Intelligent Collection and Analysis of Citizens' Reports

Giulio Angiani, Paolo Fornacciari, Gianfranco Lombardo, Monica Mordonini,
Agostino Poggi, Michele Tomaiuolo
Dipartimento di Ingegneria e Architettura
Università degli Studi di Parma
Parma, Italy
{giulio.angiani, paolo.fornacciari, gianfranco.lombardo, monica.mordonini, agostino.poggi, michele.tomaiuolo}@unipr.it

***Abstract*—The great and capillary diffusion of technology between citizens is actually creating the ideal conditions for realizing the "Smart Community" concept. In this kind of socio-technical context, it is possible to create distributed applications for the administration of cities and neighborhood, using data provided directly by citizens. In particular, it is possible to connect users and integrate their actions into the whole system, thanks to wide available instant messaging apps, used by the greatest part of mobile users. The combined use of public APIs, Web systems and automated bots allowed us to build a comprehensive framework for managing the reports sent to the local government by citizens through their already installed and well-known instant messaging apps, such as Whatsapp, Telegram and Messenger.**

In this paper we show the techniques used for retrieving and classifying texts and images of the reports, for their management by the most appropriate branch of local administration. Our results show that an automatic classification system of this kind can reach an accuracy of over the 90%.

***Index Terms*—Text analysis, Image classification, Government 2.0, Chatbot.**

I. Introduction

The widespread distribution of mobile devices and fast network coverage in recent years has created the conditions to allow all citizens to access and send information at any time of the day. The concept of "smart community" arises precisely from the possibility, for users, to interact quickly and in real time on different aspects of real life.

One of these is sending reports to local public administrations. In almost all cases, this service is offered by private companies that use proprietary apps to send and classify reports.

Typically these apps allow users to send geo-localized text and images after specifying the category to which they belong. The user who wants to use them must therefore download the app from the appropriate store, register, learn how to use it correctly. In many cases this user overhead discourages the use of the reporting system by citizens who are used to other messaging channels.

Our project has instead reversed the approach starting from the use, by citizens, of free and widely used apps (Whatsapp, Telegram and Messenger) to send the same information to public administrations. The user is therefore free to send the text and image with the app he prefers by interacting with various bots dedicated to different messaging systems.

Furthermore, the user is not responsible for categorizing the report, because this information is derived from an AI system through text-analysis, image-recognition and object-detection. Depending on the inserted text and the content of the attached image, this automatic system associates the message with one of four predefined classes (environment, lighting, maintenance and security), corresponding to branches of the local administration.

The rest of the manuscript is structured in the following way: Section I presents some of the most interesting works, conducted in this field; Section II describes the data collection and methodology used for this study; results and discussions are presented in Section III; finally, Section IV provides some concluding remarks.

Literature review

The widespread use of devices that are always connected makes it possible for citizens to actively participate in the collection of local information and news, that has been used for very different purposes. This changing landscape of technology-enabled engagement with communities, cities and spaces was first illustrated in the book *From social butterfly to engaged citizen* [1].

In [2] and [3], authors show the first results of studies carried out on the importance of citizen participation in the increase of knowledge and the sharing of information in many fields, such as the tagging of road maps, the spread of diseases [4], the dissemination of emergency news, the review of hotels and restaurants, the creation of 3D models based on user images. In [5] a mechanism of crowd sourcing and portable smart devices is used to enable real time, location based crime incident searching and reporting. This participatory approach, well presented in [6], clarifies the concept of *Government 2.0*, where knowledge is shared between citizens and institutions.

In [7] the authors explore recent research regarding the potential of ICT, social media and mobile technologies to foster citizen engagement and participation in urban planning. ICT technologies and urban planning are important aspects of the so-called *socio-technical systems*, in which the technology complexity is amplified by the organisational and procedural complexity of the application domain [8] A survey on existing approaches for fostering citizen participation is presented in [9].

In light of this, recently several e-Government services have been introduced by government and administrations in the form of conversational chatbots. Singapore's administration operates a bot [10] that can answer a broad spectrum of questions, providing citizens with links to their web portal like a traditional search engine. Another interesting case is the WienBot [11], operated by the Wien city administration. Using a rather small knowledge base it provides efficient talk capabilities but limited to its core domain. Some of limitations of the reported cases are related to the absence of a Natural Language Processing step of the contents. Infact, one of the most adavanced e-Government chatbot, the Burgeramter chatbot [12], provided by the Berlin city administration, is based on a multi-staged framework that combines Sentiment Analysis and POS-Tagging of the questions and a knowledge base. This bot provides only information and raccomendations concerning public offices and services but do not accepts citizens reports.

It should also be noted that nowadays users are , virtually, always connected to the network. Thus, they are able to send and share information in various ways: via social networks (Twitter, Facebook, Instagram, and others), or through dedicated apps, or through instant messaging systems (Telegram, Whatsapp, Messenger). For example, in [13] a mobile apps allows walkers to map urban accessibility barriers/facilities, while wandering around. In [14] there is a description of a mobile app which allows users to take geo-tagged photo of road fault reporting, attach a brief description, and submit the information as a maintenance request to the local government organisation of their city. A number of cities in USA have worked to create apps that allow users to interact in order to report grafiti or, in general, code violations, such as My-Delaware app or Boston's Citizens Connect mobile app: [15].

In these cases, the information is typically sent through natural language or self-produced images. Thus, the application needs to classify the reports based on the results of both text and image analysis. The analysis of natural language on social networks is carried out for different purposes including: (*i*) analysis of users' sentiment [16] [17] [18], (*ii*) analysis of discussed topics [19], and (*iii*) analysis of the text structure [20]. Image analysis and object-recognition are also used for several goals, such as real-time object detection [21] or 3D-modeling [22]. In our work, the object-detection process has been used for understanding what a user would send to the institutional partner.

II. Methodology

As we mentioned in the chapter I, the overall project aimed at the automatic classification of reports sent via instant messaging systems. Each sent message could contain text and images, analyzing which it was possible to automatically assign the correct category among the 4 possible ones: environment, lighting, maintenance and security. These categories are then useful to address each report to the most appropiate brach of the local administation.

As a fundamental consequence of the nature of reports, a requirement for the whole system is the ability to correctly interpret both the text and any attached image. The overall architecture of the system is shown in Figure 1. It has been realized over ActoDES, which is a software framework which adopts the actor model. In particular, it simplifies the development of complex distributed systems [23] and it already integrates modules for gathering online data from social networks and for the automatic classification of such data [24].

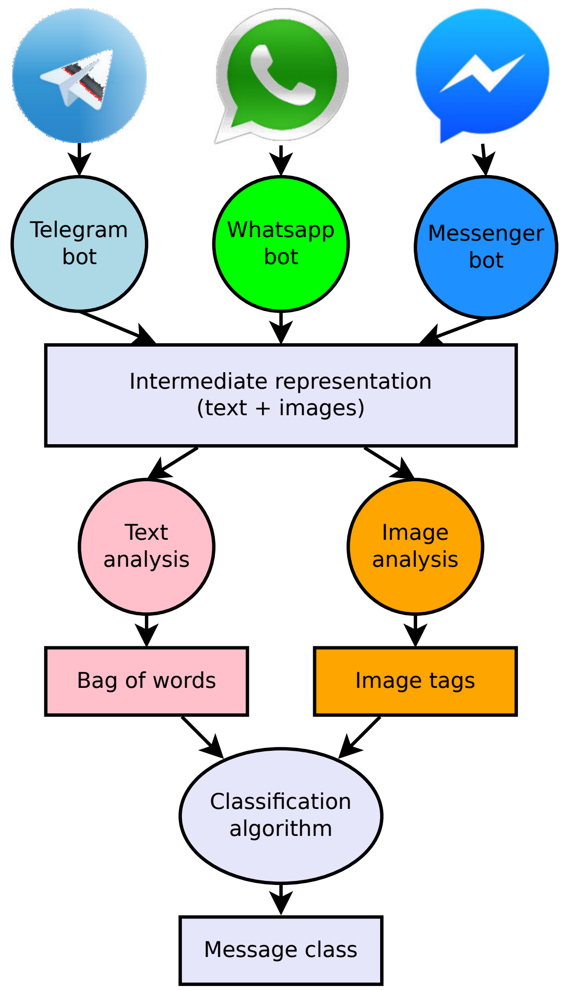

Figure 1. Data flow representation of the whole system.

The first layer includes the implemention of specific bots related to the different messaging systems and the appropriate components to make the representations of the various messages homogeneous. Then, each message is treated internally with a representation in JSON format containing information on text, images, sender user data, date and time of the message, and any information related to geolocation. Reports are then managed in a Web-based system, which we have developed for playing the role of an ad-hoc Customer Relationship Management (CRM) system, as shown in Figure 2.

A. Dataset collection

The first phase of the project is therefore focused on the definition and implementation of the text classifier. For this purpose, 7758 citizens' reports have been downloaded from the institutional websites of several Italian municipalities. These data represent the initial working dataset. These reports, publicly visible on the administrations' websites, are associated to categories directly by the users who have sent them or by the offices in charge.

Since these data reside on different systems and are managed in a non-homogeneous way, we have counted 27 categories, to which the various reports are associated. Many of these categories however differ only by designation and not by concept (i.e. "road safety" versus "road maintenance").

The first operation is therefore to reduce the numerous classes to the four chosen for the project. These four categories (Environment, Lighting, Maintenance, Security) have been identified because conceptually connected with some corresponding administrative offices that will manage the citizens' reports themselves.

However, for comparing the diffent possible types of analysis, we have limited the dataset to the massages containing both text and images. After manual analysis, we have found that the least represented category in the reduced dataset is lightening, with little more than 200 instances. We have then proceeded to balance the dataset using around 200 instances for each class, obtaining 804 instances in total, which we have used for further analysis and comparisons. Table I shows the number of reports associated with the different classes, after this selection.

Table I
NUMBER OF SELECTED REPORTS PER CLASS.

Class	Number of messages
Environment	201
Lighting	201
Maintenance	200
Security	202

1) Text analysis:

For the text analysis branch, a preprocessing operation is carried out on the text to eliminate characters not useful for classification, for example punctuation, reports without information content, emoticons. The stemming operations is applied and the text is filtered through a list of stop-words. Finally, the text is vectorized according the Bag of Words approach.

2) Training the classifier:

At this point the dataset is ready to be used for training a text classifier. As proposed in [18], the Multinomial Naive-Bayes classification algorithm is chosen for the analysis of natural language. However, we also campare it with other well known automatic classification algorithms, namely: Random Forest (RF), Support Vector Machine (SVM), and K-Nearest Neighbors (KNN).

Therefore, the next step is the creation of a four-way text classifier as in fig. 3 whose output is the class to which a message belongs. In fact, the system also emits a dictionary with the confidence values for each reported class. The process is represented in Figure 3.

An example of the classifier's output regarding confidence for each class is as follows

```
{
   "environment": 0.8731,
   "lighting": 0.1023,
   "maintenance": 0.0092,
   "security": 0.0154,
}
```

B. Image classification

The second branch of the project is focused on the analysis of the images present in the reports sent by the users. In order to associate the correct class to the entire reporting, it is necessary to understand which entities are represented within the associated image.

This is done using the Clarifai [1] [25] object-recognition service. The service is available via Rest API and a convenient Python module.

For each image associated with a message, a call is made to the service to retrieve information related to the content of the image. Among other information, the results returned by Clarifai in JSON format contain the list of *entities* (or *concepts*) associated to the image. Additionally, a probability value is associated with each entity. As an example, a list of entities like the following one can be obtained:

```
"entities":[
            {
              "name": "train",
              "value": 0.9989112
            },
            {
              "name": "railway",
              "value": 0.9975532
            },
            {
              "name": "station",
              "value": 0.992573
            },
          ]
```

Here, *value* is the confidence value of the entity indicated in *name* being represented within the image.

After analyzing the response of the external service, we construct a sequence of words using only the entities with a confidence value greater than 0.9. This sequence of entities is analyzed exactly in the same way as the bag of words obtained from the text. The same classification algorithms described

[1]Clarifai Inc., API Reference, https://clarifai.com/developer/reference/

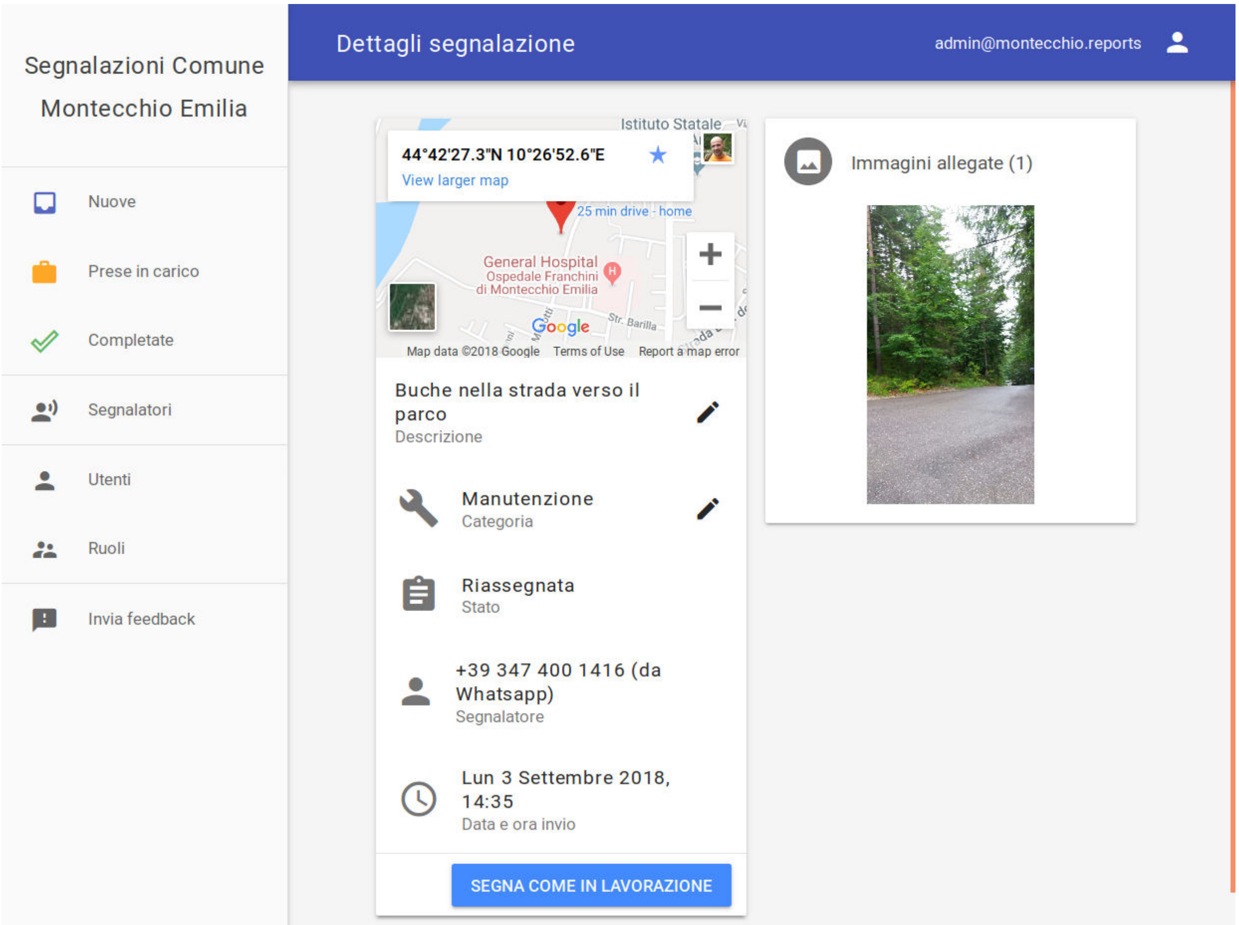

Figure 2. The Web-based interface.

above are compared also in the case of image analysis. The process is represented in Figure 4.

C. *Classification approaches*

After realizing the subsystems for the analysis of text and images, we have performed a comparison between the text classification results and that of the associated images. More precisely, we have compared three cases:

- Classification through text analysis, only (using the bag of words approach)
- Classification through image analysis, only (using the image entities)
- Classification through both text and image analyses (concatenating their lists of features)

In Section III, we show the analytical values of the accuracy of the classifier itself, using various features and algorithms. For those evaluations, the dataset is splitted and then used for both training and validating the classifier. For improving the consistency and reproducibility of results, we have adopted the well-known ten-fold Cross Validation technique.

III. Results

After creating the whole dataset, we compare the three different approaches described in the previous section. In the following, these approaches are identified as: *Text*, based only on text analysis; *Image*, based only on image analysis; and *Text+Image*, based on both text and image analyses.

We also compare various well-known classification algorithms, namely: *RF*, Random Forest; *NBM*, Naïve Bayes Multinomial; *SMO*, Sequential Minimal Optimization, based on the principles of support vectors; *KNN*, K-Nearest Neighbors. In this latest algorithm, we use K=1 to gain its best results.

It can be observed in Figure 5 and Table II that all algorithms perform better on the image entities than on the text. Moreover, all algorithms improve their results using all available features, with the exception of KNN, which is known to work better on a limited set of features.

Overall, the best classification accuracy is obtained by the NBM algorithm, using the features of both text and images. In this case, the classification is correct in over 90% of cases. However, using the RF algorithm, a very close value of accuracy can be obtained even using only the image features.

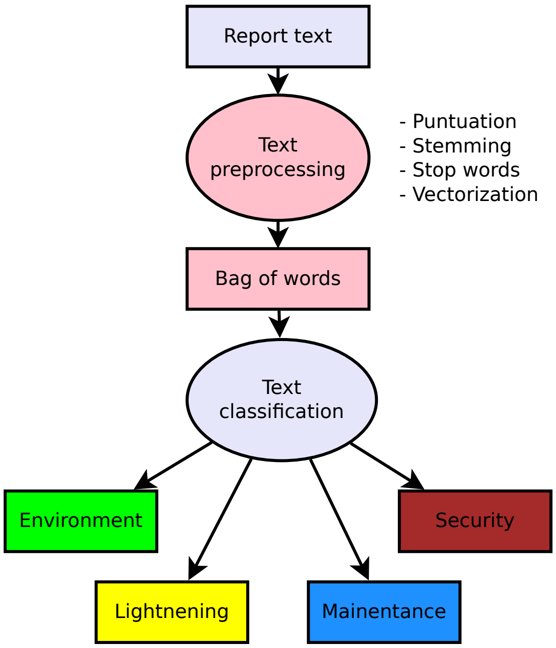

Figure 3. Text classication process.

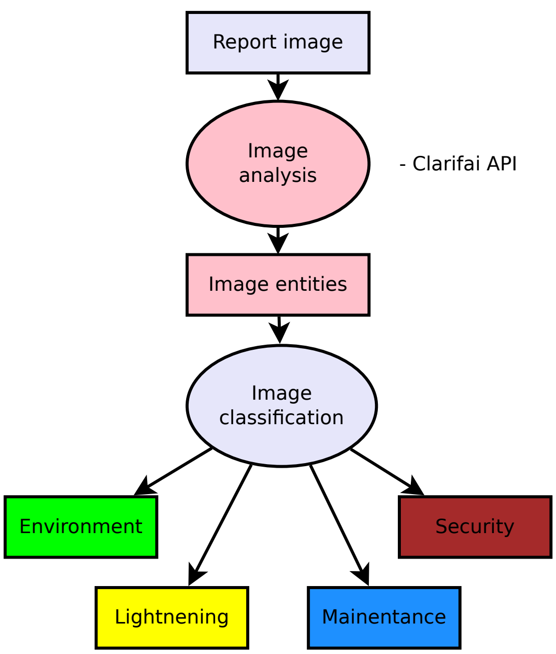

Figure 4. Image classication process.

This way, it is possible to greatly ease the burden on users, when they have to issue reports about local problems.

A. Accuracy of the classifiers

Finally, Table III represents the confusion matrix. It can be observed that few errors occurs. Among those, it is worth noting that: (*i*) 15 Environment reports are mis-classified as Security ones, possibly due to the presence of instances about unsecure parks in the dataset; and (*ii*) 14 Security reports are mis-classified as Lighting ones, since the two issues often coexhist.

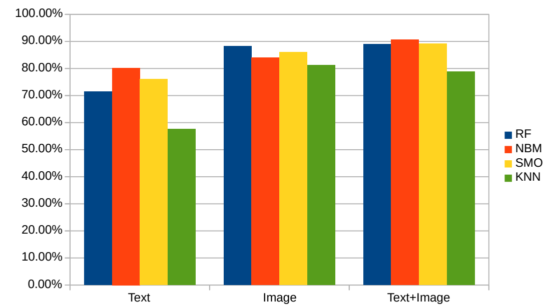

Figure 5. Accuracy of classifiers, using different features and different algorithms.

Table II
NUMERICAL ACCURACY OF CLASSIFIERS.

	Text	**Image**	**Text+Image**
RF	71.52%	88.18%	88.93%
NBM	80.22%	84.08%	90.67%
SMO	75.99%	86.07%	89.25%
KNN	57.58%	81.21%	78.85%

Table III
CONFUSION MATRIX FOR NBM ON TEXT+IMAGE; CLASSES ARE: A=ENVIRONMENT, B=SECURITY, C=MAINTENANCE, D=LIGHTING.

Classified as ↱	a	b	c	d
a	175	15	11	0
b	8	172	8	14
c	1	6	193	0
d	0	10	2	189

IV. CONCLUSIONS AND FUTURE WORKS

Our project shows an implementation of an automatic classification system for reporting citizens to public administrations through the use of instant messaging. This operation was carried out by analyzing separately the text and images of a report and then comparing the results of this analysis. The accuracy of the final classification has achieved results overall greater than 90%, using features from both text and associated images. However, also using only the entities found through the image analysis, very similar results can be obtained. These results suggest that it is possible to collect citizens' report in a very simplified way, receiving just geolocalized images, which can be classified automatically in most cases. The use of automated bots for interacting with the users allow them to correct the wrong results in a very convenient way, only when necessary.

The future developments of the project will concern different kinds of analysis, in cases of discrepancy of the classifiers. Furthermore, after the deployment of the service in some public administrations, it will be possible to carry out content analyzes that can also be based on data related to the map of the territory and the history of reports.

V. Acknowledgements

This project has been developed in collaboration and agreement with the local administration of Montecchio Emilia (Italy), following the concepts of the Government 2.0. The local administration has helped to better understand the management process for the main types of reports and has suggested some guidelines for users' operations, which can be better handled by public offices.

References

[1] M. Foth, L. Forlano, C. Satchell, and M. Gibbs, *From social butterfly to engaged citizen: urban informatics, social media, ubiquitous computing, and mobile technology to support citizen engagement*. MIT Press, 2011.

[2] A. Hermida, "Twittering the news: The emergence of ambient journalism," *Journalism practice*, vol. 4, no. 3, pp. 297–308, 2010.

[3] M. N. Kamel Boulos, B. Resch, D. N. Crowley, J. G. Breslin, G. Sohn, R. Burtner, W. A. Pike, E. Jezierski, and K.-Y. S. Chuang, "Crowdsourcing, citizen sensing and sensor web technologies for public and environmental health surveillance and crisis management: trends, ogc standards and application examples," *International Journal of Health Geographics*, vol. 10, no. 1, p. 67, Dec 2011. [Online]. Available: https://doi.org/10.1186/1476-072X-10-67

[4] M. N. K. Boulos, S. Wheeler, C. Tavares, and R. Jones, "How smartphones are changing the face of mobile and participatory healthcare: an overview, with example from ecaalyx," *Biomedical engineering online*, vol. 10, no. 1, p. 24, 2011.

[5] S. Shah, F. Bao, C.-T. Lu, and I.-R. Chen, "Crowdsafe: crowd sourcing of crime incidents and safe routing on mobile devices," in *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM, 2011, pp. 521–524.

[6] S. Chun, S. Shulman, R. Sandoval, and E. Hovy, "Government 2.0: Making connections between citizens, data and government," *Information Polity*, vol. 15, no. 1, 2, pp. 1–9, 2010.

[7] R. Kleinhans, M. V. Ham, and J. Evans-Cowley, "Using social media and mobile technologies to foster engagement and self-organization in participatory urban planning and neighbourhood governance," *Planning Practice & Research*, vol. 30, no. 3, pp. 237–247, 2015.

[8] G. Cabri, M. Cossentino, E. Denti, P. Giorgini, A. Molesini, M. Mordonini, M. Tomaiuolo, and L. Sabatucci, "Towards an integrated platform for adaptive socio-technical systems for smart spaces," in *Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE), 2016 IEEE 25th International Conference on*. IEEE, 2016, pp. 3–8.

[9] F. Salim and U. Haque, "Urban computing in the wild: A survey on large scale participation and citizen engagement with ubiquitous computing, cyber physical systems, and internet of things," *International Journal of Human-Computer Studies*, vol. 81, pp. 31–48, 2015.

[17] P. Fornacciari, M. Mordonini, and M. Tomaiuolo, "Social network and sentiment analysis on twitter: Towards a combined approach." in *KDWeb*, 2015, pp. 53–64.

[10] "5 reasons to use the gov.sg bot, march 2017. blog of singapore government," https://www.gov.sg/news/content/5-reasons-to-use-the-gov-sg-bot.

[11] "Magistrat der stadt wien. wienbot - der chatbot der stadt, june 2017," https://www.wien.gv.at/bot/.

[12] A. Lommatzsch, "A next generation chatbot-framework for the public administration," in *Innovations for Community Services*, M. Hodoň, G. Eichler, C. Erfurth, and G. Fahrnberger, Eds. Cham: Springer International Publishing, 2018, pp. 127–141.

[13] P. Salomoni, C. Prandi, M. Roccetti, V. Nisi, and N. J. Nunes, "Crowdsourcing urban accessibility: Some preliminary experiences with results," in *Proceedings of the 11th Biannual Conference on Italian SIGCHI Chapter*. ACM, 2015, pp. 130–133.

[14] M. Foth, R. Schroeter, and I. Anastasiu, "Fixing the city one photo at a time: mobile logging of maintenance requests," in *Proceedings of the 23rd Australian Computer-Human Interaction Conference*. ACM, 2011, pp. 126–129.

[15] J. Evans-Cowley, "There is an app for that: mobile applications for urban planning," *International Journal of E-Planning Research (IJEPR)*, vol. 1, no. 2, pp. 79–87, 2012.

[16] B. Liu, "Sentiment analysis and opinion mining," *Synthesis lectures on human language technologies*, vol. 5, no. 1, pp. 1–167, 2012.

[18] G. Angiani, L. Ferrari, T. Fontanini, P. Fornacciari, E. Iotti, F. Magliani, and S. Manicardi, "A comparison between preprocessing techniques for sentiment analysis in twitter." in *KDWeb*, 2016.

[19] C. C. Aggarwal and C. Zhai, *Mining text data*. Springer Science & Business Media, 2012.

[20] A. Athar, "Sentiment analysis of citations using sentence structure-based features," in *Proceedings of the ACL 2011 student session*. Association for Computational Linguistics, 2011, pp. 81–87.

[21] P. Piccinini, A. Prati, and R. Cucchiara, "Real-time object detection and localization with sift-based clustering," *Image and Vision Computing*, vol. 30, no. 8, pp. 573–587, 2012.

[22] P. Henry, M. Krainin, E. Herbst, X. Ren, and D. Fox, "Rgb-d mapping: Using depth cameras for dense 3d modeling of indoor environments," in *Experimental robotics*. Springer, 2014, pp. 477–491.

[23] F. Bergenti, A. Poggi, and M. Tomaiuolo, "An actor based software framework for scalable applications," *Lecture Notes in Computer Science (LNCS)*, vol. 8729, pp. 26–35, 2015, proc. 7th International Conference on Internet and Distributed Computing Systems (IDCS 2014); Calabria; Italy; 2014-09-22/24 [MT].

[24] P. Fornacciari, M. Mordonini, A. Poggi, L. Sani, and M. Tomaiuolo, "A holistic system for troll detection on twitter," *Computers in Human Behavior*, vol. 89, pp. 258–268, 2018.

[25] C. Inc. (2018) Clarifa.com web site. [Online]. Available: https://clarifai.com/