

Automação e estratégia: do perfil dionisíaco ao cartesiano – um estudo sobre novos Modelos de Negócio na publicidade possibilitados pelo *Big Data*

RESUMO

O trabalho analisa as inovações proporcionadas pelas novas Tecnologias de Informação e Comunicação (TICs) no marketing e na publicidade, por meio da aplicação da Mineração de Dados (*Data Mining*) e do *Big Data*. Com o *Big Data*, as empresas buscam otimizar seus processos informacionais, com a prospecção refinada de dados, encontrando motivações, tendências e padrões de comportamento do consumidor. Assim, é possível incorporar a informação obtida em grandes bancos de dados, definir suas *buyer personas* (compradoras em potencial) e transformar estas informações em estratégias mais assertivas nas campanhas para o consumidor, criando ações que estimulem a conversão de *prospects* em clientes. O estudo justifica-se devido ao aumento do fluxo de informação e da necessidade de as empresas adquirirem novas e complexas estruturas de armazenamento de dados. A pesquisa bibliográfica abordou os conceitos de Mineração de Dados e *Big Data*. Procedeu-se uma pesquisa documental sobre empresas que utilizam o *Big Data*, seguida de entrevistas com publicitários, para entender como o *Big Data* tem sido utilizado para aumentar o engajamento com o cliente.

Palavras-chave: Inovação; Mineração de Dados; KDD; *Big Data*; Marketing; Publicidade.

1 Introdução

O mundo está em processo de transformação estrutural há mais de duas décadas. Segundo Castels (2005), este processo está associado a um novo paradigma tecnológico, baseado nas tecnologias de comunicação e informação. Estas mudanças atingiram o estilo de vida do consumidor, apresentando um novo contexto de competitividade às empresas. A integração econômica e a aceleração dos ciclos produtivos estão aumentando a distância entre a coleta dos dados, sua interpretação, a decisão do cliente e a implantação da mesma. As técnicas tradicionais de utilização de dados não são mais suficientes. É preciso que as agências de publicidade, bem como os institutos de pesquisa e os veículos consigam reduzir a distância entre a automação e a estratégia, caso não queiram ser engolidas pelas empresas de tecnologia. Nesse ambiente, será necessário mais do que criatividade e estratégia. Um sistema efetivo de *Big Data*, com um bom algoritmo, realimentado pelos profissionais de todas as agências, trará um resultado mais preciso para seus clientes.

O trabalho justifica-se, primeiramente, devido ao crescimento exponencial do volume de dados armazenados associados ao *Big Data* e à necessidade de as empresas aproveitarem as novas oportunidades, potencializando este fluxo de conhecimento e utilizando-o no discurso e posicionamento de suas marcas perante o consumidor. O marketing e a publicidade foram setores muito afetados pelas Tecnologias de Informação e Comunicação (TICs) e torna-se necessário entender como os profissionais têm se reinventado neste ambiente, utilizando-se dessas técnicas e ferramentas ou mesmo da filosofia do refinamento de dados para estreitar o relacionamento com o consumidor nas suas estratégias de campanha.

Acredita-se que, neste contexto, aliado às novas tecnologias, o profissional de comunicação pode integrar essa cadeia, apresentando-se como um analista da informação, humanizando e interpretando os dados quantitativos. Esse novo especialista não apenas vai conhecer a parte técnica do modelo, como também, vai investigar as variáveis mais importantes, utilizando tais informações no desenvolvimento de estratégias de campanhas. Sua função passa a ser a de triangular dados, tornando mais complexo o trabalho do profissional. Mesmo que campanhas publicitárias de massa possuam um bom retorno e ainda sejam muito utilizadas, infere-se que a possibilidade de oferecer campanhas personalizadas sobre produtos específicos para determinados usuários aumentaria as chances de sucesso de uma campanha. Os profissionais de publicidade continuam sendo fundamentais como analistas das informações, utilizando o processo criativo para desenvolver campanhas direcionadas a clientes específicos. A mensuração de dados e a criação de estratégias de comunicação eficazes podem beneficiar e apoiar diferentes frentes que buscam solucionar problemas com o apoio da população.

Problema de Pesquisa e Objetivos

Com as ferramentas de prospecção de dados, as informações se tornam mais valiosas para as empresas, uma vez que são filtrados os conteúdos mais relevantes e utilizadas tais informações de maneira estratégica, impactando, direta e positivamente, nos principais indicadores de desempenho - *Keys Performance Indicators* (KPI's) organizacionais. Neste contexto, surge a pergunta que norteou o estudo: como as técnicas/ferramentas de *Big Data* podem ser utilizadas pelas agências de publicidade na elaboração e produção de estratégias mais assertivas de comunicação?

O artigo teve, como Objetivo Geral, pesquisar a aplicabilidade das técnicas de mensuração de dados no desenvolvimento de estratégias de comunicação mais eficientes. Especificamente, buscou pesquisar os principais conceitos, tarefas e métodos da Mineração de Dados e *Big Data*; pesquisar as estratégias de utilização do *Big Data* pelas agências em sua comunicação; e analisar de que maneiras o *Big Data* pode contribuir na produção de conteúdo direcionado.

Tecnologias de dados em auxílio às estratégias de marketing

O advento da internet e das redes sociais disponibilizou aos consumidores o acesso a todo tipo de informação e uma capacidade de comunicação sem precedentes. Segundo Kucuk (2009), o que determina a tecnologia é a sociedade; portanto, as tecnologias são sensíveis aos efeitos dos usos sociais da mesma. Neste sentido, o ambiente *online* se tornou um instigante e inevitável verificador de grau científico, no momento em que a internet apoderou-se do ambiente individual e coletivo para criar uma nova realidade social e mercadológica. A este processo de interatividade e participação do consumidor dá-se o nome de *Web 2.0*, termo que refere-se a um conjunto de novas estratégias mercadológicas e a processos de comunicação mediados pelo computador. Neste novo contexto, desenvolvimentos tecnológicos aceleraram a velocidade de acesso e navegação. Surgiram *softwares* que permitiram a produção e partilha de conteúdos (DIONÍSIO *et al.*, 2009).

Este novo ambiente incorpora recursos de interconexão e compartilhamento, levando mais pessoas a utilizarem o serviço e criando mais arquivos e redes. Nas plataformas *online*, as

empresas têm acesso direto sobre a opinião dos consumidores, o que permite o ajustamento constante dos seus produtos e serviços às necessidades e expectativas do consumidor. As mídias sociais vêm transformando o ambiente virtual em um dispositivo interativo de tecnologia da comunicação. Elas possuem um papel importante de influenciar a escolha do cliente na seleção de produtos e serviços, baseados em *feedbacks* dos clientes que apareceram nos *weblogs*, *websites*, placas *online* e outros tipos de conteúdo gerados pelo usuário (RAMANI, 2008).

Prahalad e Ramaswamy (2004) pontuam algumas mudanças comportamentais ocorridas neste ambiente relacional, em relação ao consumidor. Entre elas, cita-se o acesso variado e veloz à informação; a visão global; as redes de contatos; a experimentação e o ativismo. Para Rozwell *et al.* (2010), as redes sociais permitem às marcas constituir comunidades de consumidores, que podem identificar as marcas pelas quais têm interesse, conectarem-se a elas, seguir o que elas têm a dizer e partilhar essa informação para a sua rede de contatos. Prahalad e Ramaswamy (2004) afirmam que é preciso que as empresas repensem as suas formas tradicionais de fazer marketing e que cheguem mais perto dos seus consumidores. As empresas vêm identificando e aproveitando o potencial de negócios, sobretudo do ambiente virtual, buscando um maior e melhor relacionamento com o seu público, diminuindo possíveis reações negativas do mesmo e otimizando a promoção e gestão de suas marcas. Assim atua o *Big Data*, a partir do conhecimento do cliente, do refinamento acerca de suas preferências, da triangulação destas informações, que podem ser estruturadas ou não, conforme será explicado adiante.

Mineração de Dados – Data Mining

O termo mineração de dados é definido, de acordo com Berry e Linoff (1997, p.5), como a “exploração e a análise, por meio automático ou semi-automático, de grandes quantidades de dados, a fim de descobrir padrões e regras significativos”. Goebel e Greenwald (1999) reforçam o conceito acima, definindo-o a partir da extração de padrões ou modelos de dados observados. Goebel e Greenwald (1999) explicam que as técnicas de mineração são aplicadas em sistemas de descoberta de conhecimento em banco de dados, buscando extrair informações estratégicas escondidas em grandes bancos de dados, por meio da pesquisa dessas informações e da determinação de padrões, classificações e associações entre elas. Eles ressaltam que a mineração é considerada parte do chamado processo de Descoberta de Conhecimento em Banco de Dados, conhecido pela sigla KDD - *Knowledge Discovery in Databases*, utilizado para representar o processo de tornar dados de baixo nível em conhecimento de alto nível.

Entretanto, na literatura sobre o assunto, alguns autores divergem quanto à significação dos termos KDD e *Data Mining*. Para Wang (2005) e Han *et al.* (2006), eles são sinônimos. Já Fayyad (1996) e Cios *et al.* (2007) definem o KDD como sendo o todo o processo de descoberta de conhecimento, enquanto conceituam a Mineração de Dados como uma das atividades do processo. Porém, todos concordam que o processo deve ser iterativo, interativo e dividido em fases. A mineração de Dados é uma área multidisciplinar, com destaque para a Estatística, Banco de Dados e Aprendizado de Máquina. Na visão estatística, Hand *et al.* (2001) a definem como a análise de grandes conjuntos de dados, visando a encontrar

relacionamentos inesperados, resumindo os dados para que eles sejam úteis e compreensíveis. Na perspectiva de banco de dados, Cabena *et al.* (1998) conceituam como um campo interdisciplinar que junta técnicas de máquinas de conhecimentos, reconhecimento de padrões, estatísticas, banco de dados e visualização, para extrair informações de grandes bases de dados. Na visão do aprendizado de máquina, Fayyad *et al.* (1996) definem como um passo na Descoberta de Conhecimento, que consiste na realização da análise dos dados e na aplicação de algoritmos de descoberta que, sob certas limitações computacionais, produzem um conjunto de padrões de certos dados.

Para Fayyad *et al.* (1996), o modelo tradicional de transformação de dados em informação (conhecimento) consiste em um processamento manual de todas essas informações por especialistas que, então, produzem relatórios que deverão ser analisados. Na grande maioria das situações, devido ao grande volume de dados, esse processo manual torna-se impraticável. Ainda segundo Fayyad *et al.* (1996), o KDD (*Knowledge Discovery in Databases* ou Descoberta de Conhecimento nas Bases de Dados) busca solucionar o problema causado pela era da informação: a sobrecarga de dados. De acordo com os autores, existem vários métodos de Mineração. Entre os mais relevantes para o KDD, estão os de: Classificação, Modelos de Relacionamento entre Variáveis, Análise de Agrupamento, Sumarização, Modelo de Dependência, Regras de Associação e Análise de Séries Temporais.

A Classificação associa um item a uma ou várias classes categóricas pré-definidas, utilizando a técnica de Análise Discriminante. Os Modelos de Relacionamento associam um item a uma ou mais variáveis de predição de valores reais, consideradas variáveis independentes ou exploratórias. A Análise de Agrupamento (*Cluster*) associa um item a uma ou várias classes categóricas, determinando as classes pelos dados, independentemente da classificação pré-definida. Já a Sumarização determina uma descrição com dispersão reduzida para um subconjunto no pré-processamento dos dados. O Modelo de Dependência, que existe nos níveis estruturado e quantitativo, descreve dependências entre variáveis. O nível estruturado especifica, geralmente em forma de gráfico, quais variáveis são, localmente, dependentes. O nível quantitativo especifica o grau de dependência, usando escalas numéricas. As Regras de Associação determinam relações entre campos de um banco de dados, contribuindo para a tomada de decisão, enquanto a Análise de Séries Temporais determina características sequenciais, como dados com dependência no tempo. Seu objetivo é modelar o estado do processo, extraindo e registrando desvios e tendências no tempo.

Percebe-se, portanto, que os principais objetivos da mineração de dados visam à descoberta de relacionamentos entre dados e ao fornecimento de subsídios para que seja feita uma previsão de tendências futuras baseadas no passado. Os resultados podem ser usados no gerenciamento de informação, no processamento de pedidos de informação, na tomada de decisão, no controle de processo, entre outros e pode, conforme afirma Groth (1998), como processo de verificação e processo de descoberta. No primeiro caso, o usuário sugere uma hipótese acerca da relação entre os dados e tenta prová-la. Para tanto, são aplicadas técnicas como análises multidimensional e estatística sobre um banco de dados com informações progressas. Já no processo de descoberta, não há suposição antecipada. São utilizadas técnicas como a Descoberta de Regras de Associação, Árvores de Decisão, Algoritmos Genéticos e Redes Neurais Artificiais.

As técnicas de mineração podem ser utilizadas em várias áreas, como pontuam Viveros *et al.* (1996); Mannila (1997); Cratochvil (1999): no ambiente financeiro, na detecção de padrões sutis de fraudes, detectando, por exemplo, modelos preditores do bom perfil de um cliente; na Medicina, buscando padrões de novas doenças e identificando terapias eficazes; na Ciência, pesquisando dados genéticos; no Controle de processos e de qualidade, auxiliando no planejamento estratégico; em Instituições governamentais, buscando padrões para melhorar as coletas de taxas; no ambiente *Web*, denominado *Social Big Data*ⁱ; e no Marketing, utilizado como foco deste estudo, para descobrir preferências do consumidor e padrões de compra, aplicando-os no marketing direto de produtos e em ofertas promocionais.

No que se refere ao marketing, Bramer (2007) afirma que o atual volume de dados gerados pelos clientes pode ser utilizado em inúmeros segmentos de mercado, buscando criar eficiência no relacionamento da organização com seus clientes. Witten *et al.* (2005), Bramer (2007) e Olson *et al.* (2008) reforçam a aplicabilidade de tais técnicas na retenção e captação de clientes e no gerenciamento do relacionamento das empresas com seus clientes, sobretudo em tempos de redes sociais.

A escolha das técnicas de mineração dependerá da tarefa a ser executada, bem como dos dados disponíveis para análise. Goebel e Greenwald (1999) afirmam que muitas ferramentas disponíveis são genéricas da Inteligência Artificial ou Estatística. Elas, geralmente, operam de forma separada à fonte de dados, requerendo quantidade significativa de tempo gasto com exportação e importação de dados, pré e pós-processamento e transformação de dados. A mineração é mais usada com dados mais estruturados, como planilhas, bancos de dados relacionais e dimensionais. Já o *Big Data* trabalha com dados mais complexos e não-estruturados, que também devem estar registrados em um banco de dados. Como as escalas e os tipos de dados são diferentes, os períodos de análise e os resultados também diferem. Enquanto o *Data Mining* se refere a um processo mais pontual, que gera relatórios, apontando para questões específicas, o *Big Data* é uma análise feita de forma contínua por períodos maiores. Por esse motivo, o *Big Data* pode ser usado para fazer previsões e indicar caminhos para mudanças estratégicas na forma de gestão.

Fundamentos do *Big Data* – Volume, Velocidade, Variedade

O termo *Big Data*, de acordo com Fan e Bifet (2012), surgiu pela primeira vez em 1998, na apresentação *Big Data* e a Próxima Onda de *InfraStress*, feita pelo cientista de dados da Silicon Graphics, John Mashey (SGI). O primeiro livro citando as palavras, ainda de acordo com os autores, teria surgido também em 1998, por Weiss e Indrukya, em uma abordagem sobre mineração de dados, enquanto o primeiro trabalho acadêmico com o termo dataria do ano 2000, em um artigo de Dieboldⁱⁱ. Já para Arrigoni, segundo informa Rogenski (2013), o termo surgiu na década de 1990, na NASA, para descrever os conjuntos de dados complexos que dificultavam os limites de acesso computacional, processamento, análise, pesquisas e armazenamento.

Segundo Nesello e Fachinelli (2014, p.26), o termo teve seu marco inicial em 2010 com uma publicaçãoⁱⁱⁱ na revista britânica *The Economist*, vinculando-o à terceira Revolução Industrial

ocorrida em meados do século vinte, que teve como principais impactos a globalização, o uso da informática e o desenvolvimento de novas tecnologias. Entretanto, de acordo com estes autores, o fenômeno se iniciou na década anterior, antes do surgimento do termo e de suas concepções, “com a implementação dos sistemas de informação”. Para o cientista Albert-László Barabási (2002), o termo pode ser chamado de *Big Data*, Computação Social, Dinâmica Humana ou Ciência das Redes. O importante é dar sentido à imensa quantidade de informações coletadas sobre o indivíduo.

Barabási (2012) ressalta que a Ciência das Redes, assim como o *Open Data* (Dados Abertos), a Abordagem Sistêmica e o *Design Thinking* são as lentes filosóficas, científicas, teóricas e metodológicas pelas quais o *Big Data* deve ser pesquisado, elaborando normas, códigos de ética, modelos de negócio, de processos e indicadores. Os quatro pilares, de caráter humanista, são importantes em *Big Data*, porque, apesar da tecnologia, a analítica digital não suplanta a capacidade humana de gerar *insights*, discernir e decidir. No ano 2000, após realizar várias experiências com grande quantidade de dados digitais - documentos e dados de navegação, Barabási publicou o que pode ser reconhecido como o surgimento da Ciência das Redes, já estudadas no século dezoito pela Teoria dos Grafos^{iv}. Segundo Barabási (2002), para os matemáticos da época, as redes (estruturas formadas por nós e ligações) possuíam propriedades fixas. No século vinte, cientistas observaram outras redes não explicadas matematicamente. Como desconheciam as propriedades que governavam essas redes, imaginaram que elas se formavam e se comportavam de modo randômico. Entre as redes consideradas randômicas, cita-se a sociedade, a célula, o organismo, o ecossistema e as demais redes naturais e sociais.

Apesar das várias definições sobre o seu surgimento, todas apontam que o conceito surgiu há cerca de vinte anos, para caracterizar a geração, processamento e uso de grandes quantidades de dados e informações. O *Big Data* pode ser definido, portanto, como o termo em Tecnologia da Informação compreendido pelo conjunto de técnicas capazes de analisar grandes quantidades de dados que precisam ser processados e armazenados, por meio de ferramentas especiais que comportem este volume. A ideia é que tais dados sejam encontrados, extraídos, organizados e transformados em informações que possibilitam uma análise ampla e em tempo hábil (CETAX, 2019).

De acordo com Mayer-Schönberger e Cukier (2013), o volume de dados disponível mais do que dobra a cada dois anos e os algoritmos também se aperfeiçoam rapidamente. O *Facebook*, por exemplo, armazena, acessa e analisa mais de 50 *petabytes* (unidade de medida maior que um *terabyte*) de informações geradas pelos usuários. No *Youtube*, a cada minuto, são feitos *uploads* (descarregamentos) de 48 horas de vídeos, sendo impossível que um usuário consiga assistir a todas as mídias disponíveis. No *Twitter*, por dia, são enviados mais de quinhentos milhões de mensagens, enquanto o *Google* processa, diariamente, mais de três bilhões de pesquisas no mundo, sendo 15% inéditas. A ferramenta de busca rastreia 20 bilhões de *sites* por dia, abarcando as informações estruturadas, que possuem algum padrão ou formato que pode ser usado na leitura e extração dos dados, como dados de bancos de dados, arquivos texto (txt, XML); semi-estruturados (acompanham padrões heterogêneos e são mais difíceis de serem identificados pois podem seguir diversos padrões); e não-estruturadas, sem formato padronizado para leitura (arquivos Word, páginas de Internet/Intranet, vídeos, áudios),

armazenando 100 *petabytes* de informação. Ainda segundo Mayer-Schönberger e Cukier (2013), as técnicas de análise, até então acessíveis apenas aos laboratórios de pesquisa, grandes empresas e agências de espionagem, estão democratizadas e o *Big Data* tem se mostrado como parte da solução para problemas globais climáticos, de saúde, boa governança e do desenvolvimento econômico, tornando as redes complexas mais compreensíveis, previsíveis e controláveis.

O conceito de *Big Data* (CETAX, 2019) pode ser subdividido em cinco categorias: Volume, Velocidade, Variedade, Veracidade e Valor, sendo os três primeiros considerados os mais relevantes. O primeiro desafio é gerenciar o Volume de dados disponíveis, gerados no cotidiano, por meio da troca de mensagens eletrônicas (*e-mails* e redes sociais), transações bancárias e contatos por linhas telefônicas, por exemplo. O tráfego de dados em nuvem deve atingir 19,5 *zettabytes* (ZB) ao ano em 2021 contra 6,8 ZB em 2016, segundo a sétima versão do relatório *Global Cloud Index*, da Cisco, que avalia a movimentação em *Data Center* e tendências para a utilização em nuvem (TECMUNDO, 2018). A Velocidade é outro atributo importante, se mostrando como fator de vantagem competitiva, a partir da análise dos dados atuais, gerados em tempo mais próximo do real. No que se refere à Variedade, ressalta-se que os dados estão dispersos nas redes sociais, de forma estruturada; semi-estruturada e não-estruturada. Outro fator relevante é a Veracidade, uma vez que muitos gestores não confiam nos dados que recebem atualmente. Por fim, cita-se o Valor, pois quanto maior a riqueza dos dados, mais importante é fazer as perguntas certas nos processos de análise.

Em entrevista a Rogenski (2013), Arrigoni afirma que trabalhar com *Big Data* demanda conhecimentos em computação científica, estatística, probabilidade, *machine learning*, *text mining* (com processamento de linguagem natural), teoria dos grafos, entre outros. O desafio das empresas está em integrar os dados das estratégias de *Business Intelligence* (BI) com as de *Big Data*, analisando, separadamente, as informações e medindo outros dados não estruturados. Sobre as desvantagens de utilização do *Big Data*, cita-se o alto custo, a demora no retorno do investimento (ROI), a busca por uma demanda para implantar a tecnologia, mesmo sem ter um problema claro a ser resolvido, além o fator privacidade, visto que os sistemas têm coletado dados sobre todas as movimentações, perfis sociais etc (CETAX, 2019).

Há muitas soluções disponíveis no mercado para que as empresas utilizem a Mineração de Dados e o *Big Data*. Porém, o custo da licença de uma plataforma que processe grandes quantidades de informações não-facilmente interpretadas, manualmente, pode ser alto (TREE INTELLIGENCE, 2014) para pequenas e médias empresas, fora o custo da contratação de uma consultoria. Portanto, faz-se necessário conhecer e estudar as ferramentas gratuitas que possam substituir os serviços pagos, resultando em tomadas de decisões mais precisas. Existem várias tecnologias de *Big Data*. Segundo Tarifa (2014), entre as que não têm custo, está o *Hadoop*, plataforma de código aberto (*Open Source*), que roda sobre o servidor Apache no sistema Linux (conhecido por ser gratuito e por possuir muitas empresas trabalhando para adequar as funcionalidades para ter um trabalho mais profissional). O *Hadoop* precisa de *APIs* (*Application Programming Interfaces*), programas que interagem pela troca de informação com o núcleo principal. O *framework* do *Hadoop* é composto por dois componentes principais: o armazenamento e o processamento. O primeiro é o *Hadoop Distributed File*

System (HDFS), sistema de arquivos distribuído, projetado para armazenar arquivos muito grandes. Como explica Fonte (2013), ele não deve ser utilizado para aplicações que precisem de acesso rápido a um arquivo e sim, para aplicações em que é necessário ler uma quantidade muito grande de dados. O *HDFS* funciona em uma estrutura que necessita de dois tipos de nós de armazenamento: um nó mestre, o *Namenode* e os nós servos, denominados *Datanodes*. O *Namenode* controla o sistema de arquivos, mantendo metadados para todos os arquivos e diretórios da árvore de diretórios e arquivos do sistema. Os *Datanodes* armazenam os blocos e enviam relatórios ao *Namenode*, com as listas dos blocos que estão armazenando.

Outro modelo (*framework*) que permite realizar processamento em paralelo em um *cluster* ou agrupamento *Hadoop*, é o *MapReduce*, paradigma de programação introduzido pelo Google para processar esse grande conjunto de informações. Ele consiste em um serviço *master* denominado *JobTracker*, que gerencia a execução dos serviços, distribuindo o trabalho para os nós chamados *TaskTrackers*. (serviços *slaves*). Os *TaskTrackers* criam processos por tarefa, para evitar falhas; eles se reportam, regularmente, ao *JobTracker*, especificando que estão funcionais e trabalhando. Caso contrário, seu trabalho é enviado para outro *TaskTracker*. De acordo com Pereira (2010), assim como o *MapReduce*, o *framework* Apache Mahout possui algoritmos focados em classificação, agrupamento e recomendação. A biblioteca de aprendizagem de máquina (*Machine Learning*) possui classes que implementam algoritmos de mineração de dados e classes auxiliares.

Aplicabilidade do *Big Data Marketing* nas estratégias de mercado e na publicidade

O ambiente está mais complexo, devido ao grande número de dados existentes, tornando mais difícil a tarefa de selecionar o que é mais relevante. Paradoxalmente, quanto mais se coleta dados, a partir de experiências ou eventos pregressos registrados em bancos de dados, mais precisa e inteligente tende a se tornar a sua interpretação, uma vez que certos hábitos vão se tornando mais claros. Pesquisas mostram que a ação cotidiana do indivíduo caracteriza-se pelo hábito da repetição. Estudos conduzidos na Universidade Duke^v apontaram que grande parte dos processos de tomada de decisão não é consciente. Nas pesquisas, os participantes relataram ações heterogêneas que integram a sua rotina, todas variando a partir da força do hábito, como ler jornal, fazer exercícios e comer *fast food*. Estudos diários de amostragem por experiência com amostras de estudantes e comunidades reforçam o resultado acima, apontando que 45% dos comportamentos cotidianos tendem a se repetir no mesmo local quase todos os dias.

É nesse contexto em que aplica o *Big Data*. Além de fornecer dados, torna-se possível compreender as conexões existentes e selecionar as mais relevantes às empresas. Quanto mais conhecimento uma empresa tiver do cliente, mais personalizada será sua comunicação e mais forte tende a ser o relacionamento dele com a marca. O Marketing não é tecnologia. A Tecnologia de Marketing é apenas um meio para que se crie conexão com os clientes, gerando experiências, ações e engajamento em todos os pontos de contato com as empresas. Um benefício do *Big Data* no Marketing está no custo reduzido, uma vez que não há só ferramentas caras e sofisticadas de análise digital, conforme citado aqui. Um gestor pode pedir a seus funcionários que mantenham apenas em dia os registros dos dados, arquivando informações de projetos e clientes. Apesar de acumular menos informações que o digital, essa

ação já se qualifica como parte do *Big Data* (MUNDO DO MARKETING, 2016). O *Big Data* no marketing, Marketing por Dados ou *Data Driven Marketing* auxilia as empresas a entenderem como pensa o consumidor, enviando o conteúdo certo para o cliente certo no momento certo. Ele consegue criar equivalências entre os dados não-estruturados, promovendo uma conexão dos dados *offline* e *online*. Com o monitoramento, análise e processamento dos dados, é possível segmentar seu público, identificar oportunidades de mercado e coletar dados personalizados. É necessário fazer um diagnóstico estratégico, para conhecer esses públicos, transformando dados em informações, elaborando projetos, analisando os investimentos em marketing e comunicação e planejando campanhas.

A comunicação de uma empresa precisa ser personalizada e completa. Nesse cenário, em que a comunicação entre anunciantes e consumidores vem mudando como resultado do *Big Data*, a Publicidade precisa encontrar uma forma de responder à velocidade do processo, integrando dados *offline*, *online* e *mobile*, criando perfis de comportamento minuciosos (SINGER, 2012^{vi}) e gerando *insights* no processo, como compra de mídia, análise de tendências ou análise do resultado de campanhas. Abreu e Nicolau (2017) reforçam que as agências de publicidade estão analisando seus bancos de dados, minerando dados de centenas de bilhões de consumidores, conectando, por exemplo, informações sobre hábitos de compra dos clientes com dados demográficos, otimizando as demandas. O avanço tecnológico e a grande quantidade de canais e plataformas gerando informações tem gerado muitos desafios ao mercado publicitário, que está, sob certos aspectos, saturado. De acordo com Covaleski (2016), a necessidade de alcançar e obter uma reação do público, em um universo amplo, bem como à crescente rejeição à publicidade, devido ao seu hábito intrusivo e interruptivo, torna mais complicada e minuciosa a tarefa das empresas de decidirem onde e como alocar os orçamentos de publicidade. Nesse contexto de interseção entre os dados e o comportamento humano, as estratégias do discurso publicitário precisam adquirir um tom motivacional, racionalizando, classificando e promovendo algum tipo de adesão ao consenso social vigente.

Os profissionais de marketing também devem estar atentos às novas habilidades demandadas pelo ambiente digital, exercitando sua capacidade analítica para interpretar os dados e orientar os criativos para desenvolver ações. Um gestor de Marketing precisa ter em sua equipe um Cientista de Dados, com conhecimento em banco de dados e Estatística e um programador que tenha visão crítica sobre o uso para que um trabalho foi executado. Os que já estão no mercado devem também aprimorar-se em tecnologia, detectando as peças, *sites* e publicidade de um anunciante que impactaram um consumidor e o levaram a uma compra ou a um cadastro, atingindo o objetivo da campanha (MUNDODOMARKETING, 2016). A seguir, serão citados casos de empresas de segmentos distintos que têm utilizado o *Big Data* em suas estratégias de aproximação com o consumidor.

Empresas que utilizam o Big Data em suas estratégias

As empresas, de maneira geral, independentemente do porte, têm recorrido à tecnologia para aprimorar seus processos e evitar a obsolescência. Neste ambiente, aquelas que são orientadas a dados, ao aumento da produtividade, à experiência do consumidor, à Inteligência Artificial, à Internet das Coisas (IoT) - interconexão digital de objetos cotidianos com a internet - e ao *Big Data* tendem a criar um diferencial competitivo em relação à concorrência.

O Festival de Criatividade *Cannes Lions*, premiação mais importante da publicidade mundial, criou, em 2015, a categoria *Creative Data*, buscando premiar as campanhas que utilizam o *Big Data* como base para a sua criação. A campanha vencedora do *Grand Prix* de 2016, *The Next Rembrandt*, criada pela agência de publicidade J. Walter Thompson (JWT) holandesa, em parceria com a Microsoft para o banco ING, *produziu*, com a ajuda de inteligência artificial, dados sobre a obra do pintor Rembrandt, linguagem e técnica, permitindo que uma máquina pintasse um quadro no estilo do artista falecido. Com a ajuda de engenheiros e especialistas em arte, foram escaneadas 364 pinturas em alta resolução. Centenas de milhares de fragmentos foram analisados por algoritmos, bem como os rostos dos retratos feitos por Rembrandt. Uma impressora 3D pintou o quadro (MINDMINERS, 2017). O serviço de *streaming* Netflix também explora o conceito e técnicas de *Big Data* para prever hábitos e criar métricas de comportamento do usuário. A empresa consegue diagnosticar quantas vezes os usuários pausam um vídeo, se assistem a episódios sequencialmente e quando eles se tornam fãs de uma série (MINDMINERS, 2017). A Vivo, Claro, Renner e Itaú Unibanco também utilizam *Big Data*. A Renner, por exemplo, monitora as redes sociais de seus clientes para compreender qual caminho deve seguir na escolha das peças de cada campanha. O Pão de Açúcar utiliza sistemas que identificam os produtos mais comprados e perfil dos consumidores. O grupo implementou um programa de recompensa, com descontos exclusivos de produtos específicos (SONDA, 2017).

A American Express (AMEX), uma das mais importantes empresas americanas de serviços financeiros, passou a utilizar a análise de dados e o *Machine Learning* para detectar padrões que correspondem a transações fraudulentas. Os algoritmos aprendem o padrão de consumo do usuário, que é notificado, assim como a empresa, quando há uma transação incomum. A AMEX também oferece um aplicativo que analisa os dados de compras anteriores e oferece cupons e ofertas. A Amazon, empresa transnacional de comércio americana, criou algoritmos para levar ofertas personalizadas, prevendo o comportamento do consumidor, com o uso de *Machine Learning* e computação em nuvem (*Cloud Computing*). A empresa de aviação Delta Airlines criou um sistema que permite que o passageiro acompanhe sua bagagem (a empresa possui mais de 130 milhões de bagagens despachadas por ano). A Shell, uma das maiores petrolíferas do mundo, utiliza o *Big Data* para reduzir seus gastos de operação e minimizar riscos ambientais. A empresa monitora as áreas mais propensas de perfuração para extração de petróleo. Ela analisa as ondas sísmicas de baixa frequência abaixo da superfície da Terra, conseguindo prever o tamanho provável dos recursos de petróleo e gás (RESULTYS, 2019).

De acordo com Miquelino (2018), a Starbucks, empresa de cafés, faz análises demográficas para criar planos de Mercado e identificar pontos estratégicos para abrir lojas. Com o *Big Data*, a empresa conseguiu identificar 3 milhões dos seus 6 milhões de clientes. Por meio de aplicativos e dispositivos vestíveis (*wearables*), a marca esportiva Nike reúne dados sobre o comportamento esportiva, como distância percorrida, velocidade e rotas para treinos. Já a marca Danone nos Estados Unidos conseguiu identificar o tempo e a chegada dos produtos nas prateleiras, evitando o vencimento dos produtos perto da data de validade. A J.P Morgan, empresa do mercado financeiro, prevê tendências para compra ou venda de ações com o *Big Data Analytics*, enquanto a rede social de compartilhamento de imagens, Pinterest,

utiliza a inteligência artificial para criar um *ranking* mais rápido e conteúdo de acordo com a localização dos usuários.

Metodologia

O trabalho iniciou-se por meio de uma revisão bibliográfica (DUARTE E BARROS, 2017), apresentando os conceitos de Gestão do Conhecimento, Mineração de Dados, *Big Data*, entre outros, a partir da leitura e consulta em livros, artigos, revistas e pesquisa na internet, principalmente em *sites* especializados, uma vez que a literatura sobre o tema é, relativamente, recente. Foram referenciados estudiosos de *Big Data*, como Viktor Mayer-Schönberger, professor de governança e regulamentação da Internet na Universidade de Oxford e Kenneth Cukier, editor de dados do jornal The Economist; e em Mineração de Dados, como Fayyad (1996).

O estudo, de caráter exploratório, procedeu-se por meio de uma Pesquisa Documental, a partir da identificação de empresas que utilizam o Big Data em suas estratégias mercadológicas, aplicando-as em suas campanhas direcionadas aos clientes/consumidores. Constam empresas de médio e grande portes, de variados setores e públicos. A ideia foi mostrar que não apenas as grandes empresas podem refinar seus dados, mas que empresas de outros portes também podem utilizar *softwares* que lhes deem padrões acerca dos seus clientes/consumidores, bem como promover análises de dados não-estruturados, buscando conhecer melhor o comportamento de seus consumidores. A segunda fase do trabalho consta de uma pesquisa qualitativa (DUARTE E BARROS, 2017), por meio da realização de três entrevistas em profundidade com gestores de empresas de portes distintos, bem como de publicitários atuantes em setores estratégicos organizacionais, buscando compreender como a técnica de monitoramento de dados pode ser relevante no desenvolvimento de campanhas de fortalecimento das marcas e de relacionamento com o cliente/consumidor. Foram pontuadas as principais técnicas e ferramentas utilizadas por agências de publicidade/empresas atuantes em marketing e comunicação, para obterem maior engajamento e melhor relacionamento junto ao cliente/consumidor, no mercado contemporâneo.

Resultados da pesquisa qualitativa

As três entrevistas foram realizadas, pessoalmente, no mês de março de 2018, presencialmente e por meio de video-conferência. A amostra, por conveniência, foi obtida a partir do contato com publicitários atuantes neste mercado, donos de empresas de publicidade e marketing online e *offline*, assim como profissionais com função de gerência em núcleos digitais de agências de grande porte de Belo Horizonte. Em um primeiro momento, os entrevistados foram indagados acerca do que consistiria o *Big Data*, falando sobre sua importância na elaboração de estratégias para as empresas. Uma das entrevistadas, gerente do núcleo digital de uma agência de médio porte da capital mineira, definiu o termo como sendo a grande quantidade de dados existentes, atualmente, que consegue-se capturar, gerados a partir da interação das pessoas com a *web* e com as redes sociais, sobretudo. Em sua opinião, com o *Big Data*, é possível que as empresas – independentemente do porte ou formato (*offline* ou *online*) – possam obter *insights* para elaborar suas estratégias de trabalho. Ainda de acordo com ela, mais relevante do que a quantidade é a interpretação dos dados, buscando

formas de segmentar o que é importante, bem como a sua aplicabilidade em alguma campanha para os clientes.

Já de acordo com o dono de uma agência de comunicação de pequeno porte sediada em Belo Horizonte, nos dias atuais, com o surgimento das TIC's, é possível conhecer de forma mais detalhada os hábitos de consumo das pessoas. “Podemos rastreá-las pela internet, colhendo assim um conjunto de dados, que é o que chamamos de *Big Data*”. Segundo ele, isso o permite entender e acompanhar o comportamento do consumidor, até em tempo real. “Sabemos quais *sites* ele visita, com que frequência, quanto tempo fica em cada página, onde ele clica, o que o atrai, o que o repele”. Para ele, isto é fundamental para as empresas, que podem ajustar melhor sua oferta e estreitar o relacionamento com o seu público, oferecendo o que ele quer, onde ele precisa e na hora exata. Outro entrevistado, dono de uma agência de marketing também sediada na capital mineira, define o termo como um processo de análise de todos os dados gerados atualmente, a partir das incursões feitas pelas pessoas em dispositivos moveis. Ele completa que estes dados, isoladamente, não representam muito, sendo apenas “um grande quebra cabeça”. O que torna tais dados úteis é a possibilidade de minerar estes bilhões de códigos, identificando os que “conversam” com determinadas pessoas.

Os entrevistados também foram perguntados sobre como utilizam a análise de dados (como mecanismo de tomada de decisão para o negócio, *remarketing*, relacionamento com o cliente, estudo do perfil do consumidor etc). De acordo com a gerente, na agência em que atua, os dados são utilizados durante toda uma campanha ou em uma ação pontual para um cliente, buscando elaborar o planejamento e os conteúdos, assim como definir as mídias e segmentações. Ainda segundo ela, a agência acompanha os resultados das ações implementadas, promovendo mudanças ao longo do trabalho, se necessário. O dono da agência de publicidade afirma que acompanha a rede do social dos clientes, monitorando a satisfação dos consumidores, identificando pontos críticos, mensurando o interesse pelos produtos e serviços e interagindo com eles por uma série de ferramentas (ele não mencionou quais). Desta forma, é possível orientar os clientes, visando ao aprimoramento na gestão dos produtos/serviços, contribuindo para o aumento de suas vendas, a melhoria do atendimento e da percepção de imagem.

Já o dono da agência de marketing diz não utilizar a análise de dados em grande escala, por acreditar que esta tecnologia, por mais que seja interessante, ainda se mostre limitada para agências menores, considerando o alto custo. Ele acredita que a análise seja feita, geralmente, por empresas e agências maiores, com atuação em nível nacional. Ele afirma que cria peças de divulgação tomando o *briefing* como referência, que é menos relacionado a dados e mais focado em comportamento e em pesquisas feitas pessoalmente.

Os entrevistados também foram perguntados sobre quais ferramentas de análise de dados utilizam no trabalho e quais consideram mais efetivas. A gerente afirma que, no núcleo de redes sociais e de mídia *online*, são utilizados o *Google Analytics*, *Facebook Insights*, *Twitter Analytics*, *LinkedIn Analytics* e *RD Station*. De acordo com ela, esses dados também dão suporte aos demais planejamentos. Ela completa que a agência está finalizando a contratação de uma ferramenta de monitoramento para redes sociais, buscando analisar melhor as palavras-chaves e a percepção dos clientes em relação às marcas com que trabalham. Já para a

melhoria da usabilidade dos *sites* criados pela agência, assim como as *landing pages*^{vii}, eles utilizam o *Hotjar*^{viii}. A ideia é buscar novas ferramentas para auxiliar o trabalho diário, visando à melhora de suas entregas aos clientes. Já o dono de agência de publicidade disse utilizar as redes sociais mais conhecidas do grande público (*Facebook, Instagram, Twitter, LinkedIn*, dependendo do cliente), assim como as ferramentas de suporte do Facebook e o *Google Analytics*, para a gestão e monitoramento. Ele também cita as ferramentas de suporte com a *MLabs*^{ix} (para redes sociais) e o *RD Station*^x (para *Inbound Marketing*). Ele ressalta que o mais importante é saber o que se busca e o que fazer com as informações. Já o dono de agência de MKT afirma não utilizar tais ferramentas. Ele cita a IBM, que utiliza muitas ferramentas, devido ao seu potencial intelectual e financeiro.

Os entrevistados também foram perguntados se trabalham com coletas de dados apenas no ambiente virtual ou se se valem de metodologias centradas na experiência do usuário, com a aplicação de testes fora da internet. A entrevistada disse utilizar a coleta e ressalta que isso depende da demanda dos clientes. Para ela, o ideal é que haja sempre o cruzamento de dados gerados no ambiente virtual e fora dele. Para o dono da agência de publicidade, em alguns casos, ele recorre a empresas de pesquisa para promover a captação de dados também fora do ambiente digital. Ele ressalta que trata-se de um recurso importante para avaliar a experiência do cliente *in loco*, no ponto de venda. O dono da agência de marketing disse que realiza a maior parte de suas pesquisas fora da internet, com base nas metodologias voltadas aos grupos focais e entrevistas em profundidade. Ele justifica as escolhas, feitas devido à restrição financeira dos clientes para investir em pesquisa. O profissional considera que poderia haver maior assertividade nas ações de publicidade, caso pudesse observar seus clientes e públicos em um ambiente virtual.

Sobre que tipo de profissional deveria desempenhar o trabalho de análise de dados, a entrevistada afirma que seria interessante contar com profissionais multitarefas, opinião ratificada pelo dono de agência de publicidade. Ele exemplifica que o profissional de marketing tem atuação cada vez mais próxima da de Tecnologia da Informação e que o profissional de T.I. também deveria ter visão estratégica, ultrapassando a atuação “meramente técnica”. Ele acredita que o profissional deve investir em cursos ligados a áreas que complementem sua atividade. O profissional da agência de marketing acredita que este profissional ainda esteja em processo e que a formação do Cientista de Dados é variada. Para ele, é mais comum que este profissional advinha das Ciências Exatas – Matemática, Física ou Engenharia - do que de Ciências Humanas ou Ciências Sociais Aplicadas, devido ao seu potencial analítico para diferenciar, qualificar e categorizar estes dados. Por outro lado, ele acredita que o profissional de Humanas seja necessário para interpretar os dados de forma subjetiva, identificando, por exemplo, tendências e comportamentos dos públicos.

Segundo a entrevistada, grande parte da equipe da empresa em que atua está habilitada a utilizar as ferramentas já citadas para análise de dados, principalmente, a equipe de mídia digital, que consulta esses dados várias vezes ao dia. O dono de agência de publicidade diz contar com um analista de *Inbound Marketing*, duas pessoas que atuam no planejamento e gestão de mídias sociais, além de uma estagiária de mídias e métricas. Eles trabalham de maneira integrada, sob a direção do planejamento da empresa. O dono da agência de marketing diz que não possui tal profissional, porque este tipo de mineração não é feito nas

agências, mas de forma terceirizada, uma vez que o modelo de negócios de publicidade está mais segmentado. Todos afirmaram ter o hábito de revisar dados, buscar números de concorrentes ou de parceiros antes de lançar um produto. Para eles, é necessário observar a atuação dos concorrentes para nortear suas ações. Segundo a entrevistada, isso já era feito nas campanhas *offline*, de forma mais empírica, com pesquisa de mercado, entrevistando a concorrência e o público; agora, se tornou mais efetivo.

Ainda foi questionado aos entrevistados se suas agências desenvolvem testes e/ou protótipos para a validação de ideias e projetos. Para o profissional da agência de comunicação, sua agência realiza uma simulação de ações que deseja-se sugerir. Ele cita promoções em tempo real com os clientes, com cobertura ao vivo pelo *Facebook*. Porém, segundo ele, ao aplicar essa simulação, observou-se que o “ao vivo” não teria o impacto imaginado e restringiu-se a ação ao ambiente em que ela foi realizada, revendo a forma de sua divulgação para o público mais amplo. Ele completa que também fazem campanhas com verba separada por etapa, resguardando o investimento do cliente para que a sequência ocorra após verificação do andamento. A coleta de dados é feita no decorrer da campanha e, a partir dos resultados, há sequência. Ao final, um novo relatório é gerado.

O dono da agência de marketing diz que sua agência desenvolve os testes A/B, que não necessitam de análise de dados, permitindo a verificação sobre qual produto possui melhor desempenho em determinada situação. Segundo ele, a internet tem fornecido a possibilidade de se trabalhar com muitas possibilidades para públicos variados. Ele cita como exemplo a eleição presidencial de 2016 nos Estados Unidos, na qual foram criados mais de trinta cartazes de cada candidato, com identidades visuais distintas, incluindo fotos e legendas; os que melhor se comportaram na análise do público foram usados para a construção da campanha. Entretanto, como completa o entrevistado, essa técnica de validação de produtos utilizando o *mockup*^{xi} e lançamentos para mercados teste ainda é muito comum. Sobre os critérios utilizados pelos profissionais para definir quando um produto está pronto para ser lançado no varejo, a entrevistada disse considerar a viabilidade técnica para a execução das ações planejadas, o impacto provocado *versus* o impacto esperado durante a simulação e o investimento necessário. Segundo ela, se uma proposta ultrapassa os dois primeiros critérios e se enquadra no investimento que o cliente está disposto a realizar, pode-se dar sequência ao lançamento. Do contrário, é feita a sua revisão. Neste último caso, devem ser apresentadas situações antes do lançamento de um produto, conforme descrito a seguir: coleta do *briefing* pelo cliente, o plano de lançamento e o ciclo de vida do produto (e sua substituição, se necessário) e uma pesquisa de percepção do público. Segundo a entrevistada, existem projetos em que o trabalho das agências é muito limitado e outros em que há grande intervenção da agência no que concerne à venda daquele produto.

De acordo com o dono da agência de marketing, o *Big Data* influencia muito o consumidor. Uma vez que consegue-se identificar e minerar certos dados, é possível compreender seu público, gerando alternativas personalizadas, conforme os hábitos de comportamento de compra que ele disponibiliza na internet. Em sua opinião, as grandes empresas (agências/anunciantes) têm mais acesso a esse tipo de conteúdo, mas a tendência é que sejam criados mecanismos para facilitar a decodificação desses dados, tornando o acesso mais fácil também às agências de menor porte. Ele acredita que ao mesmo tempo em que o ambiente

online se torna cada vez mais a realidade do mercado, as experiências *offline* poderão ser também mais valorizadas. A internet serviria como espaço de captação do consumidor pelas empresas, a partir dos dados disponibilizados na *web*, enquanto criam e reforçam o relacionamento com o cliente no atendimento *offline*., por meio do pós venda, com o Serviço de Atendimento ao Cliente (SAC) e com o trabalho do Analista do Cliente (*Customer Success*), viabilizando um diferencial de mercado.

Considerações e contribuições

Este artigo pesquisa as transformações do mercado com a datatificação, analisando como os bancos de dados tornam a relação entre empresas e clientes/consumidores mais multifacetada e complexa, focalizando as mudanças a serem pensadas pelas áreas de marketing e publicidade. A mineração de dados surgiu com o objetivo principal de dar suporte à tomada de decisões nas empresas. Portanto, a aplicação de técnicas de mineração de dados em sistemas de descoberta de conhecimento em banco de dados busca a descoberta de regras e padrões em dados que trarão o conhecimento suficiente e adequado para aquelas pessoas responsáveis pela tomada de decisões na empresa.

Neste trabalho, foram abordados os conceitos de Mineração de Dados e Big Data, registrando *softwares* utilizados no processo de refinamento dos dados. O principal motivo pelo qual o *Big Data* mudará a forma como as agências de publicidade atuam está na necessidade do compartilhamento de informação. Para ser eficiente e útil para o mercado, o sistema de *Big Data* deve ser retroalimentado por todas as agências, fazendo com que todas as campanhas idealizadas e avaliadas por uma agência sejam realimentadas no sistema. Para tanto, as agências deverão mudar sua cultura e compartilhar suas estratégias, com os pontos fortes e fracos do trabalho realizado. Esta é uma grande mudança conceitual para as agências tradicionais de publicidade, que sempre agiram e se posicionaram de forma engessada em relação à concorrência. A experiência do consumidor, ao longo da jornada de consumo, configura-se como uma oportunidade para marcas melhorarem esse contato com o público, personalizando-o com a ajuda de dados e inteligência artificial. O uso tático-estratégico de informações em banco de dados é um desafio para áreas como a comunicação. Sob esta ótica, a utilização da tecnologia de mineração de dados nas redes sociais permite uma utilização mais detalhada, inteligível e ampla acerca dos clientes, propiciando que as empresas possam se mostrar mais efetivas em suas estratégias de comunicação e no reforço ao engajamento, maximizando o retorno do seu investimento e gerando maior impacto sobre os clientes, maior lealdade, intenção de compra de um boca a boca mais efetivo.

Como limitações do estudo, pontua-se o caráter qualitativo da pesquisa, que buscou identificar e analisar, de maneira subjetiva, a percepção de profissionais de comunicação acerca da importância de se utilizar técnicas de mensuração de dados na elaboração de estratégias para as marcas. Também foi feita pesquisa secundária, pontuando empresas que utilizam tais técnicas em seus planejamentos, não se atendo a mensurações de caráter quantitativo. A pesquisa também se ateve à mensuração de dados de maneira geral, não focalizando, por exemplo, os dados coletados apenas nas redes sociais. Desta forma, os resultados apresentados não podem ser generalizados para outros contextos ou áreas.

Para estudos futuros, propõe-se a análise quantitativa de dados gerados em redes sociais, como no *Twitter* ou *Linkedin*, buscando monitorar o comportamento de uma amostra de consumidores de determinado produto, mostrando a aplicabilidade do *Big Data* nas estratégias empresariais. Trata-se de um tema recente, que precisa explorar outras variáveis, talvez mostrando, também, a aplicabilidade do *Big Data* ou de técnicas de Inteligência Artificial em outros campos de conhecimento, além da comunicação. Como trabalho futuro, pode-se realizar uma comparação entre diversos algoritmos de regras de associação, realizando experimentos que integrem ferramentas de mineração de dados com ferramentas específicas para a GC, apoiando ainda mais o nível gerencial corporativo para aumentar e propiciar melhores condições na tomada de decisões de forma eficaz e eficiente. Quanto aos impactos, acredita-se que o trabalho possa servir como consulta para que as empresas busquem formas de utilizar plataformas mais assertivas de baixo custo de armazenamento em nuvem, permitindo que o *Social Big Data* possa ser utilizado por empresas de qualquer porte, obtendo maior vantagem competitiva.

REFERÊNCIAS

- ABREU, G.; NICOLAU, M.. **Big Data, publicidade e o consumidor datafocado**: o caso da série House of Cards. Cultura Mídia. Revista do Programa de Pós-Graduação em Comunicação da Universidade Federal da Paraíba. Ano X, n. 18 - jan-jun/2017. Disponível em: <<http://periodicos.ufpb.br/ojs2/index.php/cm>>. Acesso em: 23 de Abr. 2019.
- BARABÁSI, Albert-László. **Linked: the new science of networks**. Kindle Edition. Cambridge: Perseu Books Group, 2002.
- _____. **Thinking in network terms**. Edge Conversation. Set. 2012. Disponível em: <<http://edge.org/conversation/thinking-in-network-terms>>. Acesso em: 29 de Jun. 2019.
- BERRY, M.J.A.; LINOFF, G. **Data mining techniques**. New York: John Wiley & Sons, Inc. 1997.
- BRAMER, M. **Under graduate Topics in Computer Science-Principles of Data Mining**. Springer, 2007.
- CABENA, P; HADJINIAN, P; STADLER, R; JAAPVERHEES; ZANASI, A. **Discovering Data Mining: From Concept to Implementation**. Prentice Hall, 1998.
- CASTELLS, Manuel (Org.). **A Sociedade em Rede**: do Conhecimento à Ação Política. 2005. Centro Cultural de Belém. 4 e 5 de março de 2005.
- CETAX. **Big data**: o que é, conceito e definição. Disponível em: <<https://www.cetax.com.br/blog/big-data/>>. Acesso em: 25 de Jun. 2019.
- CIOS, K. J; PEDRYCZ, W; SWINIARSKI, R. W; KURGAN, L. A. **Data Mining - A Knowledge Discovery Approach**. Springer, 2007.
- COVALESKI, Rogério. **Narrativa como estratégia publicitária para ações de responsabilidade social e de políticas de consumo sustentável**. In: CASTRO, Maria Guilhermina; CAIRES, Carlos Sena; RIBAS, Daniel; PAULINHOS, Jorge (eds.). Cartografia das fronteiras da narrativa audiovisual. Porto: Ed.Universidade Católica, 2016.
- CRATOCHVIL, A. **Data mining techniques in supporting decision making**. Master Thesis - Universiteit Leiden, Leiden, 1999.
- DIONÍSIO, P.; RODRIGUES, J.V.; FARIA, H.; CANHOTO, R.; NUNES, R.C.. **Mercator**. Alfragide. Publicações Dom Quixote. 2009.

- DUARTE, Jorge; BARROS, Antônio (Org.). **Métodos e técnicas de pesquisa em comunicação**. São Paulo: Atlas, 2017. p. 39-62. 380p.
- FAN, Wei; BIFET, Albert. **Mining Big Data: Current Status, and Forecast to the Future**. SIGKDD Explorations Volume 14, Issue 2. 2012.
- FAYYAD, U; PIATETSKY-SHAPIO, G; SMYTH, P. **From Data Mining to Knowledge Discovery in Databases**. American Association for Artificial Intelligence, 1996.
- FONTE, Flávio. **O que é o Hadoop?** 2013. Disponível em: <<http://bigdatabrazil.blogspot.com.br/2013/06/o-que-e-o-hadoop.html>>. Acesso em: 10 Nov. 2018.
- GOEBEL, M.; GREENWALD, L. **A survey of data mining and knowledge discovery software tools**. ACM SIGKDD, San Diego, v. 1, n. 1, p. 20-33, 1999.
- GROTH, R. **Data mining**. Englewood Cliffs: Prentice Hall, Inc., 1998.
- HAN, J; KAMBER, M. **Data Mining: Concepts and Techniques**. Elsevier, 2006.
- HAND, D; MANNILA, H; SMYTH, P. **Principles of Data Mining**. MIT Press, 2001.
- LAS CASAS, Alexandre Luzzi. **Cocriação de valor**. Conectando a empresa com os consumidores através das redes sociais e ferramentas colaborativas. São Paulo: Atlas, 2014.
- MANNILA, H. **Methods and problems in data mining**. In: **International Conference on Database Theory**, 6., Delphi, 1997. Proceedings... Delphi: Lecture Notes in Computer Science, 1997. p. 4155.
- MAYER-SCHÖNBERGER, V; CUKIER, K. **Big Data: como extrair volume, variedade, velocidade e valor da avalanche de informação cotidiana**. Rio de Janeiro: Elsevier, 2013.
- MINDMINERS. **Entenda o que é Big Data Marketing, seus 5V's e confira 2 estudos de caso**. 11 de Mai. 2017. Disponível em: <<https://mindminers.com/blog/o-que-e-big-data-marketing/>>. Acesso em: 23 de Jun. 2019.
- MIQUELINO, Amanda. **7 cases de sucesso de empresas que usam o Big Data**. 3 de setembro de 2018. Disponível em: <<https://blog.sopj.com.br/big-data/empresas-que-usam-o-big-data/>>. Acesso e: 22 de Jun. 2019.
- MUNDO DO MARKETING. **Relatório: Tecnologia no Marketing: por que você precisa dar muita atenção?** Novembro de 2016.
- NESELLO, Priscila; FACHINELLI, Ana Cristina. **Big Data: o novo desafio para gestão**. Revista Inteligência Competitiva, São Paulo, v. 4, n. 1, p. 18-38, jan. 2014. Disponível em: <<http://www.inteligenciacompetitivarev.com.br/ojs/index.php/rev/article/view/76/101>>. Acesso em: 23 de Jun. 2019.
- OLSON, D. L; DELEN, D. **Advanced Data Mining Techniques**. Springer, 2006.
- PEREIRA, Adriano. **Mineração de dados distribuída e escalável usando Apache Mahout**. 2010. 59 f. TCC (Graduação) - Curso de Ciência da Computação, Universidade Federal de Santa Maria, Santa Maria, 2010.
- PRAHALAD, C; RAMASWAMY, V. **O futuro da competição: como desenvolver diferenciais inovadores em parceria com os clientes**. 2. Ed. Rio de Janeiro: Elsevier, 2004.
- RAMANI, Girich e KUMAR, V. **Interaction Orientation and Firm Performance**. *Journal of Marketing*, Vol. 72, pp. 27-45. 2008.
- RESULTYS. **5 Cases de empresas que usam Big Data para mudar completamente sua situação no Mercado**. 2019. Disponível em: <<https://resultys.com.br/cases-de-empresas-que-usam-big-data/>>. Acesso em: 23 de Jun. 2019.

ROGENSKI, Renato. **Uma entrevista didática sobre Big Data**. Revista Exame - Tecnologia. 25 de Jul. 2013. Disponível em: <<https://exame.abril.com.br/tecnologia/uma-entrevista-didatica-sobre-big-data/>>. Acesso em: 25 de Jun. 2019.

ROZWELL, Carol; LAPKIN, Anne; FLETCHER, Chris. **Look Beyond Marketing for Competitive Advantage With Social Media**. Gartner Research. 2010.

SONDA. Tecnologia da Informação. **Conheça 6 diferentes usos do Big Data no Brasil**. 2017. Disponível em: <<https://blog.sonda.com/big-data-no-brasil/>>. Acesso em: 28 de Jun. 2019.

TARIFA, Alexandre. **O que é big data e como usar na sua pequena empresa**. Revista Exame PME – 16 de setembro de 2014. Disponível em: <<https://exame.abril.com.br/pme/o-que-e-big-data-e-como-usar-na-sua-pequena-empresa/>>. Acesso em: 28 de Jun. 2019.

TREE INTELLIGENCE. **Big data: vale a pena criar uma equipe?** 2014. Disponível em: <<http://treeintelligence.com/big-data-vale-a-pena-criar-uma-equipe/>> Acesso em 28 de Jun. 2019.

VIVEROS, M.S. *et al.* **Applying data mining techniques to a health insurance information system**. In: VLDB CONFERENCE, 22, 1996, Bombay. Proceedings. Bombay: IIT Bombay, 1996. p. 286-295.

WANG, J, editor. Encyclopedia of Data Warehousing and Mining. Idea Group Reference, 2005.

WITTEN, I. H; FRANK, E. Data Mining - Practical Machine Learning Tools and Techniques. Elsevier, 2005

WU, J. H.; WANG, Y. M. Measuring KMS success: A respecification of the DeLone and McLean's model. Information e Management, v.43, n.6, p.728, setembro, 2006. Retrieved December 7, 2006, from ABI/INFORM Global database.

ⁱ Ver Loh *et al.* (2000); Kosala e Blockeel (2000); Ma *et al.* (2000); Mobasher *et al.* (2000).

ⁱⁱ F. Diebold. **"Big Data" Dynamic Factor Models for Macroeconomic Measurement and Forecasting**. Discussion Read to the Eighth World Congress of the Econometric Society, 2000.

ⁱⁱⁱ CUKIER, Kenneth. **Data, data, everywhere: a special report on managing information**. The Economist, 394(8671), Fevereiro 2010.

^{iv} Ramo da Matemática que estuda as relações entre objetos de um determinado conjunto. No estudo, empregam-se estruturas chamadas de grafos, onde é um conjunto não vazio de objetos denominados vértices e é um subconjunto de pares não ordenados de V.

^v **Habits - A Repeat Performance**. Publicado pela Universidade de Duke/EUA. Disponível em: <<https://dornsife.usc.edu/assets/sites/208/docs/Neal.Wood.Quinn.2006.pdf>>. Acesso em: 28 de Jun. 2019.

^{vi} **Mapping and Sharing the Consumer Genome**. The New York Times. Disponível em: <https://mobile.nytimes.com/2012/06/17/technology/acxiom-the-quietgiant-of-consumer-database-marketing.html>. Acesso em 28 de Jun. 2019.

^{vii} O termo, traduzido como "Página de Aterrissagem" ou "Página de Destino", é o primeiro espaço que o usuário da *web* acessa ao entrar em um *site*. Pode ser um anúncio ou um *banner*, que busca atrair *clients/consumidores* para gerar conversão (venda, clique etc).

^{viii} Ferramenta de inteligência e análise de *sites* que fornece dados de navegação sobre o usuário. Ele mostra, pelo mapa de calor, onde o visitante mais clicou em um portal.

^{ix} Ferramenta *web* paga de gerenciamento de redes sociais para administrar as contas das redes sociais. Oferece serviços como emissão de relatórios de desempenho.

^x Ferramenta de automação que auxilia a empresa a otimizar suas estratégias de MKT digital.

^{xi} Espécie de maquete ou de representação de um objeto ou produto no tamanho natural.