

---

***BIG DATA: Conteúdo Mediado por Computador - Um Estudo Sobre a Análise de Grandes Dados nas Estratégias de Comunicação e Marketing***<sup>1</sup>.

Christiane Rocha e SILVA<sup>2</sup>

Igor Brito SANTOS<sup>3</sup>

Gabriel Meirelles dos SANTOS<sup>4</sup>

Centro Universitário Newton Paiva, Belo Horizonte, BH

## RESUMO

O trabalho analisa a aplicação do *Big Data* no marketing e comunicação, visando à incorporação estratégica da informação de grandes bancos de dados nas campanhas ao consumidor. O *Big Data* é uma oportunidade de as empresas otimizarem seus processos informacionais, com a prospecção refinada de dados, encontrando tendências e padrões de comportamento do consumidor. 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 compreender como o *Big Data* tem sido utilizado para aumentar o engajamento com o cliente.

**PALAVRAS-CHAVE:** Mineração de Dados; KDD; Big Data; Marketing; Publicidade.

## Introdução

As técnicas tradicionais de utilização de dados já não são suficientes para traçar um perfil do comportamento do consumidor. Buscando um mapeamento mais preciso destes dados, tem sido estudada, desde o fim da década de 1980, como explicam Berry e Linoff (1997), a técnica de Mineração de Dados (*Data Mining*). 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, impactando nos indicadores de desempenho - *Keys Performance Indicators* (KPI's) organizacionais. Neste contexto, surge a pergunta que norteou o estudo: como a mineração de dados e o *Big Data* podem ser utilizados pelas empresas/marcas, na elaboração e produção de estratégias mais assertivas de comunicação e marketing?

---

<sup>1</sup>Trabalho apresentado na DT 2 de Publicidade e Propaganda, no XIX Encontro dos Grupos de Pesquisas em Comunicação, evento componente do 42º Congresso Brasileiro de Ciências da Comunicação. **APOIO: CNPq.**

<sup>2</sup> Doutora em Marketing; Professora Titular dos Cursos de Publicidade e Jornalismo do Centro Universitário Newton Paiva. Sua trajetória universitária está centrada no estudo das interfaces entre Comunicação, Marketing e Tecnologia e a Gestão de Relações Públicas. Linhas de Pesquisa: Análise de Dados, Gestão da Reputação, Estratégia e Tecnologias em Marketing. christianerocha@yahoo.com.br.

<sup>3</sup>Publicitário graduado pelo Centro Universitário Newton. Email: igorbritoss@hotmail.com

<sup>4</sup> Graduando do 6º período de Publicidade e Propaganda no Centro Universitário Newton. Email: gabriel.meirelles@newtonpaiva.br.

---

O artigo teve, como Objetivo Geral, pesquisar a aplicabilidade das técnicas de análise de grandes dados na criação de estratégias de comunicação eficientes, visando ao engajamento das marcas com o seu cliente/consumidor. Especificamente, buscou pesquisar os principais conceitos, tarefas e métodos da Mineração de Dados e do *Big Data*; pesquisar as estratégias de utilização das técnicas de mineração de dados pelas empresas em sua comunicação; e analisar de que maneiras a mineração de dados pode contribuir na produção de conteúdo para o cliente, por meio de entrevistas.

O trabalho justifica-se devido ao crescimento do volume de dados armazenados associados ao *Big Data* e à necessidade de as empresas o utilizarem no posicionamento de suas marcas. O marketing e a publicidade foram setores afetados pelas Tecnologias de Informação e Comunicação (TICs) e torna-se necessário entender como os profissionais têm se reinventado neste ambiente. Aliado às novas tecnologias, o profissional de comunicação pode apresentar-se como um analista da informação, humanizando e interpretando os dados, oferecendo campanhas personalizadas.

O trabalho iniciou-se por meio de revisão bibliográfica (DUARTE E BARROS, 2017), apresentando os conceitos de Mineração de Dados, *Big Data*, entre outros, com a leitura em livros, artigos, revistas e pesquisa na internet. O estudo, de caráter exploratório, procedeu-se por meio de Pesquisa Documental, pontuando empresas que utilizam o *Big Data* em suas campanhas ao cliente/consumidor. A segunda fase do trabalho consta de uma pesquisa qualitativa (DUARTE E BARROS, 2017), com a realização de três entrevistas em profundidade com gestores de empresas de portes distintos e publicitários, para entender como a análise de grandes dados pode contribuir na criação de campanhas que otimizem o relacionamento com o cliente/consumidor. .

### **Mineração de Dados – *Data Mining***

O termo é definido, segundo 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, o 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. A mineração é parte do chamado processo de Descoberta de Conhecimento em Banco de Dados - KDD - *Knowledge Discovery in Databases*, utilizado para representar o processo de tornar dados de baixo nível em conhecimento de alto nível.

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 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.

Para Fayyad *et. al.* (1996), o modelo tradicional de transformação de dados em informação (conhecimento) consiste no processamento manual dessas informações por especialistas que produzem relatórios a serem analisados. Na maioria das situações, devido ao 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 sobrecarga de dados. Segundo eles, há 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 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. 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. Percebe-se, portanto, que os principais objetivos da mineração de dados podem ser usados no gerenciamento de informação, na tomada de decisão, no controle de processo e ainda, como afirma Groth (1998), como processo de verificação e de descoberta.

As técnicas de mineração podem ser utilizadas em várias áreas, como pontuam Mannila (1997); Cratochvil (1999): no ambiente financeiro, na detecção de padrões sutis de fraudes na Medicina e Ciência, buscando padrões de novas doenças; no Controle de processos e qualidade, auxiliando no planejamento estratégico; no ambiente *Web*, e no Marketing, para descobrir preferências do consumidor. Witten *et al.* (2005) 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. Goebel e Greenwald (1999) afirmam que muitas ferramentas disponíveis são advindas da Inteligência Artificial ou Estatística. A mineração é mais usada com dados estruturados, como planilhas. Já o *Big Data* trabalha com dados mais complexos e não-estruturados. As escalas, tipos de dados, períodos de análise e os resultados também são diferentes. Enquanto o *Data Mining* se refere a um processo pontual, que gera relatórios, apontando para questões específicas, o *Big Data* é feito de forma contínua por períodos maiores.

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

O termo *Big Data*, segundo Fan e Bifet (2012), surgiu 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, 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 dataria do ano 2000, em um artigo de Diebold<sup>5</sup>. Para Arrigoni, segundo Rogenski (2013), o termo surgiu na década de 1990, na NASA, para descrever 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 publicação<sup>1</sup> na revista britânica *The Economist*, mas que o *fenômeno* teria se iniciado na década

---

<sup>5</sup> F. Diebold. "*Big Data*" *Dynamic Factor Models for Macroeconomic Measurement and Forecasting. Discussion Read to the Eighth World Congress of the Econometric Society*, 2000.

---

anterior, antes do surgimento do termo, com os sistemas de informação. Para Barabási (2002), o termo pode ser chamado *Big Data*, Computação Social, Dinâmica Humana ou Ciência das Redes. O importante é dar sentido à quantidade de informações.

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 etc. Para o cientista, a analítica digital não suplanta a capacidade humana de gerar *insights*, discernir e decidir. No ano 2000, após realizar experiências com grande quantidade de dados digitais, Barabási publicou o que pode ser reconhecido como o surgimento da Ciência das Redes, estudadas no século dezoito pela Teoria dos Grafos<sup>6</sup>. Segundo Barabási (2002), para os matemáticos da época, as redes (estruturas formadas por nós e ligações) tinham propriedades fixas. No século vinte, cientistas observaram outras redes não explicadas matematicamente. Como desconheciam as propriedades dessas redes, imaginaram que elas se formavam e se comportavam de modo randômico. Entre as redes randômicas, cita-se a sociedade, a célula, o organismo, o ecossistema e as redes naturais e sociais.

Apesar das várias definições sobre 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 pelas técnicas que analisam grandes quantidades de dados, por meio de ferramentas especiais (CETAX, 2019). Mayer-Schönberger e Cukier (2013) reforçam que o volume de dados disponível mais que dobra a cada dois anos e os algoritmos também se aperfeiçoam rapidamente. O *Facebook* armazena, acessa e analisa mais de 50 *petabytes* (unidade de medida maior que um *terabyte*) de informações dos usuários. No *Youtube*, por minuto, são descarregadas 48 horas de vídeos, sendo impossível assistir a tudo isso. 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, 15% inéditas e rastreia 20 bilhões de *sites* por dia, com informações estruturadas, que podem ser usadas na leitura e extração de dados, como arquivos texto (txt, XML); semi-estruturados

---

<sup>6</sup> CUKIER, Kenneth. *Data, data, everywhere: a special report on managing information*. The Economist, 394(8671), Fevereiro 2010.

(acompanham padrões heterogêneos); e não-estruturadas, sem formato padronizado para leitura (páginas de Internet/Intranet, vídeos), 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 e desenvolvimento econômico, tornando as redes mais compreensíveis, previsíveis e controláveis. O *Big Data* (CETAX, 2019) pode ser subdividido em cinco categorias: Volume, Velocidade, Variedade, Veracidade e Valor, sendo os três primeiros os mais relevantes. O primeiro desafio é gerenciar o Volume de dados disponíveis, gerados no cotidiano, na troca de mensagens eletrônicas (*e-mails* e redes sociais), transações bancárias e contatos por linhas telefônicas. 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 evolução em *Data Center* e tendências de utilização em nuvem (TECMUNDO, 2018).

A Velocidade se mostra como fator de vantagem competitiva, a partir da análise dos dados atuais, gerados em tempo mais próximo do real. Sobre a 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. 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 e outros. O desafio das empresas é integrar os dados das estratégias de *Business Intelligence* (BI) com os de *Big Data*, analisando informações e medindo outros dados não estruturados.

Sobre as desvantagens do *Big Data*, cita-se o alto custo, a demora no retorno do investimento (ROI), a busca por demanda para implantar a tecnologia, mesmo sem ter um problema e o fator privacidade, visto que os sistemas têm coletado dados variados (CETAX, 2019). Há muitas soluções para que as empresas utilizem a Mineração de Dados e o *Big Data*. Existem várias ferramentas gratuitas 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 gratuito Linux. 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 pelo armazenamento e processamento. O primeiro é o *Hadoop Distributed File System (HDFS)*, sistema de arquivos distribuído, projetado para armazenar arquivos muito grandes. Ele não deve ser utilizado para aplicações que precisem de acesso rápido, mas para aplicações em que é preciso ler muitos 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, os *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*.

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 conjunto de informações. Ele consiste em um serviço *master – 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 e se reportam ao *JobTracker*, especificando que estão funcionais. 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.

### ***Big Data Marketing* e sua aplicação nas estratégias de marketing e publicidade**

Pesquisas mostram que a ação cotidiana do indivíduo caracteriza-se pelo hábito da repetição e que parte dos processos de tomada de decisão não é consciente. Em estudos feitos na Universidade Duke<sup>7</sup>, participantes relataram ações heterogêneas em suas rotinas, todas variando a partir da força do hábito, como fazer exercícios e comer *fast food*. Estudos diários de amostragem por experiência com amostras de estudantes e comunidades reforçam tal resultado, apontando que 45% dos comportamentos cotidianos tendem a se repetir no mesmo local quase todos os dias. É nesse contexto que aplica o *Big Data*. Além de fornecer dados, torna-se possível compreender as conexões e selecionar as mais relevantes às empresas. O Marketing não é tecnologia. A Tecnologia de Marketing é apenas um meio para que se crie conexão com os clientes,

---

<sup>7</sup> *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.

gerando experiências e engajamento. Um benefício do *Big Data* no Marketing está no custo reduzido, uma vez que não há só ferramentas caras e sofisticadas de análise digital (MUNDO DO MARKETING, 2016). O *Big Data* no marketing, Marketing por Dados ou *Data Driven Marketing* consegue promover uma conexão dos dados *offline* e *online*., enviando o conteúdo certo ao cliente certo no momento certo.

Nesse cenário, 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<sup>8</sup>) e gerando *insights*. 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 informações sobre hábitos de compra dos clientes com dados demográficos. De acordo com Covaleski (2016), a necessidade atual de obter reação do público, bem como à crescente rejeição à publicidade, devido ao seu hábito intrusivo e interruptivo, torna complicada a tarefa das empresas de decidirem onde e como alocar os orçamentos de publicidade. Assim, o discurso publicitário precisa ter um tom motivacional, aderindo ao consenso social vigente. Os profissionais de marketing também devem exercitar a capacidade analítica para interpretar dados e orientar os criativos nas ações. 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).

### **Aplicações de *Big Data* no marketing e publicidade**

As empresas, independentemente do porte, têm recorrido à tecnologia para aprimorar seus processos e evitar a obsolescência. As 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. O Festival de Criatividade *Cannes Lions*, premiação mais importante da publicidade mundial, criou, em 2015, a categoria *Creative Data*, para premiar as campanhas que utilizam o *Big Data* na criação.

---

<sup>8</sup> *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.



A campanha vencedora do *Grand Prix* de 2016, *The Next Rembrandt*, criada pela agência de publicidade J. Walter Thompson (JWT) holandesa com a Microsoft para o banco ING, *produziu*, com a ajuda de inteligência artificial, dados sobre a obra do pintor Rembrandt, permitindo que uma impressora 3D pintasse um quadro no estilo do artista falecido. Com a ajuda de engenheiros e especialistas em arte, foram escaneadas 364 pinturas. Centenas de milhares de fragmentos foram analisados por algoritmos, bem como os rostos dos retratos feitos por Rembrandt (MINDMINERS, 2017).

O serviço de *streaming* Netflix também utiliza *Big Data* para prever hábitos e criar métricas de comportamento do usuário, diagnosticando quantas vezes o usuário pausa um vídeo, se assiste a episódios sequencialmente e quando se torna fã de uma série (MINDMINERS, 2017). A Vivo, Claro, Renner e Itaú Unibanco também utilizam *Big Data*. A Renner 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 (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. 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, usando *Machine Learning* e computação em nuvem (*Cloud Computing*). A Shell, uma das maiores petrolíferas do mundo, monitora as áreas mais propensas de perfuração para extração de petróleo, para reduzir gastos de operação e minimizar riscos ambientais (RESULTYS, 2019). Por meio de aplicativos e dispositivos vestíveis (*wearables*), a Nike reúne dados sobre o comportamento esportivo, como distância percorrida e rotas para treinos.

### **Resultados da pesquisa qualitativa**

As três entrevistas foram realizadas, pessoalmente, no mês de março de 2018, presencialmente e por 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 sobre o conceito de *Big Data*. A

---

gerente do núcleo digital de uma agência de médio porte da capital mineira, definiu como a grande quantidade de dados existentes, atualmente, que consegue-se capturar, gerados pela interação das pessoas com a *web* e com as redes sociais. Para ela, com o *Big Data*, é possível que as empresas – a despeito do porte ou formato (*offline* ou *online*) – obtenham *insights* para elaborar suas estratégias. Ainda segundo ela, mais relevante que a quantidade é a interpretação dos dados.

Já de acordo com o dono de uma agência de comunicação de pequeno porte sediada em BH, hoje, com as TIC's, é possível conhecer melhor os hábitos de consumo. “Podemos rastreá-las pela internet, colhendo assim um conjunto de dados, que é o que chamamos de *Big Data*”. Segundo ele, isso 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 e repele”. Para ele, isto é fundamental para as empresas, que podem ajustar sua oferta ao público, oferecendo o que ele quer, onde precisa e na hora exata. Outro entrevistado, dono de uma agência de marketing, define o termo como um processo de análise de todos os dados gerados atualmente, a partir das incursões feitas pelas pessoas em dispositivos móveis. Ele completa que estes dados, isoladamente, não representam muito, sendo apenas “um grande quebra cabeça”. O que torna os dados úteis é a possibilidade de minerá-los.

Perguntados sobre como utilizam a análise de dados, seja na tomada de decisão, no *remarketing*, no relacionamento com o cliente etc, a gerente disse que, onde 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. A agência acompanha os resultados das ações, promovendo mudanças no processo, se necessário. O dono da agência de publicidade acompanha a rede social dos clientes, identificando pontos críticos, mensurando o interesse pelos produtos/serviços e interagindo (não mencionou por quais ferramentas). 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, por ser cara. Ele afirma que cria peças de divulgação tomando o *briefing* como referência, menos relacionado a dados e mais focado em comportamento e pesquisas feitas pessoalmente. Os entrevistados também foram perguntados sobre quais ferramentas de análise de dados utilizam no trabalho e quais consideram 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*. 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, para analisar melhor as palavras-chaves e a percepção dos clientes sobre marcas. Já para a melhoria da usabilidade dos *sites* criados pela agência, assim como as *landing pages*<sup>9</sup>, eles utilizam o *Hotjar*<sup>10</sup>.

Já o dono de agência de publicidade disse utilizar as redes sociais populares (*Facebook*, *Instagram*, *Twitter*, *LinkedIn*) e as ferramentas de suporte do *Facebook* e o *Google Analytics*, para gestão e monitoramento. Ele também cita as ferramentas de suporte com a *MLabs*<sup>11</sup> (para redes sociais) e o *RD Station*<sup>12</sup> (para *Inbound Marketing*). Já o dono de agência de MKT afirma não utilizar tais ferramentas. Os entrevistados também foram perguntados se trabalham com coletas de dados só no ambiente virtual ou se se valem de metodologias centradas na experiência do usuário, com testes fora da internet. A entrevistada disse utilizar a coleta e ressalta que isso depende da demanda do cliente. Para ela, o ideal é que os dados sejam cruzados. Para o dono da agência de publicidade, às vezes, ele recorre a empresas de pesquisa para a captação de dados fora da *web*, para avaliar a experiência do cliente *in loco*, no ponto de venda. O dono da agência de marketing realiza parte de suas pesquisas fora da internet, em grupos focais e entrevistas em profundidade, devido à restrição financeira dos clientes para investir em pesquisa. Ele 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 aposta em profissionais multitarefas, opinião ratificada pelo dono de agência de publicidade. Ele exemplifica que o profissional de marketing tem atuação cada vez mais próxima da Tecnologia da Informação e diz que o profissional de T.I. também deveria ter visão estratégica, mais do que “meramente técnica”. Ele acredita que o profissional deve investir em cursos que complementem sua atividade. O

---

<sup>9</sup> 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*, cujo objetivo é atrair clientes/consumidores para gerar venda, clique etc).

<sup>10</sup> 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.

<sup>11</sup> Ferramenta *web* paga de gerenciamento de redes sociais para administrar as contas das principais redes. Oferece serviços como emissão de relatórios de desempenho.

<sup>12</sup> Ferramenta de automação que auxilia a empresa a otimizar estratégias de marketing digital.

---

profissional da agência de marketing acredita que este profissional esteja em processo e que a formação do Cientista de Dados é variada. Para ele, é mais comum que ele advinha das Ciências Exatas do que de Ciências Humanas ou Sociais Aplicadas, devido ao seu potencial analítico para diferenciar, qualificar e categorizar dados. Por outro lado, ele acredita que o profissional de Humanas seja necessário para interpretar os dados de forma subjetiva, identificando tendências e comportamentos dos públicos.

Segundo a entrevistada, grande parte da equipe da empresa em que atua está habilitada a utilizar as ferramentas 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 conta com um analista de *Inbound Marketing*, duas pessoas no planejamento e gestão de mídias sociais e uma estagiária de mídias e métricas trabalhando de forma integrada, sob a direção do planejamento da empresa. O dono da agência de marketing 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 têm hábito de revisar dados, buscar números de concorrentes ou de parceiros antes de lançar um produto. Segundo a entrevistada, isso já era feito nas campanhas *offline*, de forma mais empírica, com pesquisa de mercado, entrevistando a concorrência e o público.

Ainda foi questionado aos entrevistados se suas agências desenvolvem testes e/ou protótipos para validar ideias e projetos. Para o profissional da agência de comunicação, sua agência realiza uma simulação de ações que deseja-se sugerir, como promoções em tempo real com o cliente, 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. Eles 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. Sobre os critérios utilizados para definir o tempo de lançamento de um produto 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: coleta do *briefing* pelo cliente, o plano de lançamento e o ciclo de vida do produto (e sua substituição, se necessário) e pesquisa de percepção do público. Segundo ela, 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.

Segundo o dono da agência de marketing, o *Big Data* influencia muito o consumidor. Em sua opinião, as grandes agências/anunciantes têm mais acesso a esse tipo de conteúdo, mas a tendência é que sejam criados mecanismos para facilitar a decodificação desses dados 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, enquanto criam e reforçam o relacionamento no atendimento *offline* no pós venda, com o Serviço de Atendimento ao Cliente (SAC) e do Analista do Cliente (*Customer Success*), criando um diferencial.

### **Discussão e conclusões**

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. 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. A utilização da tecnologia de mineração de dados nas redes sociais permite uma utilização mais detalhada, inteligível e ampla acerca do cliente, propiciando que as empresas sejam mais eficazes nas estratégias de comunicação e no reforço ao engajamento.

Como limitações do estudo, pontua-se o caráter qualitativo da pesquisa, que buscou identificar e analisar, de maneira subjetiva, apenas a percepção de três profissionais da área. 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, 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. Quanto aos impactos, acredita-se que o trabalho sirva de consulta para que as empresas busquem formas de utilizar plataformas assertivas de baixo custo de armazenamento em nuvem, obtendo maior vantagem competitiva.

## REFERÊNCIAS

ABREU, G.; NICOLAU, M.. **Big Data, publicidade e o consumidor datafocado**: o caso da série House of Cards. Cultura Mimdiática. 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.

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.

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.

---

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.

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.

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.

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.

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.

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.

---