Government Information Quarterly 36 (2019) 560–570

Contents lists available at ScienceDirect

Government Information Quarterly

journal homepage: www.elsevier.com/locate/govinf

From citizens to government policy-makers: Social media data analysis

Olfa Belkahla Driss[a,*], Sehl Mellouli[b], Zeineb Trabelsi[b]

[a] COSMOS laboratory, Ecole Nationale des Sciences de l'Informatique, Ecole Supérieure de Commerce de Tunis, Université de la Manouba, Tunisia
[b] Department of Information Systems, Laval University, G1k 0A6 Quebec, PQ, Canada

ARTICLE INFO

Keywords:
Social media
Government decision
Citizen engagement
Citizen participation
Text analysis
LSA

ABSTRACT

People are more and more using social media to express themselves about the different services that their governments are delivering. They can either provide positive or negative comments on government services. Hence, it becomes important for policy-makers to have the necessary tools to extract this valuable knowledge in a comprehensive way and that they may consider in their decision-making processes. This paper provides a generic framework, based on semantic analysis of text, to extract valuable data from social media in order to provide new information for government policy-makers. The proposed framework is based on a text semantic analysis tool that collects data from social networks and extracts valuable data to be presented to government policy-makers. The proposed framework is applied to analyze Facebook posts from a page that is managed by citizens in Tunisia. This page aims to report various problems and issues occurring in Tunisian cities.

1. Introduction

Social media platforms can be considered as channels means for communication between governments and citizens (Bonsón, Torres, Royo, & Flores, 2012; Kavanaugh et al., 2012). In one hand, governments are using social media platforms as a new source of information to communicate their achievements or to engage discussions with citizens. On the other hand, citizens are using social media platforms to express their opinions and to make governments accountable. In this context, for example, several government entities (at different levels from Municipal to Federal) or citizens' groups have created Facebook (FB) pages such as those presented in (Lev-On & Steinfield, 2015). Hence, the use of social media platforms created a new source of information that government bodies need to be aware of.

These platforms bring different types of data that can be in the form of texts, images, or videos. All these forms should be treated and analyzed by governments with the hope to extract valuable data. In this paper, the focus will be on textual data. So, the challenge for governments is to find the ways to exploit social media platforms and extract valuable data that can be helpful in decision-making processes. Hence, the main research question of this paper is: How data can be brought from social media to government policy-makers?

This paper proposes a generic framework based on a text analysis tool that reads data from a FB page and then extracts valuable data that can be presented to government policy-makers. We'll apply the framework to analyze data collected from a FB page managed by citizens in Tunisia and that contains different posts related to different facts occurring in different cities in Tunisia. These facts report for example on road security, transportation problems, or government entities services delivery quality.

This paper is organized as follows. Section 2 presents the research context including e-participation and government and social networks. In Section 3, we present the natural language processing techniques in Social Media Networks. Section 4 presents the generic framework for semantic data analysis based on Latent Semantic Analysis and describes the case study. Section 5 provides experimental results. Section 6 presents the impacts on Policy Cycle and Modelling Cycle. Finally, Section 7 concludes the paper.

2. Research context

The main objective of this paper is to provide a framework to help policy-makers extract knowledge from data generated by citizens in social media platforms in order to be informed about what citizens are expressing with the hope to enlighten their future decisions. Social media are considered as one of the major means used by Governments to collect data from their citizens (Boudjelida, Mellouli, & Lee, 2016). Collecting data from citizens can be considered as a form of e-participation. So, this research is considered as part of the context of e-participation.

* Corresponding author.
E-mail addresses: olfa.belkahla@isg.rnu.tn (O. Belkahla Driss), sehl.mellouli@fsa.ulaval.ca (S. Mellouli), zeineb.trabelsi.1@ulaval.ca (Z. Trabelsi).

https://doi.org/10.1016/j.giq.2019.05.002
Received 23 June 2018; Received in revised form 12 March 2019; Accepted 13 May 2019
Available online 28 May 2019
0740-624X/

2.1. Citizens participation and governments

Citizens participation (CP) aims to reinforce citizens' involvement in decision-making processes related to significant choices that will affect their communities (Marzouki, Mellouli, & Daniel, 2017). As information and communication technologies (ICTs) have been gradually adopted in CP processes, CP evolved to become e-participation. Macintosh (2008) defines e-Participation as "the use of information and communication technologies to broaden and deepen political participation by enabling citizens to connect with one another and with their elected representatives". Actually, the integration of ICT in participatory processes is not an end to itself, but just a mean to achieve traditional goals of CP. Therefore, e-participation maintains the same goals than traditional CP while bringing new forms of communication, with the aim to increase the involvement of citizens and helping them achieve their communities' objectives. *E*-participation can thus increase the "interactions between policy-makers and citizens" and make it possible and easier for citizens to participate in policy making and policy evaluation which help to generate new options and explore new solutions (Janssen & Helbig., 2016).

There are different platforms that can be used for e-participation. As depicted in (Boudjelida et al., 2016), 41% of e-Participation activities are implemented through social networks, 24% through online forums, 22% through government's' official web sites and 13% through other online technologies. In this paper, the focus will be on e-participation through social networks.

2.2. Social networks

Social networks, such as Twitter, Facebook or Google+, are platforms to which billions of users are connected. They transformed the way people are accessing to information and the way the information is read and spread. According to an online study conducted by Reuters in 2016 (Newman, Fletcher, Levy, & Kleis, 2016) with 50,000 people in 26 countries, social networks are one of the main sources of information for 51% of the respondents, and for 12% of them, these networks are the first source of information.

Social networks are used incrementally to involve citizens by obtaining their opinions, and engaging them in political processes (Koliba, Zia, & Lee, 2011). Emergence of Social media empowered the public to play a more active role in the way their governments operate (Linders, 2012). Social media networks can be seen as the development of a collective intelligence in an online collaborative way to create content (Schoder, Gloor, & Metaxas, 2013). Unlike traditional one-to-many media, social media has transformed this hierarchy to many-to-many where everyone can become a producer of content (Porter, 2008). Chouikh, Ojo, and Belkahla Driss (2016) have generated the affordances of social media platforms to explore opportunities for collaboration between citizens themselves (C2C) and between citizens and government (C2G) in terms of public service delivery on social media platforms. For example, Eom, Hwang, and Kim (2018) examined the roles that can be played by mayors and public officials in social media networks to increase government responsiveness. So, following the growth of social media websites, a large volume of data is generated by users to share and express their opinions on topics, discuss with diverse communities, or publish posts. Social media in governments has a real role in the innovation of public sector organizations (Criado, Sandoval-Almazan, & Gil-Garcia, 2013). In this context, DePaula, Dincelli, and Harrison (2018) developed a typology of government communication on social media: democratic information provision, symbolic and presentational types of information exchanges.

Governments paid a particular attention to these networks that became a new source of information for policy-makers (Bertot, Jaeger, & Hansen, 2012). Citizens can then use these networks to express themselves on different issues related to their day-to-day situations. Citizens can do this on either dedicated spaces that are managed by governments, or use other spaces that are managed by groups of citizens. In this context, policy-makers can no longer ignore these new sources of information since the published information in these networks can spread very fast and can lead to uncomfortable situations to governments. Thus, a policy-maker should investigate the opinions of citizens on these spaces to get an insight into the opinions of the citizens and to react accordingly (Bertot et al., 2012). Hence, government agencies should not only use social media to communicate with citizens, but also adopt new approaches to attract different citizens' groups, and acquire new solutions for data retrieval (Ferro, Loukis, Charalabidis, & Osella, 2013).

With the exponential increase in the number of users and the diversity of sources, the mass of data (in different forms and that can be structured and unstructured) is constantly increasing on these networks. The challenge is then to identify, throughout this mass, the data that may be of interest to policy-makers. This paper tries to provide an answer to the following research question: How data can be brought from social media to government policy-makers?

Data on social media can be of three forms: text, image, or video. This paper will focus only on textual data. Policy-makers not only need to consider a text analysis tool but they need a generic framework with specific guidelines to acquire this data and extract knowledge from it. Meanwhile, in order to extract valuable knowledge from text, natural language processing techniques are required. In the next section, we present different natural language processing techniques that have been applied to extract data from social media.

3. Natural language processing in social media networks

The large volume of text data in social media requires the usage of techniques to extract, classify and analyze the most important information (Chen, Chiang, & Storey, 2012) (Kharde & Sonawane, 2016) (Injadat, Salo, & Bou Nassif, 2016). These techniques are referred to as Topic Modelling (TM) algorithms (Blei, 2012). TM is used to detect subjects from a collection of documents by describing each document as a mixture of latent subjects, and each subject is seen as a word distribution (Blei & Lafferty, 2009).

In what follows, we present different studies (without being exhaustive) related to text analysis of data collected from social networks. We then summarize all these studies in a table that identifies the used techniques for text analysis and the type of social networks that were used (see Table 1). All the studied text analysis tools for social media were applied in four different contexts: local Governments, other Governments levels, citizens, or the private sector.

At the local government level, the work of (Reddick et al., 2017) analyzed textual data posted on the FB page of San-Antonio municipality and that is related to solid waste management. They used the Latent Dirichlet Allocation (LDA) technique (Blei, Ng, & Jordan, 2003) with the objective to identify the topics that were raised by citizens such as for example recycling, inquiries, containers, etc. In the same vane, Ma et al. (2016) analyzed data of citizens' opinions posted on an online platform for a City in China. They also used LDA as a technique for text analysis. The objective is to provide municipal administrators a better understanding of citizens' online comments so that they could improve their decision-making in urban construction and development. In another context, Musto et al. (2015) proposed a framework, based on LDA, for text analysis collected from Twitter. It was applied to analyze posts related to the recovering of the city of L'Aquila in Italy from the earthquake of 2009. Finally, content analysis of Facebook posts in 12 German local governments was carried out in (Hofmann et al., 2013). They used the Alchemy API for sentiment analysis of the posts. They conceptualized a set of success factors of Governments' online