



Organizações & Sociedade

ISSN: 1413-585X

ISSN: 1984-9230

Escola de Administração da Universidade Federal da Bahia

Oliveira, Daniel José Silva; Bermejo, Paulo Henrique de Souza  
Mídias sociais e administração pública: análise do sentimento  
social perante a atuação do Governo Federal brasileiro  
Organizações & Sociedade, vol. 24, núm. 82, 2017, Julho-Setembro, pp. 491-508  
Escola de Administração da Universidade Federal da Bahia

DOI: 10.1590/1984-9240827

Disponível em: <http://www.redalyc.org/articulo.oa?id=400657500007>

- Como citar este artigo
- Número completo
- Mais informações do artigo
- Site da revista em [redalyc.org](http://redalyc.org)

UAEM [redalyc.org](http://redalyc.org)

Sistema de Informação Científica Redalyc

Rede de Revistas Científicas da América Latina e do Caribe, Espanha e Portugal

Sem fins lucrativos acadêmica projeto, desenvolvido no âmbito da iniciativa  
acesso aberto

## MÍDIAS SOCIAIS E ADMINISTRAÇÃO PÚBLICA: ANÁLISE DO SENTIMENTO SOCIAL PERANTE A ATUAÇÃO DO GOVERNO FEDERAL BRASILEIRO

Daniel José Silva Oliveira\*  
Paulo Henrique de Souza Bermejo\*\*

### Resumo

Este estudo procurou identificar como a análise de sentimento, baseada em textos extraídos de mídias sociais, pode ser um instrumento de mensuração da opinião pública sobre a atuação do governo, de forma a contribuir para a avaliação da administração pública. Trata-se de um estudo aplicado, interdisciplinar, exploratório, qualitativo e quantitativo. Foram revisadas as principais formulações teóricas e conceituais acerca do tema e realizadas demonstrações práticas, utilizando-se uma ferramenta de mineração de opinião que proporcionou precisão satisfatória no processamento de dados. Para fins de demonstração, foram selecionados temas que motivaram a realização da onda de protestos que envolveu milhões de pessoas no Brasil em junho de 2013. Foram coletadas, processadas e analisadas, aproximadamente, 130 mil mensagens postadas no Facebook e no Twitter sobre esses temas em dois períodos distintos. Por meio dessa investigação, observou-se que a análise de sentimento pode revelar a opinião polarizada dos cidadãos quanto à atuação do governo.

**Palavras-chave:** Análise de sentimento. Mineração de opinião. Administração pública. Mídias sociais. Opinião pública.

## SOCIAL MEDIA AND PUBLIC ADMINISTRATION: SOCIAL SENTIMENT ANALYSIS ABOUT THE PERFORMANCE OF THE BRAZILIAN FEDERAL GOVERNMENT

### Abstract


This study sought to identify as sentiment analysis, based on texts taken social media can be a measuring instrument of public opinion on the government's performance in order to contribute to the evaluation of public administration. This is an applied study, interdisciplinary, exploratory, qualitative and quantitative. The main theoretical and conceptual formulations on the subject were reviewed and conducted practical demonstrations using an opinion mining tool which provided satisfactory precision in data processing. For demonstration purposes, themes were selected that motivated the wave of protests involving millions of people in Brazil in June 2013. They were collected, processed and analyzed approximately 130,000 messages posted on Facebook and Twitter on these topics in two distinct periods. Through this analysis, it was observed that the sentiment analysis can reveal the polarized opinions of citizens about the government's performance.

**Keywords:** Sentiment analysis. Opinion mining. Public administration. Social media. Public opinion.

\**Doutorando em Administração pelo Centro de Pós-Graduação e Pesquisas em Administração da Universidade Federal de Minas Gerais. Mestre em Administração Pública pela Universidade Federal de Lavras. E-mail: djso@ufmg.br*

\*\**Doutor em Engenharia e Gestão do Conhecimento pela Universidade Federal de Santa Catarina. Professor do Departamento de Administração da Faculdade de Economia, Administração, Contabilidade e Gestão Pública da Universidade de Brasília. E-mail: paulobermejo@unb.br*

## Introdução

 grande volume de dados postados na internet, por meio das mídias sociais, está produzindo mudanças importantes na maneira pela qual as pessoas se comunicam e compartilham conhecimentos e emoções que influenciam o comportamento social, político e econômico em todo o mundo (MONTORO; MARTÍNEZ-BARCO; BALAHUR, 2012).

De acordo com Mostafa (2013), as opiniões expressas nas mídias sociais desempenham um papel importante em influenciar o comportamento da opinião pública em diversas áreas. A título de exemplo, nas áreas de política e administração pública, essas mídias têm o poder de disseminar opiniões que podem resultar tanto em melhorias nos serviços públicos quanto em protestos motivados pela insatisfação dos cidadãos com os governos (PAPACHARISSI; OLIVEIRA, 2012; O'CALLAGHAN et al., 2014).

A popularidade dessas mídias é considerável, tornando-as uma significativa fonte de informações que pode ser utilizada na melhoria dos serviços públicos (MERGEL, 2013). Portanto, além dos canais diretos em que o cidadão pode emitir sua opinião, como e-mails, ouvidorias e portais de reclamação, as mídias sociais (*blogs, microblogs, wikis, fóruns, redes sociais eletrônicas*, entre outras) podem ser utilizadas para promover serviços públicos participativos e orientados para o cidadão (SOBACI; KARKIN, 2013). Porém, o grande volume de informações que circula atualmente nessas plataformas demanda tecnologias que permitam sua análise (BONSÓN et al., 2012). Uma opção que vem ganhando espaço entre os pesquisadores é a análise de sentimento (PANG; LEE, 2008; SOBKOWICZ; KASCHESKY; BOUCHARD, 2012; OLIVEIRA; BERMEJO; SANTOS, 2015; 2016).

A análise de sentimento ou mineração de opinião pode ser considerada uma técnica automatizada de descoberta de conhecimento que visa encontrar padrões ocultos em um grande número de informações textuais, inclusive, mídias sociais (MOSTAFA, 2013). Seu objetivo é criar uma base de conhecimento, contendo opiniões (positivas, negativas e neutras) de uma forma mais estruturada e explícita que expressem sentimentos, avaliações e percepções das pessoas em relação a qualquer tema (FORTUNY et al., 2012; SOBKOWICZ; KASCHESKY; BOUCHARD, 2012).

Nesse contexto, o presente estudo foi realizado com o objetivo de avaliar se a análise de sentimento pode ser utilizada como um instrumento de mensuração da opinião pública sobre a atuação do governo, contribuindo para que ele possa atuar em consonância com o interesse público, evitando impopularidade ou mesmo conflitos. Assim, a questão que norteia este estudo é: a análise de sentimento em mídias sociais pode mensurar a opinião dos cidadãos sobre a atuação do governo e, dessa forma, contribuir para a melhoria dos serviços públicos?

Para responder à questão, esta pesquisa explorou a técnica de análise de sentimento em mídias sociais a partir de temas relevantes para a administração pública. Segundo Bonsón, Royo e Ratkai (2015) e Ferro et al. (2013), são necessários novos estudos que procurem avaliar os comentários dos cidadãos em mídias sociais, gerando conhecimento sobre a opinião da sociedade perante a atuação do governo. Este estudo procurou preencher essa lacuna.

A pesquisa se justifica, pois, no cenário político atual, faz-se necessário entender o uso de aplicativos de mídia social como canais aceitáveis para as interações entre governo e seus diversos públicos que podem, potencialmente, fazer a diferença nas percepções e sentimentos dos cidadãos em relação ao governo (MERGEL, 2013). Apesar da crescente atenção à análise de conteúdo gerado pelos usuários das mídias sociais, "os pesquisadores têm pouco conhecimento sobre como aplicar os métodos de mineração de conteúdo" (YOON; ELHADAD; BAKKEN, 2013, p. 122), faltando métricas adequadas que identifiquem, nessas mídias, os impactos das ações dos governos (MERGEL, 2013; HOFMANN et al., 2013).

Este artigo está estruturado da seguinte forma: na próxima seção, é discutida a relação entre cidadãos e governos nas mídias sociais. Na terceira seção, é feita uma correlação entre a opinião pública e as mídias sociais. Na quarta parte, a literatura recente sobre análise de sentimento foi explorada. Na quinta seção, são apresentados

alguns trabalhos importantes relacionados ao tema. Na sexta seção, são demonstrados os procedimentos metodológicos para a aplicação da análise de sentimento. Os resultados são apresentados na sétima parte e, finalmente, na oitava seção, o estudo é concluído.

## *A relação cidadão-governo nas mídias sociais*

Os cidadãos estão cada vez mais utilizando as mídias sociais para se comunicarem com seus familiares, amigos, colegas de trabalho, empresas e até mesmo com o governo (KAVANAUGH et al., 2012). Essas mídias compreendem um conjunto de tecnologias que permite a comunicação entre os cidadãos e o governo de forma colaborativa (OLIVEIRA; WELCH, 2013). Exemplo disso é o número crescente de usuários de mídias sociais interessados em vários assuntos relacionados à política e à administração pública (SOBKOWICZ; KASCHEKY; BOUCHARD, 2012). Eles cobrem todo o espectro de interessados, desde os cidadãos que expressam as suas opiniões sobre as questões do cotidiano, até políticos que utilizam esse meio para comunicar suas ideias, além de jornalistas que criticam o governo ou o próprio governo como forma de divulgar suas ações (TSYTSARAU; PALPANAS, 2012).

Até mesmo em regimes antidemocráticos o habitante do ciberespaço reúne graus de liberdade para manifestar sua opinião, fazendo com que as mídias sociais se tornem ambientes propícios para a difusão de opiniões favoráveis ou contrárias às ações dos governos (CÁCERES, 2011). Estudos recentes demonstram que as mídias sociais foram utilizadas por cidadãos de vários países na articulação de manifestações contra seus governos, como a Primavera Árabe, que foi marcada por protestos e conflitos no Oriente Médio e no norte da África, destacando-se os conflitos ocorridos no Egito e na Síria (CHEN, 2011; GLASS; COLBAUGH, 2012; LIM, 2012; PAPACHARISSI; OLIVEIRA, 2012; O'CALLAGHAN et al., 2014). Em junho de 2013, uma série de protestos também foi organizada no Brasil por meio das mídias sociais, levando às ruas milhões de pessoas em todo o país. A evolução dessa onda de protestos pode ser descrita em três etapas (SINGER, 2013). Na primeira etapa, que ocorreu entre os dias 6 e 13 de junho, a principal causa dos protestos foi o aumento do preço das passagens do transporte público na cidade de São Paulo. A repressão desmedida por parte da polícia militar durante esses protestos atraiu a atenção de um grande público, que incorporou às manifestações outras reivindicações, como combate à corrupção, melhorias nos setores de saúde e educação, entre outros. Depois, motivadas por várias causas, milhões de pessoas se organizaram por meio das mídias sociais e foram às ruas entre os dias 17 a 20 de junho, período que compreendeu a segunda etapa dos protestos. A terceira e última etapa ocorreu do dia 21 até o final do mês de junho, com protestos motivados por questões mais específicas sobre a atuação do Governo Federal e os gastos com a Copa do Mundo da FIFA de 2014. Conforme dados de uma pesquisa realizada pelo Instituto Brasileiro de Opinião Pública e Estatística (IBOPE, 2013a), em um dos protestos de maior vulto, 62% dos manifestantes souberam do evento pelo Facebook, e 75% destes convocaram outras pessoas para participar do movimento por meio das plataformas Facebook e Twitter. Outra pesquisa realizada pelo Ibope (2013b) revelou que 63% dos manifestantes tinham menos de 30 anos de idade, 43% tinham curso superior completo e apenas 15% tinham renda familiar abaixo de dois salários mínimos. Apesar do perfil dos manifestantes representar um estrato social composto predominantemente por jovens de classe média, esse exemplo demonstra que as mídias sociais são ferramentas poderosas que possibilitam a difusão de informações e a discussão de questões de interesse público com uma amplitude jamais alcançada por outros meios de comunicação (SOBKOWICZ; KASCHEKY; BOUCHARD, 2012).

Muitas organizações, principalmente as grandes corporações, já perceberam essa amplitude das mídias sociais. Cada vez mais elas investem em pesquisas para avaliar as opiniões que as pessoas expressam sobre suas marcas nessas plataformas (CAMBRIA et al., 2013). Por outro lado, a administração pública é um dos campos

que são mais claramente distantes da dinâmica do ciberespaço (CÁCERES, 2011). Para Hofmann et al. (2013), os benefícios a serem obtidos com a utilização das informações disponíveis nas mídias sociais não estão limitadas às organizações do setor privado, mas devem, também, estender-se aos governos. De acordo com Sobkowicz, Kaschesky e Bouchard (2012), até recentemente os formuladores de políticas públicas não tinham muitos indicadores disponíveis sobre a opinião dos cidadãos, e o que a maioria das pessoas sentia e pensava sobre as ações dos governos era algo inacessível. Atualmente, alguns governos já utilizam as mídias sociais com o objetivo de promover a transparência, a participação cidadã, a inovação e a colaboração, além de poderem contar com mais uma fonte de informação que contribui no processo decisório (CHUN; LUNA-REYES, 2012; MERGEL, 2013; MOSSBERGER; WU; CRAWFORD, 2013; SNEAD, 2013; MARTINS; BERMEJO; SOUZA, 2015).

Para entender como as mídias sociais podem, dessa forma, ser utilizadas para estreitar a relação entre os cidadãos e os governos, é preciso conhecer seu potencial como uma fonte de dados capaz de refletir a opinião pública.

## Opinião pública nas mídias sociais

A opinião pública tem sido discutida na literatura sob diferentes matrizes conceituais (CERVI, 2010). Em meio às várias abordagens, duas linhas de pensamento são mais facilmente identificadas: de um lado pesquisadores que conceituam a opinião pública de forma simplista, considerando-a a soma das opiniões individuais dos cidadãos (ALLPORT, 1937; CHILDS, 1939), e, de outro, aqueles que consideram outras variáveis na formação da opinião pública, por exemplo, ambiente, interação social, comunicação, emancipação, preferências políticas, necessidades, crenças e interesses (BLUMER, 1948; DEWEY, 1954; ANDRADE, 1964; NOELLE-NEUMANN, 1974; SARTORI, 1994; LIPPMANN, 2008; CERVI, 2010; PERRIN; MCFARLAND, 2011; HABERMAS, 2003; 2014).

Childs (1939, p. 328) define opinião pública como uma “força ativa ou latente derivada de um conjunto de pensamentos individuais, sentimentos e impressões, que são ponderados pelos diferentes graus de influência e agressividade das opiniões individuais dentro desse conjunto”. Nessa mesma linha, Allport (1937) acredita que a opinião pública é formada pela soma das opiniões individuais que, sejam por imposição ou influência, se destacaram perante as demais opiniões do público. Blumer (1948) supera essas posições simplistas de Allport (1937) e Childs (1939), ao considerar que até as opiniões individuais são geradas por meio da interação social e não isoladamente. Para Noelle-Neumann (1974), essa interação também é a principal característica do processo de formação da opinião pública. Nessa perspectiva, Sartori (1994, p. 125) define a opinião pública como “um público, ou uma multiplicidade de público, cujos estados de espírito difusos (opiniões) interagem com fluxos de informações relativas ao estado da *res pública*”. Para o autor, a opinião formada por essa interação é considerada pública por estar relacionada a “coisas públicas”, e não somente por ser difundida entre o público. Vale ressaltar que a interação em questão deve se balizar no debate racional, pois o público, quando tomado pelo comportamento emocional, é facilmente manipulado pelas mídias de massa unilaterais (emissoras de rádio, TV, jornais e revistas), formando, assim, apenas um sentimento coletivo, não uma opinião pública (LIPPMANN, 2008). Segundo Andrade (1964), a opinião pública se forma por meio da comunicação racional e da interação social, e não por esse sentimento coletivo produzido por uma comunicação unilateral. Outros autores também alertam sobre a influência dos meios de comunicação de massa na opinião pública (DEWEY, 1954; NOELLE-NEUMANN, 1974; PAGE; SHAPIRO; DEMPSEY, 1987; TOCQUEVILLE, 1998; HABERMAS, 2003; 2014). Habermas (2014), por exemplo, chama a atenção para as empresas de mídia que utilizam técnicas de publicidade e relações públicas com o intuito de construir uma opinião pública encenada, ou seja, um consenso fabricado. Do mesmo modo, Andrade (1964) acredita que pequenos grupos de interesse bem

organizados e capacitados para divulgar sua posição em mídias de massa exercem grande influência sobre a maioria dos atores no processo de formação da opinião pública. Aqui está situada a crítica de Tocqueville (1998), que observou o papel da imprensa como influenciadora da opinião pública na sociedade democrática por meio de seu poder “impregnado de ambiguidade”. Foi essa opinião pública influenciada pela imprensa que o autor chamou de “tirania da maioria”. De acordo com Lippmann (2008), isso agrava ainda mais a consolidação de uma opinião pública consciente, uma vez que as pessoas já têm dificuldade de se informarem de forma competente sobre questões de interesse público. Andrade (1964) concorda que mesmo com a evolução dos meios de comunicação e o progresso educacional, as pessoas nem sempre são bem informadas. Para o autor, a massa aceita por comodismo os pontos de vista dos pequenos grupos que detêm os melhores veículos de comunicação. Esse é um ponto importante na obra de Dewey (1954), que adverte sobre a importância das pessoas manterem-se informadas sobre os assuntos de interesse público, pois, do contrário, seria estabelecida uma oligarquia gerada em interesse de uma minoria. Desse modo, Andrade (1964, p. 112) conclui que “para que haja opinião pública, é preciso que predominem as considerações de ordem racional” (ANDRADE, 1964, p. 112). Nesse sentido, Habermas (2014) acredita que somente a reunião de um público formado por pessoas privadas constrói uma opinião pública com base na racionalidade do melhor argumento, mantendo-se distantes da influência do poder político e econômico e da ação das mídias de massa unilaterais. Para o autor, é a opinião pública que representa o poder de influência da sociedade para direcionar o poder político e econômico, e não o contrário. Porém, Habermas (2003) chama a atenção ao mencionar que a opinião pública não pode tomar para si o uso do poder administrativo, mas pode influenciá-lo.

Lippmann (2008) se distancia do pensamento de Habermas (2003), ao salientar que a falta de elementos para um debate racional pode colocar os cidadãos em uma posição passiva perante o poder público, que por sua vez deveria ser gerido por tecnocratas. Dewey (1954) discorda dessa posição e, assim como Habermas (2003; 2014), acredita que somente uma opinião pública formada pela interação direta dos cidadãos, respeitando-se a liberdade de expressão e associação, pode resultar em uma democracia com espírito de comunidade. Nesse caso, a função da opinião pública seria permitir a todos os cidadãos a participação ativa nas questões políticas e de interesse geral, bem como avaliar os atos governamentais (CERVI, 2010). Em contrapartida, o alinhamento entre a opinião pública e as decisões governamentais depende, de um lado, do grau de amadurecimento político-ideológico dos cidadãos e de sua capacidade de comunicação e organização e, de outro, das questões de ordem legal, política e econômica às quais o governo está submetido (NOGUEIRA, 1997; 2007).

Em suma, a opinião pública deve ser entendida como coletiva, e não apenas agregada; dinâmica, e não estática; reativa, e não unidirecional (PERRIN; MCFARLAND, 2011). Ela é formada por um processo de expressão coletiva de opiniões e tomada de decisão, dentro de uma esfera pública vibrante, em que os cidadãos podem trocar informações e ideias sobre questões públicas (STEENKAMP; HYDE-CLARKE, 2014). Segundo Habermas (2003; 2014), essa esfera pública se reproduz por meio do agir comunicativo, implicando apenas o domínio de uma linguagem natural que esteja em sintonia com a prática comunicativa cotidiana. Nesse contexto, Steenkamp e Hyde-Clarke (2014, p. 92) acreditam que as mídias sociais podem atuar como instituições comunicativas, que representam uma expansão do espaço no qual as expressões de unidade ou diversidade possam ser comunicadas em uma “esfera pública interativa”.

As mídias sociais são aplicativos baseados na internet e projetados para promover a interação social, a colaboração, a aprendizagem conjunta e a rápida difusão de informações pela sociedade (BONSÓN et al., 2012; KAVANAUGH et al., 2012). Indiscutivelmente, essas mídias tornaram-se plataformas de comunicação bastante populares, nas quais os cidadãos podem ter acesso às opiniões de outras pessoas e expressar sua própria opinião sobre qualquer tema (ZHANG et al., 2009). Dessa forma, cidadãos investidos no bem comum podem usar espaços de discussão *online* para produzir uma opinião pública informada que pode avaliar e até guiar processos



oficiais de tomada de decisão (PENNEY, 2016). Isso corrobora o pensamento de Habermas (2003), segundo o qual a opinião pública representa o poder de influência da sociedade para direcionar ou mesmo interferir nas decisões de parlamentares, governos e tribunais, sem tomar para si o poder administrativo.

A esfera pública interativa, criada por meio das mídias sociais, permite que os cidadãos se mantenham atualizados sobre qualquer assunto político ou social, no bairro, no estado, no país e em todo o mundo (EIRINAKI; PISAL; SINGH, 2012; STEENKAMP; HYDE-CLARKE, 2014). O fato de a internet estar agora ao alcance da maioria das pessoas, em razão dos *hardwares* com preços mais acessíveis, levou a uma revolução cultural, pois pessoas de todo o mundo agora são capazes de interagir uns com os outros manifestando suas opiniões livremente (MOREO et al., 2012).

As mídias sociais estão mudando a forma como as pessoas buscam informações e se comunicam (OLIVEIRA; WELCH, 2013). Percebe-se que o panorama da mídia mudou muito nos últimos anos, pois antes predominavam as mídias tradicionais unidirecionais, e agora elas estão sendo complementadas ou substituídas por meios interativos de comunicação social (SOBKOWICZ; KASCHEKY; BOUCHARD, 2012). De acordo com Fortuny et al. (2012), em um cenário imperfeito, mídias tendenciosas poderiam editar e selecionar informações com o objetivo de influenciar a opinião pública. Isso acontece quando há a formação de monopólios nos sistemas de mídias tradicionais, que impedem a manifestação das esferas comunicativas plurais, o que, consequentemente, compromete a qualidade da representação das vozes públicas (GUIMARÃES; AMORIM, 2013). Por isso, “a formação democrática da opinião e da vontade depende de opiniões públicas informais que idealmente se formam em estruturas de uma esfera pública não desvirtuada pelo poder” (HABERMAS, 2003, p. 33). Mesmo que alguém, utilizando as mídias sociais, tente influenciar a opinião pública em uma direção, haverá, da mesma forma, aqueles que tentarão influenciá-la para outra direção e, assim, o volume de mensagens concorrentes simultâneas acaba anulando-se mutuamente (BACHNER; HILL, 2014). Para Anstead e O’Loughlin (2015), embora haja claramente limitações na pesquisa de opinião pública com base em mídias sociais, a extensão de suas deficiências depende, em grande parte, da definição de opinião pública que está sendo empregada. Conforme salienta Nunomura (2013), se por um lado ainda é preciso mais estudos para conferir às mídias sociais o *status* de um termômetro da opinião pública, por outro já se percebe que elas amplificam esse conceito.

Assim, as mídias sociais podem ser consideradas fontes de informação capazes de refletir a opinião dos cidadãos sobre diversos assuntos de interesse público, inclusive política e administração pública (SOBKOWICZ; KASCHEKY; BOUCHARD, 2012). Consequentemente, nos últimos anos, tem aumentado o número de pesquisas acadêmicas e corporativas que exploram as informações contidas nessas mídias (YU; DUAN; CAO, 2013). Em contraste com os tempos mais antigos, quando encontrar fontes de informação era o principal problema para pesquisadores, organizações e indivíduos, a sociedade da informação de hoje os desafia a criarem e implementarem mecanismos para buscar, minerar e recuperar informações relevantes da enorme quantidade de dados disponíveis na web, transformando-os em conhecimento (MONTOTOYO; MARTÍNEZ-BARCO; BALAHUR, 2012).

Apesar de já existirem muitos métodos altamente precisos para analisar e extrair conhecimento relevante baseado em dados estruturados (por exemplo, tabelas ou banco de dados), a tarefa de extrair informações úteis em bases de dados não estruturados (texto, discurso, etc.), como é o caso das mídias sociais, ainda é um importante desafio (O’CONNOR et al., 2010; LIU; ZHANG, 2012). A solução para essa questão vem sendo procurada por muitos pesquisadores em um subcampo do processamento de linguagem natural<sup>1</sup> denominado análise de sentimento (FORTUNY et al., 2012).

1 *Processamento de linguagem natural é uma disciplina de inteligência artificial que lida com o tratamento automático da linguagem natural em texto ou discurso, sendo muito empregada, por exemplo, em motores de busca como Google e Bing (MONTOTOYO; MARTÍNEZ-BARCO; BALAHUR, 2012).*

## Análise de sentimento

A análise de sentimento, também conhecida como mineração de opinião, análise de atitude, análise de subjetividade e análise de favorabilidade (PANG; LEE, 2008; LANE; CLARKE; HENDER, 2012; DI CARO; GRELLA, 2013; KONTOPOULOS et al., 2013; MOUTHAMI; DEVI; BHASKARAN, 2013), é uma técnica que busca identificar, extrair e classificar informações subjetivas, como opiniões e sentimentos, sobre diversos temas a partir de textos (FORTUNY et al., 2012; YOON; ELHADAD; BAKKEN, 2013). Para Mostafa (2013), a análise de sentimento pode ser considerada como uma técnica automatizada de descoberta de conhecimento que visa encontrar padrões escondidos em um grande número de dados textuais, assim como os comentários submetidos às mídias sociais. Trata-se de uma tecnologia emergente que identifica a opinião das pessoas em relação a um tema ou objeto, sendo muito útil no monitoramento de mídias sociais para determinar automaticamente o sentimento geral de seus usuários sobre várias questões (HE; ZHA; LI, 2013; MOUTHAMI; DEVI; BHASKARAN, 2013).

O objetivo da análise de sentimento é identificar a opinião das pessoas sobre temas específicos, não se tratando somente de uma estruturação de dados (DI CARO; GRELLA, 2013). Nessa linha de pensamento, Santos et al. (2011) esclarecem que a análise de sentimento procura classificar textos não por tópicos, mas pelo sentimento ou opinião neles contidos. Seu propósito é o de criar uma base de conhecimento contendo opiniões de uma forma mais organizada e explícita (SOBKOWICZ; KASCHEKY; BOUCHARD, 2012).

Embora a opinião seja um conceito muito amplo, a análise de sentimento foca, principalmente, sentimentos positivos e negativos (MOREO et al., 2012). Ela é usada para extrair opiniões, sentimentos e subjetividade em textos não estruturados, ou seja, para identificar se as expressões indicam um parecer positivo (favorável) ou negativo (desfavorável) para o assunto (PANG; LEE, 2008). Dessa forma, um trabalho de análise de sentimento pode ser interpretado como uma tarefa de classificação em que cada categoria (positiva ou negativa) representa um sentimento (PRABOWO; THELWALL, 2009). Assim, essa análise normalmente lida com a identificação da polaridade, ou seja, determina se um texto é objetivo ou subjetivo, e se um texto subjetivo contém sentimentos positivos ou negativos, ao invés de emoções distintas, como alegria e tristeza, por exemplo (BAE; LEE, 2012). Kontopoulos et al. (2013) e Mostafa (2013) consideram, também, a opinião neutra dentro da polaridade, ao concordarem que a análise de sentimento é um processo que tem como objetivo determinar se a polaridade de um corpus textual (documento, frase, parágrafo, etc.) tende a ser positivo, negativo ou neutro.

Quanto à forma de classificar os textos, as abordagens em análise de sentimento são baseadas em orientação semântica (aprendizagem não supervisionada) ou aprendizagem de máquina (aprendizagem supervisionada), e ambas são aplicadas por meio de *softwares* específicos (KONTOPOULOS et al., 2013; YU; DUAN; CAO, 2013). A abordagem baseada em orientação semântica está centrada em listas predeterminadas de palavras positivas e negativas e, nesse caso, a identificação automatizada da polaridade do texto dependerá da frequência dos diferentes tipos de palavras que aparecerem no documento (YU; DUAN; CAO, 2013). A aprendizagem de máquina trata-se da preparação de um classificador de sentimento com base em um conjunto de treino (KONTOPOULOS et al., 2013). Um conjunto de treino é uma amostra de dados utilizada para o treinamento manual de um classificador automático, ou seja, é um conjunto de dados submetido à classificação humana para que o computador, via recursos de inteligência artificial, crie padrões sobre as características diferenciadoras de documentos (VINODHINI; CHANDRASEKARAN, 2012). Na literatura, a abordagem de aprendizagem de máquina tem apresentado melhores resultados quanto à precisão da classificação automatizada dos dados (LIU; ZHANG, 2012; KONTOPOULOS et al., 2013; YANG; YU, 2013).

Atualmente, a análise de sentimento está recebendo muita atenção por causa da grande variedade de suas aplicações diretas, como análises de produtos, serviços, perfil do público, tendências políticas, entre outros (DI CARO; GRELLA, 2013). Até mesmo questões de interesse para a segurança pública ou a qualidade de vida (p. ex., tráfego urbano, qualidade do ar) podem ser descobertas, monitoradas e mitigadas



por meio da análise dos fluxos de informação das mídias sociais, detectando padrões e tendências significativas (KAVANAUGH et al., 2012).

O fato de tantas pessoas se expressarem nas mídias sociais sobre qualquer assunto tornam as opiniões menos tendenciosas e, conseqüentemente, mais confiáveis (MONTOTO; MARTÍNEZ-BARCO; BALAHUR, 2012). Cáceres (2011) acredita que os habitantes do ciberespaço, os “novos cidadãos”, têm muito a dizer, muito a ensinar e propor muitas mudanças. Em decorrência dessas razões, as opiniões expressas nessas mídias são cada vez mais consideradas no processo decisório e na obtenção de um *feedback* imparcial sobre diversos assuntos (MONTOTO; MARTÍNEZ-BARCO; BALAHUR, 2012). Além disso, trabalhar com informações extraídas das mídias sociais nem sempre necessita de grandes investimentos por parte da organização (OLIVEIRA; WELCH, 2013).

## Trabalhos relacionados

No campo da administração pública existem poucas pesquisas sobre análise de sentimento ou mineração de opinião, mas algumas merecem destaque. Fortuny et al. (2012) publicaram o artigo intitulado “Media coverage in times of political crisis: a text mining approach”, que abordou a cobertura das mídias tradicionais sobre questões políticas na Bélgica em um período de crise no final do ano de 2011. O estudo mediu o sentimento social dos cidadãos em relação aos partidos políticos do país, extraíndo dados de notícias de grandes jornais em versões *online* por um período de dez meses. Com a pesquisa, os autores descobriram um viés na cobertura da imprensa tradicional em textos políticos. Sobkowicz, Kaschesky e Bouchard (2012), em seu artigo “Opinion mining in social media: modeling, simulating, and forecasting political opinions in the web”, investigaram como os conteúdos *online* podem ser explorados para informar aos tomadores de decisão sobre as opiniões dos cidadãos, as tendências emergentes e os impactos potenciais das iniciativas políticas. Os autores concluíram que a mineração de opinião apoia os formuladores de políticas em todas as etapas do ciclo de políticas públicas, antecipando a repercussão das ações perante a opinião pública.

No Brasil, também há poucas pesquisas sobre a aplicação da análise de sentimento em dados extraídos de mídias sociais. Nascimento et al. (2012) publicaram o estudo “Análise de sentimento de *tweets* com foco em notícias”, que foi o pioneiro no país quanto à aplicação da análise de sentimento a partir de notícias compartilhadas no Twitter. O trabalho foi uma primeira tentativa de aplicação da técnica em português brasileiro, utilizando-se dados não estruturados. Os autores tiveram como objetivo identificar o sentimento polarizado (positivo e negativo) da população a respeito de três categorias de notícias compartilhadas no Twitter: entretenimento, policial e política. A análise de sentimento foi realizada utilizando-se três classificadores automatizados (aprendizagem supervisionada). Os autores conseguiram precisão satisfatória na análise automatizada das categorias “entretenimento” e “policial”. De acordo com os autores, a categoria “política” não apresentou precisão satisfatória, em virtude do alto índice de *tweets* contendo ironia e sarcasmo, o que confundiu a ferramenta de análise.

Considerando esses trabalhos, percebe-se que a análise de sentimento baseada em informações coletadas nas mídias sociais pode ser uma valiosa forma de gerar conhecimento sobre a sociedade (TSYTSAU; PALPANAS, 2012). Em posse desse conhecimento, os agentes públicos podem, segundo Prabowo e Thelwall (2009), analisar a opinião do público no que diz respeito a serviços públicos e questões políticas. Mas algumas limitações sobre a técnica de análise de sentimento devem ser observadas. O primeiro problema importante que precisa ser resolvido é a carência de recursos substanciais para tratamento de dados em outros idiomas além do inglês (GLASS; COLBAUGH, 2012; MONTOTO; MARTÍNEZ-BARCO; BALAHUR, 2012; YANG; YU, 2013; ZHANG; HE, 2013). Tais recursos devem ser adaptados para os tipos de texto a serem analisados e para cada uma das línguas envolvidas (MONTOTO; MARTÍNEZ-BARCO; BALAHUR, 2012). Também, há outras questões linguísticas a serem consideradas, como a presença de ruídos na comunicação e a presença de linguagem informal, como gírias e abreviaturas não oficiais (MOREO et al., 2012).

## Procedimentos metodológicos

Este é um estudo aplicado, interdisciplinar e exploratório, que pode ser classificado como qualitativo no que diz respeito à coleta, classificação e análise de dados subjetivos; e quantitativo no que se refere a identificar a frequência de opiniões positivas, negativas e neutras no conjunto de dados (COLLIS; HUSSEY, 2005; RAMBOCAS; GAMA, 2013).

Para verificar se a análise de sentimento pode ser utilizada para identificar a opinião dos cidadãos sobre as ações do governo com nível de precisão satisfatório, tomamos como objeto de estudo a onda de protestos ocorrida no Brasil em junho de 2013, e identificamos a opinião dos cidadãos brasileiros sobre os fatores que motivaram tais manifestações em dois momentos distintos: seis meses e um ano após o ocorrido. A escolha desse objeto de estudo se baseou na oportunidade de coleta e análise de dados relacionadas às maiores manifestações de rua organizadas via mídias sociais no Brasil até aquele momento.

A análise de sentimento exige que o pesquisador siga algumas etapas que envolvem basicamente coleta, codificação e processamento de dados (SOBKOWICZ; KASCHEKSKY; BOUCHARD, 2012). Na literatura, não existe unanimidade quanto ao método de aplicação da análise de sentimento, uma vez que os estudos publicados sobre o assunto apresentam métodos e técnicas bastante diversificadas (LIU, 2010; LI; LIU, 2012; SOBKOWICZ; KASCHEKSKY; BOUCHARD, 2012; ROBALDO; DI CARO, 2013; YOON; ELHADAD; BAKKEN, 2013; ZHANG; HE, 2013). Os procedimentos metodológicos adotados neste estudo foram baseados nos trabalhos de Nascimento et al. (2012), Yoon, Elhadad e Bakken (2013) e Oliveira, Bermejo e Santos (2015), constituindo-se as seguintes etapas: 1) escolher os temas para análise; 2) definir os termos de busca; 3) escolher a fonte de dados; 4) selecionar aplicativo de mineração de opinião; 5) coletar dados; 6) preparar dados; 7) delimitar o conjunto de treino; 8) realizar a análise automatizada dos dados; 9) validar os resultados.

Os temas escolhidos (etapa 1) foram aqueles considerados como os principais fatores que motivaram a onda de protestos ocorrida no Brasil em junho de 2013. Esses temas foram identificados pela pesquisa Ibope (2013b) e pela Plataforma Causa Brasil<sup>2</sup>. Conforme essas pesquisas, os fatores mais citados como causa dos protestos e escolhidos como temas para análise neste estudo são: transporte público, combate à corrupção, saúde, educação, gastos com a Copa do Mundo FIFA 2014 e Governo Federal.

Os termos de busca (etapa 2) foram selecionados por meio de um conjunto de palavras-chave utilizadas no monitoramento de mídias sociais pela Plataforma Causa Brasil durante as manifestações. O critério utilizado foi o de escolher três termos ligados a cada tema. Esses termos deveriam estar diretamente relacionados aos temas e ao mesmo tempo viáveis para aplicação da técnica. Os termos selecionados estão distribuídos no Quadro 1.

**Quadro 1 – Termos de pesquisa utilizados na mineração de dados.**

<b>Temas</b>	<b>Termos de busca</b>
Transporte público	Transporte público, passe livre, preço das passagens.
Combate à corrupção	Combate à corrupção, CPI, julgamento mensalão.
Educação	Educação, salário professor, escola pública.
Saúde	Saúde, hospitais públicos, posto de saúde.
Gastos com a Copa do Mundo	Gastos com a Copa, obras da Copa, gastos com estádios.
Governo Federal	Governo Federal, Dilma Rousseff, presidenta Dilma.

Fonte: Elaborado pelos autores.

<sup>2</sup> Disponível em: <<http://www.causabrasil.com.br>>. Acesso em: 12 nov. 2013.

As fontes de dados escolhidas (etapa 3) foram as mídias sociais mais utilizadas pelos cidadãos brasileiros durante os protestos em junho de 2013. Como dito anteriormente, a pesquisa Ibope (2013b) revelou que em uma das maiores manifestações ocorridas no país, 62% dos manifestantes souberam do evento pelo Facebook, e 75% destes convocaram outras pessoas para participar por meio do Facebook e do Twitter. Outra pesquisa também realizada pelo Ibope (2013a) revelou que as mídias sociais foram os canais mais utilizados pelos cidadãos para divulgar e organizar os protestos, bem como para comentar sobre o assunto. Assim, essas duas mídias foram as fontes de dados escolhidas para a realização deste trabalho.

A seleção do aplicativo (etapa 4) para realizar a mineração de opinião baseou-se em alguns requisitos. O aplicativo deveria permitir o tratamento de dados em língua portuguesa e ser capaz de extrair e classificar automaticamente dados do Facebook e do Twitter, a partir dos termos de pesquisa definidos anteriormente, além de ser gratuito ou possuir licença para fins acadêmicos. Foram analisadas as ferramentas DiscoverText<sup>3</sup>, Rapidminer<sup>4</sup> e Scup<sup>5</sup>. Considerando os requisitos, o aplicativo adotado foi o DiscoverText. Sua escolha foi definida pela viabilidade técnica, disponibilidade gratuita para fins acadêmicos e seu desempenho satisfatório em outros estudos (BEYER, 2012; OLIVEIRA; BERMEJO; SANTOS, 2015; 2016). O DiscoverText é um aplicativo de análise de textos baseado em nuvem, capaz de capturar, filtrar e classificar um grande volume de dados estruturados e não estruturados. Por meio desse aplicativo, é possível personalizar e reutilizar classificadores, baseados em aprendizagem de máquina que permitem o uso combinado de algoritmos e codificação humana, o que aumenta a precisão nos resultados (KONTOPOULOS et al., 2013).

A coleta de dados (etapa 5) foi realizada por meio da mineração de dados, utilizando-se o aplicativo DiscoverText, entre os dias 1º e 30 de dezembro de 2013 e 1º a 30 de junho de 2014. A mineração foi realizada a partir da configuração do aplicativo, para que ele coletasse todos os *posts* do Facebook e do Twitter (dados brutos) relacionados aos termos de busca definidos no Quadro 1 durante os dois períodos estabelecidos.

Após a coleta, foi realizada a preparação dos dados (etapa 6), aplicando-se filtros que eliminaram *posts* duplicados, *links* e outros ruídos nos dados coletados, além da exclusão de postagens feitas por agências de notícias, resultando em um conjunto de dados composto apenas por menções de cidadãos brasileiros (dados processados).

O conjunto de treino (etapa 7) utilizado para o treinamento manual do aplicativo DiscoverText foi delimitado com base em uma amostra aleatória simples (COCHRAN, 2007). Para calcular a dimensão ideal do conjunto de treino, o primeiro passo foi extrair uma amostra de 100 menções sobre cada tema, compondo uma base de dados para teste contendo um total de 600 menções. A cada 60 menções classificadas manualmente, equivalente a 10% do total, foi realizada a classificação automatizada de toda a base de teste e, assim, contabilizados os acertos, até conseguir uma precisão satisfatória. A precisão foi calculada, dividindo-se o número de classificações corretas pelo total de classificações (NASCIMENTO et al., 2012; EIRINAKI; PISAL; SINGH, 2012). Após classificadas manualmente 30% das menções, a precisão da classificação automatizada chegou aos 81%. Na literatura, uma ferramenta que gera em média 80% de acertos é considerada satisfatória (NASCIMENTO et al., 2012; MOSTAFA, 2013; YOON; ELHADAD; BAKKEN, 2013; YU; DUAN; CAO, 2013). Desse modo, o conjunto de treino considerado para este estudo é de 30% das menções coletadas para cada tema, conforme distribuídos na Tabela 1.

3 Disponível em: <<http://www.discovertext.com>>. Acesso em: 15 nov. 2013.

4 Disponível em: <<https://rapidminer.com>>. Acesso em: 15 nov. 2013.

5 Disponível em: <<http://www.scup.com>>. Acesso em: 15 nov. 2013.

**Tabela 1 – Cálculo do conjunto de treino.**

Tema	Dados coletados em dezembro de 2013			Dados coletados em junho de 2014		
	Dados brutos	Dados processados	Conjunto de treino (30%)	Dados brutos	Dados processados	Conjunto de treino (30%)
Transporte público	7.874	795	239	15.194	2.482	745
Combate à corrupção	8.595	1.892	568	10.567	2.138	642
Educação	9.996	1.419	426	13.823	1.984	596
Saúde	5.178	814	245	9.874	1.895	567
Gastos com a Copa do Mundo	6.606	1.737	522	10.423	2.196	659
Governo Federal	13.608	2.642	793	19.746	3.244	974
<b>Total</b>	<b>51.857</b>	<b>9.299</b>	<b>2.793</b>	<b>79.627</b>	<b>13.939</b>	<b>4.183</b>

Fonte: Dados da pesquisa.

Conforme os dados constantes na Tabela 1, percebe-se uma grande redução do número de dados brutos para os dados processados. Isso se deve à aplicação de filtros durante a fase de preparação, que eliminou os ruídos dos dados brutos. Dessa forma, o conjunto de dados sobre cada tema foi composto apenas por menções de cidadãos brasileiros, relacionadas aos termos de busca previamente definidos. Considerando os dados processados, foi calculado o conjunto de treino para cada tema na proporção de 30% e realizada a classificação manual desses dados, que foram coletados, aleatoriamente, dentro de cada conjunto. Nessa etapa, é necessária a classificação manual de cada conjunto de treino para que se estabeleçam regras e padrões (aprendizagem de máquina) para a classificação automatizada dos dados.

Após o treinamento do aplicativo, foi realizada a classificação automatizada (etapa 8) de todos os conjuntos de dados e gerado o sentimento (positivo, neutro e negativo) sobre cada tema.

Depois de identificar o sentimento dos cidadãos perante cada tema, os resultados foram validados (etapa 9), revelando o quanto são precisos (COLLIS; HUSSEY, 2005). A precisão foi calculada com base em uma nova amostra aleatória, considerando apenas os dados que foram classificados de forma automatizada (NASCIMENTO et al., 2012; EIRINAKI; PISAL; SINGH, 2012). O tamanho da amostra foi de 30% do total de dados analisados, mantendo a mesma proporção do teste realizado no conjunto de treino. Desse modo, o cálculo da precisão da análise automatizada deveria ser o mais próximo possível do valor encontrado naquele teste, ou seja, 81%.

**Tabela 2 – Precisão da análise de sentimento.**

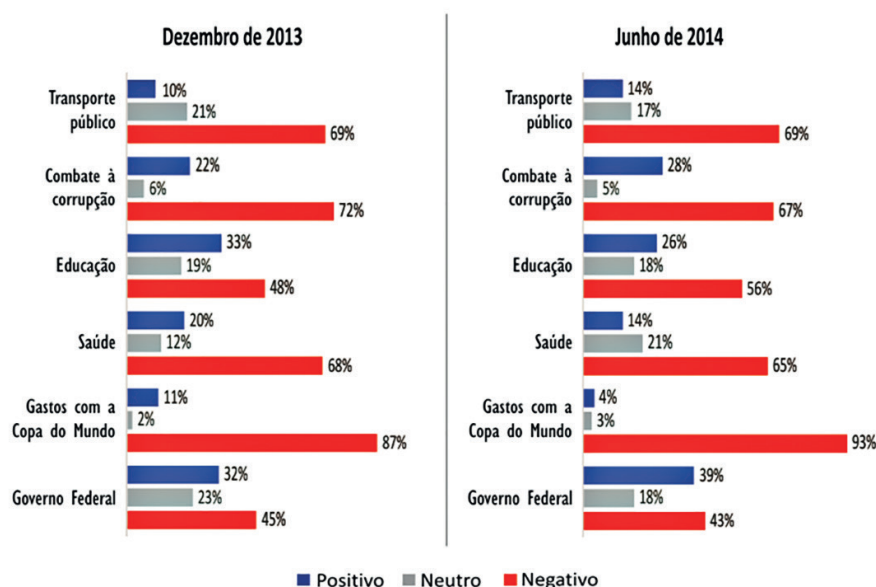
Tema	Precisão	
	Dados coletados em 12/2013	Dados coletados em 06/2014
Transporte público	76%	73%
Combate à corrupção	74%	78%
Educação	80%	82%
Saúde	82%	83%
Gastos com a Copa do Mundo	77%	76%
Governo Federal	69%	71%
<b>Média</b>	<b>76,33%</b>	<b>77,16%</b>

Fonte: Elaborada pelos autores.

A média aritmética da precisão de todos os temas foi de 76,33% para os dados coletados em dezembro de 2013 e de 77,16% para os dados coletados em junho de 2014. Dessa forma, houve uma perda de 4,66% e 3,84% de precisão nos respectivos dados, comparando-os com o resultado obtido no primeiro conjunto de teste (81%). Isso aconteceu porque a precisão da mineração de opinião pode variar de acordo com o assunto pesquisado e a qualidade dos dados (EIRINAKI; PISAL; SINGH, 2012; NASCIMENTO et al., 2012).

## Resultados

Após a classificação automatizada dos dados relacionados aos temas que motivaram a realização da onda de protestos em junho de 2013 no Brasil, pode-se observar que, mesmo após um ano, os fatores que motivaram essas manifestações continuavam gerando insatisfação nos cidadãos brasileiros, conforme demonstrado no Gráfico 1.



**Gráfico 1 – Análise de sentimento baseada nos dados coletados em dezembro de 2013 e junho de 2014.**

Fonte: Dados da pesquisa.

Ponderando a análise de sentimento realizada com dados extraídos do Twitter e do Facebook, houve predominância de opiniões negativas em todos os temas tratados. Dentre eles, os **gastos com a Copa do Mundo** foi o que gerou maior número de menções negativas (87% em dezembro de 2013 e 93% em junho de 2014). Por meio da leitura de algumas dessas menções, é possível entender melhor essa tendência negativa. Os cidadãos manifestaram-se nas mídias sociais com uma posição fortemente contra o nível de gastos públicos com os preparativos para a Copa do Mundo FIFA 2014, como podemos perceber na seguinte postagem<sup>6</sup>: *"10 bilhões gastos na construção de estádios de futebol. E o povo sem educação, saúde e segurança. E sem futuro"* (Usuário A, Twitter, 19 dez. 2013). O segundo tema com maior número de menções negativas em dezembro de 2013 foi o **combate à corrupção** (72%), caindo para 67% de menções negativas em junho de 2014. Pela leitura dessas menções,

6 Por motivos éticos, foi mantido o anonimato de todos os autores das menções reproduzidas neste trabalho.

percebe-se a falta de esperança dos brasileiros no combate efetivo à corrupção, por exemplo: *"Os corruptos cassados? Nunca serão! Cidadãos bem informados? Nunca serão! Hospitais bem equipados? Nunca serão! Nunca serão!! Nunca serão!!"* (Usuário B, Twitter, 22 dez. 2013).

Na sequência de menções negativas, no gráfico de 2013 seguem praticamente empatados os temas **transporte público** (69%) e **saúde** (68%). Seis meses depois, foi mantido o número de menções negativas sobre o tema transporte público, enquanto as menções negativas sobre o tema saúde tiveram ligeira queda, chegando a 65%. Esses resultados foram gerados a partir de menções como: *"Estou muito revoltada com a saúde pública. Como pode na cidade de Tupã não ter uma UTI Neonatal???? Minha filha teve que ser levada às pressas para outra cidade..."* (Usuário C, Facebook, 18 jun. 2014). *"Este transporte público do DF está um caos, hoje devido a superlotação do ônibus, teve uma discussão entre passageiros por falta de espaço"* (Usuário D, Twitter, 15 dez. 2013).

Os temas que apresentaram maior número de menções positivas em dezembro de 2013 foram **educação** (33%) e **Governo Federal** (32%). Isso se manteve em junho de 2014; e as menções positivas sobre o tema educação tiveram uma queda de 7%, enquanto no tema Governo Federal houve uma alta de 7%. Esse, também, foi o tema com o maior número de menções neutras em dezembro de 2013. Em junho de 2014, o tema com maior número de menções neutras foi a saúde. Essas menções foram marcadas por comentários que não expressaram claramente uma posição positiva ou negativa sobre o tema, como: *"Presidente Dilma Rousseff acaba de chegar em Porto Alegre"* (Usuário E, Twitter, 19 dez. 2013).

Como foi demonstrado na Tabela 2, o tema que apresentou menor precisão, após classificados todos os dados, foi o **Governo Federal**. Por meio da análise criteriosa de algumas menções que foram classificadas erroneamente, percebe-se que a precisão da ferramenta foi afetada por falsos-positivos, falsos-negativos e menções com alto nível de ironia, por exemplo: *"Odeio gente ignorante que critica essas bolsas do governo federal. É ser muito mesquinho e ignorante"* (Usuário F, Twitter, 14 jun. 2014). Essa menção foi considerada negativa na classificação automática, em virtude da sua carga de palavras consideradas negativas, como "odeio", "ignorante", "critica", "mesquinho", apesar de o contexto ser uma defesa ao programa de bolsas do Governo Federal. Esse é um exemplo típico de falso-negativo.

Entre os seis temas, o que apresentou maior precisão na mineração de opinião automatizada foi **saúde**. Analisando as menções coletadas sobre esse tema, percebe-se um número razoável de frases contendo opiniões claras e objetivas, por exemplo: *"Os hospitais públicos do Brasil são horríveis, entra governo e sai governo, continua a mesma coisa..."* (Usuário G, Twitter, 14 dez. 2013). Assim, conjuntos de dados que apresentam boa coerência textual podem aumentar a precisão da classificação automatizada. Por outro lado, um conjunto de dados compostos por expressões que contenham ironia, gírias e erros de ortografia podem reduzir consideravelmente a confiabilidade dos resultados (MONTORO; MARTÍNEZ-BARCO; BALAHUR, 2012; SOBKOWICZ; KASCHESKY; BOUCHARD, 2012).

## Conclusão

Este estudo apresentou as mídias sociais como um canal ativo em que os usuários manifestam suas opiniões sobre os mais variados temas, inclusive, sua satisfação ou insatisfação com as ações do governo. Diferentemente dos canais fechados, como telefone e e-mail, as informações postadas nas mídias sociais podem ser disseminadas muito rapidamente, contribuindo para a formação da opinião pública, principalmente sobre temas de grande repercussão. O desafio, no entanto, era aplicar um método de processamento e análise de dados que pudesse gerar conhecimento por meio de dados não estruturados. A proposta apresentada como forma de superar esse desafio foi a análise de sentimento.



A revisão teórica e o teste empírico demonstraram que é possível a utilização das informações postadas pelos cidadãos em mídias sociais como uma ferramenta de mensuração da opinião pública sobre vários temas. Para exemplificar a aplicação da análise de sentimento nos campos da política e da administração pública, foi apresentado um caso prático, tomando como objeto a opinião dos cidadãos sobre os principais temas que motivaram a onda de protestos ocorrida no Brasil em junho de 2013 e em dois períodos posteriores ao ocorrido. Os resultados demonstraram que os cidadãos brasileiros, um ano após os protestos, continuam emitindo opiniões negativas associadas aos fatores que os motivaram a protestar, principalmente em relação aos gastos públicos com a Copa do Mundo FIFA 2014, o combate à corrupção e o transporte público.

Diante disso, a análise de sentimento pode contribuir para revelar a opinião dos cidadãos quanto à atuação do governo e aos serviços públicos, identificando aqueles que possuem maior demanda de ampliação e/ou melhorias. A partir daí, os governos podem definir prioridades, traçar estratégias para intervir nos problemas e agir visando atender ao interesse público. Essas ações podem gerar maior nível de satisfação entre os cidadãos, o que, consequentemente, contribui na prevenção de conflitos. Principalmente quando os “holofotes” estão voltados para o país em razão da grave crise política, ou seja, um palco perfeito para protestos.

Embora os resultados deste estudo tenham apresentado uma precisão razoável na análise automatizada dos dados, algumas limitações devem ser observadas. Primeiramente, esses resultados revelam a opinião dos cidadãos que são usuários das mídias sociais investigadas e que fizeram comentários sobre os temas definidos neste estudo durante dois períodos distintos de 30 dias. Dessa forma, é inadequado generalizar os números aqui apresentados em situações que ultrapassem esse espaço temporal e o universo de usuários, que não são, necessariamente, representativos da população de um país (CERON et al., 2014). Outra limitação é a impossibilidade de identificar, por meio da mineração de opinião automatizada, a que grupo social pertencem os usuários das mídias sociais analisadas neste estudo, uma vez que elas não disponibilizam informações sobre o perfil socioeconômico de seus usuários. Além disso, o período de um ano é um intervalo de tempo curto para se ter modificações relevantes no cenário político, principalmente se considerarmos um sistema político carcomido como o do Brasil.

Por outro lado, a análise de sentimento pode ser uma alternativa menos dispendiosa do que pesquisas de opinião tradicionais, oferecendo *insights* mais rápidos sobre a opinião pública ou servir como uma fonte adicional de evidência empírica (O’CONNOR et al., 2010; BACHNER; HILL, 2014; ANSTEAD; O’LOUGHLIN, 2015).

Portanto, é importante que estudos futuros procurem avaliar novos métodos e técnicas que possibilitem revelar a opinião pública com maior precisão ou que contribuam para a efetiva participação dos cidadãos nas decisões políticas. Um caminho promissor nesse sentido é a inovação aberta.

## Referências

- ALLPORT, F. H. Toward a science of public opinion. *Public opinion quarterly*, v. 1, n. 1, p. 7-23, 1937.
- ANDRADE, C. T. S. Mito e realidade da opinião pública. *Revista de Administração de Empresas*, v. 4, n. 11, p. 107-122, 1964.
- ANSTEAD, N.; O’LOUGHLIN, B. Social media analysis and public opinion: the 2010 UK general election. *Journal of Computer-Mediated Communication*, v. 20, n. 2, p. 204-220, 2015.
- BACHNER, J.; HILL, K. W. Advances in public opinion and policy attitudes research. *Policy Studies Journal*, Urbana, v. 42, p. S51-S70, 2014.

- BAE, Y.; LEE, H. Sentiment analysis of twitter audiences: measuring the positive or negative influence of popular twitterers. *Journal of the American Society for Information Science and Technology*, New York, v. 63, n. 12, p. 2.521-2.535, 2012.
- BEYER, Y. Using discover text for large scale twitter harvesting. *Microform & Digitization Review*, London, v. 41, n. 3-4, p. 121-125, 2012.
- BLUMER, H. Public opinion and public opinion polling. *American Sociological Review*, v. 13, n. 5, p. 542-549, 1948.
- BONSÓN, E. et al. Local e-government 2.0: social media and corporate transparency in municipalities. *Government Information Quarterly*, Amsterdam, v. 29, n. 2, p. 123-132, 2012.
- BONSÓN, E.; ROYO, S.; RATKAI, M. Citizens' engagement on local governments' Facebook sites. An empirical analysis: the impact of different media and content types in Western Europe. *Government Information Quarterly*, v. 32, n. 1, p. 52-62, 2015.
- CÁCERES, J. G. Las nuevas tecnologías de información y comunicación y las políticas culturales en México: ingeniería en comunicación social del servicio de redes sociales Facebook. *Intercom*, São Paulo, v. 34, p. 175-196, 2011.
- CAMBRIA, E. et al. New avenues in opinion mining and sentiment analysis. *IEEE Intelligent Systems*, Los Alamitos, v. 28, n. 2, p. 15-21, 2013.
- CERON, A. et al. Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France. *New Media & Society*, Milano, v. 16, n. 2, p. 340-358, mar. 2014.
- CERVI, E. U. *Opinião pública e comportamento político*. Curitiba: Ibpx, 2010. 196 p.
- CHEN, T. M. How networks changed the world. *IEEE Network*, New York, v. 25, n. 6, p. 2-3, 2011.
- CHILDS, H. L. By public opinion I mean. *The Public Opinion Quarterly*, v. 3, n. 2, p. 327-336, 1939.
- CHUN, S. A.; LUNA-REYES, L. F. Social media in government. *Government Information Quarterly*, Amsterdam, v. 29, n. 4, p. 441-445, 2012.
- COCHRAN, W. G. *Sampling techniques*. New York: John Wiley & Sons, 2007.
- COLLIS, J.; HUSSEY, R. *Pesquisa em administração: um guia prático para alunos de graduação e pós-graduação*. 2. ed. Porto Alegre: Bookman, 2005.
- DEWEY, J. *The public and its problems*. Chicago: Swallow Press, 1954. 236 p.
- DI CARO, L.; GRELLA, M. Sentiment analysis via dependency parsing. *Computer Standards & Interfaces*, Amsterdam, v. 35, n. 5, p. 442-453, 2013.
- EIRINAKI, M.; PISAL, S.; SINGH, J. Feature-based opinion mining and ranking. *Journal of Computer and System Sciences*, New York, v. 78, n. 4, p. 1.175-1.184, 2012.
- FERRO, E. et al. Policy making 2.0: from theory to practice. *Government Information Quarterly*, Amsterdam, v. 30, n. 4, p. 359-368, 2013.
- FORTUNY, E. J. D. et al. Media coverage in times of political crisis: a text mining approach. *Expert Systems with Applications*, New York, v. 39, n. 14, p. 11.616-11.622, 2012.
- GLASS, K.; COLBAUGH, R. Estimating the sentiment of social media content for security informatics applications. *Security Informatics*, New York, v. 1, n. 1, p. 3, 2012.
- GUIMARÃES, J.; AMORIM, A. P. *A corrupção da opinião pública: uma defesa republicana da liberdade de expressão*. São Paulo: Boitempo, 2013. 144 p.

- HABERMAS, J. *Direito e democracia: entre facticidade e validade*. Rio de Janeiro: Tempo Brasileiro, 2003. v. 1. 354 p.
- HABERMAS, J. *Mudança estrutural da esfera pública: investigações sobre a uma categoria da sociedade burguesa*. São Paulo: Unesp, 2014. 568 p.
- HE, W.; ZHA, S.; LI, L. Social media competitive analysis and text mining: a case study in the pizza industry. *International Journal of Information Management*, Guildford, v. 33, n. 3, p. 464-472, 2013.
- HOFMANN, S. et al. What makes local governments' online communications successful? Insights from a multi-method analysis of facebook. *Government Information Quarterly*, Amsterdam, v. 30, n. 4, p. 387-396, 2013.
- IBOPE. 89% dos manifestantes não se sentem representados por partidos. *Instituto Brasileiro de Opinião Pública e Estatística*, Rio de Janeiro, abr. 2013. Disponível em: <<http://www.ibope.com.br/pt-br/noticias/Paginas/89-dos-manifestantes-nao-se-sentem-representados-por-partidos.aspx>>. Acesso em: 11 out. 2013a.
- IBOPE. Veja pesquisa completa do Ibope sobre os manifestantes. *G1*, São Paulo, 24 jun. 2013. Disponível em: <<http://g1.globo.com/brasil/noticia/2013/06/veja-integra-da-pesquisa-do-ibope-sobre-os-manifestantes.html>>. Acesso em: 12 nov. 2013b.
- KAVANAUGH, A. L. et al. Social media use by government: from the routine to the critical. *Government Information Quarterly*, Amsterdam, v. 29, n. 4, p. 480-491, 2012.
- KONTOPOULOS, E. et al. Ontology-based sentiment analysis of twitter posts. *Expert Systems with Applications*, New York, v. 40, n. 10, p. 4.065-4.074, 2013.
- LANE, P. C. R.; CLARKE, D.; HENDER, P. On developing robust models for favourability analysis: model choice, feature sets and imbalanced data. *Decision Support Systems*, Amsterdam, v. 53, n. 4, p. 712-718, 2012.
- LI, G.; LIU, F. Application of a clustering method on sentiment analysis. *Journal of Information Science*, Cambridge, v. 38, n. 2, p. 127-139, 2012.
- LIM, M. Clicks, cabs, and coffee houses: social media and oppositional movements in Egypt, 2004-2011. *Journal of Communication*, New York, v. 62, n. 2, p. 231-248, 2012.
- LIPPMANN, W. *Opinião pública*. Petrópolis: Vozes, 2008. 273 p.
- LIU, B. *Sentiment analysis and subjectivity*. 2. ed. Chicago: Handbook of Natural Language, 2010.
- LIU, B.; ZHANG, L. A survey of opinion mining and sentiment analysis. In: AGGARWAL, C. C.; ZHAI, C. (Ed.). *Mining text data*. Chicago: Springer, 2012. cap. 13, p. 415-463.
- MARTINS, T. C. M.; BERMEJO, P. H. S.; SOUZA, W. V. B. Open innovation for citizen coproduction. In: INTERNATIONAL CONFERENCE ON ELECTRONIC GOVERNMENT AND THE INFORMATION SYSTEMS PERSPECTIVE – EGOVIS, 4., 2015, Valência, *Proceedings...* Cham: Springer International Publishing, 2015. p. 177-188.
- MERGEL, I. A framework for interpreting social media interactions in the public sector. *Government Information Quarterly*, Amsterdam, v. 30, n. 4, p. 327-334, 2013.
- MONTOYO, A.; MARTÍNEZ-BARCO, P.; BALAHUR, A. Subjectivity and sentiment analysis: an overview of the current state of the area and envisaged developments. *Decision Support Systems*, Amsterdam, v. 53, n. 4, p. 675-679, 2012.
- MOREO, A. et al. Lexicon-based comments-oriented news sentiment analyzer system. *Expert Systems with Applications*, New York, v. 39, n. 10, p. 9.166-9.180, 2012.

- MOSSBERGER, K.; WU, Y.; CRAWFORD, J. Connecting citizens and local governments? Social media and interactivity in major U.S. cities. *Government Information Quarterly*, Amsterdam, v. 30, n. 4, p. 351-358, 2013.
- MOSTAFA, M. M. More than words: social networks' text mining for consumer brand sentiments. *Expert Systems with Applications*, New York, v. 40, n. 10, p. 4.241-4.251, 2013.
- MOUTHAMI, K.; DEVI, K. N.; BHASKARAN, V. M. Sentiment analysis and classification based on textual reviews. In: INTERNATIONAL CONFERENCE ON INFORMATION COMMUNICATION AND EMBEDDED SYSTEMS – ICICES 2013, Tamilnadu. *Anais...* Tamilnadu: S.A. Engineering College Chennai, 2013. p. 271-276.
- NASCIMENTO, P. et al. Análise de sentimento de tweets com foco em notícias. In: CONGRESSO DA SOCIEDADE BRASILEIRA DE COMPUTAÇÃO, 32., 2012, Curitiba. *Anais...* Curitiba, 2012. p. 16-19.
- NOELLE-NEUMANN, E. The spiral of silence a theory of public opinion. *Journal of Communication*, v. 24, n. 2, p. 43-51, 1974.
- NOGUEIRA, M. A. A dimensão política da descentralização participativa. *São Paulo em Perspectiva*, v. 11, n. 3, p. 8-19, 1997.
- NOGUEIRA, M. A. Bem mais que pós-moderno: poder, sociedade civil e democracia na modernidade periférica radicalizada. *Ciências Sociais Unisinos*, v. 43, n. 1, p. 46-56, 2007.
- NUNOMURA, E. A imprensa, o twitter e as eleições de 2010 no Brasil. *Intercom*, São Paulo, v. 36, n. 1, p. 103-126, 2013.
- O'CALLAGHAN, D. et al. *Online social media in the syria conflict: encompassing the extremes and the in-betweens*. Cornell: Cornell University, 2014.
- O'CONNOR, B. et al. From tweets to polls: linking text sentiment to public opinion time series. In: CONFERENCE ON WEBLOGS AND SOCIAL MEDIA, 4., 2010, Washington, DC. *Anais...* Washington, DC: AAAI, 2010. p. 122-129.
- OLIVEIRA, D. J. S.; BERMEJO, P. H. S.; SANTOS, P. A. Sentiment analysis, social media, and public administration. In: CEMAL, D. et al. (Ed.). *Handbook of research on democratic strategies and citizen-centered e-government services*. Hershey: IGI Global, 2015. cap. 13, p. 231-249.
- OLIVEIRA, D. J. S.; BERMEJO, P. H. S.; SANTOS, P. A. Can social media reveal the preferences of voters? A comparison between sentiment analysis and traditional opinion polls. *Journal of Information Technology & Politics*, p. 1-12, 2016.
- OLIVEIRA, G. H. M.; WELCH, E. W. Social media use in local government: linkage of technology, task, and organizational context. *Government Information Quarterly*, Amsterdam, v. 30, n. 4, p. 397-405, 2013.
- PAGE, B. I.; SHAPIRO, R. Y.; DEMPSEY, G. R. What moves public opinion?. *American Political Science Review*, v. 81, n. 1, p. 23-43, 1987.
- PANG, B.; LEE, L. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, Washington, v. 2, n. 1-2, p. 1-135, 2008.
- PAPACHARISSI, Z.; OLIVEIRA, M. D. F. Affective news and networked publics: the rhythms of news storytelling on #Egypt. *Journal of Communication*, New York, v. 62, n. 2, p. 266-282, 2012.
- PENNEY, J. Motivations for participating in 'viral politics': a qualitative case study of Twitter users and the 2012 US presidential election. *Convergence: The International Journal of Research into New Media Technologies*, v. 22, n. 1, p. 71-87, 2016.
- PERRIN, A. J.; MCFARLAND, K. Social theory and public opinion. *Annual Review of Sociology*, Palo Alto, v. 37, n. 1, p. 87-107, 2011.

PRABOWO, R.; THELWALL, M. Sentiment analysis: a combined approach. *Journal of Informetrics*, Oxford, v. 3, n. 2, p. 143-157, 2009.

RAMBOCAS, M.; GAMA, J. *Marketing research: the role of sentiment analysis*. Universidade do Porto, Faculdade de Economia do Porto, 2013.

ROBALDO, L.; DI CARO, L. Opinion mining-ML. *Computer Standards & Interfaces*, Amsterdam, v. 35, n. 5, p. 454-469, 2013.

SANTOS, L. M. et al. Twitter, análise de sentimento e desenvolvimento de produtos: quanto os usuários estão expressando suas opiniões? *Prisma.com*, Belo Horizonte, n. 13, p. 1-12, 2011.

SARTORI, G. *A teoria da democracia revisitada*. São Paulo: Ática, 1994. 324 p.

SINGER, A. Brasil, junho de 2013, classes e ideologias cruzadas. *Novos Estudos-CEBRAP*, n. 97, p. 23-40, 2013.

SNEAD, J. T. Social media use in the U.S. executive branch. *Government Information Quarterly*, Amsterdam, v. 30, n. 1, p. 56-63, 2013.

SOBACI, M. Z.; KARKIN, N. The use of twitter by mayors in Turkey: tweets for better public services? *Government Information Quarterly*, Amsterdam, v. 30, n. 4, p. 417-425, 2013.

SOBKOWICZ, P.; KASCHEKY, M.; BOUCHARD, G. Opinion mining in social media: modeling, simulating, and forecasting political opinions in the web. *Government Information Quarterly*, Amsterdam, v. 29, n. 4, p. 470-479, 2012.

STEENKAMP, M.; HYDE-CLARKE, N. The use of facebook for political commentary in South Africa. *Telematics and Informatics*, Amsterdam, v. 31, n. 1, p. 91-97, 2014.

TOCQUEVILLE, A. *A democracia na América*. São Paulo: Martins Fontes, 1998. 597 p.

TSYTSARAU, M.; PALPANAS, T. Survey on mining subjective data on the web. *Data Mining and Knowledge Discovery*, Oxford, v. 24, n. 3, p. 478-514, 2012.

VINODHINI, G.; CHANDRASEKARAN, R. M. Sentiment analysis and opinion mining: a survey. *International Journal of Advanced Research in Computer Science and Software Engineering*, v. 2, n. 6, p. 283-292, 2012.

YANG, D. H.; YU, G. A method of feature selection and sentiment similarity for Chinese micro-blogs. *Journal of Information Science*, Cambridge, v. 39, n. 4, p. 429-441, 2013.

YOON, S.; ELHADAD, N.; BAKKEN, S. A practical approach for content mining of tweets. *American Journal of Preventive Medicine*, New York, v. 45, n. 1, p. 122-129, 2013.

YU, Y.; DUAN, W.; CAO, Q. The impact of social and conventional media on firm equity value: a sentiment analysis approach. *Decision Support Systems*, Amsterdam, v. 55, n. 4, p. 919-926, 2013.

ZHANG, C. et al. Sentiment analysis of Chinese documents: from sentence to document level. *Journal of the American Society for Information Science and Technology*, New York, v. 60, n. 12, p. 2.474-2.487, 2009.

ZHANG, P.; HE, Z. A weakly supervised approach to Chinese sentiment classification using partitioned self-training. *Journal of Information Science*, Cambridge, v. 39, n. 6, p. 815-831, 2013.

**Submissão: 07/10/2015**

**Aprovação: 28/09/2016**