



HAL
open science

Categorizing Air Quality Information Flow on Twitter Using Deep Learning Tools

Brigitte Juanals, Jean-Luc Minel

► **To cite this version:**

Brigitte Juanals, Jean-Luc Minel. Categorizing Air Quality Information Flow on Twitter Using Deep Learning Tools. Lecture Notes in Artificial Intelligence, 2018, Computational Collective Intelligence, 11055, pp.109-118. 10.1007/978-3-319-98443-8_11 . halshs-01800457

HAL Id: halshs-01800457

<https://shs.hal.science/halshs-01800457v1>

Submitted on 5 Sep 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Categorizing Air Quality Information Flow on Twitter Using Deep Learning Tools

B. Juanals^{1,2}, J.L. Minel³

¹CentreNorbert Elias, Aix Marseille University- CNRS, -EHESP, Marseille, France

²UMI IGlobes, CNRS - University of Arizona, USA

³MoDyCo, CNRS - Université Paris Nanterre, France

Abstract. Environmental health is an emerging and hotly debated topic that covers several fields of study such as pollution in urban or rural environments and the consequences of these changes on health populations. In this field of intersectorial forces, the complexity of stakeholders' logics is realized in the production, use and communication of data and information on air quality. The Twitter platform is a “partial public space” that can throw light on the different types of stakeholders involved, the information and issues discussed and the dynamics of articulation between these different aspects. A methodology aiming at describing and representing, on the one hand, the modes of circulation and distribution of message flows on this social media and, on the other hand, the content exchanged between stakeholders, is presented. To achieve this, we developed a classifier based on Deep Learning approaches in order to categorize messages from scratch. The conceptual and instrumented methodology presented is part of a broader interdisciplinary methodology, based on quantitative and qualitative methods, for the study of communication in environmental health.

Keywords: Environmental health; Air quality; Instrumented methodology; Circulation of information; Mediation; Social network; Twitter

1 Introduction

Environmental health is an emerging and hotly debated topic that covers several fields of study such as pollution in urban or rural environments and the consequences of these changes on health populations. Its progressive conceptualization reveals divergent positions that may include “an extremely wide range of environmental factors related to the notion of human well-being”. The environmental factors analyzed in the work of the French School of Advanced Studies in Public Health (EHESP) fall into four broad thematic dimensions relating to polluted sites and soils, water quality and air quality and habitat; Among these factors, we focused on air quality, which was the subject of many alerts in major cities at the end of 2016 and which is becoming a national concern with the regular peaks of fine particles matter in urban areas.

From a societal perspective, air is one of the "commons" [15] that are described as natural or material resources that can be extended to information and knowledge commons [16]. Pooling and sharing these resources raises issues about their modes of governance. Air is a common good that is subject to conflicts of interest. As a result, data and information on air quality become political issues that provoke debate, even clashes, concerning the modes of production and dissemination of these data. In particular, the modes of production and interpretation of these data call for the specification of technical criteria in the distribution and control domains. The specification of measurements is a political, industrial and environmental issue. It is linked to the scientific validity of the selection of pollutants and pollution thresholds that are used. These open data make sense only in the light of standards used to interpret them and to take decisions at European, national and regional levels.

The production and dissemination of these data are regulated. They are part of European regulations and national or regional public policies, interdisciplinary cross-knowledge, militant positions as well as industrial and commercial logics. In France, environmental health (with which air quality is associated) is part of the State prerogatives; it is therefore governed by public health policies.

The European Directive on Ambient Air Quality - Clean Air for Europe - was adopted on 14 April 2008; it merged the main Directive (96/62 / EC), the first three specific directives (99/30 / EC, 2000/69 / EC, 2002/3 / EC) and the Council of Europe decision concerning data exchange (97/101 / EC). These legal constraints were gradually applied in France by a 1991 decree then by laws: the so-called Barnier law (1995) and the law on air and the rational use of energy (LAURE Act) of December 30, 1996.

In France, the Ministry of Environment and Sustainable Development no longer centralizes actions in the field of education, awareness raising, production and dissemination of information. Communication initiatives are now carried out at the regional scale defined by the administrative organization in the context of decentralization. In this perspective, the LAURE act stipulates that the State entrusts the monitoring of air quality to approved Non-governmental organizations (NGO) in charge of this surveillance. The result is the distribution of some 40 NGOs on the national territory. These organizations are "monitoring networks" characterized by a hybrid status. They bring together companies, public organizations and NGOs[1]. These NGOs are now linked by their status and their missions to the public authorities while maintaining their vocation of integrating the demands of civil society in their initiatives. These developments led the Air Quality Monitoring Associations (AASQA) to build a national federation, ATMO France, in order to become a lobby able to negotiate with national and local authorities on the subject of air quality.

In the US, the Clean Air Act (CAA) is the comprehensive federal law that regulates air emissions from stationary and mobile sources. "Among other things, this law authorizes the Environmental Protection Agency (EPA) to establish National Ambient Air Quality Standards (NAAQS) to protect public health and

public welfare and to regulate emissions of hazardous air pollutants.”¹ In June 2014, EPA proposed The Clean Power Plan, an Obama administration policy aimed at combating anthropogenic climate change, but in March 28, 2017 Donald Trump, signed an executive order mandating the EPA to review the plan and following his announcement on June 1, 2017, United States withdrawal from the Paris Agreement.

In this legal and regulatory framework, the setting up of a public debate on the implementation of a policy to monitor air quality opens up political questions about the production and the circulation of information on air quality in the public space. In this field of intersectorial forces, the complexity of stakeholders’ logics is realized in the production, use and communication of data and information on air quality. The Twitter platform is a “partial public space” that can throw light on the different types of stakeholders involved, the information and issues discussed and the dynamics of articulation between these different aspects.

The purpose of our work is to conceive, by relying jointly on methods anchored in social sciences and digital humanities, a representation of the modes of circulation and distribution of message flows on Twitter about air quality in relation with stakeholders and the content exchanged.

The outline of this paper is the following. First, in Section 2, we present the literature review and then specify the contribution of the proposed methodology to tackle this field. In Section 3, we will present the specificities and the contribution of our approach, our methodology to collect and analyze flows of tweets and the results from classification analysis. Finally, we conclude in Section 4.

2 Literature Review and Methodology

In this section, we begin with a review of the literature, and then we present the interdisciplinary and instrumented methodology we developed to conduct the analysis of messages.

2.1 Literature Review

As mentioned in [6], tweet analysis has led to a large number of studies in many domains such as ideology prediction in Information Sciences [5], spam detection [18], dialog analysis in Linguistics [2], and natural disaster anticipation in Emergency [17], while work in Social Sciences and Digital Humanities has developed tweet classifications. However, few studies aim at classifying tweets according to communication classes. They mostly rely on small reference sets analyzed by experts in Information Communication (InfoCom) rather than by Twitter users. Two exceptions worth mentioning are first the work presented in Lovejoy and Saxton [12] in which the authors (Twitter users) analyze the global behavior of nonprofit organizations on Twitter based on three communication classes: Information, Community and Action classes. The second work [11] compares the

¹ <https://www.epa.gov/>

use of the Chinese Weibo service during a 2013 smog emergency with data on Twitter concerning a weather event in North America. The data collected were hand-coded and the authors discussed the implications of their results for agencies designing social media campaigns to inform and motivate the public affected by those events, which is one of the goals of the agency we are working with.

Recently, several studies on tweet classification have been carried out in NLP. Basically, these analyses aim at categorizing open-domain tweets using a reasonable amount of manually classified data and either small sets of specific classes (e.g. positive versus negative classes in sentiment analysis) or larger sets of generic classes (e.g. News, Events and Memes classes in topic filtering). The advantage of NLP approaches is that they can automatically classify large corpora of tweets. Until recently, the most commonly used models were supervised learning, Support Vector Machine (SVM), Random Forest, Gradient Boosting Machine and Naive Bayes (NB). In supervised machine learning, features are extracted from tweets and metadata and then vectorized as training examples to build models.

2.2 An Instrumented Methodology using Convolutional Neural Networks

The main drawback of shallow supervised machine learning approaches presented above, is that they require a very time consuming step to identify linguistic or semiotic features and raise issues about the relevance of these features. Recently, new approaches based on Deep Learning techniques and especially on convolutional neural networks (convnets), which no longer require researchers to identify features, have been proposed. A second advantage of convnets is that they obtain better results in terms of accuracy than shallow machine learning systems [8,13]. It is for these reasons that we developed a classifier based on convnets. The proposed methodology, based on convnets, aims to describe and analyze the informational and communicational dynamics at work on the Twitter platform. It explores the circulation patterns of message flows and exchanges, apprehended as a dynamic process, as well as the relationships which are established between different stakeholders. Our aim is to question the forms of engagement, participation and relationships between organizations and audiences by analyzing the flow of their messages. We consider that the field of environmental health and the socio-technical device Twitter contribute to configuring the relations and the interactions between the participants. In this perspective, hashtags on Twitter may be seen as meeting points of different categories of stakeholders interested in the same themes while being anchored in different spheres - the environment and health, politics, the media, industry, the economy, etc..

The architecture of our classifier is composed of several layers as shown on figure 1. The first layer is a pre-trained word embedding as proposed by [14] with a kernel that matches the 5 words used as neighbors. A word-embedding is a distributed representation where each word is mapped to a fixed-sized vector of continuous values. The benefit of this approach is that different words with a similar meaning will have a similar representation. A fixed-vector size of 100 was

chosen. The following layers, a Conv1D with 200 filters and a MaxPooling1D are based on the works reported in [8,4]. The back-end of the model is a standard Multilayer Perceptron layer to interpret the convets features. The output layer uses a softmax activation function to output a probability for each of the three classes affected at the tweet processed (see Section 3.3.1). Finally, only the class with the highest probability is kept.

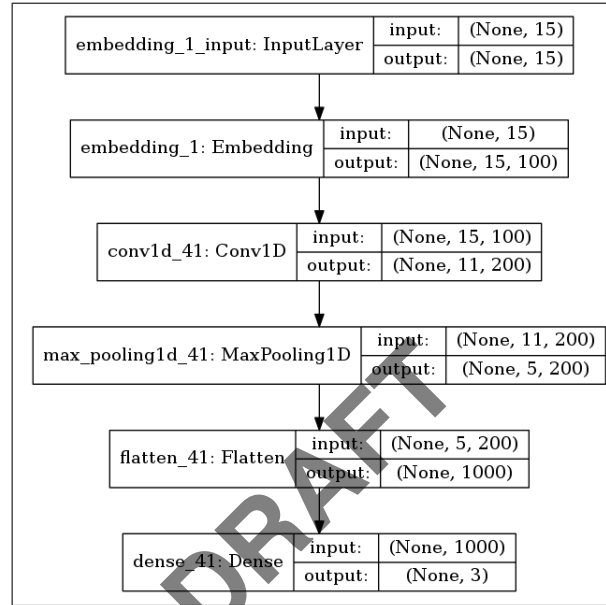


Fig. 1. Architecture of the convets classifier

We would like to point out several important features : first, although this classifier is not a very deep convets, the results are satisfactory (see section 3.2.1) and moreover it does not require significant computing resources; second, the different parameters of the architecture were determined after several trials and errors; and finally, building such a network does not require advanced programming, so it is within the reach of an interdisciplinary team.

We implemented our methodology by building a workflow based on the open access tools Gephi² (graph visualization), Neo4j³ (graph mining), Scikit-learn⁴ and Keras library (deep learning)⁵. We also developed some scripts python to manage the interoperability between all these different tools.

² <https://gephi.org/>

³ <https://neo4j.com/>

⁴ <http://scikit-learn.org/stable/>

⁵ <https://keras.io/>

3 Analyzing Tweets about Air Quality

Data Acquisition The data acquisition stage consisted in harvesting tweets with the following hashtags :the hashtag #Air and one other hashtag among the following list :#pollution, #santé (health), #qualité(quality) or #environnement(environment). Twitter maintains an Application Programming Interface (API) that returns approximately 200 features about a tweet. We developed a Python script , based on the Twarc module proposed by Ed Summers (<http://github.com/docnow/twarc>) using the search option of Twitter API.

Our analysis focuses on Twitter messages (called tweets) sent by accounts of organizations and non-institutional stakeholders. A first step was to build a terminology to describe the objects studied according to three dimensions: the message, the stakeholders, and the forms of stakeholder participation.

Concerning messages, we will call a message sent by a twitter account an 'original tweet' and an original message sent by an account different from the issuing account a 'retweet'. The current Twitter API gives access to the original tweet (and its sending account) of a retweet. The generic term tweet includes 'original tweet' and 'retweet'.

Regarding stakeholder qualification, we distinguished Twitter accounts, accounts managed by institutions (called 'organizational account'), and accounts managed by individuals (called 'private account'). This distinction was carried out by human analysts based on the description field filled out by the account holder. When this field was unfilled, the private account category was assigned.

In this paper, we limit the analysis to French tweets, by using the "lang" features in Tweeter API, sent between the first of November 2017 and the 17th of March 2018. This period of time is considered as a proof of concept and we intent to use the classifier to process all the tweets that will be sent during the year 2018.

The main figures are the following: 3027 tweets of which 30% of original tweets and 70% of retweets sent by 902 participants (405 organizational accounts, 497 private accounts). More specifically : 41% of organizational accounts and only 25% of private accounts produced original tweets. Participation for private accounts was largely limited to the action of retweeting (75% of tweets) the messages sent by the institutional partners. A peak during the week 46 in 2017 is the consequence of the COP22 event (World Climate Conference) which was the major event of the observed period of time.

3.1 Analyzing the modes of involvement and interaction between organizational and private accounts

Categorical analysis relying on automatic classification is a relevant processing method to characterize the semantics of the messages as it makes possible to analyze the modes of engagement of the stakeholders on Twitter. Automatic classification implies a prior human classification (supervised machine learning). Taking into account the size of the corpus of tweets, a human analysis would still have been possible but first, as mentioned in [10], the inter-coder minimum

reliability is usually around 0.74, and secondly, we intend to process in real time all the tweets that will be sent during the year 2018.

These two reasons argue in favour of developing an automatic classification.

In order to build a classifier, a classification analysis of the contents of a sample of 350 randomly chosen original tweets was carried out in two stages. First, a team composed of one linguist and one researcher in communication analyzed the sample of tweets to determine the classes in which to categorize the tweets. Based from three classes were identified (“informative”, “promoting”, “humorous” (in the sense of emotional or expressive)).

“Informative” tweets were those providing information concerning technical aspects of air quality or specific mitigation efforts.

If an advertisement, generally associated with a link on a Web site was present, the tweet was categorized as “promoting”.

“Humorous” tweets contained the expression of emotions (such as worry, anger, dread) or political comments. It must be noted that there were no tweets containing insults. Consequently, 350 tweets were annotated by hand by the same two experts according to the categories defined in the previous step.

In the second stage, a classifier, based on convnets (see Section 2.3), was trained on the annotated samples. It must be pointed out that none modification of the raw text of tweets was carried out, that is to say, we did not use stop words list or stemming transformation. A simple tokenizer from the Keras library was used to compute embedded vectors associated with tokens. The training loss decreases with every epoch, and the training accuracy increases with every epoch; to prevent overfitting, the training was stopped after ten epochs.

Results of the evaluation using the standard cross-validation 10-fold test [9] gave an accuracy of 0.97 in line with the state of the arts [8].

The classifier was applied to the corpus of all tweets to categorize them. Figure 2 show the findings. There is a main difference between institutional and individuals accounts concerning the “humorous” class. Private accounts sent twice as many “Humorous” tweets as institutional accounts. Similarly, institutional accounts sent much more “Promoting” tweets.

3.2 Analyzing Networks of Accounts

Observing the flow of information through the circulation of messages involves looking at the modes of stakeholder participation. They are materialized in information-communication practices. To answer this question, the analysis is based on the classification of the accounts.

In order to characterize Twitter accounts we used some of attributes proposed by [7]. The attribute “relayed” is assigned to an account if at least one of its tweets was retweeted or quoted. The “relaying” attribute is assigned to an account if the account retweeted or quoted at least one tweet. The “mentioned” attribute is assigned to an account if its Twitter account name was mentioned at least once in a tweet. And finally, the attribute “passing” is assigned to an account if it was both “relayed” and is “relaying”; this attribute is computed as the product of the “relayed” and “relaying score”.

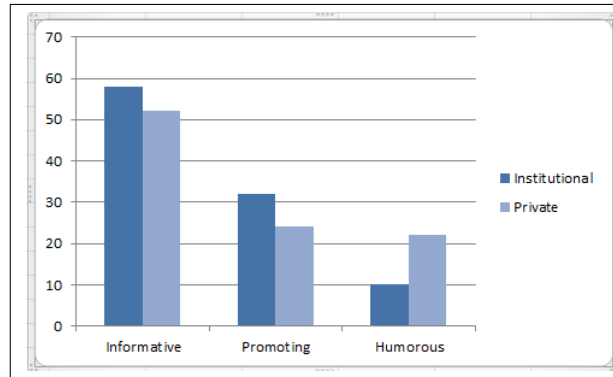


Fig. 2. Comparing tweets classification

As pointed out by [7], the value of this index is not significant in itself; it simply provides a means of comparing accounts. There are several ways to compute the importance of a node in a network [3]. Network theory proposes the notion of degree and betweenness centrality. The degree is the sum of the indegree and the outdegree. But from our point of view, an influencer must be both. Vertices with high betweenness have control over information passing between others vertices. A vertex can have quite a low degree, be connected to others that have a low degree and still have high betweenness. Vertices in roles like this are sometimes referred to as brokers because they play the role of a bridge. But from our point of view, they are not necessarily influencers. For all these reasons, we preferred to use the passing score proposed by [7], which is the product of the indegree and the outdegree. The hypothesis is that an account with a high passing score is a key influential user who actively participates in the circulation of information. We identified six top accounts and a remarkable point is the non-correlation between their passing score and their number of followers. The Kendall-correlation computed on these six passing accounts is 0.04761, i.e. is close to zero.

The list of the top six passing accounts illustrates our description of the stakeholders (see section 1) who have a hybrid status, both public and private. The analysis of these passing accounts makes it possible to identify some of their characteristics. These are all accounts of organizations with the exception of one influencer.

@Ambassad Air is the official account of an operation run by the city of Rennes, located in the West of France, and the House of Consumption and the Environment (in connection with the other associative actors and citizens) mobilizing inhabitants about the air quality in Rennes. @Anne Hidalgo is the official account of the mayor of the city of Paris who is also the chairman of @C40Cities and @AIMFrancophones. Since her election in 2014, she has promoted a very strong policy about air quality. The number of her followers (1 375 410), twenty times higher than the second passing account (in followers number), attests her

notoriety. @ARS Paca is the official account of the Regional Health Agency of the South of France. It is an autonomous public institution at the administrative and financial level. It ensures coordination between the services of the State and the Health insurance by grouping several structures dealing with public health and the organization of care'. @INERIS is the official account of the French National Institute of Industrial Environment and Risks. @RennesVilleMetropole is the official account of the city of Rennes, located in the West of France. It must be pointed out that the city of Rennes is also involved in the @Amabassad Air account. @ATMOFRANCE is the official account of the network of associations authorized to monitor air quality. French law has entrusted this network (see Section 2.1) with the responsibility for implementing air quality monitoring and public information. @Charlotte Marchandis is the private account ("My tweets are only my commitment") of an elected member of Rennes City Council, chairman of @VillesSanteOMS who ran for the French presidential election in 2017.

It is remarkable that the number of followers of these passing accounts extends over a scale of 1 to 3300 and that these key influential accounts do not share their communities of accounts.

4 Conclusions

The construction of an instrumented methodology for the analysis of the flow of messages about air quality on the Twitter platform reveals the complexity of this object of study, which led us to choose a multi-dimensional approach to be able to apprehend it. This research is part of a broader framework of ongoing work on the evolution of environmental communication in the public space.

We developed a classifier based on convolutional neural networks to categorize the flow of tweets. This classifier will be running throughout 2018 year in order to feed a National agency in charge of informing inhabitants and changing their behavior.

5 Acknowledgments

This study is partially funded by iGlobes (UMI 3157).

References

1. Boutaric, F.: Les réseaux de la qualité de l'air : des associations stratégiques de l'action publique. Développement durable et territoires pp. 1–14 (2007)
2. Boyd, D., Golder, S., Lotan, G.: Tweet, tweet, retweet conversational aspects of retweeting on twitter. In: 43rd Hawaii International Conference on System Sciences, HICSS. pp. 1–10 (2010)
3. Brandes, U., Erlebach, T.: Network Analysis. Springer Verlag (2005)
4. Brownlee, J.: Deep Learning for Natural Language Processing. Machine Learning Mystery, Vermont, Australia (2017)

5. Djemili, S., Longhi, J., Marinica, C., Kotzinos, D., Sarfat, G.E.: What does twitter have to say about ideology? In: NLP 4 CMC: Natural Language Processing for Computer-Mediated Communication. pp. 16–25 (2014)
6. Foucault, N., Courtin, A.: Automatic classification of tweets for analyzing communication behavior of museums. In: LREC 2016. pp. 3006–3013 (2016)
7. Juanals, B., Minel, J.L.: Information flow on digital social networks during a cultural event: Methodology and analysis of the "european night of museums 2016" on twitter. In: Proceedings of the 8th International Conference on Social Media & Society. pp. 13:1–13:10. #SMSociety17, ACM, New York, NY, USA (2017), <http://doi.acm.org/10.1145/3097286.3097299>
8. Kim, Y.: Convolutional neural networks for sentence classification. CoRR abs/1408.5882 (2014), <http://arxiv.org/abs/1408.5882>
9. Kohavi, R.: A study of cross-validation and bootstrap for accuracy estimation and model selection. pp. 1137–1143. Morgan Kaufmann (1995)
10. Lachlan, K., Spence, P., Lin, X., M., D.G.: Screaming into the wind: examining the volume and content of tweets associated with hurricane sandy. *Communication Studies* 65(5), 500–518 (2014)
11. Lin, X., Lachlan, K., Spence, R.: Exploring extreme events on social media: A comparison of user reposting retweeting behaviors on twitter and weibo. *Computers in Human Behavior* 65, 576–581 (2016)
12. Lovejoy, K., Saxton, G.D.: Information, community, and action: How nonprofit organizations use social media. *Journal of Computer-Mediated Communication* 17(3), 337–353 (2012)
13. Manning, C.D.: Computational linguistics and deep learning. *Computational Linguistics* 41(4), 701–707 (2015), https://doi.org/10.1162/COLI_a_00239
14. Mikolov, T., Chen, K., Corrado, G., Dean, J.: Efficient estimation of word representations in vector space. CoRR abs/1301.3781 (2013), <http://arxiv.org/abs/1301.3781>
15. Ostrom, E.: The institutional analysis and development framework and the commons. *Cornell Law Review* 95, 807–816 (2010)
16. Ostrom, E., C., H.: *Understanding Knowledge As a Commons: From Theory to Practice*. MIT Press (2007)
17. Sakaki, T., Okazaki, M., Matsuo, Y.: Tweet analysis for real-time event detection and earthquake reporting system development. *IEEE Transactions on Knowledge and Data Engineering* 25(4), 919–931 (2013)
18. Yamasaki, S.: A trust rating method for information providers over the social web service: A pragmatic protocol for trust among information explorers, and provider in formation. In: 11th Annual International Symposium on Applications and the Internet (SAINT'11). pp. 578–582 (2011)