**BIG DATA – UM NOVO CONCEITO DE TRATAMENTO E UTILIZAÇÃO DE DADOS**

Douglas Soares Braga ¹

João Goulart da Silva ¹

Matheus Calegaro Lopes da Silva ¹

Matheus Lima Ferraz ¹

Rafael Monteiro ²

O termo Big Data está cada vez mais popular, digamos que “*está na moda”*, embora por vezes ainda seja mal compreendido. O fato é que, com o crescente aumento de produção de dados, o termo vem à tona de forma natural trazendo novas perspectivas sobre o mundo de TI e novas técnicas de tomadas de decisões de negócios.

O termo se refere ao tratamento de dados de tamanho exponencialmente maior dos que estamos habituados a manusear. Além disso, as tecnologias atuais não se mostram as suficientes para tratá-los. E ainda temos a questão da velocidade com qual esses dados crescem, já que em tempos atuais, Terabytes são o que nós mesmos usamos em nossos backups pessoais hoje em dia.

Para as empresas, tratar esses dados analiticamente pode trazer grandes frutos para as suas receitas anuais. Porém, além de trazer bons frutos, o manuseio desses dados acarreta também grandes desafios para quem os usa.

Nesse ponto de vista, é necessário agir com velocidade para o tratamento dos mesmos, já que tendem a se tornar inutilizáveis de maneira subsequente.

Levando em conta essas vertentes, trabalhar com o Big Data tem sido um excelente investimento para empresas que visam ganhar vantagens competitivas no mercado coorporativo. Uma pesquisa realizada pela Bain & Company, empresa global de consultoria em negócios, coletou dados de 409 grandes empresas e apontou que as organizações adeptas do sistema Big Data prematuramente têm maior chance de superar os concorrentes. Ainda sobre a pesquisa, a B & C afirma que empresas que apostam no investimento têm duas vezes mais chances de estar entre os 25% de empresas com melhores resultados.

**GRANDE MASSA DE DADOS**

É imensurável a quantidade de dados que chega às corporações por meios como a internet. Segundo estudo feito pela IDC em 2011, estima-se que, em 2015, seja produzido em torno de 8000 Exabytes de dados, sendo que, em 2010, foram produzidos pouco mais de 2000 Exabytes. Diariamente, cerca de 15 petabytes são gerados, segundo projeções da IBM. Esse fato é explicado devido à grande demanda de informações geradas e a facilidade de acesso à Web, elevando os dados digitais à patamares econômicos de grande valor para as empresas e governos.

Esses dados chegam das mais diversas forma até os analistas: de cliques em páginas na Internet até informações não estruturadas provenientes de redes sociais. E a velocidade que podem ser coletadas, analisadas e respondidas atingiram uma nova dimensão. Porém, nem todos ainda tiram proveito dessa tendência. Segundo estudo com leitores da Harvard Business Review, apenas 23% dos entrevistados afirmou que suas organizações possuíam alguma estratégia para big data.

Especificamente, Big Data relaciona-se com a criação de dados, armazenamento, recuperação e análise que é notável em termos de volume, velocidade e variedade:

* **Volume:** Um computador comum tinha 10 gigabytes de armazenamento em 2000. Hoje, o Facebook ingere 500 terabytes de dados novos todos os dias; um Boeing 737 gera 240 terabytes de dados de voo; a proliferação de telefones inteligentes, os dados de criar e consumir; sensores embutidos em objetos do cotidiano em breve vão resultar em milhares de milhões de dados novos, constantemente atualizam dados de localização e outras informações, incluindo vídeo.
* **Velocidade:** Cliques em anúncios de diversos sites com o comportamento do usuário são capturados em milhões de eventos a todo segundo; Aalta frequência de troca dos algoritmos refletem as mudanças de mercado dentro de microssegundos; processo de troca de dados entre bilhões de máquinas; infra-estrutura e sensores que geram dados de registro em massa em tempo real; sistemas de jogos onlines suportam milhões de usuários simultâneos, cada um produzindo várias entradas por segundo.
* **Variedade:** Dados Big Data não são apenas números, datas e strings. Big Data também é dados geoespaciais, dados em 3D, áudio e vídeo e texto não estruturado, incluindo arquivos de log e de mídia social. Os sistemas de banco de dados tradicionais foram projetados para atender menores volumes de dados estruturados, menos atualizações ou, uma estrutura de dados consistente previsível. Sistemas de banco de dados tradicionais também são projetados para operar em um único servidor, fazendo com o aumento da capacidade caro e finito. Como os aplicativos evoluíram para atender grandes volumes de usuários, e como a aplicação de práticas de desenvolvimento tornaram-se ágil, o uso tradicional de banco de dados relacional tornou-se um passivo para muitas empresas, em vez de um elemento em seus negócios.

**APROVEITAMENTO DOS DADOS**

O ponto, hoje, segundo DAVENPORT ², “não é ficar deslumbrado com o volume de dados, mas sim analisá-los – convertê-los em conhecimento, inovação e valor de negócio” e que “estamos apenas arranhando a superfície do que é possível (de ser explorado e utilizado)”. A quantidade de gigabytes que o genoma humano carrega, por exemplo, não ajuda a determinar o valor que será obtido ao explorar tais dados. Ainda segundo DAVENPORT, existem três tipos de valor, sendo eles:

1. Redução de custos: gera oportunidade de lucros e investimentos em novas ferramentas, produtos e soluções;
2. Melhoria nas tomadas de decisão: adição ou aperfeiçoamento das fontes de dados a modelos explicativos e preditivos em Sistemas de Apoio a Decisão (SAD), sendo o primeiro o com maior valor, segundo analistas;
3. Melhorias em produtos e serviços: com o feedback dos usuários, principalmente nas redes sociais, fica mais fácil de se localizar um eventual erro e fazer com que os próximos produtos sejam melhores e que esteja de acordo com as vontades dos potenciais clientes;

A General Electronics, visando o potencial do big data, criou um centro de tecnologia em San Francisco somente para tratar de temas de software e big data, contratando vários cientistas para atuarem nas áreas que já movimentam grandes quantidades de dados, como a financeira e a da saúde, e também em potenciais como negócios de locomotivas, aviões e gás.

Atualmente lidamos com uma quantidade exponencial de dados, vindos de todas as partes, partes estas, jamais imaginadas a 20 anos atrás. Tudo o que fazemos na Internet gera dados, bilhões de usuários (leiam-se consumidores) deixam rastros digitais ao fazerem pesquisas no Google, editar comentários em redes sociais, efetuar compras online ou assistir vídeos no YouTube. Essa grandeza de dados está gerando um imenso banco de dados, que denomina-se Big Data. Com essa grande disponibilização de informação, empresas começam a enxergar um novo jeito de chegar até o seu cliente, com acesso seguro e rápido aos dados específicos de seu respectivo segmento. E o mais importante, aliando as ótimas execuções de tarefas com uma melhor redução de custos.

**Palavras-chave:** Big Data; Negócios; Dados; Grande Volume;

Para saber mais:

<http://www.webopedia.com/TERM/B/big_data.html>

<http://oglobo.globo.com/infograficos/bigdata/>

<http://www.sas.com/offices/latinamerica/brazil/solucoes/bigdata/>

<http://www.ibm.com/big-data/us/en/>

<http://radar.oreilly.com/2012/01/what-is-big-data.html>