Automatic Intent-based Classification of Citizen-to-Government Tweets

José L. Lavado[1], Iván Cantador[1], María E. Cortés-Cediel[2], and Miriam Fernández[3]

[1] Escuela Politécnica Superior, Universidad Autónoma de Madrid, Spain
jose.lavado@estudiante.uam.es, ivan.cantador@uam.es
[2] Facultad de Ciencias Políticas y Sociología, Universidad Complutense de Madrid, Spain
mcorte04@ucm.es
[3] Knowledge Media Institute, The Open University, United Kingdom
miriam.fernandez@open.ac.uk

Abstract. Social networking technologies offer opportunities for governments to engage with citizens. However, the inability to filter relevant citizens' messages out of the vast amount of available social media content lessens their impact. In this paper, we propose a set of categories encapsulating the different citizens' intents when directing messages to public institutions, e.g., complaining, making requests, and proposing solutions to existing problems. We present a novel artificial intelligence approach, built upon natural language processing and machine learning algorithms, that enables the categorisation of citizens' messages into such intents automatically, and at scale. Through an empirical evaluation on a Twitter dataset, we show the effectiveness of our approach in terms of categorisation performance. We also discuss the value of the presented solution, as a novel tool for governments to achieve a more effective and informed communication with citizens.

Keywords: e-participation · social networks · natural language processing · machine learning · text classification

1 Introduction

Nowadays, the implementation of e-government models in the field of public management is mainly oriented towards the production, custody and management of large-scale data [1], which are produced through technologies such as social media, IoT devices, cloud computing, and blockchain, among others [1, 18]. Data-driven governments are beginning to exploit Big and Open Linked Data through machine learning approaches and other forms of Artificial Intelligence (AI) [8], in order to reinforce policy and decision-making. They thus obtain information from citizens and other stakeholders, promote participation, provide transparency and accountability, and consequently generate public value [18].

Among the existing sources of information, social networks represent a prominent bidirectional communication channel between citizens and government. In them, citizens are not only content consumers who receive the government announcements, to which they react and freely respond according to personal ideology, interests and needs [10], but also are content providers who generate a wide range of messages targeted to government and political stakeholders.

The amount of social media content daily generated by citizens is huge and diverse, and its processing by human actors may result too costly and overwhelming. Hence, there is increasing interest and need to use computer-assisted solutions capable of automatically gathering, processing and analysing the underlying information in the citizens' messages (a.k.a. posts) on social networks.The research literature reports extensive work on mining citizen generated content. The majority of such work has focused on i) analysing social phenomena produced through the online network structures –e.g., information spreading, fake news, and opinion polarity–, and mainly originated by particular events –e.g., natural disasters, elections, and trending news [16]–, and ii) extracting the most popular topics addressed by citizens' posts in social networks, as well as the general dynamics (i.e., temporal evolution) and opinions on such topics [2]. In this paper, we are interested in the latter case. However, differently to previous work, we go beyond the extraction of topics by attempting to automatically classify citizens' posts according to their intents or purposes. That is, we aim to determine whether a post targeting government actors expresses a question, complaint or request, presents a proposal or idea to address a particular problem, spreads an announcement or news item of interest for the general public, or reflects a personal fact or opinion.

We believe this automatic classification can be very valuable for government managers and politicians in several ways. First, it would represent a mechanism to identify relevant citizen posts for which responses should be given. This may help increasing the citizens' satisfaction and engagement, who would perceive attention to their questions and requests. Hence, it may promote the openness of the public administration, and ultimately may increase the citizens' trust on a government that responds to public demands. Second, the proposed classification would allow extracting indicators about opinion on how public resources are being managed. These indicators could be used by government managers to identify problems for which new actions and public policies are needed. This may lead to increase the effectiveness and efficiency on both the management of public resources and the provision of public services, which ultimately would generate public value. Finally, the intent-oriented classification would isolate measures on current leadership perception. Taking these measures into account, political parties and leaders could make timely decisions reacting to major opinions, complaints and proposals on problematic and controversial issues.

With all the above, we propose a categorisation of intents, as well as an AI approach to automatically identify the intent of citizens' posts towards government accounts in social networks. Our approach is built upon natural language processing techniques to extract information features that represent the posts, and machine learning models to automatically classify the posts according to their feature-based representations. Conducting empirical evaluations on a Twitter dataset composed of labelled posts (i.e., tweets) targeted to the City Council of Madrid (@Madrid account in Twitter), we report promising results that show how the large-scale intent-driven categorisation of posts is a valuable method to achieve a more profound understanding of citizen-to-government messages in social networks. We also discuss challenges and opportunities of the presented research and its value, as a novel tool for governments to achieve a more effective and informed communication with their citizens.

2 Related work

Our goal is to categorise the messages that citizens explicitly direct to public institutions in social networks. Hence, we have discarded from our literature review those papers that analyse social media content generated around particular events (e.g., elections and political uprisings), where messages are not necessarily targeted to public institutions.

Among the analysed papers, we have identified two main research lines: (i) works conducting topic (or thematic) analysis of the different messages that citizens direct to their public institutions, and (ii) works attempting to understand the opinions and sentiment behind those messages. Some works use a combination of topic and sentiment analysis.

Works focused on the analysis of topics followed an approach similar to the one proposed by Driss and colleagues [3], where text analysis techniques are automatically used to extract topics. In their paper, they focused on Facebook posts from citizens of Tunisia. They found principal topics of interest, such as road security, transportation problems, and public service delivery quality.

Works that focus on the analysis of sentiment or the analysis of both topics and sentiment, can be divided into (i) those that analyse messages posted by governments and politicians [4, 20, 22], and (ii) those that analyse messages posted by citizens [7, 11, 13, 15, 17]. The first set of works shows how governments and politicians that adopt a positive tone —and undertake activities like responding directly to citizens on Twitter, sharing photos, and using exclamation points— are more likely to encourage citizen participation [22]. Also, they show that videos and images have a high positive impact on engagement, and tweets posted on weekdays obtain higher engagement than those posted on weekends [20]. The works that analyse messages posted by citizens show: (i) how users present high levels of emotionality as well as a high participation rate [11] and (ii) how, within a urban context, citizens' sentiment can be used as an indicator of perceived neighbourhood quality [7], as well as an indicator to estimate urgency of urban issues, such as overflowing trash bins and broken footpaths, among others [13].

Despite the usefulness of social media data analysis, some works [15, 17] argue that the integration of this knowledge in planning and decision-making has not been completely successful, and that a good implementation strategy is necessary to realise their full benefits. In this line, Garg and colleagues [6] proposed an automatic approach to determine which of the posts that citizens direct to institutions are “actionable”, i.e., can be acted upon by the government.

As opposed to those works that focus on the analysis of messages from governments to citizens, we focus on analysing messages that citizens explicitly post towards governments or local institutions. Giving a step forward the analysis of topics [3], sentiment [7, 11], and actionability [6], we advance the state of the art by focusing on the understanding of citizens’ intents when posting messages. We do so by proposing a novel AI-guided approach to analyse social media data automatically and at scale.

3 Classification of tweets based on their intent

Within the machine learning field, text classification (a.k.a. text categorization) refers to the task of automatically assigning a natural language text with one of a given set of *classes* (labels or categories) [19]. The classes are usually discrete values related to topics, but can also represent domain-dependent meanings, such as “spam” and “non-spam” emails, “real” and “fake” news articles, and “positive” and “negative” textual reviews. Besides, a classification problem may be binary –with two classes– or multi-class –with more than two classes.

To address this task, supervised learning assumes that a set of training data (i.e., the *training set*) has been provided, consisting of a set of *instances* (input texts) that have been labelled by hand with their correct class. On weighted feature vector representations of the training instances, a learning procedure aims to extract feature patterns and relations that allow characterizing and distinguishing instances from each class. The procedure then generates a model that attempts to meet two sometimes conflicting objectives: classifying as well as possible on the training data, and generalizing as well as possible to new (test) data [19].

In this context, the selection and extraction of features represents a key stage for the effectiveness of the final classification process. When dealing with text documents, a typical choice is to identify features with words, in the so-called *bag of words* model, and to assign each word with a real number weight equals to its TF-IDF (term frequency, inverse document frequency) value. This model, however, is not appropriate to represent tweets, due to their short length. We thus advocate for alternative features. Specifically, we propose to use both language-independent lexical and grammatical features, and social network-based features. In the next subsections, we present the classes and features considered for the proposed intent-based citizen-to-government tweets.

3.1 Proposed intent-based classes

As argued by Theocharis [21], online social network participation can be a form of political participation that should be conceptualized, identified and measured. From a revision of the literature, he considers several forms of political participation: i) posting (sharing) links to political stories or articles for others to read, ii) posting own thoughts or comments on political or social issues, iii) encouraging to take action on a political or social issue and, iv) reposting content related to political or social issues that was originally posted by someone else. Motivated by such categorization, in this paper, we focus on identifying the intent that citizens have when posting messages to their institutions. In addition, we rely on a data-driven inspection to define categories of intent. Hybrid approaches of this kind have been shown to improve rigour in exploratory studies [5]. The final ten intent-based categories extracted after this process include:

- **Complaint**. The intent is to state something that is unsatisfactory or unacceptable (e.g., "@MADRID after 1 week of calling, the city is yet not clean and the rats are taking over!! http://t.co/IiIDuaPFG9")
- **Announcement**. The intent is to make a public statement about a fact, occurrence or event (e.g., "The date, place and schedule of the Festival activities in La Latina have already been confirmed http://t.co/U0tRwKAC @madrid @madridiario")
- **News item**. The intent is to objectively inform about current events. Authors of these posts are generally media news organisations and journalists (e.g., "#oladecalor #aemet @Madrid has suffered its warmest night within the latest 100 years http://t.co/ZSjeqK6m")
- **Personal fact**. The intent is to publicise self issues and experiences (e.g., "I also support the candidature from @Madrid2020ES @MADRID #aporella")
- **Personal opinion**. The intent is to express subjective opinions about the city, its events, activities, etc. (e.g., "The activity of #emprendeenmadrid is amazing. Congratulations @MADRID and greetings from an entrepreneur")
- **Request**. The intent is to explicitly ask for something specific (e.g., "Very nice but impossible to ride a bike at normal speed #MadridRio. Please @MADRID create a bike lane with cyclist priority")
- **Notification**. The intent is to report or give notice of urban, citizenship- or government-related issues, so that the Madrid City Council can quickly act on them and help other citizens (e.g., "@MADRID can you fix this gap in c/ San Bernardino 8-10 before someone gets hurt? http://lockerz.com/s/117566458")
- **Question**. The intent is to explicitly ask for information (e.g., "@MADRID could you please give me the telephone number of the press office of the Madrid city hall")
- **Proposal**. The intent is to suggest an initiative or project. Proposals indicate broader projects and ideas than the explicit and specific demands of the request category ("There is a collection of used oil in the centre of Alicante. It would be fantastic to have something similar @MADRID")
- **Other**. The intent is unclear or does not fit into any of the previous classes.

3.2 Proposed classification features

To automatically categorise each tweet into one of the classes categories presented in the previous subsection, it is first transformed into a vector of 37 domain- and language-independent features. From them, 27 are content-based features, including:

- **Lexical features** (7): number of characters, number of words, number of exclamation marks, number of question marks, existence of a positive emoticon, existence of a negative emoticon, and existence of a vowel (or 'y') consecutively repeated 3 or more times in a word. The latter is assumed to be a signal of emphasis.
- **Grammatical features** (20): number of nouns, number of proper nouns, number of adjectives, number of verbs, number of adverbs, number of personal/possessive pronouns, number of time references (entities), and number of money-related references.

These content-based features were obtained by a computer program that makes use of the Stanford CoreNLP[4] natural language processing toolkit [12], which, as far of March 2021, allows obtaining the syntactic parsing of sentences in English, Arabic, Chinese, French, German and Spanish. For nouns, adjectives, verbs and adverbs, we also consider the number of them which were positive/negative/neutral, according to a Spanish lexicon of word opinion polarities [14].

The remainder 10 features were social network-based, including:

- **User features** (4): number of followers, number of friends (a.k.a. followees), number of posts, and number of active days in Twitter.
- **Post features** (6): number of hashtags (#), number of user mentions (@), number of hyperlinks, number of multimedia, maximum hashtag length, and existence of an explicit retweet request (i.e., "RT" abbreviation).

We discarded interaction-based features, such as the number of "likes," the number of "comments," and the number of "reposts" (i.e., retweets), since our aim is to automatically categorise tweets after they are generated. Further popularity-based signals could be used in longer term processing/analysis stages. We also discarded fine-grained grammatical features, such as the number and tense of the verbs. For instance, one may expect that first-person verbs would not appear in *news items*, and thus may represent an informative feature to characterise that class. Similarly, imperative verbs may be much frequent in *requests*, whereas conditional verbs may be predominant in *proposals*. We did not consider these features since they depend on the language in which tweets are written. Nonetheless, they could be exploited in a language-specific solution to improve classification accuracy.

4 Experiments

4.1 Dataset

As a case study to test our approach we selected the City Council of Madrid, Spain. Its Twitter account (@Madrid) has more than 700K followers, and receives a high volume of daily posts explicitly directed to it. We aimed to categorise messages posted by citizens and directed to that public institution. To

[4] Stanford CoreNLP toolkit, `https://stanfordnlp.github.io/CoreNLP/`

gather these messages, we first collected data for all the user accounts following @Madrid (670,140 at the time of collection - September 2019). The Twitter API allowed us to collect the most recent 3,200 posts for each of these accounts. We then filter those messages explicitly directed to Madrid city council (a total of 414,117 generated by 55,858 accounts) This nearly half a million messages spam a decade of tweets directed to the city council from 2009 till 2019.

Note that not all Twitter accounts engage with @Madrid with the same frequency. While some accounts have directed 1 or 2 messages towards the city council, others have directed more than 1,000 messages. Also, not all accounts belong to individuals. However, since citizens communicate with their institutions both, privately and via organised groups, neighbourhood/district representatives, and businesses, we decided to consider all account types in this work.

To obtain the necessary training data to build and evaluate our classification approach, we needed to categorise a subset of posts manually. For this purpose, we selected a random sample of 666 tweets. These tweets were manually annotated by four experts (each of them annotated a 500 sample), ensuring that each tweet received at least three annotations. All experts received explicit indications of the categories and their meaning before conducting the annotation process. In addition, an hour of debate was allocated for them to reflect on the categories and resolve possible doubts. The annotation process shows an agreement of Fleiss' κ coefficient equal to 0.98, meaning almost perfect agreement. For conflicting cases, the majority class assigned to a tweet was finally selected.

Table 1 shows statistics about the number of tweets (first row) and the frequency of feature appearance for each class in our dataset.

4.2 Classification algorithms

To validate the proposed method, we evaluated several machine learning algorithms on the generated dataset. The tested algorithms included:

- **K-Nearest Neighbours (KNN)**, which classifies an instance based on the label of its k closest (i.e., most similar) instances –called as neighbours.
- **Logistic Regression (LR)**, which estimates the probability of a certain class through a logistic function in its basic binary form, and through the one-vs-rest (class) scheme in the multi-class case.
- **Quadratic Discriminant Analysis (QDA)**, which learns a quadratic decision boundary, generated by fitting class conditional densities to the data and using Bayes' rule.
- **Decision Tree (DT)**, which builds a flowchart-like structure where each internal node represents a test on a feature, each branch represents the outcome of the test, and each leaf node represents a class label, so that the path from the root node to a leaf node represents the classification rules for the leaf. The algorithm was executed alone, and in combination with feature selection (RFECV DT) and tree pruning (AP DT) to avoid learning over-fitting.

Table 1. Feature stats for the proposed intents: complaints (COM), announcements (ANN), news items (NEWS), personal facts (PER), personal opinions (OPI), requests (REQ), notifications (NOT), questions (QUE), proposals (PRO), and others (OTH).

	COM	ANN	NEW	FAC	OPI	REQ	NOT	QUE	PRO	OTH	Total
Num. tweets	187	175	90	71	57	35	13	12	10	16	666
User mentions	331	339	169	161	129	76	18	21	22	37	1303
Hashtags	105	120	37	27	31	33	6	2	5	10	376
Links	67	145	87	26	13	22	5		3	7	375
Time refs.	48	122	34	13	5		3	3			228
Money refs.	29	5	3	3	1	1			1	1	44
Retweet refs.	2	6	5	3	1	19	1				37
Nouns	632	366	216	129	149	115	38	28	31	31	1735
Proper nouns	227	255	126	65	52	31	10	6	9	16	797
Pronouns	49	48	5	33	24	11	3	2	1	6	182
Adjectives	173	114	56	44	50	24	10	6	15	8	500
Verbs	327	126	84	65	80	58	23	10	16	16	805
Adverbs	148	66	16	26	38	16	10	7	9	7	343
Questions	39	10	1	4	1	10	4	9	3	3	84
Exclamations	25	32	2	12	10	8	4	3		1	97
Pos. emoticons		1	1	3	1		2	3			11
Neg. emoticons	2										2

- **Gaussian Process (GP)**, based on Laplace approximation, is a generalization of the Gaussian probability distribution as the basis for a sophisticated non-parametric classification.
- **Support Vector Machine (SVM)**, which aims to find a hyperplane that best splits a dataset into two classes. For the multi-class problem, it is handled by a one-vs-one scheme.
- **Bagging Ensemble (BE)**, which combines a set of randomized decision trees.

We used the algorithms implementations provided by the open-source Scikit-learn[5] machine learning toolkit. We conducted a grid search to find the best parameters for the algorithms. We provide our code and the algorithms configurations and tested parameters online together with our dataset[1].

4.3 Classification results

Due to the unbalanced distribution of the instances in the 10 classes (see the number of tweets per class in Table 1), we conducted a series of experiments where we addressed 10 binary (2-class) classification problems. Each of them aimed to distinguish the instances belonging to a particular class from the instances belonging to the other classes.

[5] Scikit-learn machine learning toolkit, `https://scikit-learn.org/`
[1] Public URL with our dataset and code, <hidden for blind review>

In addition to computing the accuracy (acc) metric, which measures the percentage of instances (i.e., tweets) correctly classified, we also computed the acc^+ and acc^- metrics, which correspond to the percentage of correctly classified instances in the minority and majority classes, respectively. As a compromise of both metrics, we considered their geometric mean $g = \sqrt{acc^+ \cdot acc^-}$. We computed average metric values from 3 independent executions of each algorithm and parameters configuration, keeping 75% of the tweets for training the machine learning models, and 25% for testing, selected randomly in each execution.

Table 2 shows the best accuracy results achieved by the evaluated algorithms on each intent-oriented classification problem. Note that the classification problems present a large unbalance between the target minority class and the majority class, ranging from $N^+ = 28\%$ of positive instances for the *complaint* class to $N^+ = 2\%$ for the *notification*, *question*, and *proposal* classes. This makes the classification problems challenging. Despite this difficulty, by exploiting the proposed domain- and language-independent features and using generic machine learning algorithms, we were able to achieve relatively high accuracy (acc^+) on identifying complaints, announcements, news items, personal opinions, and requests. The achieved classification performance is relatively high, as can be seen by comparing the acc^+ and g values against the percentage of positive instances N^+ in each class. Note that acc values alone are not informative enough, since for each intent, classifying every instance as *negative*, we would achieve an accuracy equals to N^-, but we would be wrongly classifying all *positive* instances. Regarding the effectiveness of the evaluated classifiers, Quadratic Discriminant Analysis (QDA) and Decision Trees (DT) showed the best performances, with respect to algorithms such as Support Vector Machines (SVM) and Gaussian Processes (GP), which are generally adequate for high dimensional data.

Table 2. Best accuracy (acc, acc^+, acc^-) values for each intent-oriented classification task. Geometric values g show the achieved accuracy balanced between minority (N^+) and majority (N^-) classes.

Intent	N^+	N^-	acc	acc^+	acc^-	g	**Algorithm**
Complaint	28%	72%	74.6%	66.0%	78.0%	72.0%	QDA/LR
Announcement	26%	74%	83.8%	75.0%	87.0%	81.0%	AP DT
News item	14%	86%	64.4%	61.0%	65.0%	63.0%	QDA
Personal fact	11%	89%	73.4%	44.0%	77.0%	59.0%	QDA
Personal opinion	8%	92%	83.6%	57.0%	86.0%	70.0%	QDA
Request	5%	95%	94.8%	56.0%	97.0%	74.0%	KNN
Notification	2%	98%	97.8%	33.0%	99.0%	57.0%	AP DT
Question	2%	98%	96.8%	33.0%	98.0%	57.0%	RFECV DT
Proposal	2%	98%	91.9%	33.0%	93.0%	56.0%	SVM/LR

4.4 Discriminating words and features

Table 3 shows some of the most discriminating words and features for the proposed classes. These words and features would allow governments to obtain insights about relevant topics and sentiments expressed in citizens' complaints, requests, proposals, notifications and questions, as well as to differentiate them from those commonly appearing in news items, announcements, and personal opinions and facts. Note that these informative words and features are identified automatically and exploited by the classification algorithms, regardless the language and topic of the tweets.

As illustrative examples, we first focus on user accounts mentioned in the tweets. Complaints are abundant of mentions to accounts of political parties (e.g., *Partido Popular*, *UPyD*, and *Izquierda Unida*), politicians (e.g., *Mariano Rajoy*) and accounts created to denunciate particular issues, such as political corruption (@Cobri2020) and salaries of public officials (@sueldospublicos). Announces, by contrast, have many mentions to accounts about tourism (@turismomadrid), leisure (@ociomadrid), and sports (@deportegob) in Madrid. Lastly, many requests mention @lineamadrid, the official account of the citizen attention service of the City Council. We also put our attention on proper nouns. *Ayuntamiento de Madrid* (City Council), *Ana Botella* (former mayor of the city), and *Gürtel* (a political corruption scandal) appear in complaint tweets, whereas locations are abundant in announcements, e.g., *Vallecas* (a district), *Palacio Real* (the Royal Palace) and *Caja Mágica* (a multi-purpose stadium), and in news items, e.g., Community of Madrid.

Due to the lack of space, we do not analyse other features and words reported in the table. However, we provide them in English for the reader's comprehension.

5 Conclusions

As citizens are spending more time on online social networks, generating large amounts of content, there is a need for innovative methods and tools to analyse such data. In this paper, we have presented and evaluated a novel AI approach that applies natural language processing and machine learning algorithms to automatically classify citizen-to-government posts published in social networks. Differently to previous works, which have focused on topic- and opinion-based analysis, our approach aims to classify posts based on their underlying intention or purpose, distinguishing between citizens' complaints, requests, proposals and announcements, among others. This classification represents a processing stage prior to the extraction of topics and opinions, and may help filtering and prioritising citizens' messages, and further automatising processes for more efficient and effective decision and policy making.

Despite the positive classification results achieved by our approach, there is still room for improvement. For example, more sophisticated Natural Language Processing techniques, such as language models and word embeddings [9], could be used to exploit the semantics of words and word sequences, e.g., "opinion is"

Table 3. Most discriminating words and features for each of the considered intents.

User mentions	COM	@PPopular @UPyD @iunida @marianorajoy @Cobri2020 @sueldospublicos
	ANN	@turismomadrid @ociopormadrid @deportegob
	REQ	@lineamadrid @Mango
Hashtags & Nouns	COM	mayor, euro, money, crisis, law, taxes, workers, problems, embarrassment, management, waste
	ANN	activities, schedule, workshop, visit, summer
	NEW	transport, weather, unemployment
	FAC	thanks, children
	REQ	please, signature, support, peace
Proper nouns	COM	Ayuntamiento de Madrid, Ana Botella, Gürtel
	ANN	Vallecas, Palacio Real, Caja Mágica
	NEW	Comunidad of Madrid
Adjectives	COM	intolerable, bad, worse
	ANN	free
	OPI	better
Verbs	ANN	visit, miss, offer, meet, enjoy
	FAC	wait, see
	REQ	think, leave
Adverbs	COM	now, always, never, more, less
	ANN	free, evening
	FAC	already
Time refs.	ANN	tonight, Wednesday, Thursday, Saturday
	NEW	yesterday
	FAC	today

and "really think that" could be identified as informative bigram and trigram of the *personal opinion* class. Furthermore, it could be possible to extend our approach with features from other sources of information, such as the user who creates a post and the users who are mentioned in a post (e.g., by considering their type: particular citizens, neighborhood associations, organisations, or political actors), and the nature of web resources linked in the posts (e.g., articles of online news media, personal blogs, or multimedia in social networks).

From a social inclusion perspective, and taking fairness concerns into account, we plan to investigate possible biases derived from the subset of the population posting these messages, as well as the possible biases that the classification algorithms may have depending on issues such as the users' posting activity and influence (i.e., number of followers), and ideological, political and popularity-based factors of the addressed topics.

While there is room for future work, this paper presents a unique approach to automatically, and at large scale, identify the intent from citizens' posts directed to their institutions. We hope that this novel idea, in combination with all the resources provided (code, dataset, annotations), can help researchers and practitioners to filter messages of interests among the sea of social media posts, facilitating a more effective communication among citizens and institutions.

Acknowledgements This work was conducted with financial support from the Spanish Ministry of Science and Innovation (PID2019-108965GB-I00) and the Centre of Andalusian Studies (PR137/19). José L. Lavado is partially supported by the UAM–ADIC Chair for Data Science and Machine Learning.

References

1. Charalabidis, Y., Loukis, E., Alexopoulos, C., Lachana, Z.: The three generations of electronic government: From service provision to open data and to policy analytics. In: Proceedings of the 18th IFIP WG 8.5 International Conference on Electronic Government. pp. 3–17. Springer (2019)
2. Chatfield, A.T., Scholl, H.J., Brajawidagda, U.: # sandy tweets: citizens' co-production of time-critical information during an unfolding catastrophe. In: Proceedings of the 47th Hawaii International Conference on System Sciences. pp. 1947–1957. IEEE (2014)
3. Driss, O.B., Mellouli, S., Trabelsi, Z.: From citizens to government policy-makers: Social media data analysis. Government Information Quarterly **36**(3), 560–570 (2019)
4. Espinoza, Á.O., Trujillo, A.E.: Sentiment analysis in presidential campaigns: Mexico 2018. Revista Española de Investigaciones Sociológicas **173**, 79–100 (2021)
5. Fereday, J., Muir-Cochrane, E.: Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. International Journal of Qualitative Methods **5**(1), 80–92 (2006)
6. Garg, H., Bansal, C., Kaushal, R., Thanaya, I.: Identifying actionable information from social media for better government-public relationship. In: Proceedings of the 10th International Conference on Developments in eSystems Engineering. pp. 206–211. IEEE (2017)
7. Guhathakurta, S., Zhang, G., Chen, G., Burnette, C., Sepkowitz, I.: Mining social media to measure neighborhood quality in the city of Atlanta. International Journal of E-Planning Research **8**(1), 1–18 (2019)
8. Janssen, M., Brous, P., Estevez, E., Barbosa, L.S., Janowski, T.: Data governance: Organizing data for trustworthy artificial intelligence. Government Information Quarterly **37**(3), 101493 (2020)
9. Jurafsky, D., Martin, J.H.: Speech and Language Processing. Prentice Hall (2009)
10. Kalampokis, E., Tambouris, E., Tarabanis, K.: Understanding the predictive power of social media. Internet Research (2013)
11. Lorenzo-Dus, N., Di Cristofaro, M.: # living/minimum wage: Influential citizen talk in Twitter. Discourse, Context & Media **13**, 40–50 (2016)
12. Manning, C.D., Surdeanu, M., Bauer, J., Finkel, J.R., Bethard, S., McClosky, D.: The Stanford CoreNLP natural language processing toolkit. In: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics. pp. 55–60. ACL (2014)
13. Masdeval, C., Veloso, A.: Mining citizen emotions to estimate the urgency of urban issues. Information Systems **54**, 147–155 (2015)
14. Molina-González, M.D., Martínez-Cámara, E., Martín-Valdivia, M.T., Perea-Ortega, J.M.: Semantic orientation for polarity classification in Spanish reviews. Expert Systems with Applications **40**(18), 7250–7257 (2013)
15. Nummi, P.: Social media data analysis in urban e-planning. In: Smart Cities and Smart Spaces: Concepts, Methodologies, Tools, and Applications, pp. 636–651. IGI Global (2019)

16. Panagiotopoulos, P., Barnett, J., Bigdeli, A.Z., Sams, S.: Social media in emergency management: Twitter as a tool for communicating risks to the public. Technological Forecasting and Social Change **111**, 86–96 (2016)
17. Picazo-Vela, S., Gutiérrez-Martínez, I., Luna-Reyes, L.F.: Understanding risks, benefits, and strategic alternatives of social media applications in the public sector. Government Information Quarterly **29**(4), 504–511 (2012)
18. Scholl, H.J., Rodríguez Bolívar, M.P.: Regulation as both enabler of technology use and global competitive tool: The Gibraltar case. Government Information Quarterly **36**(3), 601–613 (2019)
19. Sebastiani, F.: Machine learning in automated text categorization. ACM Computing Surveys **34**(1), 1–47 (2002)
20. Siyam, N., Alqaryouti, O., Abdallah, S.: Mining government tweets to identify and predict citizens engagement. Technology in Society **60**, 101211 (2020)
21. Theocharis, Y.: The conceptualization of digitally networked participation. Social Media+ Society **1**(2), 2056305115610140 (2015)
22. Zavattaro, S.M., French, P.E., Mohanty, S.D.: A sentiment analysis of US local government tweets: The connection between tone and citizen involvement. Government Information Quarterly **32**(3), 333–341 (2015)