



# A TENDÊNCIA DO BIG DATA

Vivemos um contexto empresarial em que todos, ou quase todos, querem “contratar soluções de *big data*”. Executivos de grandes corporações ouvem falar disso em eventos sobre tendências tecnológicas e comportamento do consumidor, e logo querem implementá-las em suas organizações.

Para começar, o termo é infeliz. *Big data* não é nem só *big* nem só *data*. *Big data analytics* (o termo correto e completo) são, segundo a EMC, dados e modelos cuja escala, distribuição, diversidade ou velocidade de criação requerem o uso de novas tecnologias de armazenamento e análise para permitir a captura do valor inserido neles.

Reconhecemos, portanto, várias dimensões por trás desse termo. “Volume”, sem dúvida, é uma delas. Não é raro falarmos em *petabytes* (um milhão de *gigabytes*) ou mesmo em *exabytes* (mil *petabytes*) para representar implantações de projetos envolvendo *big data* em grandes corporações no Brasil e no mundo, e voltados aos mais diversos mercados, como o de mídia digital, varejo supermercadista, de prospecção de petróleo, ou ainda para a previsão de microclima.

Outra perspectiva do *big data* é a “velocidade”. Um levantamento da Qmee, referente a 2014, aponta que em apenas um minuto a internet movimentou 204 milhões de *e-mails*, mais de 4 milhões de consultas no Google.com, 277 mil *tweets*, 2,46 milhões de atualizações no Facebook e 72 novas horas de vídeo disponibilizadas no YouTube.

Mas a característica que diferencia a velocidade não é seu aumento, e sim sua capacidade de ser cada vez mais assíncrona e *real time*. Em outras palavras, não

conseguimos mais controlar a velocidade com que as informações trafegam. No contexto em que precisamos tomar decisões com rapidez, para gerar valor para as organizações, isso traz um grande desafio para os modelos analíticos – eles precisam se reinventar permanentemente para se manterem úteis.

Uma terceira perspectiva do *big data* é a “variedade”. Mais de 80% dos dados da internet não são estruturados – páginas *web*, dados de *clickstream*, fotos, imagens, vídeos, textos em linguagem natural, mapas,

etc. Tipicamente, informações coletadas a todo instante por radares de trânsito, sensores de clima, câmeras de segurança, comentários e *posts* de redes sociais. Essas informações contemplam um panorama riquíssimo de significados para as empresas, que pode ser apropriado para vários contextos: avaliar como está a reputação de uma marca nas redes sociais, inferir em tempo real a melhor oferta para um usuário de filmes sob demanda no momento em que ele conclui ou desiste de assistir algum conteúdo, suspeitar do trajeto de uma pessoa no estacionamento de um *shopping center*, etc.

Modelos tradicionais de previsão tendem a se tornar lentos se considerarmos volumes muitíssimo grandes de dados.

E é aí que o *big data analytics* surge como nova grande força nas organizações: para analisar de forma coerente e rápida informações em tempo real e em quantidade não controlada, ou obter a melhor estimativa possível sobre a informação no momento em que ela está acontecendo. Afrouxamos um pouco a confiança nas estimativas tradicionais para ganharmos, significativamente, em desempenho – e isso representa, realmente, valor para as organizações.

*Big data* parece ter vindo para ficar.

ANALISAR DE FORMA  
COERENTE E RÁPIDA  
INFORMAÇÕES EM  
TEMPO REAL E  
EM QUANTIDADE  
NÃO CONTROLADA.  
ISSO REPRESENTA,  
REALMENTE,  
VALOR PARA AS  
ORGANIZAÇÕES.