

## **BIG DATA: TRANSFORMANDO GRANDE QUANTIDADE DE DADOS EM VALOR AGREGADO PARA A EMPRESA – PESQUISA BIBLIOGRÁFICA EM ANDAMENTO**

**Elisabete Tomomi Kowata<sup>1</sup>**

<sup>1</sup> Mestre em Ciência da Computação e Docente da Universidade Estadual de Goiás.

### **Introdução (Problemática e Objetivos)**

A cada dia as empresas produzem *pentabytes* de informações diariamente e apesar de terem enormes quantidades de dados, ainda é grande o desconhecimento de como tratá-las e utilizá-las para criar valor agregado para fins de competitividade.

Muitas empresas perdem a oportunidade de utilizar essas informações para resolver problemas complexos, reduzir custos e decidir melhor o direcionamento da empresa em tempos de crise.

A análise de grande quantidade de dados estão limitados a empresas de grande porte que detém informações para se manterem competitivas. Há uma frase, cuja autoria é atribuída ao Benjamin Disraeli de que “Como regra geral, o homem mais bem sucedido na vida é aquele que dispõe das melhores informações”. Com isso, as grandes empresas já estão na vanguarda utilizando as informações para buscar novos clientes e manterem a fidelidade dos já existentes.

A pesquisa tem como objetivo principal desenvolver estratégias para tornar as empresas competitivas diante das imensas quantidades de dados produzidos diariamente e transferir este conhecimento para as pequenas e médias empresas. Como objetivos específicos temos: (1) Conhecer o big data; (2) Realizar uma revisão sistemática do objeto de pesquisa; (3) Listar a importância do big data na organização; (4) Levantar oportunidades do big data nas empresas; (5) Definir etapas para a utilização do big data nas empresas; (6) Elaborar um portfólio com as soluções que as empresas podem utilizar para tornar a empresa mais competitiva.

## Referencial Teórico

É necessário primeiramente conhecer o conceito que está por trás do termo *Big Data*. A seguir apresentamos alguns desses conceitos.

O termo *Big Data* descreve um conjunto de problemas e suas soluções tecnológicas em computação aplicada com características que tornam seus dados difíceis de tratar (XEXÉO, p. 19).

Já para Davenport (2014a), o termo *Big Data* é genérico para dados que não podem ser contidos nos repositórios usuais. O autor continua referindo-se ao termo como dados volumosos demais para caber em um único servidor e não estruturado demais para se adequar a um banco de dados organizado em linhas e colunas; ou fluídos demais para serem armazenados em um *data warehouse* estático (DAVENPORT, 2014).

As principais características que descrevem o *Big Data* são os 4 Vs: volume, velocidade, variedade e valor. O volume se refere a enorme quantidade de dados e a previsão é de que sejam gerados na faixa de *petabytes* (milhares de trilhões) ou até mesmo *exabytes* (milhões de trilhões) de dados (XEXÉO, 2013). A velocidade significa que esses dados são enviados aos sistemas com taxa de bytes por intervalo de tempo muito alta, tão grande que não temos como armazená-los. A variedade significa que os dados aparecem em formas diferentes quer sejam em formatos mais livres como textos e imagens ou com estruturas específicas como, por exemplo, em redes. E por fim, qual valor está por trás dos dados armazenados e como poderá ser utilizado para os negócios?

O volume, a variedade e a velocidade de dados gerados no mundo sobressaem principalmente na América do Norte. Os norteamericados geram um volume de quase cerca de 4000 dados, seguidos pelos Europeus que chegam a gerar mais de 2000 em volume de dados, sendo que a variedade de dados é gerada de três formas: de pessoas para pessoas; de pessoas para as máquinas e de máquinas para máquinas. Na primeira variedade os dados são gerados por meio de comunidades virtuais, redes sociais e publicação na web; na segunda por arquivos, dispositivos médicos, TV digital, comércio eletrônico, etiquetas inteligentes, computadores, telefones e filmes; e na terceira por sensores, dispositivos GPS, leitores de códigos de barras, câmeras de segurança, investigação científica entre outros.

A partir dessas exposições sobre o tema de pesquisa, a questão que levantamos é: como as empresas podem tratar essas informações para desenvolver estratégias para manterem-se competitivas nesse período conturbado em que o país vive na economia?

Com este questionamento, este projeto de pesquisa tem como objetivo desenvolver estratégias para tornar a empresa competitiva diante das imensas quantidades de dados produzidos diariamente e, posteriormente transferir este conhecimento para as pequenas e médias empresas.

### **Metodologia**

Os procedimentos metodológicos utilizados para a realização deste projeto de pesquisa serão a partir de uma revisão sistemática com registro do protocolo das leituras, identificando as categorias abordadas em cada área, além da pesquisa ser bibliográfica.

A pesquisa seguirá as seguintes etapas:

1. Realização de uma revisão sistemática para identificar estudos da área de pesquisa;
2. A partir da revisão, categorizar e resumir os estudos, identificando aspectos comuns e as diferenças sobre os benefícios do *Big Data* nas empresas;
3. Será realizada uma análise dos resultados obtidos, apresentando um quadro de estudos do objeto de estudo;
4. Ao final será desenvolvido um portfólio onde estará sistematizado as oportunidades para que as empresas possam utilizá-los para tornar a empresa muito mais competitiva.

Portanto, a pesquisa é bibliográfica e será realizada uma busca de publicações na internet utilizando a metodologia de revisão sistemática e a abordagem é quali-quantitativa.

### **Resultados e Discussões**

Como resultado será elaborado um portfólio com as soluções que as empresas podem utilizar para torná-las mais competitivas e será realizada a transferência deste conhecimento para as pequenas e médias empresas, numa próxima etapa da pesquisa.

O projeto ainda está em fase de conclusão da parte bibliográfica a partir da leitura de artigos e livros da área, e os bolsistas estão fazendo estudos de ferramentas para verificar a aplicação nas empresas.

A equipe de pesquisa formada pela Coordenação e dois bolsistas participaram no dia 21 de julho de 2016 do *First Workshop on Business Statistics With Applications* realizado

pelo Instituto de Matemática e Estatística da UFG que abordou a importância e o uso da estatística na análise do *Big Data* com palestrantes estrangeiro e de empresários.

Os bolsistas Danilo Nogueira da Silva e Emanuel Raimundo Parente da Cruz elaboraram juntos um resumo expandido intitulado “Big data: descoberta e previsão de tendências para direcionar o rumo da empresa” e apresentaram no SEPE de 2016, que aconteceu de 28 de agosto a 03 de setembro de 2016 no Campus CSEH. O resumo expandido está disponível nos Anais do SEPE no link: <http://www.anais.ueg.br/index.php/sepe/article/view/7085>.

A pesquisadora também apresentou no X Seminário de Pesquisa dos Professores/SEPE com o resumo expandido “Big data: transformando grande quantidade de dados em valor agregado para a empresa” para divulgar a pesquisa entre os professores. O resumo está disponível nos Anais do SEPE no link: <http://www.anais.ueg.br/index.php/sepe/article/view/7186>.

Os bolsistas de iniciação científica, juntamente com os bolsistas voluntários participaram do curso “Business Intelligence e Big Data” promovido pelo Proin da UEG no dia 4 de maio de 2017. Nesse curso, os bolsistas conheceram os fundamentos de Big Data e utilizaram a ferramenta QlikSense para tratamento de dados.

Nas reuniões de pesquisa que foram realizadas discutimos os capítulos do livro Big Data: como extrair, volume, variedade e valor da avalanche de informação cotidiana; elaboramos a ficha de leitura e análise para registro das leituras realizadas; definimos planos de trabalho para novos bolsistas BIT e VIC para participar da seleção do próximo edital de bolsas da UEG; Discussão do tema do curso “Business Intelligence e Big Data” apresentada pela bolsista voluntária Viviane Gonçalves e apresentação do Design thinking para levantamento de oportunidades para definir um produto/serviço para ser apresentado para as empresas.

## Conclusão

Dentre os objetivos apresentados, estamos na fase de conclusão da pesquisa bibliográfica para conhecer o *big data* e várias bibliografias sobre o tema foram separadas. Finalizamos a leitura do livro *Big Data: como extrair volume, variedade, velocidade e valor da avalanche de informação cotidiana* de Victor Mayer-Schönberger e Kenneth Cukier (2013).

A reunião de pesquisa teve início no dia 2 de agosto de 2016 e foram sendo realizadas no ano de 2016 às terças-feiras no período vespertino e no ano de 2017 as reuniões foram realizadas aos sábados a tarde. Até o momento foram realizadas 14 reuniões presenciais, além de estudos paralelos para atividades de leitura, discussão e produção de textos do tema estudado.

### Referências

DAVENPORT, Thomas H. *Big data no trabalho: derrubando mitos e descobrindo oportunidades*; tradução Cristina Yamagami. 1. ed., Rio de Janeiro: Elsevier, 2014a.

DAVENPORT, Thomas H. *Dados demais!:* como desenvolver habilidades analíticas para resolver problemas complexos, reduzir riscos e decidir melhor; tradução Afonso Celso da Cunha. 1. ed., Rio de Janeiro: Elsevier, 2014b.

FEIJÓ, Bruno Vieira. *A revolução dos dados: entenda como o big data – a extraordinária quantidade de informações coletadas por novas tecnologias – pode ser usado para tornar sua empresa muito mais competitiva.* PME Exame. Setembro 2013, edição 65, p. 30-43.

MAYER-SCHÖNBERGER, Viktor; CUKIER, Kenneth. *Big data: como extrair volume, variedade, velocidade e valor da avalanche de informação cotidiana*; tradução Paulo Polzonoff Junior. 1º ed., Rio de Janeiro: Elsevier, 2013.

PETRY, André. O berço do big data: a monumental abundância de dados, sua variedade e a velocidade com que trafegam no universo digital estão revolucionando a civilização. Veja, Editora Abril, edição 2321, ano 46, n. 20, 15 de maio de 2013, p. 71-81.

XEXÉO, Geraldo. *Big data: computação para uma sociedade conectada e digitalizada.* CiênciaHoje, n. 306, Agosto 2013, p. 19-23.