

# The application of the sentiment analysis technique in social media as a tool for social management practices at the governmental level

Daniel José Silva Oliveira<sup>1</sup>

Paulo Henrique de Souza Bermejo<sup>2</sup>

José Roberto Pereira<sup>3</sup>

Daniely Aparecida Barbosa<sup>4</sup>

<sup>1</sup> Universidade Federal de Minas Gerais / Centro de Pós-Graduação e Pesquisas em Administração, Belo Horizonte / MG — Brazil

<sup>2</sup> Universidade de Brasília / Núcleo de P&D para Excelência e Transformação da Administração Pública, Brasília / DF — Brazil

<sup>3</sup> Universidade Federal de Lavras / Departamento de Administração e Economia, Lavras / MG — Brazil

<sup>4</sup> Universidade Federal de São João del-Rei, São João del-Rei / MG — Brazil

Sentiment analysis is a knowledge discovery technique developed from data mining; its purpose is to reveal people's opinions on specific topics. This is an appropriate technique to apply to unstructured data sources, such as social media, that cover information on a variety of topics (such as politics and public administration). In this context, the objective of this study was to identify whether sentiment analysis can reflect public opinion and, thus, contribute to practices of social management. Therefore, the sentiment analysis technique was applied to reveal citizens' opinions, which were expressed on Twitter and concerned some of the main social programs in force during Brazil's Rousseff government. The study consisted of a comparison between the results of the sentiment analysis and the concepts and applications involving four strategies of social media used by governments from the point of view of social management. The results revealed that the sentiment analysis technique could contribute to social management practices in the context of the network strategy.

**Keywords:** sentiment analysis; opinion mining; social media; social management; public administration.

## A aplicação da técnica de análise de sentimento em mídias sociais como instrumento para as práticas da gestão social em nível governamental

A análise de sentimento é uma técnica de descoberta de conhecimento por meio da mineração de dados, sua finalidade é revelar a opinião das pessoas sobre temas específicos. Essa é uma técnica apropriada para aplicação em fontes de dados não estruturados, como as mídias sociais, que abarcam informações sobre diversos temas, inclusive política e administração pública. O objetivo deste estudo foi identificar se a análise de sentimento pode refletir a opinião pública e, assim, trazer contribuições para as práticas da gestão social. Para tanto, a técnica foi aplicada para revelar as opiniões dos cidadãos expressas no Twitter sobre alguns dos principais programas sociais em vigor no Brasil durante o governo Dilma Rousseff. O estudo consistiu no confronto entre os resultados da análise de sentimento e os conceitos e aplicações envolvendo quatro estratégias de utilização de mídias sociais pelos governos sob a ótica da gestão social. Os resultados da pesquisa revelaram que a técnica da análise de sentimento pode contribuir para as práticas da gestão social no contexto da estratégia de rede.

**Palavras-chave:** análise de sentimento; mineração de opinião; mídias sociais; gestão social; gestão pública.

## La aplicación de la técnica de análisis de sentimiento en medios sociales como instrumento para las prácticas de la gestión social a nivel gubernamental

El análisis de sentimiento es una técnica de descubrimiento de conocimiento a partir de la minería de datos que tiene la finalidad de revelar la opinión de las personas sobre temas específicos. Esta es una técnica apropiada para aplicación en fuentes de datos no estructurados, como los medios sociales, que abarcan información sobre diversos temas, inclusive política y administración pública. En este ámbito, el objetivo de este estudio fue identificar si el análisis de sentimiento puede reflejar la opinión pública y, así, traer contribuciones a las prácticas de la gestión social. Para ello, la técnica se aplicó para revelar la opinión de los ciudadanos expresada en Twitter sobre algunos

DOI: <http://dx.doi.org/10.1590/0034-7612174204>

Article submitted on January 7, 2017 and accepted for publication on June 27, 2017.

[Translated version] Note: All quotes in English translated by this article's translator.



de los principales programas sociales vigentes en Brasil durante el gobierno de Dilma Rousseff. El estudio consistió en la confrontación de los resultados del análisis de sentimiento con los conceptos y aplicaciones que involucran cuatro estrategias de utilización de medios sociales por parte de los gobiernos bajo la óptica de la gestión social. Los resultados de la investigación revelaron que la técnica de análisis de sentimiento puede contribuir a las prácticas de la gestión social en el contexto de la estrategia de red.

**Palabras clave:** análisis de sentimiento; minería de opiniones; medios de comunicación social; gestión social; gestión pública.

## 1. INTRODUCTION

The emergence of social media has radically altered the way that individuals communicate (Kontopoulos, Berberidis, Dergiades, & Bassiliades, 2013). Citizens are increasingly using social media to communicate with family, friends, companies and even the government (Kavanaugh et al., 2012). Faced with these new possibilities of interaction, citizens no longer conform to filling only the role of passive recipients of government policies (Bonsón, Torres, Royo, & Flores, 2012). Given this trend, governments have increasingly adopted social media as a mechanism to establish relationships with their citizens (Gustafsson, 2012; Mergel, 2013b; Snead, 2013). This interaction between state and society can contribute to incorporating the opinion of civil society into public administration (Guimarães & Amorim, 2013). This is in line with social management, which, according to Cançado, Pereira and Tenório (2015), is a dialogical management action that seeks to convert communication flows into actions and political decisions. However, the challenge is to measure the results of this communication, as is measuring citizens' opinion in response to the government's actions (Mergel, 2013a).

Adopting innovative techniques to analyze social media is creating expectations about future opportunities for innovation and the emergence of tools that make using these media as a source of knowledge possible (Criado, Sandoval-Almazan, & Gil-Garcia, 2013). Sentiment analysis, also known as "opinion mining", is a promising technique that aims to identify polarized (positive, negative, or neutral) opinions on certain topics in unstructured databases (Sobkowicz, Kaschesky, & Bouchard, 2012).

In light of this, the objective of this study is to investigate how the sentiment analysis technique can contribute to the practices of social management by means of identifying the opinion of the citizens in relation to subjects that involve public interest. In this way, this study hopes to establish the possibility, even if preliminary, of the participation of civil society in the government's decision-making process. It is therefore proposed to answer the following question: How can the use of the sentiment analysis technique, based on the opinion of citizens drawn from social media, contribute to social management practices? To answer this question, data mining tools were used to identify the opinions of citizens, which were expressed on Twitter and concerned some of the main social programs developed during Brazil's Rousseff government. Subsequently, a comparison was made between the results of the sentiment analysis and the concepts and applications involving four strategies of social media use by governments from a social management perspective.

This study is justified by the lack of research on the development and analysis of methods and techniques appropriate to the practice of social management (Paes de Paula, 2010; Cançado et al., 2015). In addition, it is necessary to test new tools and methods that can measure the potential of social media to increase social participation and strengthen democratic institutions and processes (Sobaci & Karkin,

2013; Snead, 2013). According to D. J. S. Oliveira, Bermejo and Santos (2015), it is necessary to evaluate whether the sentiment analysis can contribute to promote a participatory public management, establishing a dialogical relationship between the state and society. It is intended to contribute to filling these gaps.

To do so, this article is structured as follows: In the next section, governments’ strategies of social media use are presented. The third section discusses the participatory potential offered by information and communication technologies. In the fourth section, the main concepts of social management in relation to the state and society are discussed. The fifth section describes the main approaches to the sentiment analysis technique. The methodology that was applied in this study is detailed in the sixth section. In the seventh section, the results are presented, and finally, the conclusion of this study is presented in the eighth section.

## 2. GOVERNMENTS’ STRATEGIES FOR USING SOCIAL MEDIA

The strategic adoption of social media tools (such as blogs, microblogs, wikis, and electronic social networks) by governments is changing the landscape of public agencies and bureaucracies worldwide (Mergel, 2013a). In the United States, the government uses social media to promote transparency, participation, and collaboration, which were predominant features of the Obama administration’s “Open Government Initiative” (Chun & Luna-Reyes, 2012; Criado et al., 2013; Meijer & Thaens, 2013; Mergel, 2012, 2013a, 2013b; Snead, 2013). Other studies show that social media is being used by governments in many other countries, such as China (Zheng, 2013), Turkey (Sobaci & Karkin, 2013), Mexico (Picazo-Vela, Gutiérrez-Martínez, & Luna-Reyes, 2012), Sweden (Gustafsson, 2012), the Netherlands (Bekkers, Edwards, & Kool, 2013), and other countries of the European Union (Bonsón et al., 2012). The focus of social media use by these governments varies widely, from accountability to public security and natural disaster prevention (Meijer & Thaens, 2013; Kavanaugh et al., 2012).

In recent years, research about the process of social media adoption by governments has demonstrated the application of different strategies. Mergel (2012) and Meijer and Thaens (2013) present four types of strategies, as shown in Box 1.

### BOX 1 GOVERNMENTS’ STRATEGIES FOR USING SOCIAL MEDIA

Type of strategy	Government-citizen relations	Role of social media
Push strategy	Citizens as an audience for government information	Social media as a broadcasting channel
Pull strategy	Citizens as sensors for the government	Social media as a channel for citizen input
Network strategy	Citizens as co-producers of government policies	Social media as an interactive channel
Transaction strategy	Citizens as partners in government activities	Social media as an interactive channel

Source: Based on Mergel (2012) and Meijer and Thaens (2013).

In governments that use the push strategy, social media is predominantly used to convey information to citizens without interaction, like in traditional mass media (Mergel, 2012). In the pull strategy, however, governments use social media to attract citizens' attention so that they can provide feedback through comments (Meijer & Thaens, 2013). In governments adopting the network strategy, social media is not only used to control or direct messages to its citizens, but also as a strategic tool for sharing information and creating knowledge through interaction and discussion of real issues with citizens (Mergel, 2013c). Finally, in governments that use the transaction strategy, social media is used for real transactions between governments and citizens, that is, there is the direct provision of public services through social media applications (Meijer & Thaens, 2013).

According to Chun and Luna-Reyes (2012), the use of social media by governments can be considered a technological innovation, as well as a transforming agent. Thus, it generates civic engagement through social participation initiatives that promote more democratic decisions. These topics will be covered in the next section.

### 3. THE PROMOTION OF SOCIAL PARTICIPATION

Cançado et al. (2015) believe that a representative democracy in Brazil can open up more space for citizen participation as the processes of communicative action expand throughout society. In this sense, it is necessary to look for new ways to overcome the existing level of democracy (Dahl, 2001) through participatory democracy, which can be defined as “[...] a pyramidal system, with direct democracy at the base and democracy by delegation at other levels” (Paes de Paula, 2005, p. 160). Thus, the shift of representative democracy towards participatory democracy is within the reach of society, especially with the emergence of new information and communication technologies (Cançado et al., 2015).

Today, these technologies (especially internet-based tools) have the potential to contribute significantly to the promotion of large-scale social participation, breaking down the barriers of representative democracy (Chun & Luna-Reyes, 2012; Criado et al., 2013; Sobaci & Karkin, 2013). These technologies contribute to the association of citizens towards common goals; this, according to Habermas (2003), elevates society to a level where it can be heard by its rulers.

The defense of social participation does not imply a constant public deliberation on all the subjects, imposing to the citizens a permanent and exclusive involvement in the public sphere because “[...] the important one is that people deal with the central questions and that the fundamental information is available” (Guimarães & Amorim, 2013, p. 130). The challenge is to build a model in which the whole is represented by its parts, “[...] overcoming the geospatial and functional fragmentation typical of classical bureaucracies” (Keinert, 2007, p. 97). This is a challenge that has already been set out by Alexis de Tocqueville in his book *Democracy in America*, which was originally published in 1835. He stated that “[...] so far no political form had been found to promote the development and prosperity of all classes that make up society” (Tocqueville, 1998, p. 180). Currently, a field of knowledge emerges that proposes a new form of management in an attempt to fill this secular gap: social management.

#### 4. SOCIAL MANAGEMENT

According to Cançado et al. (2015), social management considers society to be the protagonist in relations between the state and the market. This is a dialogical management process where decisions are made by everyone involved; in other words, “[...] the social adjective describing the management noun is perceived as the privileged space of social relations where everyone has the right to speak, without any type of coercion” (Tenório, 2006, p. 1146).

To understand this universe of social management, it is necessary to first better understand some of the key terms that are often related to the topic in the literature, such as: civil society, public sphere, communication flows, deliberative citizenship, and alternative knowledge (Cançado et al., 2015; Fischer et al., 2006; Guimarães & Amorim, 2013; Habermas, 2003; Keinert, 1994, 2007; Natividade, Pereira, & V. A. R Oliveira, 2011; Tenório, 1998, 2006, 2008).

According to Habermas (2003, p. 100), *civil society* is composed of “[...] movements, organizations and associations, which capture the echoes of social problems that resonate in private spheres, condense them and then transmit them to the sphere Public policy.” For this author, civil society has its own opinions that are capable of influencing political decisions in a democratic context. According to Cançado et al. (2015, p. 101), “[...] social management is located in the intersection of relations between State, Market and Society and becomes capable of converting the communication flows of public spheres into political actions and decisions from of its organization in civil society.” They believe that the main tasks of social management is to systematize alternative knowledge, to structure specialized technical assessments aimed at the common good, and to strengthen the public sphere by providing quality information to it. From this perspective, it is considered that social media has contributed greatly to the development of social management since it provides opportunities for society to participate in political issues, taking into account the formation of public opinion.

Keinert (2007) draws attention to an important issue in considering that, although the *public sphere* is historically linked to the state, its roots lie in society itself. According to Habermas (2003), the public sphere is a network of communication of contents, positions and opinions where the communication flows (based on the natural language) are condensed into specific public opinions. In this way, it is up to the public sphere to reinforce the pressure exerted by the problems and not simply identify them, seeking to problematize them and dramatize them until they are assumed by the rulers (Habermas, 2003). From this perspective, social management, when occupying the intermediation space between the state, society, and market, must identify, understand, and propose solutions to the problems of society (Cançado et al., 2015; Tenório, 2008).

The *communication flows* of the public sphere are closely linked to the domains of private life, and they make it possible for civil society to identify new problems before the state’s agents even do (Habermas, 2003). In this way, social management, based on a communicative action, would be a way to expand the possibilities of social emancipation (Cançado et al., 2015). According to Habermas (2003, p. 22), this communicational network forms “[...] arenas in which a more or less rational formation of opinion can and will take place on subjects relevant to the whole society”. Therefore, the formation of a public communication system is fundamental to renew the discursive basis of popular sovereignty and give a voice to those who are socially or culturally marginalized and disadvantaged (Guimarães & Amorim, 2013). In this context, Habermas (2003, p. 84) considers that “[...] the integration of a



highly complex society is not effected through a paternalistic system that ignores the communicative power of the citizens”.

Thus, social management practices are related to the principles of *deliberative citizenship*, which constitutes a participatory process of reflection and decision based on an understanding between the parties concerning decision making within the public sphere (Cançado et al., 2015; Tenório, 2008). In general terms, deliberative citizenship means that the legitimacy of decisions must originate in “[...] discussion processes guided by the principles of inclusion, pluralism, participatory equality, autonomy and the common good” (Tenório, 2008, p. 160). In addition to these principles and specialized knowledge, alternative knowledge, which is formulated by the knowledge and information of the general public, forms what Habermas (2003) calls the community of interpreters. Just because the public is composed of laity and “[...] public communication in a common speech understandable to all does not necessarily mean an obscuring of the essential questions or reasons that leads to a decision” (Habermas, 2003, p. 107).

After clarifying the concept of social management, how its principles can be put into practice still must be determined. In this sense, Cançado et al. (2015) consider the need to create management instruments that can effect social management proposals, corroborating Guimarães and Amorim (2013, p. 138) when they proposed “[...] the formation of a democratic public opinion as a horizon, it invites the democratic imagination to create the conditions of possibility of its maximum realization in the historical context in which it is exercised”.

Under this circumstance, this study proposes to test the sentiment analysis technique as a possible way to identify important information for the decision-making process of public administration from the dialogical relationship between the government and civil society. This is done from the point of view of social management.

## 5. SENTIMENT ANALYSIS

Also known as opinion mining and favorability analysis (Pang & Lee, 2008; Kontopoulos et al., 2013), sentiment analysis is a subfield of natural language processing that is concerned with extracting, classifying, and analyzing opinions on various topics in large volumes of textual data (Yoon, Elhadad, & Bakken, 2013). Its purpose is to classify texts not by topic, but by the sentiment or opinion contained in the database (Di Caro & Grella, 2013). This is a technique that has, as its basic task, the automatic classification of databases containing polarized (positive, negative, or neutral) opinions on previously defined themes (Yang & Yu, 2013; Pang & Lee, 2008; Sobkowicz et al., 2012). With the aid of information technology, sentiment analysis drastically reduces the need to read a large number of documents to extract opinions (Yu, Duan, & Cao, 2013).

Sentiment analysis seeks to meet the challenge of identifying and extracting subjective information from large volumes of unstructured data through a combination of data mining, natural language processing, information retrieval, and knowledge management techniques (Mouthami, Devi, & Bhaskaran, 2013). Currently, the sentiment analysis technique is applied under two main approaches: supervised learning and unsupervised learning. Supervised learning requires the manual sorting of database samples (or training set) to create text-based patterns that will serve as a parameter for the automatic classification of the entire database with the aid of machine learning algorithms (Pang

& Lee, 2008). Unsupervised learning, however, is based on semantic orientation based on a set of words previously constructed, containing positive and negative terms, so that automated polarity identification is determined by the frequency of these terms in the text — without the need for a sample of the database to be previously classified (Yu et al., 2013). The advantage of supervised learning (compared to unsupervised learning) is that, when sorting data samples, words, or sentences belonging to the same domain and context of the database (which were submitted to the automated classification), the information will be incorporated into the classifier, providing greater precision in the results (Liu & Zhang, 2012). For this reason, the supervised learning approach prevails in most research (Kontopoulos et al., 2013; Yang & Yu, 2013).

For Yu et al. (2013), there are three main reasons for choosing sentiment analysis as a research approach: It 1) converts large volumes of unstructured data into information that allows predictions on specific issues, 2) constructs models to aggregate collective opinions and reveals useful information about a population's behavior to predict future trends, and 3) makes it possible to gather information about people's opinions on various topics.

The sentiment analysis technique can be applied to various types of textual data, both online and offline. However, the rise of social media such as blogs, microblogs and social networking sites has sparked the interest of several researchers in exploring the opinions of users of these media through the sentiment analysis (Mouthami et al., 2013). According to Zhang and He (2013), the fact that citizens are increasingly using such media to express their opinions on varied subjects has made social media one of the main sources of data for sentiment analysis. The growth in the number of searches involving sentiment analysis has brought significant progress to the opinion tracking techniques of social media, especially for those who use Twitter as a source of data (Yu et al., 2013).

## 6. METHODOLOGY

The present study is an applied, interdisciplinary, and exploratory research that can be classified as qualitative regarding the process of data classification and classified as quantitative when verifying the frequency of positive, negative, and neutral opinions contained in the database (Rambocas & Gama, 2013). The methodological procedures adopted were based on the studies of Yoon et al. (2013) and D. J. S. Oliveira et al. (2015, 2017), which consisted of the following steps: 1) define the object of study, 2) choose the data source, 3) define keywords for data mining, 4) select the opinion mining application, 5) collect and prepare the data, 6) delimit the training set, 7) perform automated classification, and 8) validate the results.

The object of study (step 1) was citizens' views expressed on social media about the major social programs currently being maintained in Brazil by the federal government. The four social programs<sup>1</sup>, considered by the Brazilian press as “flags” of the Rousseff government (Carazzai & Britto, 2014), selected for analysis were:

---

<sup>1</sup> More information about each of the social programs is available in the “Plans and Programs” section of Portal Brasil: [www.brasil.gov.br/planos-e-programas](http://www.brasil.gov.br/planos-e-programas).

- *Bolsa Família* (Family Allowance): a program of direct income transfer that benefits families in poverty and extreme poverty throughout the country;
- *Minha Casa, Minha Vida* (My House, My Life): a housing finance program that provides discounts and benefits and that decreases housing insurance values, facilitating home ownership;
- *Mais Médicos* (More Doctors): a program aimed to improve the services for public health users through investments in the infrastructure of hospitals and health facilities, as well as providing Brazilian and foreign doctors to regions where there are shortages; and
- *Programa Nacional de Acesso ao Ensino Técnico em Emprego - Pronatec* (National Program for the Access to Technical Education and Employment): a program that aims to increase the supply of free courses in vocational and technological education in Brazil.

The data source (step 2) was chosen based on three criteria: it should be a popular social media platform in Brazil; the posts of the users should be, predominantly, in the form of text; and there should be technical feasibility for mining and data analysis. Thus, the social media platform chosen was Twitter<sup>2</sup>, which can be accessed on both desktops and mobile devices, where its users share information and opinions through short messages, or tweets, limited to 140 characters (Deller, 2011; Bae & Lee, 2012; Larsson & Moe, 2011). Because of this limitation, tweets tend to be more concise, expressing users' opinions more objectively, which makes processing such data more effective than the long posts allowed on other social media (Kontopoulos et al., 2013).

The next step (step 3) was to define the keywords used for performing data mining. Aiming to collect as many posts as possible on each of the four chosen social programs, the keywords used in data mining were defined in two distinct patterns based on the advanced search rules of Twitter. The first pattern was based on the use of the names of the four social programs in quotation marks, since this searches for only the exact expression, reducing noise in the data. The second standard was defined through the official hashtags<sup>3</sup> used in the sharing of information regarding these programs. In this way, the keywords used in data mining were defined according to Box 2.

## BOX 2 KEYWORDS USED IN DATA MINING

Social Programs	Hashtags
"Bolsa Família"	#BolsaFamília
"Minha Casa, Minha Vida"	#MinhaCasaMinhaVida
"Mais Médicos"	#MaisMédicos
"Pronatec"	#Pronatec

**Source:** Elaborated by the authors.

<sup>2</sup> Twitter: [www.twitter.com](http://www.twitter.com).

<sup>3</sup> *Hashtag*: Word or phrase that precedes the # symbol and classifies or categorizes the text that accompanies it, as in a tweet (Merriam-Webster, 2014).



The selected software (step 4) for conducting this study was DiscoverText<sup>4</sup>. Applications such as Rapidminer<sup>5</sup> and Scup<sup>6</sup> were analyzed, but DiscoverText was chosen due to its technical feasibility, free availability for academic purposes, and performance in other studies (e.g., Beyer, 2012; D. J. S. Oliveira et al., 2015, 2017). DiscoverText is a cloud-based text analysis application that captures, filters, and classifies a large amount of structured and unstructured data. It facilitates the capture of social media data, provides well-integrated analytical tools to explore and encode the data (Beyer, 2012). Using this application, it is possible to customize and reuse machine learning classifiers that allow the combined use of algorithms and human coding (supervised learning), which increases the accuracy of the results (Kontopoulos et al., 2013).

Data collection (step 5) was performed between July 3 and August 3, 2014 through the DiscoverText application, using the keywords defined in Table 2. After the data collection, the data were prepared by applying filters to eliminate noise, such as tweets in other languages, hyperlinks, and duplicate data.

The size of the training set (step 6) considered for this work was defined as 30% of the total tweets collected on each one of the chosen social programs, since D. J. S. Oliveira et al. (2015) identified that the manual classification of 30% of the data collected is sufficient to obtain satisfactory accuracy in the automated classification using DiscoverText. Thus, the training set was calculated according to Table 1.

**TABLE 1** TRAINING SET CALCULATION

Social Programs	Raw Data	Prepared Data	Training Set (30%)
Bolsa Família	12,598	2,033	610
Minha Casa, Minha Vida	15,537	2,525	758
Mais Médicos	16,641	2,342	703
Pronatec	14,236	3,276	983
<b>TOTAL</b>	<b>59,012</b>	<b>10,176</b>	<b>3,054</b>

**Source:** Elaborated by the authors.

According to Table 1, there was a large difference between the raw data and the prepared data. This is due to the application of filters that eliminated duplicate tweets and other noise in the raw data — as well as the exclusion of posts made by news agencies — resulting in a dataset composed only of tweets composed by Brazilian citizens.

<sup>4</sup> DiscoverText: [www.discovertext.com](http://www.discovertext.com).

<sup>5</sup> Rapidminer: <https://rapidminer.com>.

<sup>6</sup> Scup: [www.scup.com](http://www.scup.com).

After being calculated, each training set was manually classified to establish rules and standards (machine learning) so that the application could classify all the data in an automated way. After the manual classification of each training set, the automated classification of all the data sets (step 7) was performed, generating the sentiment (positive, negative, or neutral) toward each of the social programs.

After the automated classification, the results were validated (step 8) to determine how accurate they were. The accuracy of the results was calculated based on a simple random sample of the automated data (Cochran, 2007). The sample size maintained the same proportion of the training set (30% of the data). The accuracy calculation was based on the model of Eirinaki, Pisal, and Singh (2012), dividing the total number of correct classifications by the total number of each sample, as shown in Table 2.

**TABLE 2 ACCURACY OF AUTOMATED DATA ANALYSIS**

	A	B	C	D
Programas Sociais	Data classified automatically	Sample to calculate the accuracy (A × 30%)	Data classified correctly	Accuracy (%) $C \div B \times 100$
Bolsa Família	1,423	427	318	74.47%
Minha Casa, Minha Vida	1,767	530	441	83.20%
Mais Médicos	1,639	492	385	78.25%
Pronatec	2,293	688	592	86.04%

**Source:** Elaborated by the authors.

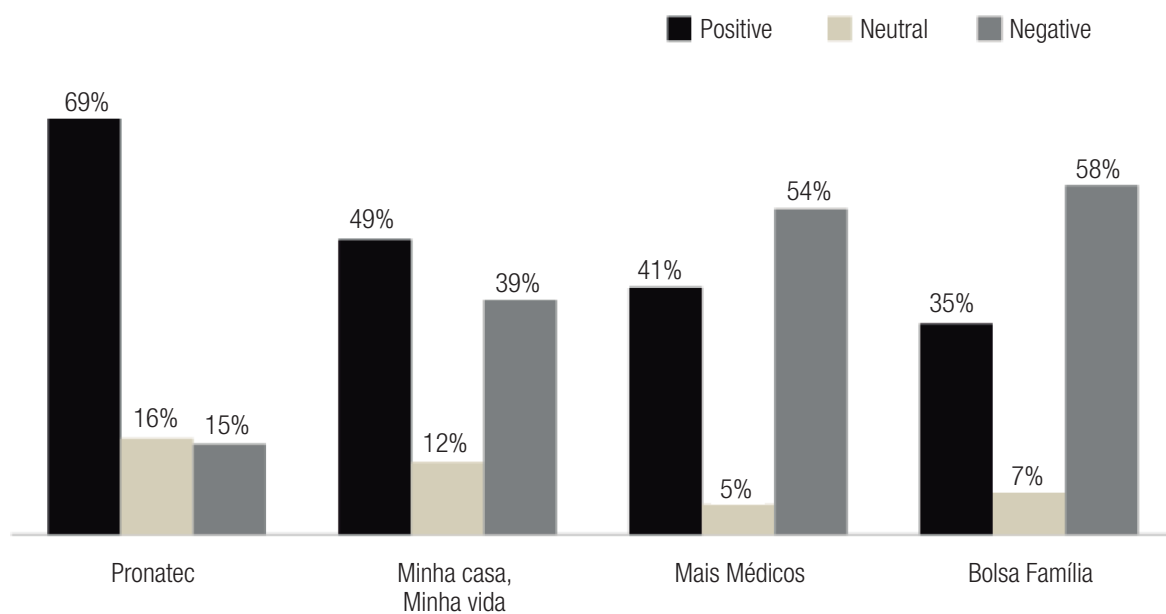
Column A of Table 2 represents the automated classified data, which is the result of the subtraction between the prepared data and the training set that was manually sorted (see Table 1). Column B refers to the sample for the calculation of precision, which was defined as 30% of the data in column A. Column C corresponds to the amount of sample data correctly sorted by the application. Column D represents the accuracy of the automated classification, which was obtained by dividing the number of correct classifications (column C) by the total of each sample (column B).

The accuracy achieved by the application used in this study can be considered to be satisfactory, since in the literature a tool that generates an average of 80% hits is considered satisfactory (D. J. S. Oliveira & Bermejo, 2017; Mostafa, 2013; Yoon et al., 2013; Yu et al., 2013). Even though some datasets had an accuracy below 80%, it can be stated that the reliability achieved by the arithmetic mean of the four presented datasets is 80.49%.

## 7. RESULTS AND DISCUSSION

The results of the automated classification of tweets related to social programs revealed that for *Pronatec* and *Minha Casa, Minha Vida*, there was a predominance of positive opinions, while for *Mais Médicos* and *Bolsa Família*, opinions were mostly negative, as shown in Graph 1.

**GRAPH 1** POLARIZED OPINION OF BRAZILIANS REGARDING FOUR SOCIAL PROGRAMS



Source: Elaborated by the authors

Based on sentiment analysis, the social program that presented the highest number of positive tweets was the *Pronatec* program (69%). This was also the program with the highest number of neutral tweets (16%). The *Minha Casa, Minha Vida* program had 49% positive tweets compared to negative tweets (39%). With predominantly negative opinions, the *Mais Médicos* and *Bolsa Família* programs had 54% and 58% negative tweets, respectively. However, these social programs were those that presented less precision in the automated classification. One reason for this was false-positive classification due to tweets with high irony loads, slang, or misspellings (Montoyo, Martínez-Barco, & Balahur, 2012). Some examples of tweets classified in this study are shown in Box 3.

**BOX 3**      **EXAMPLES OF TWEETS CLASSIFIED BY POLARITY**

Polarity	Tweets in Portuguese	Tweets in English
Positive	“somos todos #Pronatec! haha amoooo!” (User A, Twitter, 07/17/2014).	“we are all #Pronatec! haha I love!”
Negative	“o Minha Casa, Minha Vida, do jeito que é feito, não consegue solucionar o déficit habitacional do Brasil” (User B, Twitter, 07/24/2014).	“Minha Casa, Minha Vida, the way it is done, cannot solve the housing deficit in Brazil”
Neutral	“Vale a pena ler pelo depoimento de médicos e pacientes: Como está o programa Mais Médicos no RS um ano depois” (User C, Twitter, 07/27/2014).	“This report by doctors and patients is worth reading: How is the Mais Médicos program in RS a year later”
False-positive	“Vou dormir pra ter energia e aumentar o PIB amanhã... e ajudar a pagar a bolsa família...” (User D, Twitter, 07/23/2014).	“I sleep to have energy and increase GDP tomorrow... and help pay the Bolsa Família...”

Source: Elaborated by the authors.

To determine whether sentiment analysis can contribute to social management practices, the concepts and applications involved in the four strategies of social media used by governments (Meijer & Thaens, 2013; Mergel, 2012), social management (Cançado et al., 2015; Tenório, 2006), and the sentiment analysis technique (Pang & Lee, 2008; Sobkowicz et al., 2012) were examined (Box 4).

**BOX 4**      **RELATIONSHIPS BETWEEN STRATEGIES OF SOCIAL MEDIA USE, SOCIAL MANAGEMENT, AND SENTIMENT ANALYSIS**

Strategies of social media use	Does it meet the principles of social management?	Can it be applied to sentiment analysis?
Push strategy	No	No
Pull strategy	No	Yes
Network strategy	Yes	Yes
Transaction strategy	Yes	No

Source: Elaborated by the authors.

According to Box 4, sentiment analysis contributes to social management practices by promoting coproduction of policies in governments that adopt a network strategy for social media. The principles of social management include inclusion, pluralism, dialogue, autonomy, participatory equality, social emancipation, and the common good (Cançado et al., 2015; Tenório, 1998, 2006, 2008). This is a dialogical management action focused on non-state public interest capable of converting

communication flow in the public sphere — including alternative knowledge — into political actions and decisions (Cançado et al., 2015). According to Mergel (2012) and Meijer and Thaens (2013), the push strategy, in which citizens are receivers of information, and pull strategy, in which citizens are government sensors, do not meet the requirements for social management practices because both are monological forms of communication. Network strategy, in which citizens are co-producers of government policies, and transaction strategy, in which citizens are partners in government activities, are directly related to principles of social management because these strategies establish dialogical relationships between state and society. Thus, push and pull strategies tend to have more centralized arrangements and limitations in relation to the public (Meijer & Thaens, 2013).

This research found that push and transaction strategies cannot be applied to sentiment analysis. Push strategy communication occurs unilaterally, from government to citizens, through the dissemination of information in official profiles of government agencies; it is not feasible to use sentiment analysis in this case because official publications are informative rather than opinionated. In transaction strategy, citizens act as government partners. They contribute, through social media, to the dissemination and organization of activities related to campaigns and programs promoted by the government, such as healthcare, education, environment, and public safety programs. Thus, sentiment analysis is not adequate for transaction strategy because it identifies opinions among citizens on certain subjects but does not contribute to the execution of public services.

In contrast, pull and network strategies can be applied to sentiment analysis. Using the pull strategy, sentiment analysis measures citizen opinions on various government-related issues. When using network strategy, sentiment analysis contributes to coproduction of government policies. According to Meijer and Thaens (2013), network strategy decisions are decentralized and serve multiple coproduction goals between governments and citizens because citizen opinions are evaluated when creating public policy.

Social media platforms are effective tools for exercising freedom of expression, which is fundamental for citizens to participate more actively in public management and to pressure and direct government decisions (Habermas, 2003). These media allow the sharing of large volumes of information and opinions that can be used in governmental decision-making processes (Mergel, 2013b). The challenge, in this case, is to transform large amounts of unstructured data into knowledge that can influence government decisions (Mergel, 2013a), based on the social management model of Cançado et al. (2015).

## 8. CONCLUSION

This study proposed that the sentiment analysis, performed by mining data extracted from social media, can be used to measure public opinions to contribute to social management practices. The theoretical review revealed that social media has great potential to increase social participation and strengthen democracy, but transforming large amounts of unstructured data into knowledge is necessary. The analysis results accurately identified the number of positive, negative, and neutral opinions about the four social programs adopted as the objects of this study. Two social programs that created predominantly positive opinions were *Pronatec* and *Minha Casa, Minha Vida*, while negative opinions were more common for the programs *Mais Médicos* and *Bolsa Família*. These results



revealed that the sentiment analysis technique can contribute to social management practices by measuring opinions in civil societies on various topics of public interest. The results reveal informal public opinions that are ideally formed in structures of a public sphere not corrupted by the power (Habermas, 2003), using Twitter as an alternative and independent data source that enabled the citizens to exercise their freedom of expression (Dahl, 2001). Thus, analysis of feeling can be used to study the relationship between state and society by systematizing and converting communication flow into knowledge that can be incorporated in civil society's opinion and the body politic (Guimarães & Amorim, 2013). It is civil society that institutionalizes discourse capable of influencing political decisions (Cançado et al., 2015); thus, it is necessary to establish public management that does not centralize decision-making power in the state and that considers the complexity of relationships between all actors involved by encouraging various channels of participation (Paes de Paula, 2005).

Finally, the analysis of feeling is an advantageous technique for governments adopting the network strategy for social media and social management practices. However, this study had some limitations. First, the results of the sentiment analysis were restricted to the opinions of citizens during the period of data collection, specifically opinions of Twitter users who were not necessarily served by the social programs used in this study. This must be considered to avoid creating a tyranny of the majority. Another limitation was reduced precision of the automated classification of texts due to high irony load. The presence of irony in texts is a significant challenge to sentiment analysis (Montoyo et al., 2012). To overcome these limitations, future research is needed to test the effectiveness of the sentiment analysis technique in different situations and to search for new technologies that can give voice to citizens.

## REFERENCES

- Bae, Y., Lee, H. (2012). Sentiment analysis of twitter audiences: measuring the positive or negative influence of popular twitterers. *Journal of the American Society for Information Science and Technology*, 63(12), 2521-2535.
- Bekkers, V., Edwards, A., & Kool, D. (2013). Social media monitoring: responsive governance in the shadow of surveillance? *Government Information Quarterly*, 30(4), 335-342.
- Beyer, Y. (2012). Using discovertext for large scale twitter harvesting. *Microform & Digitization Review*, 41(3-4), 121-125.
- Bonsón, E., Torres, L., Royo, S., & Flores, F. (2012). Local e-government 2.0: social media and corporate transparency in municipalities. *Government Information Quarterly*, 29(2), 123-132.
- Cançado, A. C., Pereira, J. R., & Tenório, F. G. (2015). *Gestão social: epistemologia de um paradigma* (2a ed.). Curitiba: Editora CRV.
- Carazzai, E. H., & Britto, P. (2014, 3 de julho). *Muda melhor quem sabe mudar*, diz Dilma, em clima de campanha. Folha de S. Paulo, Recuperado de <www1.folha.uol.com.br/poder/2014/07/1480765-muda-melhor-quem-sabe-mudar-diz-dilma-em-clima-de-campanha.shtml>
- Chun, S. A., & Luna-Reyes, L. F. (2012). Social media in government. *Government Information Quarterly*, 29(4), 441-445.
- Cochran, W. G. (2007). *Sampling techniques*. Nova York: John Wiley & Sons.
- Criado, J. I., Sandoval-Almazan, R., & Gil-Garcia, J. R. (2013). Government innovation through social media. *Government Information Quarterly*, 30(4), 319-326.
- Dahl, R. A. (2001). *Sobre a democracia*. Brasília: Editora da UnB.
- Deller, R. (2011). Twittering on: audience research and participation using Twitter. *Participations*, 8(1), 216-245.
- Di Caro, L., & Grella, M. (2013). Sentiment analysis via dependency parsing. *Computer Standards & Interfaces*, 35(5), 442-453.
- Eirinaki, M., Pisal, S., & Singh, J. (2012). Feature-based opinion mining and ranking. *Journal of Computer and System Sciences*, 78(4), 1175-1184.
- Fischer, T., Melo, V. P., Carvalho, M. R. D. O., Jesus, A. D., Andrade, R. A., & Waiandt, C. (2006). Perfis visíveis na gestão social do desenvolvimento. *Revista de Administração Pública*, 40(5), 789-808.
- Guimarães, J., & Amorim, A. P. (2013). *A corrupção da opinião pública: uma defesa republicana da liberdade de expressão*. São Paulo: Boitempo.
- Gustafsson, N. (2012). The subtle nature of Facebook politics: swedish social network site users and political participation. *New Media & Society*, 14(7), 1111-1127.
- Habermas, J. (2003). *Direito e democracia: entre facticidade e validade*. Rio de Janeiro: Tempo Brasileiro.
- Kavanaugh, A. L., Fox, E. A., Sheetz, S. D., Yang, S., Li, L. T., Shoemaker, D. J., . . . Xie, L. (2012). Social media use by government: from the routine to the critical. *Government Information Quarterly*, 29(4), 480-491.
- Keinert, T. M. M. (1994). Os paradigmas da administração pública no Brasil (1900-92). *Revista de Administração de Empresas*, 34, 41-48.
- Keinert, T. M. M. (2007). *Administração pública no Brasil: crises e mudanças de paradigmas* (2a ed.). São Paulo: Annablume.
- Kontopoulos, E., Berberidis, C., Dergiades, T., & Bassiliades, N. (2013). Ontology-based sentiment analysis of Twitter posts. *Expert Systems with Applications*, 40(10), 4065-4074.
- Larsson, A. O., & Moe, H. (2011). Studying political microblogging: Twitter users in the 2010 swedish election campaign. *New Media & Society*, 14(5), 729-747.
- Liu, B., & Zhang, L. (2012). A survey of opinion mining and sentiment analysis. In C. C. Aggarwal, & C. X. Zhai (Ed.), *Mining text data* (Cap. 13, pp. 415-463). Chicago: Springer.
- Meijer, A., & Thaens, M. (2013). Social media strategies: understanding the differences between North American police departments. *Government Information Quarterly*, 30(4), 343-350.

- Mergel, I. (2012). The social media innovation challenge in the public sector. *Information Polity*, 17(3-4), 281-292.
- Mergel, I. (2013a). A framework for interpreting social media interactions in the public sector. *Government Information Quarterly*, 30(4), 327-334.
- Mergel, I. (2013b). Social media adoption and resulting tactics in the U.S. federal government. *Government Information Quarterly*, 30(2), 123-130.
- Mergel, I. (2013c). *Social media in the public sector: a guide to participation, collaboration and transparency in the networked world*. San Francisco: John Wiley & Sons.
- Merriam-Webster. (2014). Definition of hashtag. *Merriam-Webster Online Dictionary*, Oxford. Recuperado de <www.merriam-webster.com/dictionary/hashtag>.
- Montoyo, A., Martínez-Barco, P., & Balahur, A. (2012). Subjectivity and sentiment analysis: an overview of the current state of the area and envisaged developments. *Decision Support Systems*, 53(4), 675-679.
- Mostafa, M. M. (2013). More than words: social networks' text mining for consumer brand sentiments. *Expert Systems with Applications*, 40(10), 4241-4251.
- Mouthami, K., Devi, N. K., & Bhaskaran, M. V. (2013). Sentiment analysis and classification based on textual reviews. In *Anais da International Conference on Information Communication and Embedded Systems (ICICES)* (pp. 271-276), Chennai: IEEE.
- Natividade, E. A., Pereira, J. R., & Oliveira, V. A. R. (2011). Gestão de política pública de trabalho e renda: uma análise do Programa Nacional de Incubadoras de Cooperativas Populares. In J. R. Pereira (Ed.), *Gestão social de políticas públicas* (Cap. 3, pp. 93-142). Lavras: Editora Ufla.
- Oliveira, D. J. S., & Bermejo, P. H. S. (2017). Mídias Sociais e Administração Pública: Análise do sentimento social perante a atuação do governo federal brasileiro. *Organizações & Sociedade*, 24(82), 491-508.
- Oliveira, D. J. S., Bermejo, P. H. S., & Santos, P. A. (2015). Sentiment analysis, social media, and public administration. In D. Ćemal, K. Ejub, R. Dragan, & S. Boban (Ed.), *Handbook of research on democratic strategies and citizen-centered E-government services* (Cap. 13, pp. 231-249). Hershey: IGI Global.
- Oliveira, D. J. S., Bermejo, P. H. S., & Santos, P. A. (2017). Can social media reveal the preferences of voters? A comparison between sentiment analysis and traditional opinion polls. *Journal of Information Technology & Politics*, 14(1), 34-45.
- Paes de Paula, A. P. (2005). *Por uma nova gestão pública: limites e potencialidades da experiência contemporânea*. Rio de Janeiro: Editora da FGV.
- Paes de Paula, A. P. (2010). Por uma nova gestão pública: reinserindo o debate a partir das práticas possíveis. In Cunha, A. S., Medeiros, B. A., & Aquino, L. M. C. (Ed.), *Estado, instituições e democracia: república* (Vol. 9, Cap. 15, pp. 505-530). Brasília: Ipea.
- Pang, B., & Lee, L. (2008, janeiro). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2), 1-135.
- Picazo-Vela, S., Gutiérrez-Martínez, I., & Luna-Reyes, L. F. (2012). Understanding risks, benefits, and strategic alternatives of social media applications in the public sector. *Government Information Quarterly*, 29(4), 504-511.
- Rambocas, M., & Gama, J. (2013). *Marketing research: the role of sentiment analysis*. Porto: Universidade do Porto.
- Snead, J. T. (2013). Social media use in the U.S. executive branch. *Government Information Quarterly*, 30(1), 56-63.
- Sobaci, M. Z., & Karkin, N. (2013). The use of twitter by mayors in Turkey: tweets for better public services? *Government Information Quarterly*, 30(4), 417-425.
- Sobkowicz, P., Kaschesky, M., & Bouchard, G. (2012). Opinion mining in social media: modeling, simulating, and forecasting political opinions in the web. *Government Information Quarterly*, 29(4), 470-479.
- Tenório, F. G. (1998). Gestão social: uma perspectiva conceitual. *Revista de Administração Pública*, 32(5), 7-23.
- Tenório, F. G. (2006). A trajetória do Programa de Estudos em Gestão Social (Pegs). *Rev. Adm. Pública*, 40(6), 1145-1162.

Tenório, F. G. (2008). *Tem razão a administração?* Ensaios de teoria organizacional (3a ed.). Ijuí: Editora da Unijuí.

Tocqueville, A. (1998). *A democracia na América* (4a ed.). Belo Horizonte: Itatiaia.

Yang, Dong-Hui, & Yu, G. (2013). A method of feature selection and sentiment similarity for chinese micro-blogs. *Journal of Information Science*, 39(4), 429-441.

Yoon, S., Elhadad, N., & Bakken, S. (2013). A practical approach for content mining of tweets. *American Journal of Preventive Medicine*, 45(1), 122-129.

Yu, Y., Duan, W., & Cao, Q. (2013). The impact of social and conventional media on firm equity value: a sentiment analysis approach. *Decision Support Systems*, 55(4), 919-926.

Zhang, P., & He, Z. (2013). A weakly supervised approach to chinese sentiment classification using partitioned self-training. *Journal of Information Science*, 39(6), 815-831.

Zheng, L. (2013). Social media in Chinese government: drivers, challenges and capabilities. *Government Information Quarterly*, 30(4), 369-376.

### Daniel José Silva Oliveira



<https://orcid.org/0000-0002-0475-6564>

PhD candidate in administration at the Federal University of Minas Gerais (UFMG); master's degree in public administration from the Federal University of Lavras (UFLA); bachelor's degree in public administration from the same university; specialist in public management at the Federal University of São João del Rei (UFSJ). E-mail: djsso@ufmg.br

### Paulo Henrique de Souza Bermejo



<https://orcid.org/0000-0003-2680-8342>

Master's degree in production engineering, specializing in applied intelligence, and a PhD in engineering and knowledge management, both from the Federal University of Santa Catarina (UFSC); postdoctoral fellow in innovation at Bentley University in Massachusetts; adjunct professor IV (i.e., on the permanent board) at the University of Brasília (UnB), with a post in the Administration Department of the Faculty of Accounting and Economics Administration (FACE/UnB) and a coordinating position in the R&D Nucleus for Excellence and Transformation of the Public Sector (NEXT/UnB); permanent professor and adviser in the post-graduate program in administration (master's and doctorate) at the UnB; holds a CNPq Scholarship for Technological and Scientific Development, category DTC-A. E-mail: paulobermejo@unb.br

### José Roberto Pereira



<https://orcid.org/0000-0003-1570-2016>

PhD in sociology from the UnB, a master's degree in administration from the UFLA; bachelor's degree in rural administration from the same university; full professor in public administration and social management at the UFLA; Scholarship Producer in Technological Development and Innovative Extension at CNPq; advisor for the professional master's in public administration program, in the academic master's program, and in the doctorate in administration program at the UFLA. E-mail: jpereira@dae.ufla.br

### Daniely Aparecida Barbosa



<https://orcid.org/0000-0001-7998-8581>

Specialist in public management from the Federal University of São João del-Rei (UFSJ); bachelor's degree in public administration from the UFLA. E-mail: danybarbosamg@gmail.com