



XVII Encontro Nacional de Pesquisa em Ciência da Informação (XVII ENANCIB)

GT 8 – Informação e Tecnologia

PROPOSTA DE APLICAÇÃO DE TÉCNICAS DE BIG DATA EM COMÉRCIO ELETRÔNICO: IDENTIFICAÇÃO E UTILIZAÇÃO DO VALOR DA INFORMAÇÃO

A PROPOSAL FOR APPLYING BIG DATA TECHNIQUES IN E-COMMERCE SCENARIOS: IDENTIFYING AND USING THE INFORMATION VALUE

Mauricio Leonardo Plaza-Carvajal¹, Leonel Paes Furtado², Douglas Dyllon Jeronimo de Macedo³, Moisés Lima Dutra⁴

Modalidade da apresentação: Pôster

Resumo: O comércio eletrônico pode ser considerado como um potencializador de negócios e transações em ambientes Web. Contudo, como todo negócio, precisa de uma correta identificação do seu mercado consumidor, no intuito de reduzir incertezas de produção, fidelizar clientes e melhorar processos internos e o próprio negócio. Neste contexto, o valor da informação exerce um papel-chave, visto que é o insumo básico para a identificação do mercado consumidor e para a tomada de decisão dentro do ambiente organizacional. Big Data é um paradigma que permite o processamento de uma grande quantidade de dados e informações, de maneira não convencional, possibilitando a recuperação da informação a partir de fontes não alcançáveis e/ou não processáveis por humanos ou pelos sistemas tradicionais de armazenamento, coleta e análise de dados. Existe, portanto, uma enorme sinergia a ser explorada no que tange ao gerenciamento do valor da informação que é recuperada dentro deste paradigma. Este trabalho propõe a aplicação de técnicas de Big Data para identificar e utilizar o valor da informação em cenários de comércio eletrônico. Trata-se de uma pesquisa exploratória e aplicada, que utiliza abordagem quali-quantitativa, faz parte de um projeto de pesquisa e está sendo desenvolvida no escopo de duas dissertações de mestrado. Por fim, conclui-se que a aplicação de técnicas de Big Data em cenários de comércio eletrônico, por meio da descoberta do valor de informações não recuperadas pelos processos tradicionais, pode aportar melhorias significativas na correta identificação do mercado consumidor deste tipo de transação.

Palavras-chave: Comércio Eletrônico. Big Data. Valor da Informação.

¹ Universidade Federal de Santa Catarina

² Universidade Federal de Santa Catarina

³ Universidade Federal de Santa Catarina

⁴ Universidade Federal de Santa Catarina

Abstract: *E-commerce activities can boost business and transactions in Web environments. However, like every business, a correct identification of the consumer market is a requirement for reducing production uncertainties, retaining customers, and improving internal processes and the own business. In this context, the information value plays a key role, since it is the basic input for the identification of the consumer market and for decision making within the organizational environment. Big Data is a paradigm that allows processing a large amount of data and information in an unusual manner, allowing the retrieval of information from sources that are unreachable and/or not processable by humans or traditional storage, collection and data analysis systems. There are, therefore, huge opportunities to explore regarding to managing the information value that is retrieved within this paradigm. This work proposes the application of Big Data techniques to identify and use the information value in e-commerce scenarios. This is an exploratory and applied research, which uses a qualitative and quantitative approach, is part of an ongoing research project and is being developed in the scope of two master thesis dissertations. Finally, we can conclude that the application of Big Data techniques in e-commerce scenarios, through the discovery of the information value not retrieved by traditional processes, can provide significant improvements in the correct identification of the consumer market for this kind of business.*

Keywords: *E-Commerce. Big Data. Information Value.*

1 INTRODUÇÃO

O uso intensivo da Internet e dos dispositivos móveis promoveram um aumento expressivo no uso da informação, ao permitir que o usuário desta passasse a utilizá-la conforme a sua conveniência. Um desses objetivos de utilização da informação é o consumo por meio do comércio eletrônico. Entretanto, e apesar dos avanços tecnológicos, as vendas no ambiente Web ainda estão distantes dos resultados positivos almejados. Segundo Chiara (2015), entre 2014 e 2015, o Brasil perdeu 14 posições, saindo do 7º para o 21º lugar numa lista dos 30 principais países com maior atratividade no comércio eletrônico", despencando no ranking global do varejo online. Ramuski (2009) afirma que para qualquer organização independente de seu porte, é importante conhecer o comportamento de seus consumidores. Para Pelissaro (2009), entender o comportamento do consumidor e os elementos motivadores e inibidores que os levam a optar pelo varejo presencial ou online continuam sendo um dos desafios primordiais para o marketing das organizações.

Superar este desafio, fazendo uso dos inúmeros mecanismos tecnológicos à disposição no mercado aparenta ser relativamente fácil. Entretanto, o consumidor que existe por trás dos monitores, *tablets* e *smartphones* ainda é um total desconhecido por parte das organizações. Afinal que informações são valorizadas por este consumidor virtual? Qual é o impacto das comunidades virtuais no processo de compra? Que percepções são procuradas e valorizadas no momento de busca por informações que induzam ao consumo? Que palavras são as mais digitadas nos navegadores? Que produtos o consumidor “*curte*”? Existem sinônimos para um mesmo produto por faixa etária ou classe social? Segundo Cardoso *et al.* (2016), “para se destacar nesse mercado é necessário inovar e

chamar a atenção do consumidor de forma certa, e a Internet e os meios digitais são objetos centrais nessa mudança de comportamento”.

A partir da necessidade de diferenciação por meio de informação, observa-se oportunidade para aplicação de técnicas de Big Data, paradigma caracterizado por exigir tecnologias e técnicas capazes de extrair valor a partir de grandes quantidades de dados, com velocidade de processamento e *feedback* de resposta, preferencialmente em tempo-real, com aplicação nos mais diversos tipos de dados, sejam estes estruturados, semiestruturados ou não-estruturados. Desta forma, neste trabalho nos propusemos a responder à seguinte questão: **O gerenciamento do valor da informação em cenários de comércio eletrônico pode ser auxiliado por técnicas de Big Data?** Esta pesquisa objetiva aplicar de técnicas de Big Data para identificar e utilizar o valor da informação em cenários de comércio eletrônico.

2 O VALOR DA INFORMAÇÃO EM AMBIENTES BIG DATA

O termo Big Data está diretamente relacionado aos “Vs” que o caracterizam. Na literatura, pode-se encontrar até cerca de uma dezena e meia de características relacionadas ao conceito de Big Data, chamados de “Vs”, porque as características mais citadas pelos autores começam exatamente pela letra ‘V’: volume, velocidade, variedade, veracidade e valor. O volume se refere à quantidade de dados; a velocidade está relacionada à agilidade de acesso ou de processamento de dados; a variedade se refere aos diferentes de tipos e formato de dados; a veracidade diz respeito ao grau de confiança dos dados; e o valor, característica referente aos *insights* e novos conhecimentos, até então implícitos, por meio da extração de informações valiosas de grandes conjuntos de dados (MAYER-SCHÖNBERGER; CUKIER, 2013; LAURILA, 2012;). O ‘valor’ Big Data permite, portanto, a geração de informações úteis para os negócios, suporte para a tomada de decisões ou *insights* para novos produtos e serviços (BELLO-ORGAZ; JUNG; CAMACHO, 2016; MONOSTORI *et al.*, 2015; LI *et al.*, 2015; DAVENPORT, 2014; BABU; SASTRY, 2014).

Para Furtado, Dutra e Macedo (2016), quando o termo “valor” é referido em contextos Big Data, observa-se foco em tomada de decisão baseada em dados, tanto quando aplicadas em geração de *insights* para novos produtos e serviços, quanto em proposições de redução de custos. Outra proposição que caracteriza o “valor” em ambientes Big Data é a participação indireta de clientes, em interações *on line* (por exemplo, comentários em redes sociais), que geram dados e são utilizadas para análises e criação de valor. Um estudo apresentado por Buhalis e Foerste (2015 p.151), revela que algumas empresas baseadas em marketing buscam envolver dinamicamente os usuários para co-criação por meio de suas experiências e com base na otimização das condições externas e internas. Para Fawcett e Waller

(2014) as empresas estão monitorando o comportamento do cliente como nunca antes, utilizando-se destas análises para a construção de perfis, que são usados para desenvolver novos produtos, gerir portfólios de produtos, redesenhar exposição de produtos, promoções alvo e revisão de preços.

Conseqüentemente, a aplicação do paradigma Big Data resulta na transformação de conjuntos de dados armazenados em informações de valor. Este processo se dá por meio de análises de dados, com apoio de técnicas e tecnologias capazes de suportar o volume, a variedade e a velocidade que as formas tradicionais não suportam.

3 COMÉRCIO ELETRÔNICO E TEORIA DA CAUDA LONGA

O comércio eletrônico pode ser definido pelo uso da Internet para a compra e venda de produtos, serviços ou informações (ALBERTIN, 2010; DA SILVEIRA COELHO; OLIVEIRA; ALMERI, 2013). De acordo com Boone e Kurtz (2009), o comércio eletrônico é a principal função da Internet, sendo que aproximadamente quatro de cada cinco sites são, de alguma forma, dedicados a esta atividade. Contudo, para Barth (2004), apesar do investimento e pesquisa em estruturação tecnológica dos websites serem importantes, é fundamental conhecer o processo pelo qual os consumidores coletam informações acerca do que pretendem comprar. Desta forma, a obtenção de informações em tempo real sobre o consumidor poderá permitir ofertas personalizadas e preços que oscilem a depender da hora e do local onde se encontra este consumidor, tendo por base uma rastreabilidade em tempo-real.

Para Volkart e Fagundes (2015), ao contrário da loja física, que possui limitação de espaço, na qual os produtos vendidos precisam ser selecionados, as lojas virtuais possuem prateleiras infinitas para exporem os produtos que desejarem, sendo que, além dos mais vendidos, possuem também “espaço” para oferecer produtos mais segmentados. Ou seja, aumenta-se a participação coletiva das vendas dos produtos de nicho, criando uma cauda mais longa na distribuição das vendas (DIAS; DA ROSA; CASAGRANDE, 2016), conforme evidenciado na Figura 1.

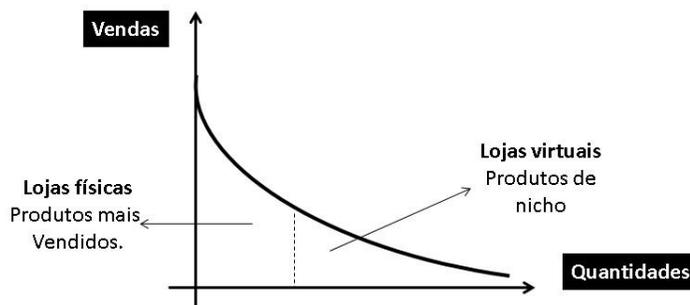


Figura 1. Gráfico da Cauda Longa.
Fonte: Os autores.

Anderson (2006) *apud* Dias, Da Rosa e Casagrande (2016) caracteriza a Cauda Longa como sendo as distribuições estatísticas que possuem uma população com alta frequência seguida por uma de baixa frequência de maneira gradativa e decrescente. Os eventos desta população no final da cauda têm probabilidade de ocorrência baixa. Entretanto, apresentam uma população maior ao longo da cauda. A Teoria da Cauda Longa permite efetuar uma correlação entre a variabilidade de produtos não disponíveis no varejo presencial, ou seja, produtos de nicho e o retorno financeiro das lojas virtuais que, em tese, deverão disponibilizar o máximo destes produtos de nicho em ambiente Web. Neste contexto, acredita-se que as lojas com foco em comércio eletrônico deveriam utilizar todas as ferramentas tecnológicas disponíveis no mercado para traçar um perfil do seu público consumidor e oferecer uma elevada quantidade de produtos e informações sob medida, permitindo que esta *customização* seja compartilhada nas suas redes sociais e comunidades virtuais ampliando o interesse pelos produtos de nicho entre os seus pares. Ou seja, segmentar e massificar no intuito de aumentar as vendas e oferecer – via comércio eletrônico – produtos não disponíveis nas lojas físicas.

4 ASPECTOS METODOLÓGICOS

Esta pesquisa é exploratória e aplicada, e utiliza uma abordagem quali-quantitativa. A abordagem quantitativa está presente no tratamento dos dados coletados pelas técnicas de Big Data, conforme evidenciado no Quadro 1; e a qualitativa está associada à avaliação do valor da informação a ser gerado a partir desses dados.

Camada	Técnicas	Ferramentas
Repositório de Dados e Metadados	NoSQL	Cassandra, MongoDB, CouchDB, HBASE, Hadoop HDFS, Redis, riak, etc.
	SQL	PostgreSQL, IBM DB2, MySQL, Oracle, Microsoft SQL Server, Microsoft Access, etc.
ETC (Extração, Transformação e Carga)	-	Informatics, Information Server (IBM), Oracle Warehouse Builder, Hadoop Mapreduce, Pentaho, Microsoft SSIS, etc.
Análise Descritiva Análise Prescritiva Análise Preditiva	Mineração de Dados	Linguagem R, Microsoft SSIS, Hadoop Mapreduce, Pentaho, Spark, rapidminer, WEKA, Orange, KNIME, NLTK, etc.
	Mineiração de Textos	
	Data Analytics	
	Métodos Estatísticos	
Apresentação	-	Ferramentas customizadas, Pentaho, Cognos, MicroStrategy, Oracle Discoverer, Brio, etc.

Quadro 1. Técnicas e Ferramentas Big Data para o Gerenciamento do Valor da Informação.
Fonte: Os autores.

Este estudo faz parte de um projeto de pesquisa em andamento, do qual parte dos objetivos estão sendo desenvolvidos no escopo de duas dissertações de mestrado do Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Santa Catarina (PGCIN/UFSC). Estão previstas as seguintes etapas: (i) levantamento bibliográfico sobre os conceitos associados; (ii) definição de técnicas e ferramentas Big Data; (iii) definição de um cenário de comércio eletrônico, no qual será feito um estudo de caso; (iv) aplicação das técnicas de Big Data a partir das ferramentas escolhidas; e (v) avaliação dos resultados.

5 CONSIDERAÇÕES PARCIAIS

A partir da literatura já levantada e consultada, e das ferramentas Big Data já analisadas, pode-se afirmar que a proposta de aplicação destas em cenários de comércio eletrônico – para gerenciamento do valor da informação – mostrou-se viável e potencialmente profícua. Com próximas etapas deste trabalho, serão desenvolvidas as etapas (ii) e (iii) descritas acima, que tomarão como base a revisão de literatura já feita.

AGRADECIMENTOS

Este trabalho foi parcialmente financiado pelo CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), no escopo da Chamada Universal MCTI/CNPq – 14/2013.

REFERÊNCIAS

ALBERTIN, Alberto Luiz. **Comércio Eletrônico: modelo, aspectos e contribuições de sua aplicação**. 6. Ed. São Paulo, Atlas, 2010.

ANDERSON, Chris. **A Cauda Longa: A nova dinâmica de marketing e vendas – como lucrar com a fragmentação dos mercados**. Rio de Janeiro: Elsevier, 2006.

BABU, MS Prasada; SASTRY, S. Hanumanth. **Big data and predictive analytics in ERP systems for automating decision making process**. In: Software Engineering and Service Science (ICSESS), 2014 5th IEEE International Conference on. IEEE, 2014. p. 259-262.

BARTH, Célia Amaral Souza Marengo. **O Comportamento do consumidor na compra no varejo eletrônico**. 2004. Tese de Doutorado - FGV EAESP, São Paulo, 2004.

BELLO-ORGAZ, Gema; JUNG, Jason J.; CAMACHO, David. **Social big data: Recent achievements**

and new challenges. Information Fusion, v. 28, p. 45-59, 2016.

BOONE, Louis E.; KURTZ, David L. **Marketing Contemporâneo.** 8ª ed. Rio de Janeiro: Livros Técnicos e Científicos, 2009.

BUHALIS, Dimitrios; FOERSTE, Marie. **SoCoMo marketing for travel and tourism: Empowering co-creation of value.** Journal of Destination Marketing & Management, v. 4, n. 3, p. 151-161, 2015.

CARDOSO, Moisés; DAROLT, Everton; MENDES-DA-SILVA, Paulo Eduardo. **O Comportamento do consumidor nas plataformas digitais: a pesquisa por informação no segmento da moda.** Razón y Palabra, v. 20, n. 93, p. 807-823, 2016.

CHIARA, Marcia de. **Brasil cai 14 posições no ranking global de comércio eletrônico.** Jornal O Estado de São Paulo. 07 Abr 2015. Disponível em: <<http://economia.estadao.com.br/noticias/geral,brasil-cai-14-posicoes-no-ranking-global-de-comercio-eletronico-imp-,1664945>>. Acesso em 21 jul 2016.

DA SILVEIRA COELHO, Lidiane; OLIVEIRA, Rafaela Carvalho; ALMÉRI, Tatiana Martins. **O crescimento do e-commerce e os problemas que o acompanham: a identificação da oportunidade de melhoria em uma rede de comércio eletrônico na visão do cliente.** Revista de Administração do UNISAL, v. 3, n. 3, 2013.

DAVENPORT, Thomas. **Big data at work: dispelling the myths, uncovering the opportunities.** Harvard Business Review Press, 2014.

DIAS, Ibsem Agrello; DA ROSA, Marcelo Medeiros; CASAGRANDE, Jacir Leonir. **Um referencial teórico sobre o mercado de e-commerce sob a perspectiva do fenômeno cauda longa.** Revista da UNIFEFE, v. 1, n. 15, p. 85-98, 2016.

FAWCETT, Stanley E.; WALLER, Matthew A. **Supply chain game changers—mega, nano, and virtual trends—and forces that impede supply chain design (ie, building a winning team).** Journal of Business Logistics, v. 35, n. 3, p. 157-164, 2014.

FURTADO, Leonel; DUTRA, Moises; MACEDO, Douglas. **Characterizing the Value Creation in Organizations That Implement Big Data Environments.** 2016. Disponível em: <<http://ebooks.iospress.nl/volumearticle/45478>>. Acesso em: 4 out. 2016

LAURILA, Juha K. *et al.* **The mobile data challenge: Big data for mobile computing research,**” in Proc. MDC Workshop, 2012.

MAYER-SCHÖNBERGER, V.; CUKIER, K. **Big Data: A Revolution that will transform how we live, work, and think.** New York: Houghton Mifflin Harcourt, 2013.

PELLISSARO, Renato Cortopassi. **Elementos motivadores e inibidores no comportamento de compra no varejo online.** 2009. Tese de Doutorado. Universidade de São Paulo.

RAMUSKI, C. L. Comportamento do consumidor digital. In LAS CASAS, A. L. (Org.). **Marketing**

Móvel: Tendências e oportunidades no marketing eletrônico. São Paulo: Saint Paul, 2009.

VOLKART, Viviana Klein; FAGUNDES, Dorneles Sita Fagundes Sita. **Estratégias de marketing como ferramenta para fidelizar clientes do comércio eletrônico: Estudo de caso de uma empresa varejista do município de Igrejinha/RS.** Revista de Administração de Empresas Eletrônica-RAEE, v. 1, n. 1, 2015.