UNIVERSITÀ DEGLI STUDI FIRENZE

DISEI

DIPARTIMENTO DI SCIENZE PER L'ECONOMIA E L'IMPRESA

WORKING PAPERS – BUSINESS

A series on Accounting, Finance, Management, Marketing and Organizational Studies

Big Data for smart cities and citizen engagement: evidence from Twitter data analysis on Italian municipalities

SILVIA BLASI, EDOARDO GOBBO, SILVIA RITA SEDITA

Working Paper N. 1/2022

DISEI, Università degli Studi di Firenze
Via delle Pandette 9, 50127 Firenze (Italia) www.disei.unifi.it

Big Data for smart cities and citizen engagement: evidence from Twitter data analysis on Italian municipalities

Blasi, S.[1]; Gobbo, E.[2], Sedita, S.R.[3]

Abstract

Smart cities are increasingly keen to establish a fruitful conversation with their citizens, to better capture their needs, and create virtual platforms for stimulating co-creation processes between government and users, with the final objective of increasing the quality of life and well-being. Social media applications provide an opportunity for dialogic communication, where, for a relatively low cost, a large amount of information reaching a wide audience can be published and exchanged in real time, fueling opportunities for citizens' engagement.

This study is based on a social media listening method, through a twitter data mining, which enabled disentangling different components of citizen engagement (popularity, commitment and virality) for a sample of Italian municipalities. In addition, we executed a deep analysis of the types of communication artifact exchanged and, through a content analysis of the tweets published by followers of the municipalities' accounts, we identified main areas of interests of the social media conversations. Our results are based on the analysis of online conversations engaged by followers of twitter accounts of a sample of 28 Italian municipalities, chosen among the most active and densely populated.

We show that municipalities tend to use the twitter account as a channel of communication to inform the population about a variety of topics, such as transports and public works, among the others. The volume of activity and number of followers (audience) vary from one municipality to the other. There is generally a negative relationship between the density of the population of a municipality and citizens' engagement: smaller municipalities show a higher citizens' engagement; the biggest ones, like Roma, Milan, Turin, Naples, are laggards. We finally conducted a city profiling process, which provides a representation of key citizens' segments in terms of engagement. Policy makers could find in our work useful tools to increase citizens' listening capacity.

Keywords: smart cities, e-government, twitter, web scraping, social media listening, we-government

JEL Code: M10, M38

[1] Department of Business Administration, University of Verona, via Cantarane, 24 - 37129, Verona, Italy. Email: silvia.blasi@univr.it

[2] Department of Economics and Management, University of Padova – Email: edoardogobbo@gmail.com

[3] *Corresponding Author.* Department of Economics and Management, University of Padova, Via del Santo 33 – 35123, Padova, Italy. Email: silvia.sedita@unipd.it

1. Introduction

Since the beginning of this millennium, the concept of smart cities has gained traction amongst businesses, governments, media, and academia (Kitchin, 2015; Zheng et al., 2020). This term has been mainly used to refer, on the one hand, to the use of ICTs (Information and Communication Technologies) to stimulate economic development and, on the other hand, to the extensive embedding of digital platforms into the fabric of the city to augment urban management towards e-governance (Caragliu et al, 2011; Kitchin, 2014; Nijkamp and Cohen-Blankshtain, 2013; Samuel et al., 2020). This approach finds support in the OECD and EUROSTAT Oslo Manual (2005), which highlighted the importance of ICT as driver for urban innovation. Lately, the concept of smart city is morphing toward a more comprehensive understanding of city planning and development, which is inclusive of social and environmental concerns (Zheng et al., 2020). However, the concept of smart city is still anchored on a 'one-size fits all' and efficient view of technology. In particular, the development of new technologies and new ways of communication, alongside the increasing population density in urban centers, represents a unique opportunity for new technologies, and specifically the Big Data, in developing smart cities. Through Big Data, city management and citizens are given access to a wealth of real-time information about the urban environment upon which to base decisions, actions, and future planning in a collaborative atmosphere, reducing inequalities and social polarization (Engin et al., 2020). The wide usage of Big Data technologies helps achieving new frameworks and paths for planning smart cities, enabling the creation and supporting the development of social and relational capital (Coe et al., 2001).

Smart cities initiatives are not limited to optimize traffic patterns, parking management, efficient lighting, improvements of public works etc., rather are oriented to fuel citizen engagement as critical to accomplish a smart and inclusive growth trajectory (Lee, 2019). Smart cities picture themselves as creating technologies, techniques and visions that are scientifically grounded, objective, pragmatic and apolitical (Kitchin, 2015; Vanolo, 2014). However, this technologically grounded model of city development is generating several social and environmental concerns, whose impact and meaning are often overlooked (Hollands, 2008).

On the one hand, smart cities using big data can generate benefits in transport, health and safety and pollution management (Belanche et al, 2016; Söderström et al., 2014). Through digital mapping it is possible to identify areas and times of maximum congestion, or to evaluate the safety of places, or, finally, to create optimization processes aimed at reducing emissions, saving people liters of water and reducing solid waste production (Cheela et al., 2021). On the other hand, leaving aside issues such as privacy and social control, which are often covered, albeit with little results, in the columns of national and international newspapers, there are more and more scholars, policy makers and city managers voicing their concern for the negative effect of this model (Kummitha and Crutzen, 2017). Even Florida (2017), who is one of the main proponents and admirer of this hi-tech model of cities, urges policy makers and city managers to take corrective actions. This is because smart city strategies are producing inequalities, gentrification, segregation, and progressive worsening of both natural and man-made environment. The dream of a smart city and of a smart development has translated into a fragmented patchwork: large part of the wealth generated has spatially concentrated into few small residential areas sporadically lived by a global creative class constantly moving from one place to the other; on the contrary, large part of the poverty has been localized in vast areas, surrounding the firsts, of social, cultural, and environmental degradation. It is for this reason, according to Florida (2017), that political situation, such as Brexit, has come to place.

In the light of this debate, existing academic contributions have devoted great attention to the technological development of smart cities, with the introduction of the Web 2.0 and the implementation of crucial technologies in our daily lives, nevertheless just few studies have been conducted in understanding how local governments effectively interact with their citizens through the usage of such technologies. However, we believe that the relationship between municipalities and citizens is a crucial point in this scenario. It is essential that the administration has a clear understanding of the smart city model that is most capable of transforming the community in a positive sense, through digital technology: a model that is economically sustainable, i.e. based on the ability of a territory to activate collaboration between public, private and civil society, from the socio-cultural point of view, i.e. focused on the daily needs of citizens. The smart city must first and foremost be "human" and for this to be the case it is necessary that informed administrators and active citizens share a common vision.

Driven by this scientific curiosity, we aim to answer the following research questions: Which are the main areas of interactions between local governments and citizens? How do municipalities, through citizens engagement, face the transition towards smart cities? Which are the profiles of citizens that interact with municipality managers?

In this paper we contribute to the existing literature by providing a general overview of the use of social media networks for transferring and exchanging information at the city level from Italian citizens and municipalities. We focus our analysis on Twitter conversations, since this is the preferred social media for the interaction city-citizens. In addition, we explore which factors influence citizens engagement considering municipalities' size, Twitter activity, audience, content, and media types. Finally, we conducted a city profiling process, which provides a representation of key citizens' segments in terms of engagement. The results show that municipalities use their Twitter account as a channel of communication to inform the population about transports, public works, etc. The volume of activity and number of followers (audience) change from one municipality to the other. There is a negative relationship between the municipality size and its citizen's engagement: smaller municipalities have a higher citizens engagement; the bigger ones, like Roma, Milano, Torino, Napoli, on the other hand should improve this aspect. If we look at the composition of citizens, three groups of cities emerge: Melting pot, Lagging and Change Makers. They are characterized by different levels of engagement.

The structure of the paper is as follow. Section 2 presents the theoretical framework, Section 3 introduces the data, Section 4 explains the methodology used, Section 5 presents the results. Finally, Section 6 offers the discussion and some conclusive remarks.

2. Theoretical Framework

2.1 Big Data and Smart Cities

In recent years, cities are undergoing significant changes. There are many reasons for this: people's habits and lifestyles are changing, new needs are arising, and new problems need to be solved (i.e., climate change, improving the quality of life). Urban centers are therefore becoming the focus of this change, with the consequent rethinking of urban planning in technological terms. To be intelligent, provide efficient services and manage infrastructures in an advanced way, a city must assess real needs and provide concrete answers. To do this, it is essential to start with data (Franke & Gailhofer, 2021). In an increasingly computerized world, in fact, the role of data collection and analysis is becoming fundamental for organizing and managing workplaces, businesses and even cities (Al Nuaimi et al., 2015; Allam and Dhunny, 2019; Lim et al., 2018; Neirotti et al., 2014; Soomro et al., 2019).

A data-driven approach requires technologies capable of collecting and analyzing data (Del Vecchio et al., 2018). Data collection can be carried out in a variety of ways, for example by installing sensors on lampposts, at urban infrastructures or in buildings, although potentially any object or device connected to the network becomes a possible source of data collection. Constant monitoring would make it possible to estimate the health of urban infrastructures, such as roads, sewers, and energy distribution systems. Collecting and analyzing data not only allows immediate action to be taken when a problem is detected, but also prevents breakdowns, malfunctions, or damage through regular, targeted maintenance. Other factors that can be monitored by sensors are air and noise pollution, providing sufficient information to safeguard the urban environment. Sensors and data also help in emergency situations by collecting and disseminating information in real time. The design of new services must also be based on a data-driven approach, for example by analyzing the habits of citizens, the way they move around and live in the city. This opens a range of possibilities in the fields of transport, traffic management, security, design, and development of green and public spaces. Data can also help in the more efficient management of energy and water systems, distributing resources where and when they are needed, without waste or malfunctioning.

More specifically, as regards transport, using Big Data every type of transport (car, bus, bicycle) can be monitored. It means identifying areas and times of maximum congestion and possible alternative routes (Zannat & Choudhury, 2019). The response of cities could be both strategic and instantaneous. Data analysis would suggest the direction that urban development should take; it would indicate the best positioning of bus stops and stations (for an organic network that interweaves different means of transport, from trams to bike sharing) (Batty, 2013; Liu et al., 2017); it would encourage traffic lights to work more efficiently and advise which urban and suburban public lines should be strengthened and which should be lightened (Hancke et al., 2012; Schlingensiepen et al., 2016). By building an 'operations center', the data could then provide a real-time response: every driver, thanks to increasingly advanced navigation systems, could receive information on traffic jams and accidents (Mehmood et al., 2017; Tranos and Mack, 2019; Ziemke et al., 2021).

Regarding health and safety, Big Data analysis can create a map of crimes. It would identify the riskiest areas, beyond the clichés (Catlett et al., 2019; Rathore et al., 2016). Understanding where infractions and crimes are concentrated would not only allow the concentration of the police forces but would indicate methods of prevention that can be (always with the data) tested and corrected. The combination artificial intelligence-Big Data, in fact, has not only the capacity to take a photograph of the existing situation, but through models and algorithms, it has predictive power.

Finally, regarding pollution and the management of energy and water systems, cities through the optimal use of Big Data could reduce emissions and drastically decrease solid waste production (Miles et al., 2018; Mizuki et al., 2012; Toma et al., 2019). Using sensors and artificial intelligence, pollutant production, electricity and water networks can be monitored. This activity leads benefit for the environment, but also for the household, business, and government coffers.

Currently, only few municipalities have platforms or systems for live monitoring and inferring of urban process parameters. Data collection exercises are often costly and difficult to replicate, however there is an increased demand on municipalities for incorporating smart technologies that collect the required data and analyze them. Through advanced sensing and computational capabilities, data are gathered and real-time evaluated to extract the information, which is further converted to usable

knowledge: the whole process will help enhancing the decision making of city management and citizens to turn the city into a smart city (Jin, et al., 2014).
Since big data applications are complex to be set and adopted by citizens and urban planners, it is crucial to get information about the main topics of interest that could be hosted by the city digital platform, and the most effective modalities of interaction between citizens and government. Social media, in this sense, could highlight unusual events and emergencies as soon as they occur or even before. Governments should therefore have social media analysis tools as well as alerting tools, so that they can quickly gather critical information from the different platforms on which citizens and local stakeholders are posting. The local government should make full use of the information available in the social media ecosystem to optimize the quality of life for its citizens. This analysis should be preliminary to an effective implementation of big data applications in cities. Our work is exactly oriented to explore existing dynamics of interaction exploiting data coming from Twitter, by monitoring the civil participation in the conversations launched through the social media accounts of the municipalities. Results will inform of possible future trajectories of e-government.

2.2 E-government and citizen engagement

Many municipalities, through the introduction of e-democracy initiatives, are seeking to increase citizen involvement in decision-making, in order to increase consensus and legitimacy of public action. The term e-democracy refers to the participation of citizens in the activities of local public administrations and their decision-making processes using new communication technologies (Kneuer, 2016). The innovative use of ICT allows the opening of new spaces for dialogue between citizens and administration that complement and reinforce traditional forms of participation (Chadwick, 2003; Coleman & Norris, 2005). Involvement in public activities offers citizens the opportunity to familiarize with the initiatives of the administration and to express opinions and points of view that the public administration can take into consideration when making decisions. Such involvement not only fosters acceptance of public choices, but also makes the administration more open and reliable to citizens, builds a sense of trust in public authorities, improves the quality of public policies, and strengthens administration/citizen relations (Anttiroiko, 2003; Macintosh, 2008; Spirakis et al., 2010).

The information, consultation and active participation of citizens provide the administration with a better basis for formulating public policy and enable it to become a learning organization. The involvement of citizens in the different phases of the policy life cycle can be an important resource for gathering more information and alternative solutions from civil society, and for anticipating needs and requirements not expressed through the classic channels of representative democracy (Balestrini et al., 2017; De Lange & De Waal, 2017). New technologies are a valuable support tool to provide citizens with all the information necessary for informed participation (information level), to activate dialogue mechanisms (consultation level) and to reach shared decisions (active participation level). Information, consultation, and active participation increase the transparency of government and give it greater responsibility. Under these conditions, strengthening the relationship between government and citizens encourages active citizenship and promotes its integration into society (Falco & Kleinhans, 2018; Macintosh, 2004). Likewise, it stimulates citizens' engagement in the public sphere, encouraging them, for example, to participate in political debates, vote, join associations, etc. (Chadwick, 2008; van den Hoven, 2005).

Within a context where municipalities struggle to involve citizens, the spread of social media and the use of technologies have certainly favored the massive dissemination of information and generated new modes of interaction and collaboration that allow the involvement of citizens and organizations in the production of public services (co-production of public services or citizen coproduction) (Skoric et al., 2016; Verma et al., 2017). Social media, in fact, have become in a short time more and more pervasive in public administrations, becoming a central element of the e-government process (Bryer, 2013; Marino & Lo Presti, 2018). The use of social media in the public sector has also attracted the attention of the academic world, which has seen in the new modes of communication and interaction via the Internet the possibility of rethinking the traditional boundaries between individuals, organizations, communities, and the different levels of public administration (Bertot et al., 2010; Bertot et al., 2012; Jaeger & Bertot, 2010). We are seeing, in fact, a significant evolution of the relationship between the different actors in the public sector in which citizens are no longer customers, but partners of the administration for the definition of policies, and to produce public services. This involvement is justified by the possibility of increasing the quality of public services not only through greater citizen participation in decision-making, but also through greater control over resources and outcomes (Karakiza, 2015; Loeffler & Martin, 2015).

The development of new technologies and the use of social media has changed not only the tools for interaction between the public and private sectors but has also modified citizens' expectations regarding the quality and timing of service delivery. We move from the concept of 'e-government', in which the citizen was the final consumer, to 'we-government' in which the same becomes a partner in the production of public services. The main purpose of our research is to provide a general overview of how municipalities in Italy use social networks, in our case we decided to choose Twitter, as a social media that allows municipalities to interact with citizens. Moreover, the research aims to find out what factors influence citizen engagement considering the size of municipalities, Twitter activity, audience, content, and media types.

3. Data

Our research is based on Twitter data mining. Twitter uses an algorithm that pushes content into the user timeline (Asadi and Agah, 2017). The algorithm evaluates the relevance of each tweet based on how much recent the tweet is or whether it contains media, the eventual past interactions with the author of the tweet, and tweets the user found engaging in the past. The sample for this study is taken from the Istituto Nazionale di Statistica (ISTAT) which includes data for the 28 largest Italian municipalities ranked by number of citizens. These municipalities include those with a population exceeding 99,000 inhabitants and had an official website. The Twitter account of each municipality was collected by following the icon link on the municipality's official website or by searching. As last step before the web scraping activity, we evaluate the activity of each Twitter account. We selected only those accounts that post at least one tweet in the previous 6 months (the threshold limit was then set on April 2019, given that the data collection started in October 2019). Moreover, we included an additional threshold: each account needed to have issued at least 1,000 tweets since joining Twitter. We set these thresholds to avoid biases due to inactive accounts or fake accounts that could have affected our analysis. We then obtained our sample, composed by 28 municipalities among the largest Italian cities with an active Twitter account.

The data collection took place between October and November 2019 using Twint. In[4], a Python library which allows to perform Twitter scraping. Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. It is extremely attractive in the field of Rapid Application Development because it offers dynamic typing and dynamic binding options. In order to assess the citizen engagement per municipality, we utilized a framework developed by Bonsón, et al., (2019). As can be detected from the Table 1, for assessing the citizen engagement the authors introduced three different metrics, such as: "Popularity" which is based on the number of "likes" received by each municipality, "Commitment" which is based on the number of "replies" and "Virality" which refers to the number of retweets obtained. Once obtained the three different metrics per 1,000 followers the "Engagement" can be easily achieved by summing the components up.

[4] Twint.In is an advanced Twitter scraping tool written in Python language that does not use Twitter's API, allowing you to scrape a user's followers, following, tweets and much more while evading most API limitations (Bonsón et al., 2019). Twint.In utilizes Twitter's search operators to let scraping tweets from specific users. The tweets then were automatically scraped in a csv format and stored in different folders.

Table 1: Metrics for citizen engagement

Metrics	Code	Calculation
Popularity	P1	Number of tweets favorited/total tweets
	P2	Total number of times favorite/total tweets
	P3	(P2/total number of followers)*1000
Commitment	C1	Number of tweets commented/total tweets
	C2	Total number of comments/total tweets
	C3	(C2/number of followers)*1000
Virality	V1	Number of tweets retweeted/total tweets
	V2	Total number of retweets/total tweets
	V3	(V2/number of followers)*1000
Engagement		P3+C3+V3

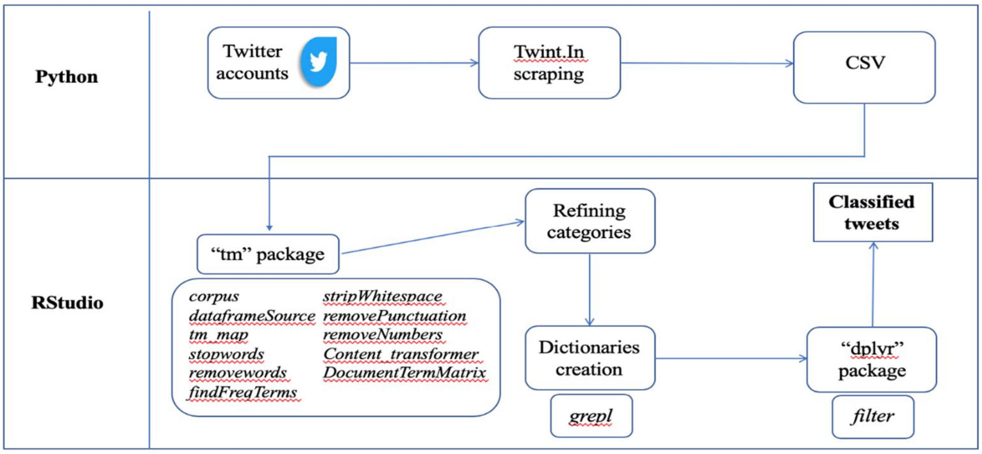
Source: Bonsón, et al., 2019, p. 484.

We collected 316,801 tweets coming from 28 different municipalities. We used the Python software (Van Rossum and Drake Jr, 1995) to scrape every tweet from each account (Zacharias and Poldi, 2018), and the R software to analyze them. Subsequently, we performed a qualitative and quantitative data analysis. The qualitative analysis, thought a content analysis, is aimed to identify the content of the tweet and the media type. The quantitative analysis, through statistical tools, investigates the metrics related to the followers (retweets, replies, likes) to understand the level of citizens' engagement.

4. Methodology

The proliferation of social networks, such as Twitter, offers new and alternative measurement approaches (Schwartz and Ungar, 2015). In this paper, through content and text mining analysis we examined all the tweets coming from different municipalities for understanding the rate of citizen engagement between the Italian population and their municipalities. The content analysis process is composed by three steps: 1) sampling and data collection, 2) coding; 3) content analysis. The Figure 1 provides a graphic explanation of the process.

Figure 1: Flowchart of the content analysis process.



Source: adaptation of Bonsón, et al., 2019, p.484

As it is possible to see in Figure 1, the first part of the content analysis (the Pyton one) refers to the data collection process (described above in the "data" section"); the second part (the RStudio one) describes the coding process and the analysis of the tweets. Before starting to code the tweets, we need to pinpoint the tweets "Content Type". The identification of the Content Type is based on an existing list created by Torres and Pina (2001), and subsequently adapted by several authors (Bonsón et al., 2019, 2015; Martí et al., 2012). We also revised and adapted our list according to the most frequent words appeared in the tweets after the analysis was accomplished. Following Bonsón et al. (2019) we set automatic vocabularies in order to assess the different content type of words. Setting automatic vocabularies represents a crucial phase of the content and text mining analysis. This technique guarantees more transparency and capture more connections between words (Schwartz and Ungar, 2015). The "Content Type" list was further refined according to the most frequent words published in the tweets provided by local governments, using RStudio package "tm"[5]. As it emerges in Figure 1, the "tm" package is composed by several functions. *Corpus* is composed by a set of text document and it represents the building block of this analysis. *Data Frame Source* function analyze a great number of tweets at the same time, considering each tweet as an independent document. *Tm_map* function cleans the Corpus of all "useless" information (e.g. remove punctuation, strip white space, identify and delete all the most common words (articles, verbs, etc.). The corpus was then transformed in a *DocumentTermMatrix* which employ spars e matrices for the Corpus. Thanks to this process we were able to identify the Content Type (Table 2)

[5] The "tm" package is a text- mining tool which offers some powerful functions which aids in text processing steps

Table 2: Content categories.

Content type	Tweet content	Code
Cultural	Cultural activities and events.	1
Employment and Education	Employment and training schemes. Public education.	2
Environment	Environmental concerns.	3
Security and Health	Citizen Protection and security. Health service.	4
Sport	Sport activities and events	5
Transport and Public Works	Public and private transport. Public works in the city. Town Planning.	6
Others		7

Source: Bonsón, et al., 2019, p. 484.

Thanks "grepl", we created a dictionary for each category, characterized by the words that composed each "Content Type". This stage on one hand reduces subjectivity, on the other hand gives an overlap error, since there are tweets that share more than one category (Bonsón et al., 2019). After the coding phase, using the RStudio library "dplyr", we automatically examined and categorized all the tweets.

5. Results

5.1 Descriptive analysis

The Twitter accounts of the municipalities under analysis have different number of followers: the account with the largest audience has 468,600 followers, while the account with the lowest audience has just 1,316 followers. This also imply a huge difference regarding the total number of tweets published. The average activity (number of tweets for single municipality) amounts to 11,314 whilst the average amount of followers per municipality is 56,795 (Table 3). The average amount of audience is clearly biased by the number of followers belonging to the largest cities in Italy (Roma, Milano, Torino and Napoli).

Table 3: Number of followers and tweet

	Average	Median	Maximum	Minimum	Std. Deviation
Audience	56,795.57	9,169.00	468,600.00	2,220.00	111,446.45
Activity	11,314.36	7,733.50	48,402.00	1,316.00	10,321.33

From the analysis of the Content Type categories emerge that municipalities tend to adopt Twitter account as a channel of communication to inform the population about the transports, public works and new town planning ideas, given that 26% of the tweets fall under this category. The second content category type, represented under the name “others” (25%), covers a diversity of topics that had not been included in our category selection. Cultural (19%) and Sport (13%) appear as crucial topics used by the municipalities for engaging with the population (Table 4).

Table 4: Percentage of Content Type

Content Type	Percentage
Transport and Public Works	26%
Others	25%
Cultural	19%
Sport	13%
Security and Health	7%
Environment	5%
Employment and Education	5%

The preferred media type adopted by the Italian municipalities is the website link (40%), which nowadays is roughly included in one out of two tweets produced. Although, the massive presence on the social network of photos and videos, the plain text still seems to be one of the most preferred ways of communication adopted by municipalities (28%). Photos and videos or the combination of web links- photos and videos appear with percentage of 18% and 12%, respectively (Table 5).

Table 5: Percentage of Media Type

Media Type	Percentage
Web link	40%
Text	28%
Photo/Video	18%
Web links Photo/Video	12%
Others	2%

In order to assess the citizen engagement between the Italian municipalities and the population, we used the metrics proposed by Bonsón et al. (2019, 2015). Table 6 shows the main results.

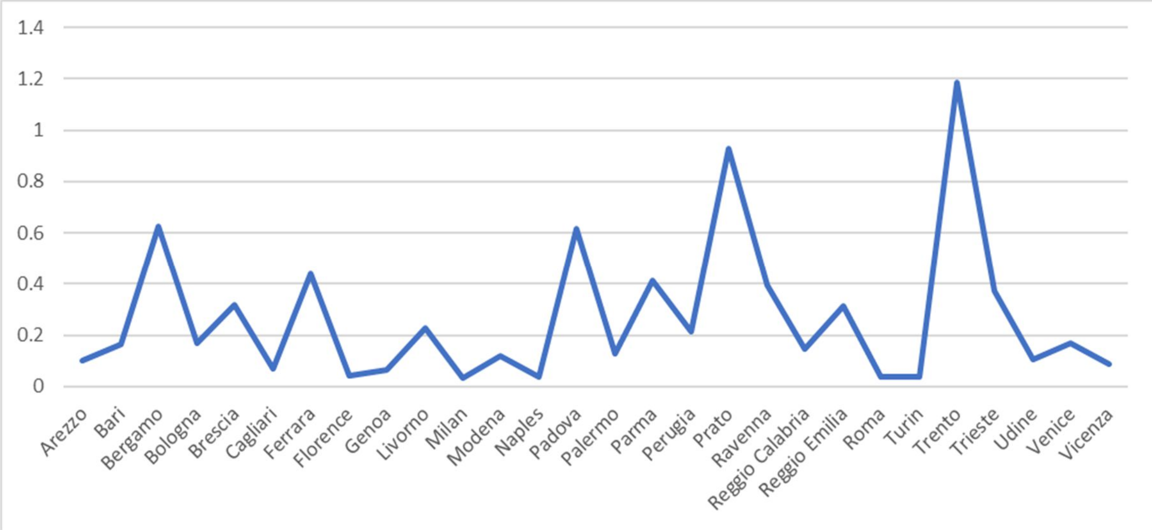
Table 6: Metrics for stakeholder engagement

	Code	Max	Average	Min	Std.deviation
Popularity	P1	0.86943824	0.42809801	0.14086629	0.188947599
	P2	11.8008553	2.16548154	0.16705465	2.568666337
	P3	0.75714314	0.14798961	0.013935	0.165891054
Commitment	C1	0.49788214	0.11194724	0.01363485	0.111304662
	C2	0.91118785	0.17493106	0.01407826	0.232629249
	C3	0.03635865	0.00988196	0.00181481	0.007941579
Virality	V1	0.93790275	0.46104294	0.0287071	0.214761821
	V2	5.9154787	1.52004359	0.1354617	1.435520691
	V3	0.45039497	0.12046044	0.01233172	0.118683037
ENGAGEMENT	(P3+C3+V3)	1.24389676	0.27833201	0.02808153	0.292515669

On average roughly 42% of the tweets are liked as represented by P1, 11% of the tweets are replied (C1) and 46 % of the tweets are retweeted (V1) by citizens. Moreover, as can be noticed looking at P2, C2 and V2 the average number of likes and retweets per tweet are high, 2,17 (P2) and 1,52 (V2) respectively, while the average number of replies per tweet is low (0,17; C2). According to the methodology proposed by Bonsón et al. (2015), dividing P2, C2 and V2 by the number of followers and multiplying it by 1,000, we obtained the average number of likes, replies and retweets per tweet per 1,000 followers. The results show that citizens prefer to interact with municipalities through "likes" rather than answers and retweets. Finally, it is possible to see that the average citizen engagement Italian index is about 27,83%.

By applying the methodology proposed by Bonsón, et al., (2015) we noticed that smaller municipalities have higher levels of citizen engagement. The municipalities with the highest levels of citizen engagement are Trento, Prato, Bergamo, Ferrara and Padova. Perhaps, this unexpected result is due to the higher levels of closeness to the municipality that citizens from smaller cities tend to show (Figure 2).

Figure 2: Citizen engagement per municipality (ranked by population)



5.2 Statistical analysis

For the purpose of our study, we are also interested in identifying the factors associated with citizen engagement. More specifically, we wanted to understand whether the audience (number of followers) and the activity (number of tweets) of municipalities might have an impact on the engagement, and if the municipality size was somehow linked with it. The great numbers showed by certain municipalities regarding audience and activity might bias the idea of citizen engagement as it is, indeed, smaller municipalities and more generally small villages, tend to care much more about their cities compared to those living in the largest ones. Engagement stands for collaborating with the local governments in enhancing old services and providing new ones for reaching a common wellness. As can be noticed from the table 7 the analysis shows a significant negative relationship among citizen engagement, activity, and audience. This implies that increasing the level of activity and audience generates a low level of engagement. Table 7 also shows that there is no relationship between the population, and therefore the size of the municipality, and the level of engagement of citizens. We apply the Spearman's coefficient for understanding the relationships among the variables. The Spearman's correlation between two variables is equal to the Pearson's correlation between the rank values of those two variables; while the Pearson's correlation gives linear relationships, the Spearman's correlation assesses monotonic relationship. In other words, the Spearman's correlation between two variables will be higher when the observations have similar rank and lower when observations have a dissimilar rank between two variables.

Table 7: Relationship between activity, population, audience, and citizen engagement.

Dependent variable	Independent variable	Spearman's coefficient	Significance
Engagement	Activity	-0.6551724**	0.0002183
	Population	-0.4499179	0.01716
	Audience	-0.7192118**	0.000002701

*** Significant at p < 0.01 (2-tailed)*

After analyzing the relation between activity, population, audience, and citizen engagement, we investigate the relationships between media types and citizens reactions on Twitter (likes, replies and retweets). Table 8 summarizes the descriptive statistics that shall light about the ways in which citizens interact with municipalities on Twitter.

Table 8: Descriptive statistics of media types and citizens engagement.

Media Type		Likes	Replies	Retweets
Web links	Mean	0.63	0.09	2.01
	Std.Deviation	3.07	0.61	4.01
Text	Mean	0.68	0.78	0.61
	Std.Deviation	2.90	0.19	3.53
Photo/ Video	Mean	4.09	0.28	4.12
	Std.Deviation	13.34	0.90	8.69
Web links - Photo/Video	Mean	3.01	0.25	2.39
	Std.Deviation	4.46	0.98	8.92
Others	Mean	3.90	0.59	2.49
	Std.Deviation	16.11	1.45	11.16
Total	**Mean**	**12.31**	**1.99**	**11.62**
	Std. Deviation	**39.88**	**4.13**	**36.31**
	N	**936,713**	**77,342**	**732,192**

Looking at Table 8, it emerges that all those tweets that contain photos or videos generate much more likes (4.09) and retweets (4.12) compared to the other media types. This outlines a common global trend, indeed due to the massive usage of smartphones and other devices, photos and videos seem to be a more direct way for interacting with the others: social networks widely used, such as Instagram and Snapchat that are based on sharing photos and videos represent two of the greatest examples of that trend. While, when dealing with plain text tweets citizens tend to reply with a higher frequency compared to other media types (0.78): this can be explained by the fact that this media type helps in starting open directly dialogues more easily with municipalities and thus increasing the engagement rate. However, the frequency appears low and thus, once again, suggests that Twitter users do not use the reply button as much as they do with the like one and the retweet one, probably due to its own feature of not being as "immediate" as likes and retweets.

Finally, we analyze how citizen engagement varies according to different content type, defined in Table 2. Table 9 presents the main results. We can see that each tweet generates on average 14.21 likes, 1.71 replies and 12.41 retweets. In Italy citizens interact with municipalities by pushing the like button while scrolling Twitter, rather than using the reply button and the retweet one, even if the average number of retweets per tweet is considerably high (12.41). From our evidence, the *Sport* category is the content that generates more likes (3.78) and retweets (2.87), *Environment* is the content that obtains more replies. Summarizing, looking at Table 9, we can see that different content types generate different levels of engagement among citizens.

Table 9: Descriptive statistics of content types and citizens engagement.

Content Type		**Likes**	**Replies**	**Retweets**
Transport and Public Works	Mean	2.03	0.24	1.23
	Std.Deviation	10.89	0.61	2.07
Others	Mean	2.01	0.32	2.22
	Std.Deviation	5.79	0.54	6.98
Cultural	Mean	2.33	0.14	2.79
	Std.Deviation	6.65	1.78	6.72
Sport	Mean	3.78	0.21	2.87
	Std.Deviation	11.48	0.43	8.42
Security and Health	Mean	0.75	0.24	0.21
	Std.Deviation	5.63	0.56	4.89
Employment and Education	Mean	1.22	0.11	1.06
	Std.Deviation	6.78	0.66	4.98
Environment	Mean	2.09	0.51	2.03
	Std.Deviation	6.32	0.71	5.95
Total	**Mean**	**14.21**	**1.71**	**12.41**
	Std. Deviation	**53.54**	**5.29**	**40.01**
	N	**1,180,445**	**97,092**	**774,69**

5.3 City profiling

In order to profile cities, and to investigate how citizens engagement varies across cities, we conducted a cluster analysis, which is oriented to separate cities according to some specific features. The cluster analysis is a set of statistical techniques designed to identify groups of units similar to one another in relation to a set of characteristics taken into consideration according to a specific condition. The objective is to combine heterogeneous units in several subsets that tend to be homogeneous and mutually exhaustive. The statistical units are, in other words, subdivided into a certain number of groups according to the level of "similarity" evaluated, starting from the values that a series of chosen variables assumes in each unit (Fabbris 1990).

Given the theoretical attention to the 3T's approach launched by Florida (2002) in identifying common development patterns across places, we decided to isolate specific features of cities considering three indicators: Tolerance (as the percentage of immigrants on the overall population), Creative class (as the percentage of people employed in the creative industries), Technology (as percentage of people employed in high-tech industries). We then performed a hierarchical cluster analysis based on Ward distance (Everitt 1979, Johnson 1967). This made it possible to visualize the data structure through a dendrogram (or tree diagram) – See Figure 3, which facilitated the authors' choice regarding the number of groups to select (3).

Figure 3. Dendrogram

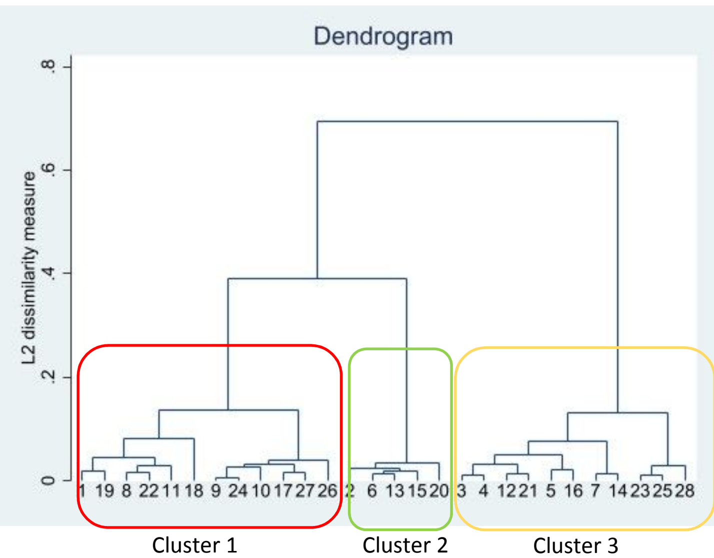


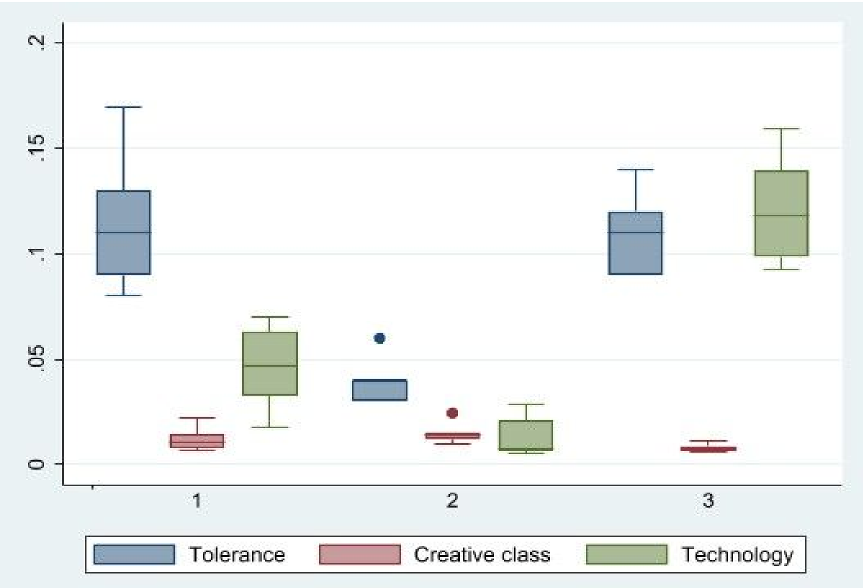
Figure 3 shows that the first cluster (in the red frame) groups 12 municipalities, the second (in the green frame) 5 municipalities, and the third (in the yellow frame) 11 municipalities.

Table 10 summarizes the values of the clustering variables (Tolerance, Creative Class, and Technology) for the three clusters, and Figure 3 represents graphically the distribution of the three indicators by cluster, through a box-plot analysis.

Table 10: Mean values of the clustering variables by cluster

	Mean	**Std. err.**	**[95% conf. interval]**	
Tolerance				
Cluster 1	0.1125	0.007797	0.096502	0.128498
Cluster 2	0.04	0.005477	0.028762	0.051238
Cluster 3	0.11	0.005222	0.099285	0.120715
Creative Class				
Cluster 1	0.011872	0.001373	0.009055	0.014689
Cluster 2	0.014964	0.002541	0.00975	0.020179
Cluster 3	0.007519	0.000433	0.006631	0.008407
Technology				
Cluster 1	0.046643	0.004769	0.036859	0.056427
Cluster 2	0.013757	0.004618	0.004282	0.023232
Cluster 3	0.12251	0.007244	0.107647	0.137372

Figure 4: Box-plot of the clustering variables by cluster



The box-plot analysis reveals interesting features of each cluster. Cluster 1 is formed by municipalities where Tolerance has a dominant role, creating the pre-conditions of a city governance oriented to cope with diversity as a potential for future development. We could label these municipalities as “Melting pot cities”. Cluster 2 is formed by municipalities that show low levels of all the three indicators, we could label these municipalities as “Lagging cities” in terms of potential for future development. Cluster 3 is formed of municipalities that show high levels of both Tolerance and Technology indicators, we could label these municipalities as “Change makers”, because of the great potential in terms of use of the technologies and ability to cope with diversity issues.

We are now able to measure the mean values of citizen engagement in each cluster, which is summarized in Table 11.

Table 21: Mean value of citizen engagement by cluster

Number of obs = 28

	Mean	**Std. Err.**	**[95% conf.**	**interval]**
Melting pot cities	0.29175	0.108674	0.068769	0.514731
Lagging cities	0.1084	0.024264	0.058614	0.158186
Change makers	0.319182	0.060621	0.194798	0.443566

Not surprisingly, Table 11 reveals that the “change makers” cities present the highest level of citizen engagement, followed by the “melting pot” and the “lagging”. The cultural ground of the cities impacts the willingness of citizens to be active participants in the city planning, thus offering interesting insights on the possibility to activate in these municipalities bottom-up governance platforms, which could help municipalities to become inclusive and intelligent cities.

6. Discussion and Conclusions

Our research has attempted to deepen the present understanding of how Big Data technologies can be applied to smart cities development, through the usage of social media. We provided an overview on how governments and local municipalities are implementing such technologies for enabling citizens to actively participate in a two-way relationship with city government. E-government, through Internet of Things technologies, represents a unique opportunity for helping citizens to improve societies’ engagement with governments. We then tried to explain how Big Data and data mining process can be used to improve and shape citizens engagement in a sample of 28 Italian municipalities. To perform our studies, we executed a content analysis on 316,801 tweets retrieved from the official Twitter accounts of our sample of Italian municipalities. The content analysis allowed us to move beyond classical statistical analyses on social media, mapping the reaction of citizens to a variety of thematic areas proposed by the municipalities where they live. We built on Bonsón, et al., (2015, 2019), which performed studies on citizens engagement on Western European municipalities (through Facebook) and Andalusian municipalities (through Twitter) and extended their results by offering a city profiling analysis. By doing so, we were not only able to retrieve information on types and intensity of engagement, but also to observe how city features relate to citizens' engagement.

The results show an heterogenous situation of the local governments in Italian municipalities, regarding audience (total followers) and activity (total tweets). This evidence supports the idea that it is not possible to apply a one-size-fits-all approach to smart city development. Regarding citizen engagement, likes are the most frequent way for Italian citizens to interact with their local city council account, on the other hand the percentage on replies per tweet appears low, outlining how nowadays citizens tend not to spend much time in giving feedbacks to municipalities on different contents and topics. Our findings show that the media type category that generates more engagement (in terms of likes and retweets) is Photos/Videos. Transport and Public Works and Cultural are the content categories most frequently re-tweeted, whereas the content type category with the highest levels of engagement is Sports.

Understanding which media types generate more engagement could be a useful tool in order to stimulate citizens' participation in city government. For instance, if a municipality is willing to receive feedbacks about certain topics, issuing a plain text tweet could generate more replies, on average, than issuing tweets supported by other media types. On the other hand, if a municipality is willing to reach many citizens, it should issue tweets including photo, video, and web links. We therefore observe that there is a tradeoff between reach and richness. Contrary to Bonsón, et al., (2019) in Italy citizens interact with municipalities by pushing the like button while scrolling Twitter, rather than using the reply button and the retweet one. Thus, there are some country specificities that are needed to be considered for maximizing the efforts of municipalities in interacting with citizens.

Finally, looking at the results of the city profiling process, it emerges that it is important to detect the cultural ground of the cities, following the Florida's approach, which suggests that the development of places is strictly connected to some not economic indicators such as Tolerance, Talent (Creative Class plus Skilled Individuals), and Technology. In our work we decided to monitor three specific aspects of the human capital available in the city, which were the percentage of immigrants

(Tolerance), the percentage of creative employees (Creative Class[6]), and the percentage of high-tech employees (Technology). We were able to classify cities in three groups: Melting pot, Lagging, and Change makers. The three groups present different levels of citizens engagement, where the municipalities with a high level of Tolerance and Technology are the ones where citizens are more active. Consequently, city planning should take in account the cultural tissue that characterizes citizens, being aware of specific features that enhance the possibility to well introduce mechanisms of bottom-up city governance. The success of smart cities initiatives is linked not only to the ability of policy makers, but also to the city profile, in terms of its citizens, which can be associated to a variety of social ad relational capital.

6.2 Limitation and further research

A series of limitations together with recommendations for future research must be acknowledged before concluding. The first limitation regards the size of the sample: 28 Twitter accounts cannot guarantee the generalization of the results; future research could be oriented to extend the analysis to a larger sample, if not including all the Italian provinces. Moreover, the study could be replicated using data coming from other social media networks, such as Facebook and LinkedIn, even if gathering data from them could be difficult, because of their privacy protection tools. Furthermore, future research could benefit from the usage of an automated content analysis, classifying tweets through a dictionary of the most frequent words, thus reducing subjectivity issues. Another relevant limitation to be acknowledged is the way in which citizen engagement was assessed. We used likes, replies and retweets, without accounting for the sentiment, which takes in consideration positive or negative impressions related to a specific content. Nevertheless, we were not willing to capture the feelings of citizens, but their level of activity in general. A final note is due to mention the ethic drawbacks of this type of analysis, in case it is used by policy makers and city governments, because of the risk of abusing of big data on individual preferences, incurring in what Bauman and Lyon (2015) refers to as the "surveillance era".

[6] We decided to measure only the Creative Class, because we observe high correlation between Skilled Individuals and Technology in our sample.

References

Al Nuaimi, E., Al Neyadi, H., Mohamed, N., & Al-Jaroodi, J. (2015). Applications of big data to smart cities. *Journal of Internet Services and Applications*, *6*(1), 25.

Allam, Z., & Dhunny, Z. A. (2019). On big data, artificial intelligence and smart cities. *Cities*, *89*, 80–91.

Anttiroiko, A.-V. (2003). Building Strong E-Democracy-The Role of Technology in Developing Democracy for the Information Age. *Communications of the ACM*, *46*(9), 121–128.

Asadi, M., & Agah, A. (2017). Characterizing User Influence Within Twitter. *International Conference on P2P, Parallel, Grid, Cloud and Internet Computing*, 122–132. Springer.

Balestrini, M., Rogers, Y., Hassan, C., Creus, J., King, M., & Marshall, P. (2017). A City in Common: A Framework to Orchestrate Large-scale Citizen Engagement around Urban Issues. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2282–2294. https://doi.org/10.1145/3025453

Batty, M. (2013). Big data, smart cities and city planning. *Dialogues in Human Geography*, *3*(3), 274–279. https://doi.org/10.1177/2043820613513390

Bauman, Z., & Lyon, D. (2015). *Sesto potere: la sorveglianza nella modernità liquida*. Bari: Gius. Laterza & Figli Spa.

Belanche, D., Casaló, L. V, & Orús, C. (2016). City attachment and use of urban services: Benefits for smart cities. *Cities*, *50*, 75–81.

Bertot, John C., Jaeger, P. T., & Grimes, J. M. (2010). Using ICTs to create a culture of transparency: E-government and social media as openness and anti-corruption tools for societies. *Government Information Quarterly*, *27*(3), 264–271. https://doi.org/10.1016/j.giq.2010.03.001

Bertot, John Carlo, Jaeger, P. T., & Hansen, D. (2012). The impact of polices on government social media usage: Issues, challenges, and recommendations. *Government Information Quarterly*, *29*(1), 30–40. https://doi.org/10.1016/j.giq.2011.04.004

Bonsón, E., Perea, D., & Bednárová, M. (2019). Twitter as a tool for citizen engagement: An empirical study of the Andalusian municipalities. *Government Information Quarterly*, *36*(3), 480–489.

Bonsón, E., Royo, S., & Ratkai, M. (2015). Citizens' engagement on local governments' Facebook sites. An empirical analysis: The impact of different media and content types in Western Europe. *Government Information Quarterly*, *32*(1), 52–62.

Bryer, T. A. (2013). Designing social media strategies for effective citizen engagement: a case example and model. *National Civic Review*, *102*(1), 43–51. Retrieved from https://go.gale.com/ps/i.do?p=AONE&sw=w&issn=00279013&v=2.1&it=r&id=GALE%7CA329365390&sid=googleScholar&linkaccess=fulltext

Caragliu, A., Del Bo, C., & Nijkamp, P. (2011). Smart Cities in Europe. *Journal of Urban Technology*, *18*(2), 65–82. https://doi.org/10.1080/10630732.2011.601117

Catlett, C., Cesario, E., Talia, D., & Vinci, A. (2019). Spatio-temporal crime predictions in smart cities: A data-driven approach and experiments. *Pervasive and Mobile Computing*, *53*, 62–74. https://doi.org/10.1016/j.pmcj.2019.01.003

Chadwick, A. (2003). Bringing E-Democracy Back In. *Social Science Computer Review*, *21*(4), 443–455. https://doi.org/10.1177/0894439303256372

Chadwick, A. (2008). Web 2.0: New Challenges for the Study of E-Democracy in an Era of Informational Exuberance. *I/S: A Journal of Law and Policy for the Information Society*, *5*. Retrieved from https://heinonline.org/HOL/Page?handle=hein.journals/isjlpsoc5&id=19&div=&collection=

Cheela, V. R. S., Ranjan, V. P., Goel, S., John, M., & Dubey, B. (2021). Pathways to sustainable waste management in Indian Smart Cities. *Journal of Urban Management*, *10*(4), 419–429. https://doi.org/10.1016/j.jum.2021.05.002

Coe, A., Paquet, G., & Roy, J. (2001). E-Governance and Smart Communities. *Social Science Computer*

Review, *19*(1), 80–93. https://doi.org/10.1177/089443930101900107

Coleman, S., & Norris, D. F. (2005). A New Agenda for e-Democracy. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1325255

De Lange, M., & De Waal, M. (2017). Owning the City: New Media and Citizen Engagement in Urban Design. In *Urban Land Use* (pp. 109–130). https://doi.org/10.1201/9781315365794-13

Del Vecchio, P., Mele, G., Ndou, V., & Secundo, G. (2018). Creating value from social big data: Implications for smart tourism destinations. *Information Processing & Management*, *54*(5), 847–860.

Engin, Z., van Dijk, J., Lan, T., Longley, P. A., Treleaven, P., Batty, M., & Penn, A. (2020). Data-driven urban management: Mapping the landscape. *Journal of Urban Management*, *9*(2), 140–150. https://doi.org/10.1016/j.jum.2019.12.001

Everitt, B. S. (1979). Unresolved problems in cluster analysis. *Biometrics*, 169–181.

Fabbris, L. (1990). *Analisi esplorativa di dati multidimensionali, cleup editore*. Padova.

Falco, E., & Kleinhans, R. (2018). Beyond technology: Identifying local government challenges for using digital platforms for citizen engagement. *International Journal of Information Management*, *40*, 17–20. https://doi.org/10.1016/j.ijinfomgt.2018.01.007

Florida, R. (2002). *The rise of the creative class* (Vol. 9). Basic books New York.

Florida, R. (2017). *The new urban crisis: How our cities are increasing inequality, deepening segregation, and failing the middle class-and what we can do about it*. Basic Books.

Franke, J., & Gailhofer, P. (2021). Data Governance and Regulation for Sustainable Smart Cities. *Frontiers in Sustainable Cities*, *3*. https://doi.org/10.3389/frsc.2021.763788

Hancke, G., Silva, B., & Hancke, Jr., G. (2012). The Role of Advanced Sensing in Smart Cities. *Sensors*, *13*(1), 393–425. https://doi.org/10.3390/s130100393

Hollands, R. G. (2008). Will the real smart city please stand up? *City*, *12*(3), 303–320. https://doi.org/10.1080/13604810802479126

Jaeger, P. T., & Bertot, J. C. (2010). Transparency and technological change: Ensuring equal and sustained public access to government information. *Government Information Quarterly*, *27*(4), 371–376. https://doi.org/10.1016/j.giq.2010.05.003

Johnson, S. C. (1967). Hierarchical clustering schemes. *Psychometrika*, *32*(3), 241–254.

Karakiza, M. (2015). The Impact of Social Media in the Public Sector. *Procedia - Social and Behavioral Sciences*, *175*, 384–392. https://doi.org/10.1016/j.sbspro.2015.01.1214

Kitchin, R. (2014). The real-time city? Big data and smart urbanism. *GeoJournal*, *79*(1), 1–14.

Kitchin, R. (2015). Making sense of smart cities: addressing present shortcomings. *Cambridge Journal of Regions, Economy and Society*, *8*(1), 131–136.

Kneuer, M. (2016). E-democracy: A new challenge for measuring democracy. *International Political Science Review*, *37*(5), 666–678. https://doi.org/10.1177/0192512116657677

Kummitha, R. K. R., & Crutzen, N. (2017). How do we understand smart cities? An evolutionary perspective. *Cities*, *67*, 43–52.

Lee, N. (2019). Inclusive Growth in cities: a sympathetic critique. *Regional Studies*, *53*(3), 424–434. https://doi.org/10.1080/00343404.2018.1476753

Lim, C., Kim, K.-J., & Maglio, P. P. (2018). Smart cities with big data: Reference models, challenges, and considerations. *Cities*, *82*, 86–99.

Liu, Y., Weng, X., Wan, J., Yue, X., Song, H., & Vasilakos, A. V. (2017). Exploring Data Validity in Transportation Systems for Smart Cities. *IEEE Communications Magazine*, *55*(5), 26–33. https://doi.org/10.1109/MCOM.2017.1600240

Loeffler, E., & Martin, S. (2015). Citizen engagement. In *Public Management and Governance* (pp. 301–318). https://doi.org/10.4324/9781315693279-35/CITIZEN-ENGAGEMENT-ELKE-LOEFFLER-STEVE-MARTIN

Macintosh, A. (2004). Using Information and Communication Technologies to Enhance Citizen Engagement in the Policy Process. In *Promise and Problems of E-Democracy* (Vol. 9789264019, pp. 19–142). https://doi.org/10.1787/9789264019492-3-en

Macintosh, A. (2008). *E-Democracy and E-Participation Research in Europe*. https://doi.org/10.1007/978-0-387-71611-4_5

Marino, V., & Lo Presti, L. (2018). From citizens to partners: the role of social media content in fostering citizen engagement. *Transforming Government: People, Process and Policy*, *12*(1), 39–60. https://doi.org/10.1108/TG-07-2017-0041/FULL/PDF

Martí, C., Royo, S., & Acerete, B. (2012). The effect of new legislation on the disclosure of performance indicators: The case of Spanish local governments. *International Journal of Public Administration*, *35*(13), 873–885.

Mehmood, R., Meriton, R., Graham, G., Hennelly, P., & Kumar, M. (2017). Exploring the influence of big data on city transport operations: a Markovian approach. *International Journal of Operations and Production Management*, *37*(1), 75–104. https://doi.org/10.1108/IJOPM-03-2015-0179/FULL/PDF

Miles, A., Zaslavsky, A., & Browne, C. (2018). IoT-based decision support system for monitoring and mitigating atmospheric pollution in smart cities. *Journal of Decision Systems*, *27*(sup1), 56–67. https://doi.org/10.1080/12460125.2018.1468696

Mizuki, F., Mikawa, K., & Kurisu, H. (2012). Intelligent Water System for Smart Cities. *Hitachi Review*, *61*(3), 147–151.

Nijkamp, P., & Cohen-Blankshtain, G. (2013). The Importance of ICT for Cities: e-Governance and Cyber Perceptions. In *Advances in Spatial Science* (Vol. 72, pp. 295–308). https://doi.org/10.1007/978-3-642-32141-2_13

Oecd, & Eurostat. (2005). *Oslo Manual*. Retrieved from https://ec.europa.eu/eurostat/documents/3859598/5889925/OSLO-EN.PDF

Rathore, M. M., Ahmad, A., Paul, A., & Rho, S. (2016). Urban planning and building smart cities based on the Internet of Things using Big Data analytics. *Computer Networks*, *101*, 63–80. https://doi.org/10.1016/j.comnet.2015.12.023

Samuel, M., Doctor, G., Christian, P., & Baradi, M. (2020). Drivers and barriers to e-government adoption in Indian cities. *Journal of Urban Management*, *9*(4), 408–417. https://doi.org/10.1016/j.jum.2020.05.002

Schlingensiepen, J., Nemtanu, F., Mehmood, R., & McCluskey, L. (2016). Autonomic Transport Management Systems—Enabler for Smart Cities, Personalized Medicine, Participation and Industry Grid/Industry 4.0. In *Studies in Systems, Decision and Control* (Vol. 32, pp. 3–35). https://doi.org/10.1007/978-3-319-19150-8_1

Schwartz, H. A., & Ungar, L. H. (2015). Data-driven content analysis of social media: a systematic overview of automated methods. *The ANNALS of the American Academy of Political and Social Science*, *659*(1), 78–94.

Skoric, M. M., Zhu, Q., Goh, D., & Pang, N. (2016). Social media and citizen engagement: A meta-analytic review. *New Media & Society*, *18*(9), 1817–1839. https://doi.org/10.1177/1461444815616221

Söderström, O., Paasche, T., & Klauser, F. (2014). Smart cities as corporate storytelling. *City*, *18*(3), 307–320.

Soomro, K., Bhutta, M. N. M., Khan, Z., & Tahir, M. A. (2019). Smart city big data analytics: An advanced review. *WIREs Data Mining and Knowledge Discovery*, *9*(5), e1319. https://doi.org/10.1002/widm.1319

Spirakis, G., Spiraki, C., & Nikolopoulos, K. (2010). The impact of electronic government on democracy: e-democracy through e-participation. *Electronic Government, an International Journal*, *7*(1), 75. https://doi.org/10.1504/EG.2010.029892

Toma, Alexandru, Popa, & Zamfiroiu. (2019). IoT Solution for Smart Cities' Pollution Monitoring and the Security Challenges. *Sensors*, *19*(15), 3401. https://doi.org/10.3390/s19153401

Torres, L., & Pina, V. (2001). Publicprivate partnership and private finance initiatives in the EU and Spanish local governments. *European Accounting Review*, *10*(3), 601–619.

Tranos, E., & Mack, E. (2019). Big data: A new opportunity for transport geography? *Journal of Transport*

Geography, *76*, 232–234. https://doi.org/10.1016/j.jtrangeo.2018.08.003
van den Hoven, J. (2005). E-democracy, E-Contestation and the Monitorial Citizen*. *Ethics and Information Technology*, *7*(2), 51–59. https://doi.org/10.1007/s10676-005-4581-4
Van Rossum, G., & Drake Jr, F. L. (1995). *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam.
Vanolo, A. (2014). Smartmentality: The smart city as disciplinary strategy. *Urban Studies*, *51*(5), 883–898.
Verma, R. K., Kumar, S., & Ilavarasan, P. V. (2017). Government portals, social media platforms and citizen engagement in India: Some insights. *Procedia Computer Science*, *122*, 842–849. https://doi.org/10.1016/j.procs.2017.11.445
Zacharias, C., & Poldi, F. (2018). TWINT - twitter intelligence tool. Retrieved from https://github.com/twintproject/twint
Zannat, K. E., & Choudhury, C. F. (2019). Emerging Big Data Sources for Public Transport Planning: A Systematic Review on Current State of Art and Future Research Directions. *Journal of the Indian Institute of Science*, *99*(4), 601–619. https://doi.org/10.1007/S41745-019-00125-9/TABLES/3
Zheng, C., Yuan, J., Zhu, L., Zhang, Y., & Shao, Q. (2020). From digital to sustainable: A scientometric review of smart city literature between 1990 and 2019. *Journal of Cleaner Production*, 120689.
Ziemke, D., Charlton, B., Hörl, S., & Nagel, K. (2021). An efficient approach to create agent-based transport simulation scenarios based on ubiquitous Big Data and a new, aspatial activity-scheduling model. *Transportation Research Procedia*, *52*, 613–620. https://doi.org/10.1016/j.trpro.2021.01.073