



## **BIG DATA ANALYTICS: apresentação do estudo de caso da webmotors**

### ***BIG DATA ANALYTICS: presentation of the webmotors case study***

Jander César Fernandes Júnior - jandercesarfjunior@gmail.com

Giuliano Scombatti Pinto - giuliano.pinto@fatectq.edu.br

Faculdade de Tecnologia de Taquaritinga (FATEC) – São Paulo – Brasil

### **RESUMO**

Este artigo contém a apresentação de um estudo de caso sobre a Webmotors com o objetivo de mostrar na prática a eficiência que o *big data analytics* traz para uma empresa juntamente com uma maior intimidade em relação às preferências do cliente. Exibe preliminarmente uma exploração bibliográfica, com a ideia de variados autores sobre as metodologias apresentadas. Também apresenta desafios, benefícios e os próximos passos a serem tomados do estudo de caso exibido.

**Palavras-chave:** *Big data analytics*. Eficiência. Metodologias.

### **ABSTRACT**

This article contains a presentation of a case study on Webmotors to demonstrate the efficiency of analyzing large data for a company with great intimacy regarding customer preferences. It shows preliminarily a bibliographic forest, with an idea of several authors about the methodologies. Thus, the presentation of the challenges, the benefits and the correct steps are taken from the study of the case presented.

**Keywords:** Big data analytics. Efficiency. Methodology

## **1 INTRODUÇÃO**

A quantidade de dados gerados pela humanidade nos últimos anos aumentou de forma exponencial (Galdino, 2016). Segundo uma pesquisa recente (IBM, 2013), no ano 2000, 25% (vinte e cinco por cento) dos dados eram digitalizados. No ano de 2007, esse número saltou para 93% (noventa e três por cento), e no ano de 2013, foi para 98% (noventa e oito por cento). Segundo Galdino (2016), esse crescimento, devido principalmente a fatores como aumento do acesso a dispositivos eletrônicos e a popularização da internet, está gerando uma revolução no tratamento de dados.



Galdino (2016) diz ainda que esse valor, obtido através de correlações entre dados, pode se dar através de descoberta de padrões, preferências de usuários, aumento no número de vendas em determinada época do ano, descoberta de cura de doenças, entre diversos outros benefícios aplicáveis a diversas áreas de estudo.

Há alguns anos, um número cada vez maior de organizações, de diversos portes e segmentos, vem empregando a filosofia de *Big Data Analytics* como ferramenta de apoio estratégico (HEKIMA, 2017).

Todos esses indicativos representam para as corporações a possibilidade de tomar decisões mais precisas e, sobretudo, antecipadas à concorrência. Decisões que, em um cenário de extrema competitividade, podem ser o divisor entre o sucesso e a estagnação do seu negócio (HEKIMA, 2017).

O objetivo deste trabalho é expor uma pesquisa bibliográfica exploratória, mostrando os benefícios trazidos pela análise do *Big Data* e as possíveis melhorias nos processos de trabalho das organizações, fornecendo *insights* valiosos sobre as tendências de mercado, comportamento de seus clientes e suas expectativas, apresentando, posteriormente o estudo de caso da Webmotors.

De acordo com Cervo e Bervian (1996), o método pode ser entendido nas ciências como o conjunto de processos que o espírito humano deve empregar na investigação e demonstração da verdade. Segundo Gil (2009), uma pesquisa é um procedimento racional e sistemático que tem como objetivo proporcionar respostas aos problemas que são propostos.

Como métodos de procedimento, esta pesquisa classifica-se como pesquisa bibliográfica, pois procura explicar um problema a partir de referências teóricas publicadas em documentos (CERVO; BERVIAN, 1996).

Esse artigo está organizado em três seções:

- a) Referencial teórico: apresenta a revisão bibliográfica das metodologias utilizadas.
- b) Apresentação do estudo de caso da webmotors.
- c) Considerações finais: conclusão.

Para tal, foram usadas diversas fontes de pesquisa, como livros, revistas, vídeos explicativos, sites, artigos publicados.

## **2 FUNDAMENTAÇÃO TEÓRICA**

### **2.1 O valor da informação**

Um gerente eficaz em qualquer área de negócios deve entender que as informações são um dos recursos organizacionais mais valiosos e importantes. Ainda assim, o termo é constantemente confundido com o termo “dado” (STAIR E REYNOLDS, 2009).

Stair e Reynolds (2009) dizem ainda que dados são compostos por fatos básicos, como o nome e a quantidade de horas trabalhadas em uma semana de um funcionário, número de peças em estoque ou pedidos. Quando esses fatos são organizados ou arrançados de maneira



significativa, eles se transformam em informações. Informação é um conjunto de fatos organizados de modo a terem valor adicional, além do valor dos fatos propriamente ditos.

Mas para ter esse valor adicional, é de fundamental importância para uma organização a qualidade da informação adquirida e/ou processada, pois esse aspecto reflete na efetividade do processo de tomar decisões em uma empresa (CAIÇARA JÚNIOR, 2011). Para Rezende e Abreu (2000), as informações de qualidade devem ser comparativas, confiáveis, geradas em tempo hábil e no nível de detalhe adequado.

Um conceito que gera muitas controvérsias entre os estudiosos refere-se ao valor de uma informação (CAIÇARA JÚNIOR, 2011). Para Rezende e Abreu (2000), a informação tem valor altamente significativo e pode representar grande poder para quem a possui, indivíduos ou instituição. Ela está presente em todas as atividades que envolvem pessoas, processos, sistemas, recursos e tecnologias.

Para Stair e Reynolds (2009), o valor da informação está diretamente ligado a como ela auxilia os tomadores de decisões a atingir seus objetivos organizacionais. O valor da informação poderia ser medido pelo tempo necessário para tomar uma decisão ou pelo aumento dos lucros de uma companhia.

## 2.2 *Big Data Analytics*

O *Big Data* representa uma tendência em tecnologia que está abrindo caminho para um novo método de compreensão do mundo e do processo decisório de negócios. Essas decisões são tomadas com base em quantidades muito grandes de dados estruturados, não estruturados e complexos, cujo processamento usando banco de dados básico e ferramentas de gestão de *warehouse* é muito difícil (ISACA, 2013). *Big data* refere-se, principalmente, aos conjuntos de dados que são muito grandes ou com rápidas mudanças para serem analisados com técnicas de banco de dados relacionais tradicionais ou multidimensionais ou ferramentas de software comumente usadas para capturar, gerenciar e processar os dados em um tempo razoável (ISACA, 2013).

De acordo com Siewert (2013), *Big Data* é definido genericamente como a captura, gerenciamento e a análise de dados que vão além dos dados tipicamente estruturados, que podem ser consultados e pesquisados através de bancos de dados relacionais. Frequentemente são dados obtidos de arquivos não estruturados como vídeo digital, imagens, dados de sensores, arquivos de logs e de qualquer tipo de dados não contidos em registros típicos com campos que podem ser pesquisados.

O objetivo do *Big Data* é propiciar dados e informações que possam ser analisados visando subsidiar tomadas de decisão (FERNANDES E ABREU, 2009). Esses autores ainda relatam que a tomada de decisão é possível em função não somente do volume de dados, da velocidade de captura dessas informações, das fontes variadas de informações e de novos softwares para fins de modelagem dessas informações. É possível, por exemplo, ver uma tendência de crescimento da venda de um produto em função de comentários favoráveis no Facebook. Este tipo de análise é o que está sendo denominado *data analytics*.

De acordo com a ISACA (2013), esses dados são coletados para serem analisados para descobrir padrões e correlações que podem não ser inicialmente evidentes, mas que podem ser úteis na tomada de decisões de negócio. Ainda diz que esses dados são muitas vezes dados



peçoais que são úteis a partir de uma perspectiva de marketing para entender as preferências e antipatias de potenciais compradores e para analisar e prever seu comportamento de compra.

Segundo a ISACA (2013), os dados pessoais podem ser classificados como:

- Dado voluntário - criado e explicitamente compartilhado por indivíduos (Exemplo: Perfis de redes sociais)
- Dado observado - capturado ao gravar as ações dos indivíduos (Exemplo: dados de localização no uso de *smartphones*)
- Dado inferido – dado sobre indivíduos com base em informação voluntariada ou observada (Exemplo: classificações de crédito)

O principal objetivo da análise do Big Data é contribuir para que as empresas tomem melhores decisões de negócio. As análises dos dados podem permitir uma abordagem de marketing que dá a empresa uma melhor compreensão de seus clientes, entendimento esse que influenciará os processos internos e aumentará o lucro, o que proporciona a vantagem competitiva que a maioria das empresas está buscando (ISACA, 2013).

De acordo com a HEKIMA (2017), *Big Data Analytics* é o trabalho analítico e inteligente de grandes volumes de dados, estruturados ou não-estruturados, que são coletados, armazenados e interpretados por *softwares* de altíssimo desempenho. Tratar-se do cruzamento de uma infinidade de dados do ambiente interno e externo, gerando o que chamam de “bússola gerencial” para tomadores de decisão.

Segundo a HEKIMA (2017), essas são algumas das fontes usadas por um *software* de *Big Data Analytics*:

- Dados extraídos de ferramentas de Inteligência de Negócios (*Business Intelligence* – BI);
- Arquivos de log de servidores *web*;
- Conteúdos de mídias sociais;
- Relatórios empresariais;
- Textos de e-mails de consumidores à empresa;
- Indicadores macroeconômicos;
- Pesquisas de satisfação;
- Estatísticas de ligações celulares capturadas por sensores conectados à “internet das coisas”;
- Bases de dados das empresas de cartão de crédito;
- Programas de fidelidade;
- *Reviews* de produtos nos sites das empresas.

### 2.2.1 Tipos de análise de Big Data

Quando considerado o volume de dados obtidos pelo *Big Data Analytics*, pode-se imaginar a infinidade de caminhos analíticos que podem ser traçados para tomadas de decisão.



No entanto, segundo HEKIMA (2017), há quatro tipos de análise de Big Data que se destacam pela usabilidade e potencialidade de seus resultados, que são:

- **Análise preditiva:** Pode-se defini-la como uma análise de possibilidades futuras. A partir da identificação de padrões passados em sua base dados, esse tipo de análise permite aos gestores o mapeamento de possíveis futuros em seus campos de atuação. A ideia é deixar de tomar decisões baseadas unicamente na intuição, conseguindo estabelecer um prognóstico mais sólido para cada ação. Para que isso seja possível, a análise preditiva utiliza a mineração de dados (*data mining* - De forma geral *Data Mining* pode ser definido como o método de exploração de grandes volumes de dados com objetivo de identificar padrões e informações preciosas para o negócio das empresas), dados estatísticos e dados históricos para conhecer as futuras tendências.
- **Análise prescritiva:** É uma forma de definir qual escolha será mais efetiva em determinada situação. Dentro de uma indústria ou setor, o valor dessa análise se dá pela capacidade de numerar determinados padrões e filtrá-los por especificidades, obtendo um cenário bastante fiel da situação e como cada intervenção responderá (HEKIMA, 2017).
- **Análise descritiva:** Compreensão em tempo real dos acontecimentos é o que define a análise descritiva. É a mineração de dados na base da cadeia de *Big Data*. É uma maneira de visualizar os dados, entender como uma *database* se organiza e o que significa para o presente sem necessariamente relacioná-la com padrões passados ou futuros.
- **Análise diagnóstica:** A análise diagnóstica tem como objetivo compreender de maneira causal todas as suas possibilidades. Como uma espécie de relatório expandido, quando feita em uma base de dados volumosa, esse tipo de análise permite ainda entender a razão de cada um dos desdobramentos das ações adotadas e, a partir disso, mudar estratégias ineficazes ou reforçar as funcionais (HEKIMA, 2017).

*Big Data Analytics* opera de várias formas diferentes. Quem os analisa é responsável por aplicá-los da melhor maneira possível. Cada tipo de análise tem seu próprio escopo e sua própria finalidade. Por isso é preciso compreendê-las bem (HEKIMA, 2017).

### 2.3 Vantagens e riscos trazidos pelo *Big Data Analytics*

Ao extrair e combinar resultados de todas as fontes já citadas, soluções de *Big Data Analytics* podem entregar informações extremamente importantes para a melhoria do desempenho de qualquer tipo de empresa no mercado. O ponto de partida é a viabilização de análises profundas de um público por região. Ao saber quais são os produtos mais procurados e desejados em um determinado local, é possível direcionar as vendas para aquele lugar e até mesmo melhorar o sistema de logística para tornar o atendimento mais eficiente (HEKIMA, 2017).

Ainda de acordo com a HEKIMA (2017), graças às previsões concedidas pelos estudos dos dados analisados do *Big Data*, uma empresa pode tomar caminhos diferentes e mais eficientes daqueles que seus concorrentes escolheram percorrer. Sendo assim, é possível

oferecer promoções específicas, aumentar o estoque de determinado produto mais procurado ou então diminuir a oferta de um produto que já não traz mais tanto lucro para a empresa.

As empresas que dominam a disciplina emergente do gerenciamento de *Big Data* podem obter recompensas significativas e se diferenciar de seus concorrentes. O uso adequado de *Big Data* vai além de coletar e analisar grandes quantidades de dados; também exige a compreensão de como e quando usar os dados ao tomar decisões cruciais (ISACA, 2013).

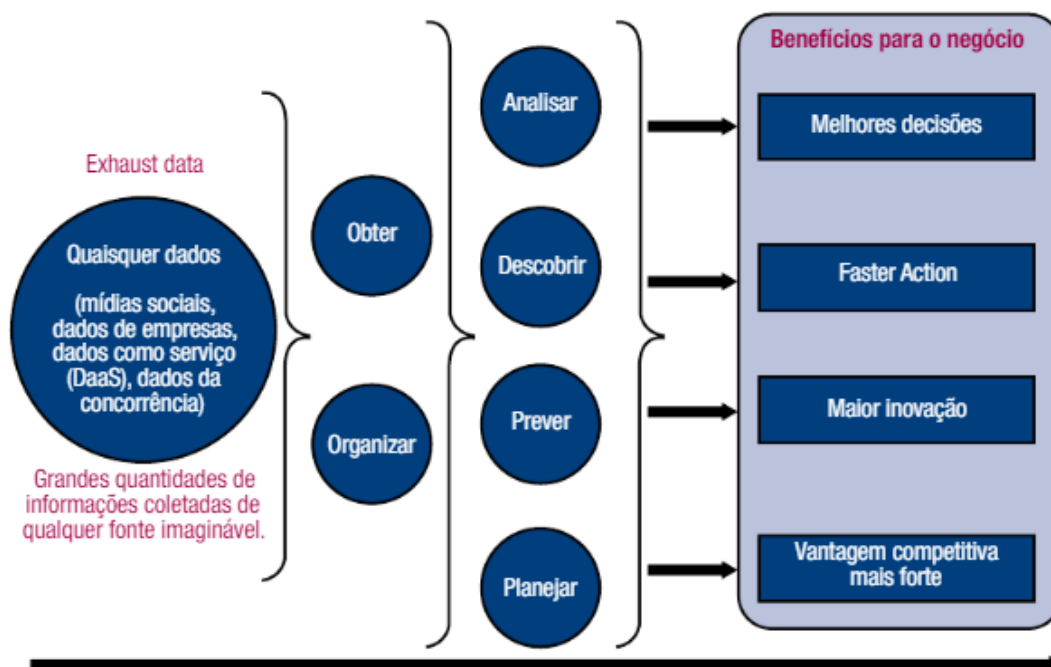
Os benefícios financeiros podem ser concretizados quando os processos de gerenciamento de dados estão alinhados com a estratégia da empresa, o que pode exigir o envolvimento da alta gestão para definir a direção e supervisão das decisões principais (ISACA, 2013).

Análise de *big data* pode impactar positivamente:

- Desenvolvimento de produto
- Desenvolvimento do mercado
- Eficiência operacional
- Experiência e lealdade do cliente
- Previsões de demanda de mercado

A Figura 1 apresentada a seguir, mostra o processo para acesso de *insights* comerciais específicos da empresa, o que permite ver graficamente como são obtidas as previsões fornecidas pelo tratamento dos dados armazenados no *Big Data*.

Figura 1 - Analisando os insights comerciais específicos da empresa



Fonte: ISACA, 2013.

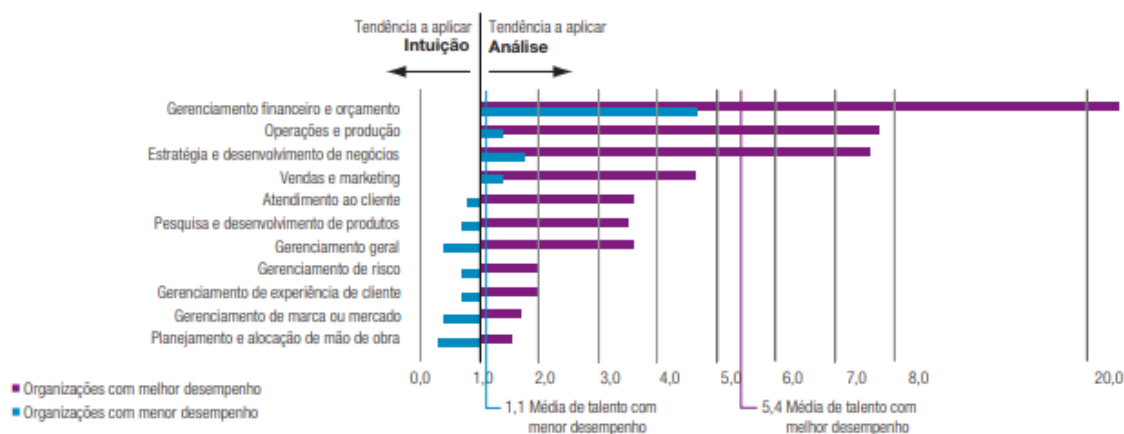
Uma análise conjunta do IBM (*Institute for Business Value*) e do *Massachusetts Institute of Technology* (MIT) identificou que empresas que investem em análise de dados



possuem uma visão melhor do seu negócio, conseguindo uma receita 33% maior do que seus concorrentes, crescimento do lucro doze vezes maior e retorno sobre o investimento de capital 32% maior (IBM, 2011).

A partir dessa pesquisa, conclui-se que o lucro é muito superior se utilizadas as informações geradas pela análise do *Big Data*, se comparado apenas à intuição dos gerentes de alguma organização, como é apresentado a seguir na Figura 2:

**Figura 2 Análise: A marca do alto desempenho**



Fonte: IBM, 2011.

De acordo com a ISACA (2013), apesar de o *Big Data* trazer vantagem competitiva a partir do capital investido, isto também acarreta riscos significativos. Afinal, agora a empresa tem enormes quantidades de dados para serem analisados.

Então surgem perguntas, como:

- Onde os dados serão armazenados?
- Como os dados serão protegidos?
- Como utilizar os dados de forma segura e legal?

Os riscos que devem ser gerenciados são:

- Perda de dados “tóxicos” armazenados como informações privadas ou de custódia, contas de clientes, números de cartão de crédito, fórmulas da empresa, etc.
- Utilização de informações obtidas por meio de redes sociais, visto que pode haver problemas sobre questões de privacidade ou falta de consenso jurídico.
- Segurança da Informação.
- Qualidade dos dados obtidos, visto que podem ser dados imprecisos, incompletos ou manipulados de forma fraudulenta.
- Capacidade da infraestrutura para suportar o *Big Data*.
- Qualidade e capacidade do responsável por capturar, armazenar ou analisar os dados.

As vulnerabilidades de dados são especialmente sérias para as empresas que se baseiam em dados pessoais que são gerados ou podem ser modificados pelo público. Por exemplo, os dados de mídia social podem ser uma fonte de grande valor para avaliar o sentimento do cliente, acompanhando a eficácia das campanhas de *marketing* e aprendendo



mais sobre os consumidores. No entanto, utilizar este tipo de dados pessoais exigirá abordar incertezas atuais e pontos de tensão (ISACA, 2013).

### 3 APRESENTAÇÃO DE ESTUDO DE CASO DA WEBMOTORS

O presente trabalho expõe a apresentação de um estudo de caso com dados que foram coletados de um trabalho já realizado pela plataforma de serviços de computação em nuvem AWS (*Amazon Web Services*), aplicado a Webmotors (site de veículos), no intuito de visualizar e analisar as informações obtidas através de seus usuários. Para isso, foi utilizada a ferramenta *Amazon Redshift* para reestruturar o modelo de *Big Data* da empresa. Depois de revisar a arquitetura de BI e *Analytics*, o banco de dados foi modelado em alta performance na nuvem da AWS.

Após o processo de transferência para a nuvem, o *Big Data Analytics* foi adotado para processar o grande volume de dados e gerar *insights*, como consta no embasamento teórico do trabalho, onde a ideia principal do mesmo é mostrar que o objetivo desse trabalho analítico é contribuir para que as empresas tomem melhores decisões de negócio e a partir da análise dos dados obtidos pelo *Big Data*, possam atender seus clientes de forma diferenciada e proporcionar vantagem competitiva no mercado.

#### 3.1 Sobre a Webmotors

A Webmotors é um portal onde são disponibilizados classificados para venda de veículos – com mais de 200 mil ofertas por mês –, programas de manutenção preventiva veicular e produtos específicos direcionados especialmente às revendas de veículos e empresas atuantes no segmento automobilístico, como fabricantes de veículos e autopeças, e outras empresas que atuam nesse segmento. O portal agora investe em mais uma frente estratégica do negócio com o *Big Data*, para extrair informações muito valiosas ao setor. Extrair inteligência para ajudar a entender o uso do portal, melhorar a experiência do usuário e ampliar a oferta aos clientes com informações importantes sobre o mercado eram as necessidades cruciais da Webmotors ao buscar as soluções da *Amazon Web Services*.

#### 3.2 O Desafio

Segundo o coordenador de TI da empresa, Denis Dainese, a *Webmotors* estava com dificuldades para visualizar e analisar as informações que tinha, e a primeira constatação da gestão da equipe foi que era necessário corrigir esse problema: era primordial conseguir extrair e cruzar determinadas informações que as ferramentas usadas até então não permitiam.

Segundo Denis, não era possível consolidar informações de audiência para estudar tendências. Diz também que não era possível fazer análises dos logs de busca, nem para descobrir qual era a cor de carro mais buscada, ou a palavra-chave mais digitada no campo aberto de busca.

Com o *Amazon Redshift*, a Webmotors reestruturou seu modelo de *Big Data*. Em pouco mais de dois meses a empresa revisou toda a arquitetura de BI e *Analytics*, modelando o banco de dados de alta performance na nuvem da AWS.





Denis diz ainda que após reestruturar todo o BI e criar o banco de dados colunar, com o suporte da AWS, foi montado um *cluster* de *Big Data*, para que se tivessem condições de processar esse nível de informação e tirar os *insights*.

### 3.3 Os benefícios

Apesar do pouco tempo de projeto, a Webmotors já consegue extrair muito valor ao negócio com a nova arquitetura de *Big Data*, e em pouco mais de dois meses consegue criar novos produtos a partir das informações.

“Nós já temos painéis em tempo real sobre o portal, então eu já consigo ver o faturamento em tempo real, olhar toda a parte de *match* entre clientes e vendedores de veículos. Consigo saber a quantidade de propostas que eu estou gerando naquele período, fazer comparações com o período anterior. Então o projeto de *Big Data* foi muito rápido e ajudou a obter todos esses indicadores, e a processar essa informação, que para nós é muito valiosa” (WEBMOTORS, 2018).

### 3.4 Próximos Passos

O projeto continua evoluindo internamente, com o uso de produtos analíticos ainda mais avançados, e *machine learning*. A Webmotors está desenvolvendo seu *LiveMarket*, um novo módulo com informações mais precisas para os clientes: portfólio de carros, melhores cotações, melhor valor de venda para o carro, e vai monetizar as análises de informações que antes da solução de *Big Data* eram trabalhosas ou mesmo inviáveis. “Nosso cliente agora vai saber com mais precisão qual a melhor forma dele comprar ou qual a melhor forma dele vender” (WEBMOTORS, 2018).

## 4 CONCLUSÃO

As empresas experientes e profissionais de TI sabem que otimizar seu uso de *Big Data* como um recurso irá gerar valor real de negócios. Uma abordagem de governança e gerenciamento abrangente é necessária para concretizar tais benefícios e gerenciar o risco associado à coleta, análise e armazenamento de informações confidenciais e as implicações de recursos envolvidos (ISACA, 2013).

Devido à quantidade de dados que são coletados pelas empresas atualmente, o *Big Data* vem ganhando cada vez mais importância para as empresas que buscam uma vantagem competitiva no mercado. Vantagem essa que pode ser obtida através da análise desse volume gigantesco de dados fornecidos pelos diversos meios aqui já abordados.

A ideia de tornar mais próxima a relação de uma empresa com seus clientes é possível graças à análise desses dados e identificação de padrões que outras ferramentas não conseguem realizar, ajudando os tomadores de decisão a escolher os fatos em vez de apenas seguirem a intuição, o que resulta numa clientela mais fiel e satisfeita. Visto que nos dias atuais, o tempo é cada vez mais valorizado e o objetivo do *Big Data Analytics* é fornecer velocidade de captura de dados e a sua análise profunda, a tendência é de que as empresas que apostam nesta nova tecnologia tomem a frente do mercado, fornecendo promoções específicas a cada pessoa que se torna seu cliente, rapidez na hora da compra de um produto ou serviço e



às vezes até antecipando os possíveis desejos do mesmo, o que gera mais confiança e preferência pela empresa.

## REFERÊNCIAS

- CAIÇARA JÚNIOR, C. **Sistemas integrados de gestão ERP**. 3. ed. Curitiba: Ibepex, 2008.
- CERVO, A.L.; BERVIAN, P.A. **Metodologia científica**. 4.ed. São Paulo: Makron Books, 1996.
- FERNADES, A, A; ABREU, V. F. **Implantação a governança de TI: da estratégia a gestão dos processos e serviços**. 4. ed. Rio de Janeiro: Brasport, 2014.
- GALDINO, N. **Big data: Ferramentas e aplicabilidade**. XIII SEGet Simpósio de excelência em gestão e tecnologia: 2016.
- GIL, A. C. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2002.
- HEKIMA. **Big data analytics**. Janeiro de 2017. Disponível em: <<http://www.bigdatabusiness.com.br/>>. Último acesso em: setembro de 2018.
- IBM. **Soluções analíticas e otimização de negócios: A nova vantagem competitiva**. Junho de 2011.
- ISACA. **Big data: impactos e benefícios**. Março de 2013.
- REZENDE, D.A.; ABREU, A. F. de. **Tecnologia da informação aplicada a sistemas de informações empresariais: o papel estratégico da informação e dos sistemas de informação nas empresas**. São Paulo: Atlas, 2000.
- SIEWERT, A. B. **Big data in the cloud: data velocity, volume, variety veracity**. IBM Developers Works. July 2013.
- STAIR, R. M; REYNOLDS, G. W. **Princípios de sistemas de informação**. São Paulo: Cengage Learning, 2009.
- WEBMOTORS. **Estudo de caso da Webmotors**. Disponível em: <<https://aws.amazon.com/pt/solutions/case-studies/Webmotors>>. Último acesso em: Setembro de 2018.