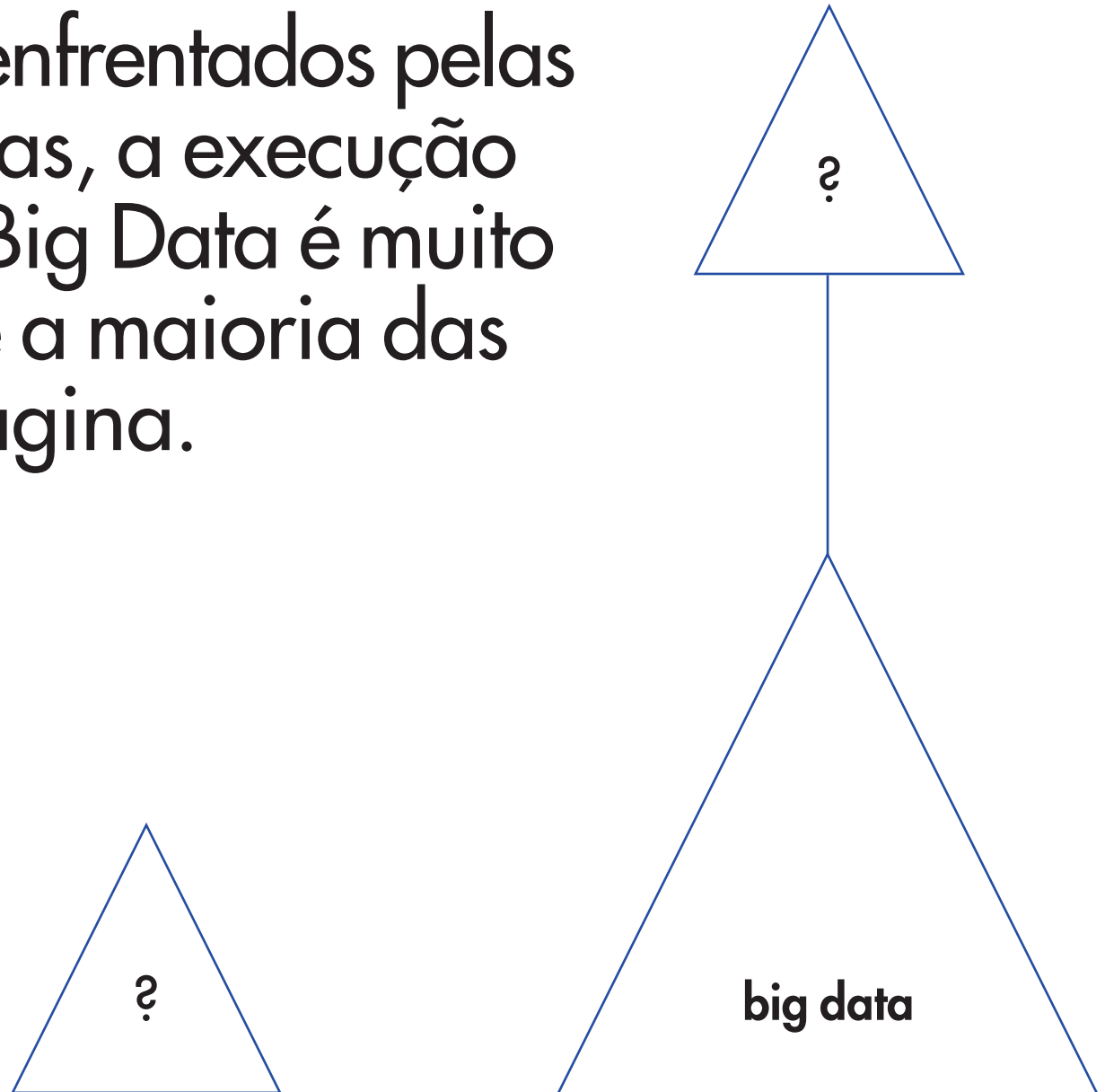
-  Compreender clientes e suas relações com a empresa.
-  Oferecer a eles as melhores recomendações de produtos.
-  Aprender como eles interagem em vários canais.
-  Fornecer aos gestores de riscos painéis de análises em tempo real.
-  Aperfeiçoar o gerenciamento de inventário.
-  Gerir com mais eficiência as relações com os fornecedores.
-  Distribuir produtos com base na demanda em tempo real.





A lista de oportunidades tangíveis que o Big Data representa é interminável.

Embora o Big Data tenha as respostas para alguns dos desafios mais importantes enfrentados pelas empresas modernas, a execução de um projeto de Big Data é muito mais difícil do que a maioria das organizações imagina.

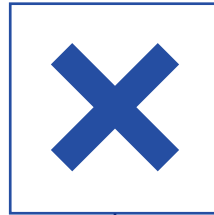


Na verdade,  
mais da

*metade*

de todos os projetos  
de Big Data não são  
concluídos.<sup>1</sup>

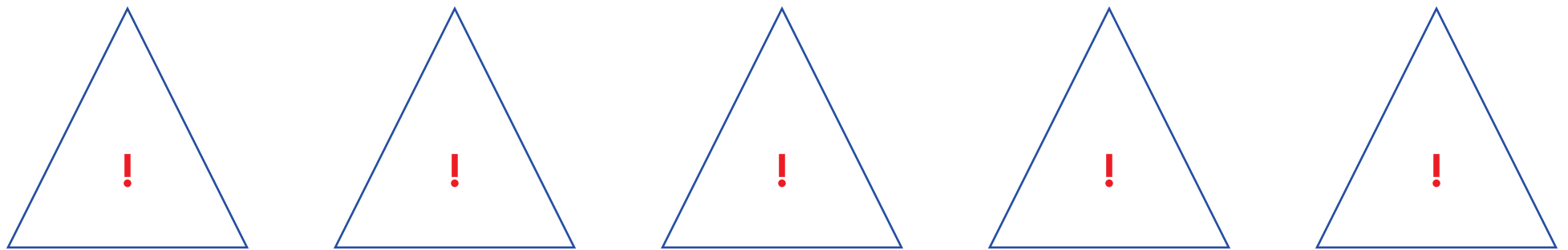




Isso acontece por uma série de motivos, entre eles, aumento nos custos, atrasos excessivos, objetivos vagos e falta de talento para lidar com dados.<sup>2</sup>



Porém, com bastante frequência, a principal ameaça aos projetos de Big Data passam despercebidos:





# Dados duplicados, inconsistentes, incompletos, cheio de erros, não confiáveis.



Duplicados porque são oriundos de vários sistemas.



Inconsistentes porque são armazenados em formatos diferentes.



Incompletos por conta de entradas de dados não verificadas.

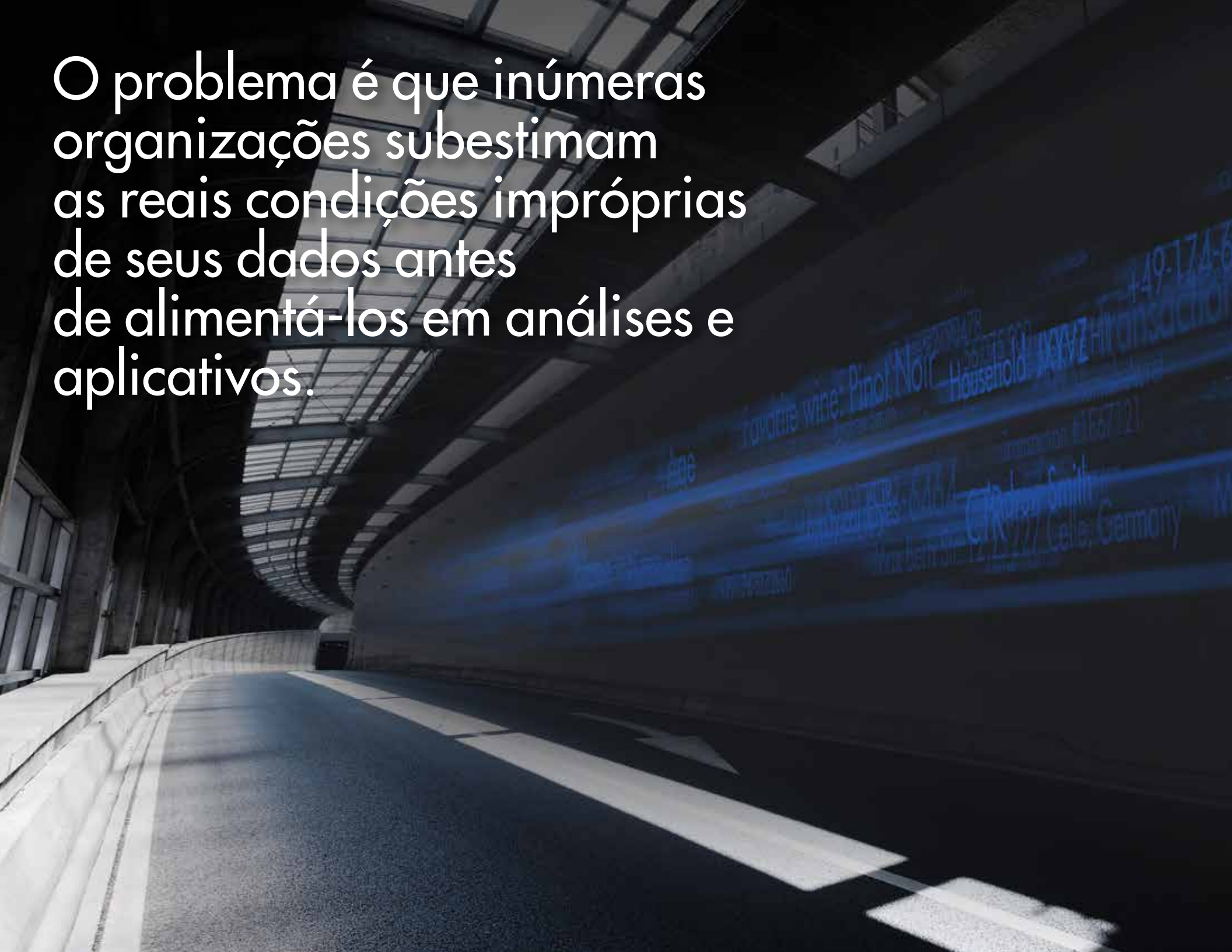


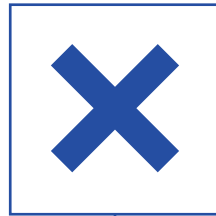
Cheios de erros porque nunca foram limpos.



E não confiáveis por você trabalhar com versões conflitantes da verdade que podem ou não estar atualizadas.

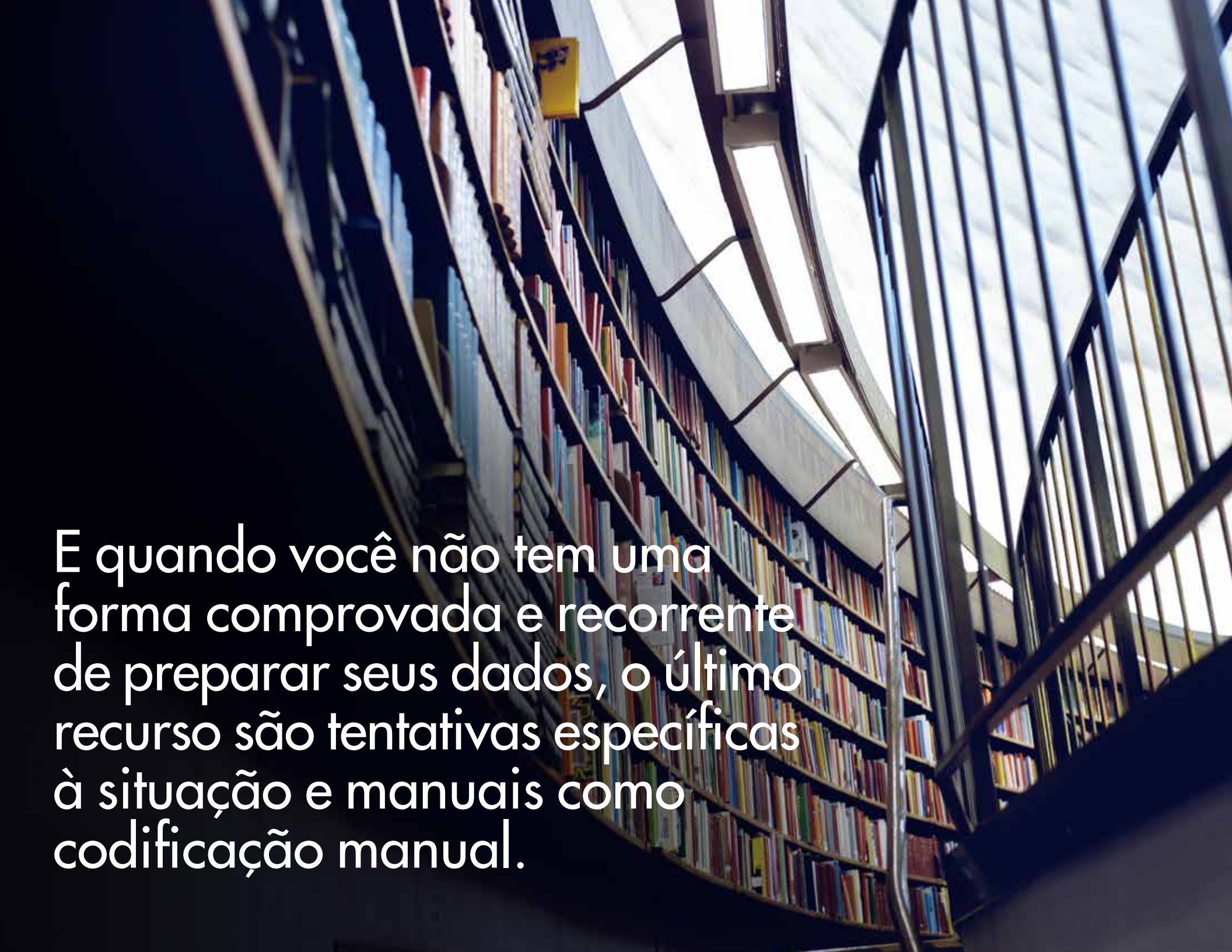
O problema é que inúmeras organizações subestimam as reais condições impróprias de seus dados antes de alimentá-los em análises e aplicativos.





Portanto, independentemente do nível de sofisticação da tecnologia de análise, a falha é inevitável quando os dados envolvidos não estão prontos.

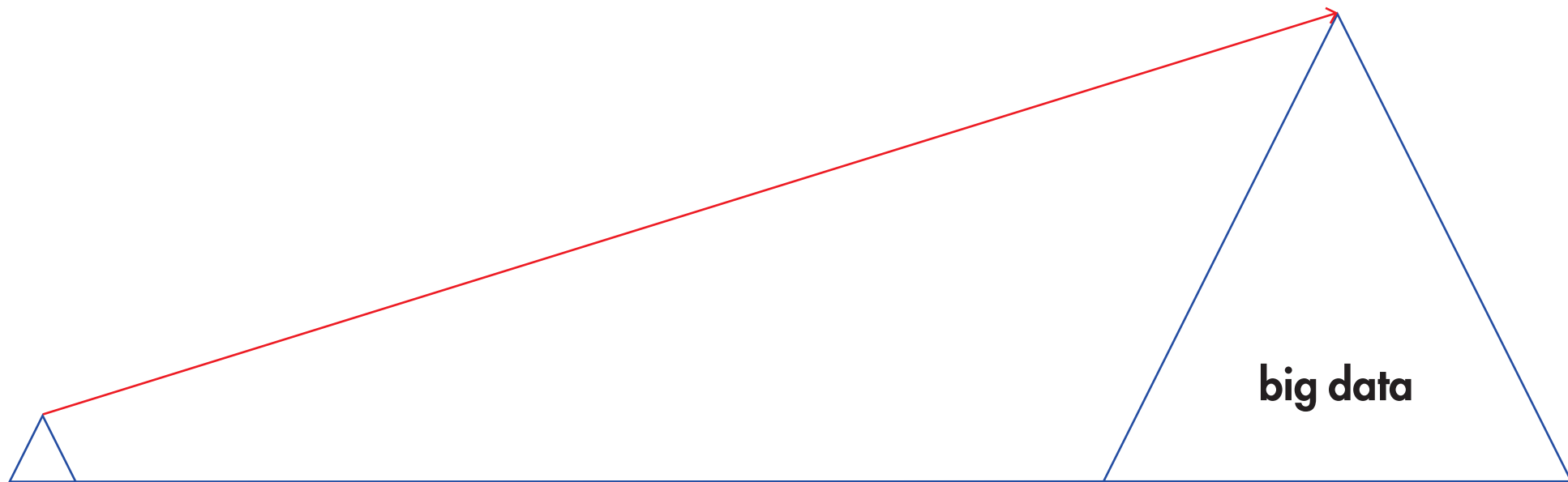


A low-angle, perspective view of a modern library. The image shows curved bookshelves filled with books, extending into the distance. A glass skylight is visible at the top, allowing natural light to illuminate the space. A metal railing is visible on the right side, suggesting a mezzanine or upper level. The overall atmosphere is bright and organized.

E quando você não tem uma forma comprovada e recorrente de preparar seus dados, o último recurso são tentativas específicas à situação e manuais como codificação manual.

E enquanto essa estratégia pode até funcionar em uma escala menor, o trabalho se torna precário quando aplicado à escala e complexidade do Big Data.

Os resultados nunca são satisfatórios.



# 1

Seus cientistas de dados brilhantes, caros e difíceis de encontrar passam a maior parte do seu tempo fazendo o trabalho sujo de limpeza e preparação dos dados (data wrangling).<sup>3</sup>

E apenas uma parcela muito pequena do seu tempo se dedicando à função para que foram contratados: produzir insights relevantes.

# 2

Seus executivos têm que tomar decisões com base em dados não confiáveis que eles recebem. Mas se eles não podem confiar nos dados, não deveriam confiar na análise e certamente não podem se dar ao luxo de confiar nos "insights".

Então a decisão impulsionada por dados que você imaginou que representava um investimento, na verdade, culminou nada mais do que uma decepção.

# 3

Tentativas manuais para preparar o Big Data não apenas atrasam os projetos e custam muito, como são notadamente insustentáveis. Eles não são dimensionados, estão cheios de erros e os processos são tão longos que fica difícil repeti-los.

De repente seu programa de Big Data começa a parecer apenas um produto de custo alto e lento, além de um fardo não confiável para sua empresa.



*Mas há boas  
notícias:*



Existe uma maneira melhor  
de lidar com o Big Data.

Ela começa com o  
Gerenciamento de **D**ados **M**estre.  
(**MDM** abreviado.)

MDM

Se você já conhece, talvez queira saber  
mais sobre como ele está incrementando  
a tomada de decisões de Big Data.

**Leia o**  
**e-Book.**

MDM é uma disciplina que se infiltra em seu projeto de Big Data com algumas qualidades fundamentais:



Ele dá a você uma visão 360° dos seus dados mais importantes sobre clientes e produtos, fornecedores e locais.

MDM é uma disciplina que se infiltra em seu projeto de Big Data com algumas qualidades fundamentais:



Ele gerencia todos os seus dados em perfis ricos e "mestres" que reconciliam dados duplicados automaticamente. Ou seja, ele sabe que precisa mesclar seus registros de "J. Robinson" e "Jodie Robinson" em uma única versão da verdade.

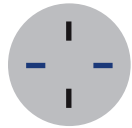


MDM é uma disciplina que se infiltra em seu projeto de Big Data com algumas qualidades fundamentais:



Ele mantém automaticamente todos os dados limpos, preciso e atualizados, inclusive transferindo todas as coisas boas para suas análises e aplicativos, melhorando seus processos e tomada de decisões dentro de sua empresa.

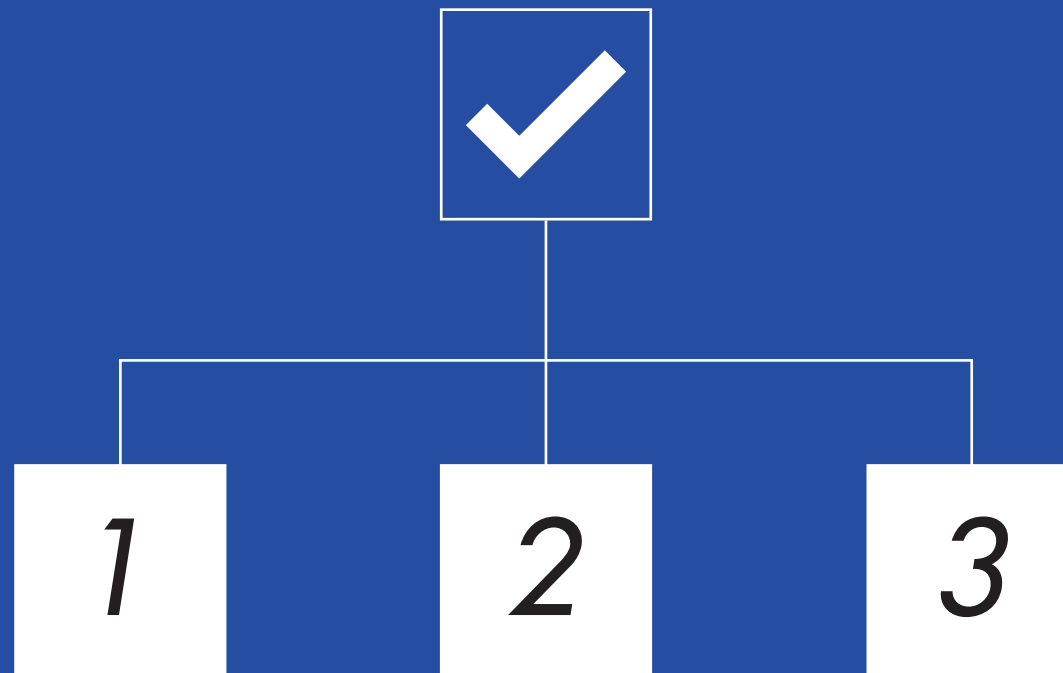
MDM é uma disciplina que se infiltra em seu projeto de Big Data com algumas qualidades fundamentais:



E porque ele gerencia esses perfis mestres, ele também mostra a relação entre eles. Dessa forma você poderia obter uma visão completa de todas as suas interações entre Jodie e, por exemplo, um de seus produtos. Ou perceber que Jodie é casada com um de seus clientes, George.

*É simplesmente  
a maneira mais  
eficiente de tornar  
o Big Data em  
dados excelentes.*

Isto é fantástico  
por três motivos:



1

Seus cientistas de dados brilhantes, caros e difíceis de encontrar podem se concentrar em produzir os insights relevantes (em vez de perder seu tempo com tarefas de limpeza e higiene de dados).



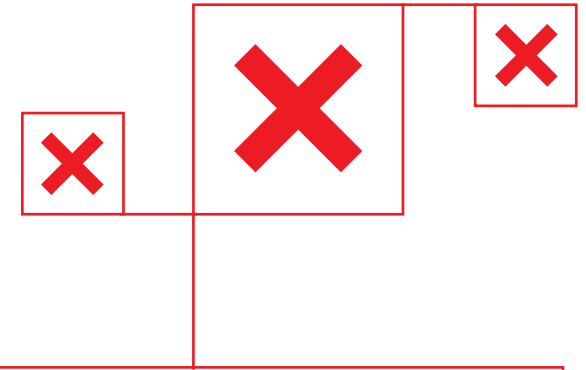
2

Seus executivos podem tomar decisões com base em dados confiáveis (em vez de cruzar os dedos e torcer para aqueles gráficos estejam certos).

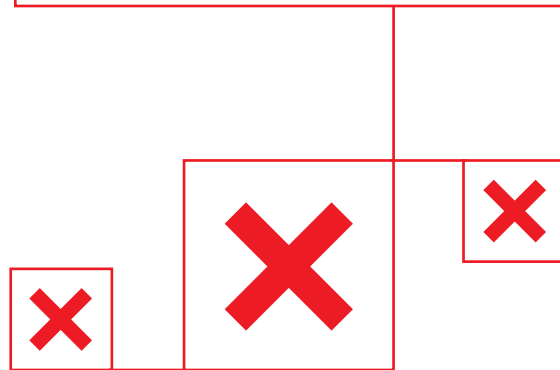
3

E seus projetos de Big Data incorporam a confiabilidade e a dimensionalidade necessárias para produzir os insights que sua organização precisa (em vez de gastar uma fortuna e passar meses sempre que precisar aumentá-lo).

Ou seja, ele livra o seu projeto de Big Data do principal vilão dos seus projetos de Big Data:



*Dados corrompidos.*







E proporciona uma visão  
limpa e conectada de tudo  
o que realmente importa.



*É assim que  
você domina  
o Big Data para  
VENCER.*

É assim que você  
se transforma em

*decision ready.*

# *Leitura adicional.*

Para saber como o MDM atua em prol do Big Data, leia:  
Gerenciamento de dados mestres em um mundo de Big Data.  
Transforme sua empresa em decision ready.



**Leia agora.**

# *Sobre a Informatica.*

Somos a Informatica e estamos ajudando empresas de todos os portes a lidarem com Big Data para transformarem em decision ready. A nossa solução MDM proporciona visões 360 de seus clientes, produtos, fornecedores e locais, além de capacitarem seus usuários de negócios.



**Vamos  
conversar.**



# Fontes.

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)