

FUNDAMENTOS DE BIG DATA

Sergio Eduardo Nunes



SOLUÇÕES
EDUCACIONAIS
INTEGRADAS



Características de cenário de *big data*

OBJETIVOS DE APRENDIZAGEM

- Ao final deste texto, você deve apresentar os seguintes aprendizados:
- > Descrever as características de cenários de *big data* (5 Vs).
 - > Definir os desafios que um cenário de *big data* impõe a uma infraestrutura computacional.
 - > Relacionar diversos cenários de *big data* e suas características.

Introdução

Com o avanço das tecnologias ligadas a redes de comunicação, o advento da internet móvel, a maior oferta de acesso à internet por parte das operadoras de telecomunicações e o aumento dos serviços convergentes para a rede mundial de computadores, a quantidade de dados gerados aumentou proporcionalmente. Porém, para que, de fato, esses dados possam ser convertidos em informações, é necessário que ocorra um tratamento, visto que se trata de grandes volumes de dados.

Neste capítulo, você vai compreender como identificar os tipos de dados gerados pelos mais diferentes meios. Você vai verificar no que consistem os 5 Vs do *big data* e porque esses conceitos são considerados orientadores de qualidade para que se utilize o *big data* de forma adequada. Em um segundo momento, você vai compreender os desafios encontrados para se operacionalizar o *big data* nas empresas e, dessa forma, gerar *insights* interessantes para que os gestores possam utilizar as informações nas tomadas de decisão. Por fim, você vai estudar os possíveis cenários nos quais podem ser aplicadas as técnicas do *big data*.

Introdução ao *big data* (5 Vs)

As técnicas de *big data* se tornaram um grande diferencial competitivo para as empresas, mas não têm aplicação somente nessa área. Um exemplo de uso do *big data* foi a eleição presidencial dos Estados Unidos de 2016, em que foram utilizados os termos mais pesquisados para que se conseguisse atingir o maior número de delegados para alguns presidenciáveis. Mas o que, de fato, vem a ser o *big data*? Segundo *Morais et al. (2018)*, o termo *big data* se refere a uma base de dados com uma quantidade enorme de dados, que pode estar alocada em múltiplos servidores, e não necessariamente esses dados precisam estar estruturados (como ocorre nos bancos de dados do tipo relacional).

Tipos de dados em um sistema de gerenciamento de banco de dados

Os dados dentro do sistema de gerenciamento de banco de dados (SGBD) podem estar organizados de três formas: estruturados, não estruturados ou semiestruturados, conforme descrito a seguir.

Dados estruturados

São aqueles com tamanhos definidos em seu desenvolvimento. Em grande parte, correspondem a números, datas e palavras. Geralmente são utilizados em bancos de dados do tipo relacional, em que a estrutura precisa estar definida por tipos e tamanhos, o que define o relacionamento por meio das chaves. Observe a Figura 1, em que é apresentada uma forma estruturada dos dados.

```
4 ▼ CREATE TABLE Meus_Contatos (  
5     Id int PRIMARY KEY auto_increment,  
6     Nome varchar(40),  
7     DataNasc date NOT NULL,  
8 );
```

Figura 1. Exemplo de dados estruturados.

Nesse exemplo, o atributo `Id` foi definido para receber um número inteiro; apesar de não ter sido determinado o tamanho, o SGBD adota o padrão de comprimento 11. Já o atributo `Nome` é do tipo `varchar` de tamanho 40. Por fim, o atributo `DataNasc` é do tipo `data`, sendo que o seu formato pode variar conforme

a linguagem de programação de banco de dados — por exemplo, dia/mês/ano, ano/mês/dia etc. Esses tipos de dados são encontrados em planilhas eletrônicas (Excel), arquivos XML (eXtensible Markup Language), arquivos CSV (comma-separated values), arquivos JSON (JavaScript Object Notation), entre outros.

Dados não estruturados

São dados que não possuem formatos e cujo tamanho pode variar. São bastante utilizados, principalmente aqueles relacionados a imagens. Os dados não estruturados são encontrados em fotografias, vídeos, imagens de satélites, dados científicos e mídias sociais, entre outros. A Figura 2 traz um exemplo de dados não estruturados.

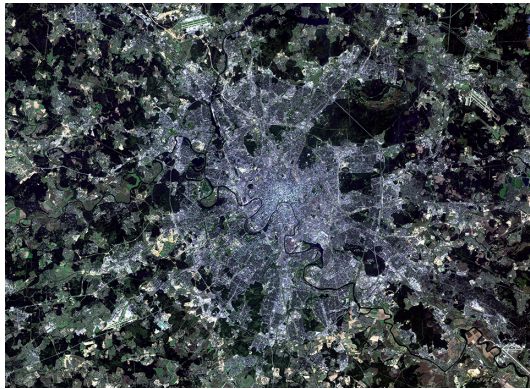


Figura 2. Exemplo de dados não estruturados.

Fonte: Voran/Shutterstock.com.

Com isso, ao se armazenar um dado não estruturado, é possível fazer análises numéricas, por meio do mapeamento dos vértices e das arestas, entre outras tratativas. Esses tipos de dados são encontrados em arquivos de texto, vídeos do YouTube, imagens do Instagram, mensagens SMS (*short message service*, ou serviço de mensagens curtas), dados de geolocalização, conversas de *chat*, entre outros.

Dados semiestruturados

São considerados por especialistas da área como um meio termo entre os dados estruturados e os dados não estruturados. O seu uso está ligado a

aplicações *web*, em que os dados são convertidos em *tags* (etiquetas — termo utilizado na linguagem de marcação utilizada por navegadores). A Figura 3 traz um exemplo de dados semiestruturados.

```
1 <?xml version="1.0" encoding="ISO-8859-1"?>
2 <recado>
3   <para>Aluno</para>
4   <de>Professor Serginho</de>
5   <título>Vamos estudar?</título>
6   <corpo>Fundamentos de Big Data</corpo>
7 </recado>
```

Figura 3. Exemplo de dados semiestruturados.

Nesse exemplo, foram utilizadas as *tags* do XML. As *tags* são bem definidas quanto ao tipo de dado que vai ser recebido, diferentemente do HTML (HyperText Markup Language, ou Linguagem de Marcação de Hipertexto), em que as *tags* são predefinidas, como `<form> </form>` (para formulários), `<p> </p>` (para parágrafos), entre diversas outras. Esses tipos de dados são encontrados em redes sociais (*posts*, *likes*, *tweets* etc.), mídias (fotos, áudio e vídeo), arquivos HTML, *e-mail*, entre outros.

Os 5 Vs do *big data*

Agora que você conhece os tipos de dados utilizados em um SGBD, poderá compreender a forma como o *big data* é estruturado seguindo os conceitos dos **5 Vs**, apresentados a seguir com base em Barbieri (2011).

- **Volume:** se refere ao conceito principal do *big data*, em que a quantidade de dados a ser coletada e tratada representa um grande volume. Esses dados são advindos de diversas fontes, como redes sociais, motores de busca da internet, *e-commerce*, entre outras. Este talvez tenha sido o grande diferencial em relação aos outros bancos de dados: o tratamento de grandes volumes de informações de dados, com formas diferentes (estruturados, não estruturados e semiestruturados).
- **Velocidade:** como os dados são gerados em grandes volumes, o seu processamento deve efetuar o tratamento da mesma forma, caso contrário, ocorrem filas e a consequente degradação do serviço. Para

tanto, o servidor (*hardware*) deve possuir uma capacidade de processamento adequada às necessidades. O SGBD também deve permitir que grande quantidade de dados seja recebida e processada.

- **Veracidade:** esse talvez seja um desafio encontrado no *big data*, pois os dados coletados nem sempre vão representar uma informação na qual se possa ter total confiabilidade. Para isso, ao se estruturar uma análise de dados, faz-se necessário verificar as fontes, os dados tendenciosos e as datas de publicação.
- **Variedade:** os bancos de dados que trabalham dentro do conceito de *big data* devem possuir técnicas que possibilitem o tratamento dos mais diferentes tipos de dados, como números, *tags*, imagens, textos etc. Isso se refere a dados dos três tipos: estruturados, não estruturados e semiestruturados.
- **Valor:** mostra a significância dos dados coletados e tratados para a organização, de forma que possam trazer informações relevantes, que possam proporcionar um diferencial de mercado ou, ainda, auxiliar os gestores na tomada de decisão. Esse último ponto das características do *big data* é exatamente o que as empresas buscam e precisam, pois os dados de valor podem ser convertidos em informações importantes para a companhia.



Fique atento

As tecnologias encontradas no *big data* devem sempre estar alinhadas com as legislações vigentes no país. Isso porque existe um limiar entre mineração de dados e espionagem digital. Embora sejam conceitos bem distintos, o *big data*, conforme a sua aplicabilidade, pode configurar um ato ilícito.

Mas, de que forma que o 5 Vs, de fato, são utilizados na aplicação do *big data*? Inicialmente, você deve analisar o cenário no qual será aplicado e utilizado o *big data*. Com isso, será possível compreender o volume e a velocidade com que esses dados são gerados, sendo necessário observar a variedade dos dados e utilizar filtros para se garantir a veracidade desses dados coletados. Por fim, deve-se garantir que esses dados sejam de grande significância e valor para a empresa.

Como você pode perceber, o *big data* possui uma estrutura bem organizada e definida. O seu potencial de mercado faz com que seu domínio e sua compreensão cada vez mais sejam encontrados entre os colaboradores

especializados em análise de dados. Esse movimento se dá devido à evolução das tecnologias relacionadas a banco de dados, análise de dados, mineração de dados e *big data*.

Os desafios do *big data*

Você percebeu a grande capacidade tecnológica e o potencial de negócio que estão ligados ao *big data*? Pois bem, para que tudo isso possa realmente funcionar, é necessário que exista uma infraestrutura de tecnologia da informação (TI) adequada ao cenário que se deseja explorar. Ainda, há a necessidade de profissionais que saibam operacionalizar, desenvolver e configurar os *software* e toda a infraestrutura física para mineração e tratamento dos dados.

Existem diversos desafios para se colocar em prática o *big data*, de forma segura e que gere informações confiáveis. Embora já existam equipamentos, infraestruturas e demais tecnologias desenvolvidas para prover o serviço, ainda assim, existem entraves e restrições que necessitam ser ajustados e contornados. A seguir, você vai estudar os aspectos que se apresentam como desafios aos profissionais de TI que desejam utilizar o *big data*.



Exemplo

Imagine que você é dono de uma escola de programação. De que forma você vai chegar nos seus clientes? Quem são e onde estão os seus clientes? O *big data*, por meio da mineração de pesquisas em buscadores, *tags*, histórico de pesquisa em *sites* de cursos, entre outros meios, proverá dados para serem tratados, analisados e, por fim, utilizados na divulgação de sua empresa.

Infraestrutura de redes de computadores

As redes de computadores possuem diversos equipamentos intermediários, servidores, meios guiados e não guiados de transmissão, entre outras particularidades. Porém, toda essa infraestrutura de redes possui uma limitação quanto à taxa de processamento de dados. Além disso, as redes de comunicação atuais possuem características de redes convergentes, em que serviços que antes eram utilizados por outros meios acabaram convergindo para a rede mundial de computadores. Um exemplo disso são os serviços

de televisão, que hoje são consumidos em sua grande maioria por vídeo *on demand* e IPTV (Internet Protocol Television).

Segundo Tanenbaum (2011), esse cenário contribuiu significativamente para o aumento da vazão de dados (*throughput*), sendo esse um grande desafio para o *big data*, que abrange a mineração dos dados advindos de serviços providos na rede mundial de computadores. Tais limitações não são impeditivos para que o *big data* possa ser operacionalizado, porém, em alguns casos, a velocidade com que esses dados serão minerados pode variar conforme o cenário que está sendo trabalhado.



Saiba mais

Se você quer entender um pouco mais de redes de computadores, acesse o *síte* da Teleco, um centro de conhecimentos sobre telecomunicações, redes convergentes e redes de computadores.

Extração e tratamento dos dados

Devido à quantidade de dados gerados, um dos grandes desafios encontrados nas técnicas de *big data* consiste em identificar quais dados são interessantes e essenciais para a empresa. O processo de coleta deve possuir filtros que permitam separar ao máximo quais dados realmente devem ser incluídos no tratamento de dados.

Outro desafio nesse quesito está relacionado ao tipo dos dados coletados. Como visto anteriormente, os dados podem ter diferentes formatos, sendo eles: estruturados, não estruturados e semiestruturados. Quando se tem uma coleta em que existam dados não estruturados ou semiestruturados, é necessário que ocorra um pré-processamento desses dados; somente depois disso é possível iniciar uma análise. Esse ponto pode variar conforme o SGBD, pois, se o SGBD for do tipo NoSQL (Not Only Structured Query Language), será possível utilizar os dados de forma não estruturada ou, ainda, semiestruturada.

Embora esse seja um desafio ao *big data*, você não deve esquecer que a essência das técnicas de mineração de dados é exatamente permitir que dados de diferentes fontes, formatos e nichos possam ser capturados segundo um filtro que atenda às necessidades de uma empresa e, conforme os objetivos, possam gerar informações relevantes aos gestores. O principal intuito da mineração de dados é garantir uma extração de dados eficaz, que, por meio dos tratamentos, possa ser útil.

Proteção e segurança

Aqui, existem dois pontos de vista que são um grande desafio para as empresas e para os profissionais de TI, descritos a seguir.

- **Proteção dos dados da empresa:** normalmente, as empresas possuem mecanismos de proteção contra invasões ou, ainda, formas de proteger os dados quando são transmitidos. Isso pode ser um impeditivo nas tarefas de captação dos dados, pois, devido aos métodos criptográficos, os dados capturados não podem ser utilizados para análise de dados (*data analytics*).
- **Proteção das informações do *big data*:** uma vez que os dados tenham sido coletados e tratados, estarão prontos para análise. Esses dados já filtrados requerem proteção, pois refletem um precioso bem para a empresa, consistindo em informações relevantes para os negócios.

Os aspectos relacionados à segurança da informação sempre se mostraram como uma grande preocupação. São inúmeros os ataques que ocorrem diariamente, podendo um incidente ser irremediável no que tange à sobrevivência da empresa. Tanenbaum (2011) define que os investimentos em segurança da informação são necessários não só para a proteção de dados e equipamentos, mas para a garantia de continuidade dos negócios como um todo.

Mão de obra

Muitos segmentos do mercado de trabalho sofreram com algumas recessões que prejudicaram o País financeiramente. Porém, a TI apresentou um aumento na demanda quando os outros nichos do mercado encontraram uma possível saída para a crise ao empreenderem no mercado eletrônico. Apesar dessa grande oferta de mercado, ainda existe uma grande dificuldade de encontrar profissionais especializados em algumas tecnologias.

Os profissionais especializados em *big data* e *data analytics* são raros no mercado. Isso se deve possivelmente ao nível de dificuldade em aprender, configurar, desenvolver e programar as tecnologias necessárias para que se possa, de fato, utilizar o *big data* tendo em vista as necessidades das empresas. Outro desafio nesse ponto é encontrar profissionais de TI que, além de possuírem competências e habilidades no desenvolvimento de tecnologias

de banco de dados e *big data*, saibam estreitar seu relacionamento com os gestores, permitindo que os dados gerem *insights* confiáveis e proporcionem uma tomada de decisão estratégica.

Infraestrutura de sistema de gerenciamento de banco de dados

Esse desafio está muito mais ligado aos aspectos financeiros do que com as limitações impostas pela capacidade dos equipamentos. Isso porque, com a quantidade cada vez maior de dados, é necessário ter uma infraestrutura para o SGBD adequada às necessidades da empresa. Para tal, são necessários *data centers*, que permitem ter em sua estrutura servidores com uma grande capacidade de armazenamento.

Uma das tecnologias mais adotadas pelas empresas é a migração de serviços de armazenamento para a nuvem. Hoje, existem diversas empresas que possuem soluções acessíveis e interessantes. Os motivos para essa mudança envolvem os aspectos técnicos descritos a seguir.

- **Segurança:** normalmente, os provedores disponibilizam métodos de segurança por padrão nos pacotes de serviços.
- **Disponibilidade:** esse ponto pode variar conforme o pacote contratado. O SLA (*Service Level Agreement*) é o acordo de nível de serviço em que é determinado o tempo de disponibilidade dos serviços. Um exemplo é um servidor que tenha um SLA de 99,999%, ou seja, o serviço poderá ficar indisponível apenas 5,26 minutos por ano.
- **Interoperabilidade:** os serviços em nuvem, em grande parte, permitem que diferentes equipamentos e tecnologias possam utilizar sistemas de forma transparente. Essa interoperabilidade é muito interessante no que tange ao *big data*, uma vez que os dados podem ser originados de diversas fontes, localizações e equipamentos, com características distintas.
- **Armazenamento:** é um ponto muito interessante, pois existe a possibilidade de se contratar um espaço desejado ou pagar por espaço consumido em determinado tempo. Como as necessidades do *big data* são sempre de grande capacidade de armazenamento de dados, a solução em nuvem é uma das saídas mais interessantes.

Esses aspectos fazem com que cada vez menos empresas escolham montar uma infraestrutura de banco de dados — a maioria opta por comprar soluções

de empresas especializadas em armazenamento em nuvem. Fica a critério dos profissionais de TI adequar e configurar os serviços conforme as necessidades dos gestores e da empresa.

Gerenciamento do crescimento dos dados

Estima-se que, diariamente, o planeta produza por volta de 2,5 trilhões de bytes de dados. É claro que não existe um único computador que seria capaz de processar todos esses dados. Ao se planejar a captação dos dados das mais diferentes fontes, tem-se a necessidade de utilizar técnicas como as descritas a seguir.

- **Compactação:** abrange técnicas e ferramentas tecnológicas que possibilitam diminuir o espaço de armazenamento de um grande grupo de dados. Uma analogia dessa técnica é a utilização de *software* de compactação do tipo RAR.
- **Desduplicação:** trata-se de uma técnica que visa a “limpar” os dados que foram captados. A ideia é utilizar técnicas que permitam excluir os dados duplicados na tabela — dessa forma, a base a ser analisada terá uma taxa menor de processamento.
- **Hierarquização:** é uma forma de organizar os dados minerados de forma que as ferramentas possam fazer a análise dos dados estruturadamente. Essas técnicas ajustam os dados, a fim de diminuir o seu tempo de processamento, para minimizar o tempo de retorno das devolutivas aos gestores.

Nesse contexto, são utilizadas ferramentas especializadas em *big data*, como Hadoop, Spark, *data analytics software*, *business intelligence software*, inteligência artificial, entre outras.

Resistência organizacional

Além dos desafios tecnológicos, existem aqueles ligados aos problemas administrativos encontrados em grande parte das organizações. Esse tipo de desafio pode apresentar três pontos principais, descritos a seguir.

- **Falta de alinhamento organizacional:** ocorre quando existe um distanciamento entre as necessidades gerenciais e o que, de fato, a equipe de TI está desenvolvendo. Essa falta de alinhamento organizacional é um dos fatores que mais podem causar fracasso na operacionalização do *big data*.

- Falta de entendimento: os gerentes têm que estar alinhados com a equipe de TI. Esse estreitamento é necessário para que os filtros possam ser construídos pela equipe de desenvolvimento. Assim, os dados serão extraídos conforme a análise necessária para a tomada de decisão de forma mais assertiva.
- Resistência gerencial: novas soluções ou tecnologias podem causar certa resistência. Cabe ao setor de TI demonstrar o quanto os dados gerados pelo *big data* podem ser úteis para que os gestores possam tomar decisões com dados atualizados e tratados.

Assim, percebe-se que, apesar da potencialidade que o *big data* apresenta como ferramenta competitiva nas empresas, são muitos os desafios. Porém, os profissionais de TI possuem bases de conhecimento bem sólidas, que permitem alcançar os resultados esperados, e podem contar com uma infraestrutura com diversas opções de equipamentos e serviços. Um ponto importantíssimo está ligado à grande oportunidade de mercado que os profissionais especializados e qualificados têm diante da necessidade das empresas.

Aplicação do *big data*

O *big data* surgiu como uma tecnologia inovadora e com alto potencial para que as empresas possam garantir o seu espaço em um mercado cada vez mais competitivo. Mas, em quais segmentos o *big data* pode ser aplicado? Quais são os *insights* esperados pelos gestores? Conhecer exemplos de algumas aplicações dessa tecnologia pode proporcionar ao profissional de TI maior assertividade no momento de aplicar as técnicas do *big data* em diversos cenários.

Seria impossível discutir todos os cenários em que o *big data* é aplicável; porém, ao longo desta seção, você vai verificar alguns segmentos nos quais essa tecnologia já é largamente utilizada. O intuito desses exemplos é servir como um guia para as demais aplicações do *big data* em diversos outros cenários, a fim de se proporcionar o aumento da eficiência e o alcance de resultados que levem a empresa a um diferencial competitivo.

Antes de iniciar as discussões, vamos definir o que são **cenários** no contexto do *big data*. Cenário significa “Conjuntura; os fatos ou situações que acontecem num dado momento: cenário econômico; cenário cultural; cenário histórico” (DICIO, c2020, documento *on-line*). No contexto do *big data*, os cená-

rios analisados demonstram a conjuntura encontrada em alguns segmentos do mercado. Em sua maioria, englobam serviços consumidos por grande parte da população, setores de empresas e serviços públicos.

Ramo empresarial

Segundo Dumbill (2013), com o grande volume de informações geradas, o *big data* despertou o interesse de diversas empresas. Quando aplicado no ramo empresarial, permite observar os aspectos descritos a seguir.

- **Comportamentos e tendências:** o *big data* permite compreender o comportamento dos consumidores e a tendência de consumo, apontando quais produtos ou serviços podem ser direcionados a determinado nicho ou grupo. Uma aplicação está ligada, por exemplo, ao ramo imobiliário, em que é possível compreender se as pessoas que moram na capital, ao se aposentarem, preferem o interior ou o litoral para morar. Isso permite ao setor imobiliário enviar opções de imóveis para os clientes e atingir um maior número de vendas.
- **Estratégia de *marketing*:** grande parte da aplicação comercial do *big data* está ligada às estratégias de *marketing*, pois ele permite realizar análises de dados e direcionar as ofertas de produtos e serviços de forma mais assertiva. Isso permite aos gestores direcionar propagandas e demais materiais publicitários para potenciais clientes.
- **Melhoria de produtos e serviços:** os *feedbacks* fornecidos por consumidores de produtos e serviços em redes sociais, *sites* de reclamação etc. são utilizados de forma a proporcionar *insights*, que são informações úteis para equipes de desenvolvimento e concepção de produtos.



Exemplo

Por exemplo, repare que, ao fazer uma pesquisa de algum produto ou serviço no Google, o Facebook, após um tempo, começa a exibir no histórico de notícias exatamente o que você havia pesquisado. Outro exemplo é a Amazon, em que, ao entrar na página de venda de livros, são exibidos os títulos relacionados ao que você estava buscando.

Segundo Morais *et al.* (2018), o *big data*, com o passar dos anos, vai moldar a forma como as empresas competem, em um mercado acirrado e com poucas chances de erros. Isso ocorrerá por meio da identificação de padrões de

comportamentos, fazendo com que as empresas se tornem mais eficientes e dinâmicas. Isso tudo pode se tornar um diferencial não só em nível comercial, mas também do ponto de vista estratégico.

Área da saúde

Segundo Chiavegatto Filho (2015), existem duas potenciais aplicações do *big data* na saúde, descritas a seguir.

- **Medicina de precisão:** visa a corrigir as deficiências na prescrição de medicamentos que funcionam bem para determinados pacientes e para outros não. Além disso, já vem sendo utilizada a fusão entre *big data*, Internet das Coisas e aprendizagem de máquina para a redução de acidentes vasculares encefálicos, diabetes, entre outros.
- **Prontuários eletrônicos:** esse tipo de prontuário já é utilizado em grande parte do sistema de saúde, porém, é pouco explorado pelo *big data*. Essa ferramenta computacional auxilia os profissionais da saúde na prescrição de medicamentos, em diagnósticos baseados em históricos similares e no acesso a documentos em diferentes centros de saúde.

Nem sempre os dados são utilizados de forma correta a fim de se prevenir algumas epidemias. Um caso recente se deu quando empresas ligadas à saúde no Brasil analisaram, nos anos de 2007 e 2008, as pesquisas americanas relacionadas à gripe H1N1, abrangendo temas como epidemia, medicamentos e demais assuntos relacionados. Porém, em 2009, as informações proporcionadas pelo *big data* não foram suficientes para prevenir uma epidemia de H1N1 no Brasil, que acabou vitimando diversas pessoas nos cinco continentes.

Os dados são gerados a todo momento e, se bem utilizados, podem beneficiar não somente os profissionais de determinado segmento, mas a população como um todo. Quando a tecnologia do *big data* foi concebida, grande parte da sua aplicação estava ligada à indústria e ao comércio; com o tempo, outras áreas do conhecimento foram se beneficiando dos resultados. É claro que, para que a aplicação eficaz do *big data* ocorra, existe a necessidade de atuação de profissionais de ambas as áreas; por exemplo, no caso da saúde, é necessário que médicos e profissionais de TI encontrem formas de utilizar os dados gerados.

Trânsito em grandes cidades

A aplicação do *big data* no trânsito ocorre por meio da chamada gestão de tráfego, que já conta com testes e estudos relacionados ao tema. Em cidades onde o número de veículos é grande, como nas capitais, as pessoas gastam boa parte do tempo no deslocamento entre pontos dentro do município, como: casa–trabalho, trabalho–faculdade, faculdade–casa. O nível de degradação física e mental originado por esse número de deslocamentos faz com que as atividades não sejam executadas com todo o potencial do colaborador no seu trabalho, por exemplo.

Por meio da captação de dados relacionados a eventos, estreias, tendências de compras, entre outros, por exemplo, o *big data* é capaz de gerar informações de possíveis pontos de congestionamento, permitindo aos gestores responsáveis pelo trânsito e pela mobilidade urbana tomarem decisões quanto à fluidez dos veículos nas ruas e avenidas das cidades. Essas técnicas vêm sendo utilizadas timidamente pelo Poder Público, porém, com o aparecimento dos resultados positivos, o número de cidades deve gradativamente aumentar frente aos benefícios do *big data*.

Comportamento dos consumidores

Uma das primeiras aplicações do *big data* está relacionada ao comportamento de consumo, a encontrar possíveis compradores e públicos-alvo de interesses e às demais tendências comportamentais. Nesse setor, o *big data* pode representar um potencial ganho financeiro para as empresas.



Exemplo

Imagine que você tenha uma fábrica de celulares e saiba que o tempo médio de vida das baterias varia entre 3 e 4 anos. Caso uma base de dados de grandes lojas e supermercados fosse compartilhada, possibilitando conhecer os clientes que compraram um celular da sua marca nesses locais, isso possibilitaria direcionar a esses clientes propagandas de baterias novas a cada 3 a 4 anos.

Atualmente, as empresas de telecomunicações, mais especificamente com produtos ligados à telefonia e à internet móvel, utilizam a análise de dados para compreender o comportamento do consumidor. Dessa forma, podem ser enviadas ofertas de produtos e serviços alinhadas ao perfil de consumo do cliente, assim como promoções personalizadas.



Saiba mais

O Marco Civil da Internet (Lei nº. 12.965, de 23 de abril de 2014) foi um avanço constitucional para a proteção das pessoas e das empresas. Com a sua promulgação, alguns parâmetros que não contavam com definições constitucionais foram amplamente descritos e explorados. O Comitê Gestor de Internet no Brasil disponibiliza na íntegra o Marco Civil da Internet em seu *site* oficial.

Investimentos

Esse cenário talvez represente uma das maiores aplicabilidades do *big data*. Grande parte das corretoras possuem sistemas inteligentes que efetuam análises cambial, de taxas de mercado, de variação de ações e moedas, de balanços financeiros das empresas, entre outras. Isso se dá por meio de ferramentas integradas a diversas tecnologias. Para compreender o seu funcionamento, observe a Figura 4.

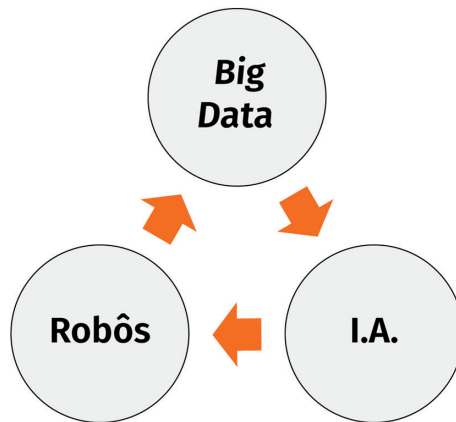


Figura 4. Funcionamento de sistemas inteligentes no mercado de ações.

O funcionamento desses sistemas inteligentes se baseia na mineração de dados e no respectivo tratamento desses dados (*big data*), de forma que a inteligência artificial de uma aplicação possa tomar a melhor decisão. Isso, por sua vez, permite que os robôs investidores executem as melhores opções de mercado, como: compra de ações, *tradings*, câmbio de moedas, entre outras possíveis operações do mercado financeiro. Essas operações efetuadas são recapturadas pelas ferramentas do *big data*, e o processo é reiniciado.

Você já deve ter ouvido aquela expressão “o céu é o limite” — ela se encaixa muito bem no contexto dos cenários em que o *big data* pode atuar. Não existe uma limitação para determinado segmento. Desde que seja possível, de alguma forma, gerar e captar dados, será possível fazer com que o *big data* se torne uma importante ferramenta de estratégia de mercado. Contudo, deve sempre haver um cuidado dos profissionais ligados a essa área do conhecimento em desenvolver os métodos mais adequados para cada caso.

Referências

BARBIERI, C. *BI2 - business intelligence: modelagem e qualidade*. Rio de Janeiro: Elsevier, 2011.

CHIAVEGATTO FILHO, A. D. P. Uso de big data em saúde no Brasil: perspectivas para um futuro próximo. *Epidemiol. Serv. Saúde*, v. 24, n. 2, p. 325-332, 2015. Disponível em: <https://www.scielo.br/pdf/ress/v24n2/2237-9622-ress-24-02-00325.pdf>. Acesso em: 31 ago. 2020.

DICIO. *Cenários*. [S. l.]: Dicio, Dicionário Online de Português, c2020. Disponível em: <https://www.dicio.com.br/cenarios>. Acesso em: 31 ago. 2020.

DUMBILL, E. Making sense of big data. *Big Data*, v. 1, n. 1, 2013. Disponível em: <http://online.liebertpub.com/doi/pdf/10.1089/big.2012.1503>. Acesso em: 8 jun. 2016.

MORAIS, I. S. et al. *Introdução a big data e internet das coisas (IoT)*. Porto Alegre: SAGAH, 2018.

TANENBAUM, A. S. *Redes de computadores*. 5. ed. São Paulo: Pearson Education do Brasil, 2011.

Leituras recomendadas

BAND JORNALISMO. *Big data: faltam profissionais de análise de dados*. [S. l.]: YouTube, 2020. 1 vídeo. Disponível em: https://www.youtube.com/watch?v=_lkFtmGEPQQ. Acesso em: 6 jul. 2020.

BRASIL. *Lei nº 12.965, de 23 de abril de 2014*. Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil. Brasília, DF: Presidência da República, 2014. Disponível em: <https://www.cgi.br/lei-do-marco-civil-da-internet-no-brasil/>. Acesso em: 6 jul. 2020.

BULIAN, B. S.; ALENCAR, C. A. A importância do big data para a compreensão dos hábitos de compra dos consumidores, aumentando as vantagens competitivas no comércio varejista. *Revista Científica Multidisciplinar Núcleo do Conhecimento*, v. 3, n. 2, p. 18-41, 2017.

COSTA, C. P. N.; COSTA, R. S.; CONRADO, L. M. S. Big business: o emprego do big data na melhoria da gestão de negócios. *Revista Científica Multidisciplinar Núcleo do Conhecimento*, v. 6, n. 3, p. 29-40, 2020.

TELECO. [S. l.]: Teleco, c2020. Disponível em: <https://www.teleco.com.br/>. Acesso em: 6 jul. 2020.



Fique atento

Os *links* para *sites* da *web* fornecidos neste capítulo foram todos testados, e seu funcionamento foi comprovado no momento da publicação do material. No entanto, a rede é extremamente dinâmica; suas páginas estão constantemente mudando de local e conteúdo. Assim, os editores declaram não ter qualquer responsabilidade sobre qualidade, precisão ou integralidade das informações referidas em tais *links*.

Conteúdo:



SOLUÇÕES
EDUCACIONAIS
INTEGRADAS