

USO DO *BIG DATA* EM CAMPANHAS POLÍTICAS ELEITORAIS

THE USE OF BIG DATA IN ELECTORAL POLITICAL CAMPAIGNS

Cleide Luciane Antoniutti⁶³
Sarita Albagli⁶⁴

Resumo: Visa-se compreender e analisar o uso do *big data* em campanhas políticas eleitorais. Sistematiza-se, na primeira parte, uma revisão atualizada da literatura sobre a emergência do *big data* no contexto da explosão informacional e da datificação. Nesse cenário, as redes sociais digitais configuram-se como ambientes propícios de circulação de informação, em que os “rastros digitais” deixados por seus usuários constituem a matéria prima principal de projetos de *big data*. Em seguida, como base empírica, o trabalho apresenta e discute a experiência da campanha de Barack Obama à reeleição à presidência dos Estados Unidos, em 2012. Essa campanha, que ficou conhecida como Obama for America – OFA, transformou-se em um verdadeiro *case* de utilização de ferramentas analíticas de *big data* em estratégias eleitorais.

Palavras-chave: *Big data*. Datificação. Campanha Eleitoral. Barack Obama.

Abstract: The paper aims to understand and analyse the use of big data in electoral political campaigns. In the first part, it systematizes an updated revision of the literature on the emergence of big data within the information explosion and datafication context. In this scenario, digital social media are suitable environments for information circulation, in which digital traces left by their users constitute raw material for big data projects. Then, as its empirical basis, the paper presents and debates the experience of Barack Obama's campaign for his reelection as President of the United States, in 2012. That campaign, known as Obama for America – OFA, became a case in the use of big data analytical tools for electoral strategies.

Keywords: Big data. Datafication. Electoral Campaign. Barack Obama.

1 INTRODUÇÃO

Este trabalho apresenta resultados parciais de pesquisa cuja finalidade é compreender e analisar o uso do *big data* em campanhas políticas eleitorais. Os novos usos das análises de dados no contexto das disputas eleitorais revelam que, cada vez mais, dados e informação constituem o grande diferencial das estratégias políticas da atualidade.

O termo *big data*, ainda pouco difundido no Brasil e antes tratado apenas entre profissionais de Tecnologia da Informação, é agora cada vez mais problematizado em outras esferas e áreas, principalmente, nas ciências sociais, ganhando destaque em projetos de comunicação e marketing político. Além disso, o tema *big data* começa a entrar em discussões e análises acadêmicas, mostrando-se um importante objeto de investigação,

⁶³ Doutoranda -PPGCI/IBICT-UFRJ- lutti.antoniutti@gmail.com.

⁶⁴ Pesquisadora e Professora- IBICT-UFRJ – sarita.albagli@gmail.com.

principalmente para a Ciência da Informação, pois o uso de dado e informação, que sempre foram objetos de estudo dessa área, é também o que sustenta o conceito de *big data*.

O trabalho sistematiza, na primeira parte, uma revisão atualizada da literatura em torno do conceito de *big data*, abordando o fenômeno da datificação no contexto da explosão informacional. Nesse cenário, as mídias sociais digitais configuram-se como ambientes precípuos de circulação de informação, em que os “rastros digitais” deixados por seus usuários constituem a matéria prima principal de projetos de *big data*. Abordam-se também as questões que envolvem as práticas de monitoramento e vigilância de informações tanto pelo aspecto de gerenciamento de informações para obtenção de conhecimento, como pelo aspecto da privacidade e ética dos usos de dados de usuários das redes sociais digitais.

Em seguida, como base empírica, o trabalho apresenta e discute a experiência da campanha de Barack Obama à reeleição à presidência dos Estados Unidos, em 2012. Essa campanha, que ficou conhecida como *Obama for America* (OFA) transformou-se em um verdadeiro *case* de utilização de ferramentas analíticas de *big data* em estratégias eleitorais. A campanha é considerada por especialistas como pioneira no uso do *big data* em ações de comunicação e marketing político.

O trabalho busca então responder às seguintes indagações: o que de novo o uso do *big data* vem trazendo para as estratégias eleitorais? Qual o aprendizado com esse uso na campanha pela reeleição de Barack Obama? Que implicações o uso do *big data* pode trazer em relação à privacidade de dados dos eleitores?

2 EXPLOSÃO DA INFORMAÇÃO E DATIFICAÇÃO

Mesmo antes da difusão da internet, já se apontava a existência de uma verdadeira “explosão informacional”, sobretudo na área científica e tecnológica (ARAÚJO, 1996). O desenvolvimento e a difusão das tecnologias da informação e comunicação, nas últimas décadas, principalmente com a ampliação do uso de dispositivos móveis e de redes sociais digitais, contribuiu ainda mais para o exponencial crescimento e o acúmulo do volume de dados⁶⁵, no que se tem chamado de “datificação”. (MAYER-SCHÖNBERGER; CUKIER, 2013). Essa mudança quantitativa, segundo Mayer-Schönberger e Cukier (2013) levou a uma

⁶⁵ O volume de informação cresce vertiginosamente nos ambientes da internet. A cada minuto são enviados 168 milhões de e-mails, 600 novos vídeos são inseridos no Youtube, mais de 500 mil telefonemas são feitos pelo Skype. Pesquisas revelam que, até o ano de 2003, foram gerados 5 *exabytes* de informações. Esse mesmo volume é gerado em apenas dois dias. Calculou-se ainda que, a partir de 2012, o volume de informações dobraria a cada dois anos até chegar a 8 *zettabytes* em 2015 (TAURION, 2013).

mudança qualitativa: a revolução da informação residiria menos em máquinas capazes de calcular dados, do que nas possibilidades de novos usos desses dados em diferentes domínios – econômico, político e social.

Muitos consideram que o volume de informações hoje disponíveis excede a capacidade humana de explorá-las, exigindo cada vez maior esforço e tempo para acessá-las e processá-las com os objetivos desejados. Acredita-se que o excesso de informação pode ocasionar a perda de controle e a imobilidade em utilizá-la efetivamente (TERRA, 2000; BAUMAN, 2011; VAIDHYANATHAN, 2011; SCHONS, 2007).

Cabe aqui diferenciar digitalização e datificação⁶⁶. Digitalização é o processo de converter todo tipo de informações analógicas para código binário, por meio de processos computacionais. Conforme ressaltam Mayer-Schönberger e Cukier (2013) a revolução computacional precedeu a revolução digital.

Já datificação refere-se à transformação em dados as informações que, de algum modo, possam ser quantificadas, seja no mundo digital ou analógico, dados esses que podem ser usados de diferentes modos, frequentemente imprevisíveis (MAYER-SCHÖNBERGER; CUKIER, 2013). A digitalização vem contribuindo para “turbinar” a datificação, ante a ampla adoção da internet e das mídias digitais, com a digitalização de textos, músicas e vídeos. É importante ressaltar que, separados, os dados não mostram muita coisa. Quando aglutinados e sistematizados, aí sim podem revelar padrões antes não enxergados.

Inversamente à perspectiva tradicional, na Ciência da Informação, da linearidade causal entre dado, informação e conhecimento, a datificação, na era do *big data*, não há uma hierarquia prévia entre dados, informações e conhecimento, em que o conhecimento é decorrente da informação, que por sua vez, deriva de um conjunto de dados⁶⁷. A hierarquia reversa também pode ser considerada, o conhecimento ao ser explicitado, difundido através de meios físicos ou virtuais, se torna uma informação ou um conjunto delas, que ao serem desmembradas se tornam um conjunto de dados. Os dados que compõem o *big data* podem ser estruturados (armazenados em banco de dados tradicionais) ou não estruturados (incluindo áudio, imagens, vídeos e sensores), vindos principalmente de mídias sociais da internet.

⁶⁶ Datificação de informações existem muito antes do surgimento da internet e da digitalização de documentos como observam Mayer-Schönberger e Cukier (2013).

⁶⁷ Davenport e Prusak (1998) consideram dados como registros sem significados, que se transformam em informações ao adquirirem algum significado.

Como observam os especialistas, a datificação atinge hoje o âmbito das relações interpessoais e das próprias intersubjetividades transformando “elementos intangíveis do cotidiano [...] em dados que podem ser usados para outros fins” (MAYER-SCHÖNBERGER; CUKIER, 2013, p. 63), deixando assim verdadeiros “rastros digitais” (LEMOS, 2005). Na internet o rastro acompanha toda ação-navegação, incluindo busca, simples cliques em links, downloads, produção ou reprodução de um conteúdo, suscetível de ser capturado e recuperado (BRUNO, 2012), para diversas finalidades, principalmente de ordem comercial e publicitária, mas também política.

Vale destacar que a informação não é um objeto de interesse recente, somente pelo aspecto e dimensão digital que ela tem. A diferença agora é que ela passa a assumir papel cada vez mais importante nas tomadas de decisões nos âmbitos econômico, social e principalmente político. Dessa forma destacam-se os pensamentos de Castells:

O que caracteriza a atual revolução tecnológica não é a centralidade de conhecimentos e informação, mas a aplicação desses conhecimentos e dessa informação para a geração de conhecimentos e de dispositivos de processamento/comunicação da Informação em um ciclo de realimentação e seu uso (CASTELLS, 1999, p.50-51).

Desde 2007, notícias e protestos contra a vigilância de dados e a violação de privacidade por parte de diversos serviços da internet têm ganhado destaque, principalmente no campo acadêmico. Essas discussões se tornaram mais amplas com as denúncias feitas pelo ex-técnico da Agência Nacional de Segurança (NSA) Edward Snowden das vigilâncias cometidas pelo governo americano por meio do Prism⁶⁸ (Programa de Vigilância), que tinham por finalidade rastrear dados⁶⁹ e ações de governos e pessoas de diversos países, nos mais variados ambientes de informação e comunicação.

Esse episódio coloca em evidência as discussões sobre o conceito de privacidade, principalmente nas redes sociais digitais, preocupação esta que, embora legítima, é limitada se não for acompanhada de uma problematização mais ampla dos processos característicos da vigilância digital (BRUNO, 2013). É preciso observar que a popularização das redes sociais

⁶⁸⁶⁸ PRISM é um dos programas do sistema de vigilância global da [NSA](#) que foi mantido secreto desde 2007 e até sua revelação na imprensa em 7 de junho de 2013. Sua existência veio à público por meio de publicações feitas pelo jornal britânico [The Guardian](#), com base em documentos fornecidos por [Edward Snowden](#).

⁶⁹ Nos Estados Unidos o monitoramento de dados e informações de cidadãos é legal. O Prism foi autorizado nas emendas de 2008 ao Ato de Vigilância de Inteligência Estrangeira – ao contrário dos programas de vigilância expedidos sem mandado na gestão de George W. Bush. O programa tem autorização do Congresso para funcionar até 2017.

intensificou as discussões sobre a exposição da privacidade e da intimidade nas tecnologias de comunicação. Poster (1990) apontava pela primeira vez, que nas novas tecnologias de comunicação não seríamos apenas disciplinados à vigilância, mas também à participação. Diversos estudos (Fuchs, 2011; Rose, 1998; Riesman, 1995) parecem concordar que detalhes cotidianos da vida privada, traços comportamentais e fluxos subjetivos como gostos, crenças, opiniões, nunca foram tão amplamente visíveis e deliberadamente publicizados, segundo uma necessidade de espetacularização do ser. Para Bruno (2013) a exposição da vida pessoal desperta questões relativas aos limites entre o público e o privado, ao sentindo de intimidade, bem como às práticas de vigilância.

Segundo Bruno (2013) é preciso ter em mente que a noção de privacidade está em disputa. Para a autora, a definição tradicional, fundada nos princípios de proteção ao anonimato, à solidão e ao segredo, não é suficiente no contexto da vigilância digital. “A insuficiência reside no princípio de identidade aparente ou passível de averiguação, pois muitas vezes pode-se prescindir dessa identificação” (BRUNO, 2008, p.11). Por outro, a questão da privacidade implica reforçar a sua dimensão coletiva, e não simplesmente privada e particular. “Essa dimensão coletiva só pode ser problematizada se analisarmos os processos mais amplos de monitoramento, classificação e controle da informação individual, apreendendo o modo como os indivíduos e as identidades estão sendo produzidos, bem como as implicações ético-políticas em jogo” (BRUNO, 2008, p.11).

Para Bruno (2013) a história recente dos processos de vigilância tem estreitas relações com as tecnologias de informação e de comunicação. Desde a convergência da informática com as telecomunicações, as redes de comunicação distribuída, como a internet, ampliam significativamente as possibilidades de monitoramento, coleta e classificação de dados pessoais.

Num estrato mais superficial e explícito, há as informações pessoais e publicações que divulgamos voluntariamente na web (postagens em blogs, dados de perfil e conversações nas redes sociais). Mas além e aquém desse nível declarativo e sua respectiva inscrição, uma série de outras ações – navegação, busca, simples cliques em links, downloads, produção ou reprodução de conteúdo – deixam vestígios mais ou menos explícitos, suscetíveis de serem capturados (BRUNO, 2013, p. 123).

Dessa forma, cabe afirmar que as dinâmicas da vigilância na internet estão intimamente atreladas às formas de participação dos usuários e aos embates que lhes correspondem. Assim, se constituem uma fonte valiosa de informação ou conhecimento, podendo ser usadas para diversas finalidades e ações, no caso deste trabalho, em estratégias políticas eleitorais. Toda vigilância, como se sabe, implica não apenas a observação de

indivíduos e populações, mas na produção de um saber que permita governar as suas condutas (FOUCAULT, 2013; ROSE, 1998).

4 CONTEXTUALIZANDO BIG DATA

O termo *big data* foi usado pela primeira vez em 1997, em trabalho apresentado por Michael Cox e David Ellsworth, pesquisadores da *National Aeronautics and Space Administration* (Nasa) norte-americana, na *8th IEEE Visualization '97 Conference*, organizada pelo Instituto de Engenheiros Eletricistas e Eletrônicos (IEEE) (COX; ELLSWORTH, 1997). Eles anteviam os novos desafios na área de computação para lidar com grandes volumes de dados. Desde então, diversas definições têm sido dadas ao termo.

Mas afinal o que é *big data*? As definições encontradas convergem para o entendimento de que *big data* se refere ao volume de dados frequentemente não estruturados, cujo tamanho excede a capacidade de processamento convencional dos sistemas e arquiteturas de banco de dados, estando além da habilidade de ferramentas de *software* comuns de captura, armazenamento, gerenciamento e análise (MCKINSEY GLOBAL INSTITUTE, 2011; MANYIKA; CHUI; BROWN *et al*, 2011; DUMBILL, 2012). Nesse meio são proclamados como componentes necessários para a definição de *big data* os 5 Vs: volume, velocidade, variedade, veracidade, valor. Alguns substituem o termo veracidade por viabilidade, mas ambos dizem respeito à importância da autenticidade de dados e da análise dessa informação para que façam sentido (KANASHIRO *et al*, 2013).

Taurion (2013) ao se referir sobre o *big data* também chama a atenção para a característica da velocidade em que os dados são gerados. “Dados não tratados e analisados em tempo hábil, são dados inúteis, pois não geram informação. Na prática estima-se que cerca de 90% dos dados digitais disponíveis não estejam sendo adequadamente aproveitados” (TAURION, 2013). Dar sentido a massa de elementos não estruturados disponíveis na internet, a partir de cruzamentos e interpretação, é um esforço no sentido de buscar convertê-la em informação passível de agregar valor estratégico seja com objetivos mercadológicos ou não, isto é, gerar conhecimento e inteligência a partir de grandes quantidades de dados complexos. Para Bruno (2012) uma das promessas do *big data* envolve, inclusive, a revelação de correlações entre dados jamais imaginadas, tanto por aqueles que os geram, quanto por aqueles que os capturam.

Sobre o uso da massa de dados, Xexéo (2013, p.19) observa que “[...] guardar cada vez mais rápido é o desafio de hoje”. Já Mayer-Schonberger e Cukier (2013) observam que o segredo do valor dos dados está na aparente ilimitada reutilização em potencial: o custo/benefício. “Coletar dados é essencial, mas não o bastante, já que a maior parte do valor dos dados está no uso, não na detenção” (MAYER-SCHONBERGER; CUKIER, 2013, p.85).

Como já se discutiu, uma das possibilidades de uso de *big data* é por meio do monitoramento de mídias sociais digitais (ARRIGONI, 2013), dependendo do tamanho da captação, da complexidade de processamento e da análise desses dados. Esse tipo de uso, segundo especialistas, vem ganhando destaque no contexto de campanhas políticas, por alguns motivos. Destacam-se aqui alguns destes motivos: 1º porque permite o acompanhamento em tempo real do que os eleitores estão falando e pensando; 2º porque ajuda a conhecer o desempenho dos candidatos entre determinados grupos focais; 3º pode antecipar crises e 4º fornece informações potencialmente de interesse de diferentes áreas de um mesmo comitê de campanha. A esse respeito, Oliveira (2014) observa que, no contexto de uma campanha eleitoral, soluções analíticas baseadas em fontes de dados ricas e heterogêneas, incluindo redes sociais, podem ajudar a responder perguntas fundamentais que são estabelecidas para definir a direção e a estratégia do marketing eleitoral em todas as suas dimensões. Por exemplo, a partir da implementação de modelos estatísticos com o apoio de ferramentas de análise adequadas é possível identificar com alto nível de precisão onde e quem são os eleitores indecisos e quais as ações e recursos mais eficazes para influenciá-los a votar a favor de um determinado candidato ou contra (seu)s adversário(s). O conhecimento e o uso dessa informação na forma e no tempo certo da campanha pode ser decisivo no resultado final da eleição, tanto em confrontos para cargos majoritários como em disputas de vagas para cargos proporcionais.

Para ilustrar este trabalho apresenta-se aqui como parte empírica a campanha de reeleição do democrata Barack Obama intitulada *Obama for America* (OFA), que usou a analítica e inteligência de dados, a partir das informações dos eleitores americanos, com propósitos específicos de obter vantagens competitivas eleitorais em relação ao candidato adversário, o republicano Mitt Romney. Barack Obama foi eleito para o segundo mandato com 65.917,257 votos (51,01%) e o ex-governador de Massachusetts Mitt Romney recebeu 60.932,235 votos que corresponderam a 47,16% dos eleitores americanos.

5 BIG DATA NA CAMPANHA DE OBAMA

Segundo especialistas, a campanha política de Barack Obama intitulada *Obama for America* (OFA) de 2012, além do uso intensivo das mídias digitais teve outro diferencial, que entrou para a história das campanhas eleitorais. Essa campanha contou com o uso de tecnologias e inteligência analítica de dados, sendo largamente baseada em análise de *big data*, com destaque para técnicas de segmentação de mercado e de relacionamento com

eleitores. Usou-se na campanha a garimpagem de dados, que utiliza algoritmos⁷⁰ complexos para identificar padrões ocultos em enormes bancos de dados e converter essa informação em conhecimento útil para a tomada de decisões.

Mas qual eram os objetivos da campanha em 2012 com o uso do *big data*? A campanha de Obama girava em torno da: identificação e convencimento de pessoas certas a fazerem doações; aumento do comparecimento às urnas e busca de apoios à campanha, mantendo a fidelidade e ampliando o voto daqueles eleitores que votaram em Obama em 2008.

Para atingir a esses objetivos foram usadas as seguintes estratégias: (1) técnicas usadas por empresas para atrair consumidores, o *microtargeting*⁷¹ - marketing direcionado e microssegmentação; (2) compra de dados em empresas como Experian e a Acxiom; (3) enriquecimento de dados com informações de redes sociais (em especial Twitter e o Facebook); (4) segmentação dos eleitores de acordo com o perfil de navegação no site oficial da campanha; (5) criação de *cluster* de acordo com os gostos e características das pessoas (perfil político, etc); (6) direcionamento da campanha certa para a pessoa certa (*one to one*) e ações de comunicação individualizadas.

As redes sociais mostraram-se importantes ambientes de coleta de informações digitais para formar bancos de dados. Esses dados foram organizados e associados a outros dados, coletados de diversas fontes, como por telefone, mala direta, empresas de análise de consumo, entre outros.

A datificação das informações contidas nas redes sociais era transferida diretamente para o banco de dados da computação em nuvem, que funcionava como um servidor central⁷² utilizado pela equipe de análise de dados da campanha. Ao seguir a página oficial do presidente (www.barackobama.com), o eleitor automaticamente transferia diversas informações importantes diretamente à equipe de campanha, como idade, localização,

⁷⁰ Segundo Cormen um algoritmo pode ser visualizado como uma ferramenta para resolver um problema computacional bem especificado, onde este problema especifica em termos gerais o relacionamento entre o conjunto de valores como entrada e o conjunto de valores da saída desejada (CORMEN *et al*, 2002, p. 03).

⁷¹ *Microtargeting* é uma estratégia de marketing que usa os dados do consumidor e dados demográficos para identificar indivíduos ou pequenos grupos de indivíduos da mesma opinião e influenciar seus pensamentos ou ações.

⁷² “*Inside The Cave*” é o documento que relata como foi montada a campanha de Barack Obama em 2012.

escolaridade e profissão. Segundo Lopes (2012) o objetivo era montar uma infraestrutura de *data mining*⁷³ (mineração de dados) capaz de mostrar os eleitores-alvo e a melhor forma de fidelizá-los com a figura do presidente e suas propostas de campanha. As informações obtidas pelo sistema de informação montado a partir daí ajudaram o comitê a encontrar líderes comunitários, defensores das ideias de Obama e pessoas-chave na articulação política em cada cidade e local dos Estados Unidos.

5.1 Estrutura da Campanha OFA com o *big data*

Segundo documento oficial da campanha “*Inside the Cave*” (LEOHY, 2013) as estratégias de analítica de dados e de marketing político de segmentação começaram a ser desenvolvidas pelo Partido Democrata, anos antes das campanhas de Obama à presidência dos EUA. Foram realizados cadastros de pessoas favoráveis às suas ideias e a seus candidatos. Entre 2006 e 2008, o banco de dados dos democratas foi aperfeiçoado por uma empresa chamada *Voter Activation Network* (VAN) de Somerville, Massachusetts. A VAN, contratada pelo Comitê Nacional Democrata, reuniu um banco de dados dos 50 estados americanos e criou maneiras simples para os partidários acessarem os dados de formas limitadas e controladas por meio da internet. O MyBO (site oficial da campanha – My Barack Obama) e os sites de outros candidatos democratas em outras disputas fizeram links de grande eficiência com esse banco de dados recém-aperfeiçoado.

Ao final da primeira campanha de Obama em 2008 foram coletados 13 milhões de endereços de simpatizantes, sendo que, destes, três milhões efetuaram doações para ajudar na campanha. A segmentação das bases permitiu que comunicados específicos fossem enviados por e-mail e SMS (cerca de 1 milhão de cadastrados), com o objetivo de envolvê-las em diferentes áreas da campanha.

O especialista em dados Dan Wagner, que respondia como diretor de segmentação do Diretório Nacional do partido (DNC) tornou-se responsável pela coleta e análise de informações dos eleitores individuais por mala direta e telefone, trabalho que se iniciou em 2009 (ISSENBERG, 2012). O especialista pediu ao departamento de tecnologia do DNC para desenvolver software que pudesse transformar esses dados em tabelas, o que ele chamou de Gerente de Pesquisas (*Survey Manager*).

⁷³ *Data mining* é o processo de explorar grandes quantidades de dados à procura de padrões consistentes como regras de associação ou sequências temporais para detectar relacionamentos sistemáticos entre variáveis, detectando assim novos subconjuntos de dados (Wikipédia, 2014).

Como aponta Issenberg (2012) durante os 18 meses finais da campanha de 2012, um grupo de análise de dados coordenado por Dan Wagner coletou e combinou milhares de dados sobre todo o público votante dos Estados Unidos. A equipe de 2012 tinha cinco vezes o número de analistas de dados do que a eleição anterior (2008), sendo composta principalmente por cientistas de dados, estatísticos, matemáticos, especialistas em mídias digitais e tecnologia. Além disso, contava com o especialista em *big data*, Rayid Ghani.

A campanha de Obama estruturava-se em duas frentes e muitos dos profissionais vinham de empresas como Google, Facebook ou Amazon. A campanha digital comandada por Teddy Golf e Joe Rospars, cuja função era cuidar do público da campanha, tendo como atividades a arrecadação de fundos, o gerenciamento de e-mails e de redes sociais. E outra tecnológica e analítica, comandada pelo diretor de Inovação e Integração, Michael Slaby⁷⁴. Para trazer mais pessoas especializadas em mídias digitais, a campanha buscou, em junho de 2011, em suas listas de doadores, pessoas com perfis e habilidades em tecnologia da informação, dados e *design*. A cada perfil encontrado eram disparados *e-mails* com mensagem de engajamento profissional à campanha. O brasileiro estrategista em marketing digital Ricardo Cappra⁷⁵ era um dos integrantes da equipe de comunicação digital, que ajudou a gerenciar e comunicar a marca Obama nas mídias digitais. Além dele, outro brasileiro também trabalhou como voluntário. Maurício Moura⁷⁶, especialista em *microtargeting*, analista político e professor da *The George Washington University*, com experiências em campanhas eleitorais brasileiras colaborou também na análise de dados e informações para a campanha de reeleição de Obama.

5.2 Dado e informação na campanha: da coleta a análise

A equipe da campanha OFA conhecia o nome de cada um dos 69.456.897 americanos que votaram no presidente Barack Obama em 2008. O desafio em 2012 era garantir que ao menos esse mesmo número de americanos votaria novamente para sua reeleição e, além disso,

⁷⁴ Slaby ficou conhecido como o “guru” digital de Barack Obama, chefe de tecnologia da campanha que levou o então senador a seu primeiro mandato presidencial, em 2008.

⁷⁵ Ricardo Cappra já trabalhou em projetos estratégicos para marcas como Barcelona FC, Obama (como integrante da equipe que elegeu o presidente em 2008), Microsoft, Blackberry, entre outros. É professor de pós-graduação de Digital Branding em várias unidades do ensino. É vice-presidente da Associação Internacional de Investigadores em Branding, que se propõe a estudar o comportamento das marcas (FERNANDES, Revista Expansão, 2012).

⁷⁶ Maurício Moura é responsável pela analítica de dados (*big data*) da campanha eleitoral de 2014, para governo do Rio de Janeiro, do candidato Luiz Fernando Pezão (PMDB).

registrar e mobilizar novos eleitores, especialmente nos estados americanos que apresentavam crescimentos demográficos.

Toda a captação de dados na campanha para a reeleição foi integrada, reunindo-se um volume gigantesco de dados sobre eleitores americanos, chegando a 4 gigabytes processados por segundo, 8,5 bilhões de requisições ao banco de dados e 180 terrabytes de armazenamento em três data centers.

Mas de onde vinham esses dados? De que forma? Segundo documento oficial da campanha OFA “*Inside the Cave*” (LEOHY, 2013) os dados vinham de diversas fontes tanto da web como fora dela, incluindo campanhas passadas, empresas especializadas (como Experian e Acxiom). Os dados também eram coletados por cabos eleitorais, que faziam visitas nas casas dos eleitores indecisos - eles usavam *smartphones* para as entrevistas, através do *app* os visitantes conseguiam saber todos os dados dos eleitores e encaminhavam instantaneamente para o banco de dados central da campanha-, foram usadas informações também de redes sociais e dos assinantes do aplicativo *Obama for America* (barackobama.tumblr.com/). O banco de dados da campanha era abastecido com informações sobre os modos como as pessoas abriam o e-mail, tuitavam ou publicavam no Facebook e sobre mudanças demográficas. Seus “rastros digitais” eram datificados e armazenados em bancos de dados centrais.

Pesquisa⁷⁷ realizada na época pela *Global Strategy Group* (GLOBAL STRATEGY GROUP, 2012) apontava que a internet oferecia aos candidatos a oportunidade real para alcançar os eleitores indecisos. Segundo a pesquisa em, 2012 eram 90% das casas com acesso à internet em alta velocidade, além do uso mais disseminado de redes sociais do que em 2008, o que de certa forma facilitava ainda mais o contato e o direcionamento da campanha com os eleitores americanos.

Para analisar os dados coletados foram criados algoritmos⁷⁸ e aplicativos. Os métodos eram baseados em pesquisas eleitorais, partindo de um extenso banco de dados que comparava as informações de diversos institutos americanos na previsão real dos resultados. Os dados ficavam armazenados em uma plataforma de computação em nuvem, proporcionada através da internet pela Amazon.com (AWS), que oferecia diversos serviços para atender aos

⁷⁷ De acordo com a pesquisa, 500 eleitores foram entrevistados na Flórida, Ohio, Nevada, Virgínia e Wisconsin. A pesquisa foi realizada entre 02 de outubro e 07 de outubro de 2012 (GLOBAL STRATEGY GROUP, 2012).

objetivos da campanha. Ferramentas de análises como *Google Analytics*⁷⁹ foram fundamentais a partir de então, contribuindo para a otimização rápida, em tempo real, especificamente durante os debates presidenciais e vice-presidenciais. Usando relatórios em tempo real, a equipe da campanha percebeu rapidamente os termos de busca relacionados aos debates e trouxe a maioria dos espectadores para o site de Obama. Assim, eram colocados anúncios de busca nestes termos no Google para direcionar os usuários a fatos sobre a posição do presidente em questões-chave.

O objetivo da tarefa de *analytics* é executar a análise preditiva dos dados por meio da execução de *mining* (minerações). De acordo com Oliveira (2013) e Tavares (2014), inicialmente são tratados os dados com o uso de técnicas estatísticas, para separação e reunião de conjuntos (denominado de fase de *discovery*).

O diretor de Marketing Digital da OFA, Nate Lubin, em entrevista ao *Google Analytics Case Study* (GOOGLE, 2013), credita ao *Google Analytics* parte do sucesso da campanha de Obama. Segundo ele, a ferramenta foi essencial para que a equipe pudesse obter informações e tomar decisões que influenciaram o curso das estratégias da campanha. "Ao longo da campanha, o *Google Analytics* nos ajudou a democratizar nossos dados na web, e permitiu que todas as nossas equipes tomassem decisões importantes rapidamente" (LUBIN, 2013). Nas observações feitas à campanha OFA verificou-se que boa parte das estratégias para obter doações para a campanha baseou-se no uso de *big data*. Por exemplo, foi possível identificar que havia em Nova York um público de alto poder aquisitivo possivelmente interessado em participar em um jantar com celebridades, especialmente com a atriz Sara Jessica Parker. Daí nasceu um concurso de doação para a campanha, em que quem doasse poderia jantar com a atriz e com Obama. A estratégia também foi aplicada com o ator George Clooney, em que mulheres de 40 a 42 anos manifestavam interesse em um jantar na companhia da celebridade. Em torno de 150 pessoas participaram do jantar na residência do ator em Los Angeles, Califórnia, cuja entrada custou 40 mil dólares. O evento arrecadou US\$ 15 milhões para a campanha de Obama.

Do mesmo modo, foram identificados possíveis eleitores doadores de quantias pequenas (até U\$ 200.00), que representavam 95% dos apoiadores. Para isso foram criadas

⁷⁹ *Google Analytics* é um serviço gratuito e é oferecido pela [Google](#) no qual, ao ativar-se o serviço por intermédio de uma [conta do Google](#), e ao cadastrar-se um [site](#) recebe-se um código para ser inserido na página cadastrada e, a cada exibição, [estatísticas](#) de visitação são enviadas ao sistema e apresentadas ao dono do site. Foi criado principalmente como um auxílio aos [webmasters](#) para otimizar seus sites para campanhas de [marketing](#) (GOOGLE, 2013).

experiências com *sites* e plataformas para aparelhos móveis como *smartphones* e *tablets*. “Criamos um sistema de “doação rápida” em que, para doar uma segunda vez, bastava um clique. Queríamos que as pessoas tivessem uma experiência o mais simples possível”, afirma Slaby em uma entrevista para uma revista brasileira⁸⁰ (TIRABOSCHI, 2013). Nessa campanha Obama arrecadou quase US\$ 700 milhões de forma online, de aproximadamente quatro milhões de eleitores.

Outro exemplo de segmentação e individualização da campanha, proporcionados pelo *big data* foi a identificação de eleitores católicos contrários ao aborto, que automaticamente ficariam excluídos do envio de *e-mails* sobre os direitos das mulheres ao aborto. Cada assunto era tratado de acordo com o perfil do eleitor, a partir das análises de informações que vinham do banco de dados unificados.

No entanto, essas estratégias de análise de dados de eleitores foi muito criticada na época por especialistas americanos. Andrew Rasiej⁸¹ observava que o grande problema do levantamento de toda essa base de dados durante a eleição era o risco de desrespeitar os direitos ligados à privacidade dos eleitores. Essas observações reforçam debates teóricos sobre vigilância, controle e privacidade em relação aos usos que se fazem das informações e dados de eleitores, principalmente vindas de ambientes de sites de relacionamentos digitais, conforme já discutidos neste trabalho. No caso da campanha de Obama, o monitoramento, processamento e a armazenagem, bem como os usos dos dados levam a preocupações no que se refere à privacidade. Dados que antigamente vistos isoladamente poderiam gerar pouca informação, se combinados podem gerar muita informação sobre pessoas e seu provável comportamento – quase sempre à revelia das mesmas.

6 CONSIDERAÇÕES FINAIS E APONTAMENTOS FUTUROS

Este trabalho evidenciou que o uso de *big data* mostra-se cada vez mais relevante no âmbito de campanhas políticas eleitorais, demonstrando que o uso de informação e dado é elemento essencial, usados como estratégias para disputas políticas. No entanto, cada vez mais os indivíduos passam a ser monitorados e vigiados não só por empresas com objetivos comerciais, mas também por políticos com intenções eleitorais, por meio dos “rastros

⁸⁰ Revista IstoÉ online, entrevista postada em 11 de junho de 2013. “Os governos deveriam oferecer oportunidades para as pessoas participarem”.

⁸¹ Fundador do *Personal Democracy Forum* (personaldemocracy.com).

digitais” (LEMOS, 2005; BRUNO, 2013) em que os eleitores/usuários deixam principalmente em suas ações e mensagens postadas em mídias sociais como Facebook e Twitter.

A campanha de 2012 do presidente norte-americano Barack Obama à reeleição é apontada como inovadora e estratégica, neste sentido, ao privilegiar o uso de tecnologias e análises de grandes volumes de dados de eleitores, capturadas de vários ambientes, principalmente das mídias sociais digitais. O *big data* possibilitou uma segmentação bem mais precisa do público alvo/eleitor americano, alcançando o nível do indivíduo, reduzindo assim o papel de estratégias políticas massivas baseadas somente em pesquisas políticas tradicionais. Com o *big data*, a equipe de Obama identificou o perfil do eleitor indeciso, o que ele precisava ouvir, que argumentos o fariam sair de casa e votar. As pesquisas eram diárias, via internet, com dezenas de milhares de eleitores, podendo-se perceber, precisamente, as mais tênues flutuações. Jamais se conduziu uma campanha eleitoral com tanta informação. Nunca tantas decisões foram tomadas com tanta tranquilidade.

Muito além da simples divisão do eleitorado por sexo, idade e região, para se ter uma segmentação sólida é necessário haver o maior número possível de informação acerca do eleitorado, suas características, gostos, necessidades e costumes. E a partir do monitoramento e gerenciamento de informação e dados, isso se tornou possível. Por outro lado, o uso do *big data* também implica questões éticas e de privacidade dos eleitores, porque se apropria de dados e informações de diversas fontes e ambientes. Como observa Bruno (2014), a participação do usuário é hoje um motor e um capital valioso da internet, mas há tensões nos mais diferentes setores, principalmente na política.

Dessa forma, a campanha de Obama de 2012 – *Obama for America* (OFA) - pode ser estudada sob diversos pontos de vista. Seja como uma campanha de ações de inteligência analítica de dados, utilizando-se de elementos chave do mundo dos negócios, ou como uma campanha que se aproveita de todas as informações dos eleitores americanos, exercendo ações de “monitoramento”, vigilância, coletando dados pessoais. Notamos que se tratou de uma vigilância que operou mais com os sistemas informacionais, sobre dados e rastros e mais com a finalidade de antecipar tendências, preferências e interesses políticos e não com a finalidade de corrigir e reformar. Assim, o campo das campanhas políticas talvez seja aquele em que pode se analisar mais diretamente a importância do monitoramento da informação e dado. Aqui, torna-se possível conhecer desejos dos eleitores, as críticas feitas por eles e testar a receptividade a posturas e ideias. Além disso, é também possível monitorar adversários políticos passando a conhecer melhor seus pontos de vantagem e desvantagem com o eleitorado. No caso da campanha OFA, o sistema de análise de dados, que se alimentavam e

geravam informações em alta velocidade, em busca de tendências, cenários, motivações e reações, permitiram inúmeros *insights* para a equipe da campanha, mas as principais foram o alcance de cifras recordes de arrecadação; definição e direcionamento com precisão dos meios nos quais se investiu em propaganda, especialmente nos Estados americano onde a vitória de Obama era mais incerta, possibilitando a personalização da campanha de acordo com seus públicos e tipos de eleitores. Essa capacidade de obtenção e processamento de informações sobre os eleitores foi decisiva ao desenvolver uma campanha nacional como uma eleição local, em que os interesses individuais eram conhecidos e abordados e a partir daí saíam suas estratégias de comunicação política na busca pelo convencimento do voto do eleitorado. Como afirma Figueiredo (2008) em comunicação política mais do que quantidade, o que vale é a qualidade da mensagem. Foi esse o fator diferencial da campanha à reeleição de Obama, ao utilizar de forma inédita a tecnologia com o marketing político no cruzamento de dados de uma enorme base de dados unificada chamada *big data*.

No que se refere ao contexto brasileiro, em que esta pesquisa se desdobrará, o *big data* como soluções analíticas em disputas eleitorais constituem, ainda, uma estratégia mercadológica pouco explorada no Brasil e sem grandes discussões no campo científico. Embora tenha aumentado a busca por dados e informações públicas, vindas principalmente de monitoramentos feitos em sites de relacionamentos como Facebook e Twitter, existem poucas evidências de utilização desses dados em modelos de análises mais sofisticados (*big data*), que possam antecipar e influenciar o voto do eleitor com menor custo e maior eficácia. Para alguns especialistas da área de marketing político e com experiências em grandes campanhas⁸², somente comitês com grandes estruturas poderão investir, ao menos a médio prazo, na análise dos grandes conjuntos de dados e utilizar diversas tecnologias para apoiar suas estratégias de comunicação e marketing eleitoral. Outros⁸³ observam que é necessário levar em conta, que a experiência americana de 2012 com *big data*, transformada em *case* de marketing político eleitoral, pode não ser adequada ao perfil do eleitorado, às reais condições estruturais dos comitês e ao sistema eleitoral brasileiro.

⁸² Entrevista concedida à pesquisa, em junho de 2014, por Renato Pereira, sócio proprietário da agência de comunicação e marketing Prole, do Rio de Janeiro.

⁸³ Entrevista concedida à pesquisa, em julho de 2014, por Carlos Manhanelli, professor e consultor na área de comunicação e marketing político. Autor de diversos livros na área de campanhas eleitorais brasileiras.

Chega-se a conclusões iniciais de que os estudos que contemplam *big data* e seu uso em campanhas políticas eleitorais são recentes ou praticamente inexistentes. O *case* de reeleição do presidente Barack Obama, apresentado nesse trabalho, serve de exemplo para a pesquisa que está em andamento no contexto brasileiro, principalmente com foco na disputa eleitoral de 2014. Espera-se nessa perspectiva, deixar contribuições que enriqueçam o arcabouço teórico no campo da Ciência da Informação, bem como em outras áreas em que o tema possa ser contemplado. Mas cabe ressaltar que é difícil tirar conclusões consistentes e imediatas de um processo tão novo como o uso do *big data* em campanhas eleitorais, mesmo num contexto internacional.

REFERÊNCIAS

- ARAÚJO, V.M.R.H. de; FREIRE, I.M. **A rede Internet como canal de comunicação, na perspectiva da Ciência da Informação**. Transinformação, v.8, n.2, 1996.
- ARRIGONI, R. **Big data, a ciência do grande**. Disponível em: <<http://convergecom.com.br/tiinside/21/06/2013/big-data-a-ciencia-do-grande/#.Uoy0-sRJ6s1>> Acesso em 10 de março de 2014.
- AWS Case Study: **Obama for America Campaign 2012**. Harper Reed. Disponível em: <<http://aws.amazon.com/pt/solutions/case-studies/obama/>>. Acesso em 20 de maio de 2014.
- BAUMAN, Z. **44 cartas do mundo líquido moderno**. Rio de Janeiro: Jorge Zahar Editor, 2011.
- BRUNO, F. **Monitoramento, classificação e controle nos dispositivos de vigilância digital**. Revista Famecos – Porto Alegre nº 36 – Agosto de 2008.
- BAUMAN, Z. **Máquinas de Ver, Modos de Ser: vigilância, tecnologia e subjetividade**. Porto Alegre: Sulina, 2013.
- BAUMAN, Z. **Rastros Digitais sob a perspectiva da teoria ator-rede**. Revista FAMECOS. Porto Alegre, 2012, V. 19 n3 pp. 681-704. Disponível em: <<http://revistaseletronicas.pucrs.br/ojs/index.php/revistafamecos/article/viewFile/12893/8601>>. Acesso em 23 de julho de 2014.
- CASTELLS, M. **A sociedade em rede**. São Paulo: Paz e Terra, 1999.
- CORMEN. T. H. *et al.* **Algoritmos: teoria e prática**. Rio de Janeiro: Elsevier, 2002.
- COX, M.; ELLSWORTH, D. **Application-controlled demand paging for out-of-core visualization**. Proceedings of the 8th IEEE Visualization '97 Conference. 1997. Disponível em: <http://www.evl.uic.edu/cavern/rg/20040525_renambot/Viz/parallel_volviz/paging_outofcore_viz97.pdf>. Acesso em 21 jul 2014.

DAVENPORT, T; PRUSAK, L. **Conhecimento empresarial:** como as organizações gerenciam o seu capital intelectual. Rio de Janeiro: Campus, 1998.

DUMBILL, E. **What is Big Data?** In: O'Reilly Media Inc. Big Data Now: current perspectives. O'Reilly Media:California. 2012. Disponível em: <<http://www.oreilly.com/data/free/files/big-data-now-2012.pdf>>. Acesso em: 03 maio 2014

FIGUEIREDO, R. **Marketing político em tempos modernos.** Rio de Janeiro: Konrad-Adenauer – Stiftung, 2008.

FERNANDES, V. **O estrategista das marcas.** Revista Expansão. Maio 2012.

FOUCAULT, M. **Vigiar e Punir.** 41. ed. Petrópolis. Vozes, 2013.

FUCHS, C. **Internet and Surveillance.** The Challenges of Web 2.0 and Social Media. New York: Routledge, 2011.

GOOGLE Analytics Blog. **Obama for America uses Google Analytics to democratize rapid, data-driven decision making.** Disponível em <<http://analytics.blogspot.com.br/2013/08/obama-for-america-uses-google-analytics.html>> Acesso em 17 jun 2014.

ISSENBERG, S. **How President Obama's campaign used big data to rally individual voters.** Disponível em: <<http://www.technologyreview.com/featuredstory/509026/how-obamas-team-used-big-data-to-rally-voters/>>. Acesso em 05 de maio de 2014.

KANASHIRO, M.; BRUNO, F.; EVANGELISTA, R.; FIRMINO, R. **Maquinaria da privacidade.** RUA [online]. 2013, no. 19. Volume 2.

LE MOS, A. **Cibercultura e mobilidade:** a era da conexão. In: XXVIII Congresso Brasileiro de Ciências da Comunicação –Rio de Janeiro/RJ, 2005. Disponível em: <<http://www.intercom.org.br/papers/nacionais/2005/resumos/r1465-1.pdf>> Acesso em: 03 de jun. 2013.

LEOHY, M. P. **“Inside The Cave” Documents Obama Campaign Tech Superiority Over Gap.** Disponível em: <www.breitbart.com/Big-Government/2013/01/07/Inside-the-Cave-Documents-Obama-Campaign-Tech-Superiority-Over-GOP>. Acesso em 10 de março de 2014.

LOPES, A. **Com quantos Harper Reeds se disputa uma eleição nos EUA - See more at.** Disponível em:<<http://webinsider.com.br/2012/10/08/com-quantos-harper-reeds-se-disputa-uma-eleicao-nos-eua/#sthash.BMzedc29.dpuf>>. Acesso em 10 de março de 2014.

LUBIN, N. **Obama for America uses Google Analytics to democratize rapid, data-driven decision making.** Google Analytics. Case Study, 2013. Disponível em: <<http://static.googleusercontent.com/media/www.google.si/pt-BR/si/analytics/customers/pdfs/obama-2012.pdf>>. Acesso em 03 de abril de 2014.

MANYIKA, J.; CHUI, M.; BROWN, B.; et al. **Big Data:** The next frontier for innovation, competition, and productivity. McKinsey Global Institute, 2011.

MAYER-SCHONBERGER, V.; CUKIER, K.; **Big Data:** Como Extrair Volume, Variedade, Velocidade e Valor da Avalanche de Informação Cotidiana. Editora Campus: Rio de Janeiro, 2013.

MACKINSEY GLOBAL INSTITUTE. **Big Data:** The next frontier for innovation, competition and productivity. Maio de 2011. Disponível em: <http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation>. Acesso em 10 de jan de 2014.

OLIVEIRA, A. **Data Science and Data Analytics.** Palestra apresentada no 1º. EMC Summer School on Big Data. EMC/NCE/UFRJ. Rio de Janeiro. 2013

POSTER, M. **The mode of information.** Post-structuralism and social context. Chicago: University of Chicago Press, 1990.

RIESMAN, D. **A multidão solitária.** São Paulo: Ed. Perspectiva, 1995.

ROSE, N. **Inventing Our Selves:** Psychology, Power and Personhood. Cambridge: Cambridge University Press, 1998.

SCHONS, C. H. **O volume de informações na internet e sua desorganização: reflexões e perspectivas.** Inf, Londrina, v. 12, n.1, jan./jun. 2007.

TAVARES, E. **BIG DATA in Business.** 12 de maio de 2014. Palestra apresentada no 2o. EMC Summer School on Big Data. EMC/NCE/UFRJ. Rio de Janeiro. 2014. Disponível em: <http://2014.emcbigdataschool.nce.ufrj.br/images/presentations/Apresentacao_Elaine_Tavares.pdf>. Acesso em 30 maio 2014.

TAURION, C. **Big Data:** nova fronteira em gerenciamento de dados. Disponível em: <www.ibm.com/developerworks/mydeveloperworks/blogs/ctaurion/entry/big_data_nova_frenteira_em_gerenciamento_de_dados?lang=en>. Acesso em 17 jan. 2013.

TAURION, C. **Big Data.** Brasport. 2013. Edição Digital disponível pelo Kindle.

TERRA, J. C. C. **Gestão do conhecimento:** o grande desafio empresarial. São Paulo: Negócio Editora, 2000. 283p.

TIRABOSCHI, J. **Os governos deveriam oferecer oportunidades para as pessoas participarem.** Entrevista IstoÉ online, 11 jun 2013. Disponível em: http://www.istoe.com.br/reportagens/306023_OS+GOVERNOS+DEVERIAM+OFERECER+OPORTUNIDADES+PARA+AS+PESSOAS+PARTICIPAREM+>. Acesso em 20 de maio de 2014.

VAIDHYANATHAN, S. **A googlelização de tudo (e por que devemos nos preocupar):** a ameaça do controle total da informação por meio da maior e mais bem-sucedida empresa do mundo virtual. Sao Paulo: Cultrix, 2011.

WORD ECONOMIC FORUM. **Big Data, Big Impact: New possibilities for International Development.** Disponível em <<http://www.weforum.org/reports/big-data-big-impact-new-possibilities-international-development>>. Acesso em 20 junho 2014.

XEXÉO, G. Big data: computação para uma sociedade conectada e digitalizada. Ciência Hoje, v. 51, n. 306, ago. 2013.