Open data governance: civic hacking movement, topics and opinions in digital space

Mara Maretti[1] · Vanessa Russo[1] · Emiliano del Gobbo[2]

Abstract

The expression 'open data' relates to a system of informative and freely accessible databases that public administrations make generally available online in order to develop an informative network between institutions, enterprises and citizens. On this topic, using the semantic network analysis method, the research aims to investigate the communication structure and the governance of open data in the Twitter conversational environment. In particular, the research questions are: (1) Who are the main actors in the Italian open data infrastructure? (2) What are the main conversation topics online? (3) What are the pros and cons of the development and use (reuse) of open data in Italy? To answer these questions, we went through three research phases: (1) analysing the communication network, we found who are the main influencers; (2) once we found who were the main actors, we analysed the online content in the Twittersphere to detect the semantic areas; (3) then, through an online focus group with the main open data influencers, we explored the characteristics of Italian open data governance. Through the research, it has been shown that: (1) there is an Italian open data governance strategy; (2) the Italian civic hacker community plays an important role as an influencer; but (3) there are weaknesses in governance and in practical reuse.

Keywords Open data · Influencers · Civic hacking · Twitter · Semantic network analysis method · Online focus group

✉ Mara Maretti
ma.e.maretti@gmail.com; mara.maretti@unich.it

Vanessa Russo
russov1983@gmail.com

Emiliano del Gobbo
emidelgo@gmail.com

1   Department of Law and Social Sciences, University "G. D'Annunzio" of Chieti-Pescara, Chieti, Italy

2   Department of Neuroscience, Imaging and Clinical Sciences, University "G. D'Annunzio" of Chieti-Pescara, Chieti, Italy

🖄 Springer

1 Introduction

The expression 'open data' relates to a system of informative and freely accessible databases that public administrations make generally available online in order to develop an informative network between institutions, enterprises and citizens.

Open data (OD) can be regarded as an integral part of the open government movement (Yu and Robinson 2011; Kundra 2012; Ubaldi 2013; Gerunov 2015). This informational system derives from the need for public bodies to share with citizens the data they collect in the course of their activities. In 2009, open government data (OGD) strategies started to emerge in the political agenda of various governments: the 2009 Obama Executive Order[1] is an example. The 2010 Digital Agenda of the European Commission[2] and the 2013 G8 Open Data Charter[3] are further examples of how political institutions consider OD to be an important resource that can: contribute to a more transparent and efficient government; strengthen democracy, by encouraging participation from citizens; and contribute to economic growth, by facilitating entrepreneurship, innovation and scientific discoveries, thereby improving citizens' lives and contributing significantly to job creation (Huijboom and Van Den Broek 2011; Jetzek 2013; Jung and Park 2015; Donker and Van Loenen 2017). In addition to these social, political and economic benefits, we can also mention operational and technical benefits, such as the ability to reuse data; optimisation of administrative processes, improvement of public policies, access to external problem-solving capacity, and fair decision-making (by enabling comparison, easier access to data and discovery of data); creation of new data derived from combining data; external quality checks of data (validation); sustainability of data (no data loss); and the ability to merge, integrate and mesh public and private data (Janssen et al. 2012).

The Open Data Maturity Report 2018 (Cecconi and Radu 2018) published by the European Data Portal, defines a series of indicators selected to measure OD maturity across Europe. These indicators cover the level of development of national policies promoting OD. The majority of the 28 European countries analysed demonstrate a solid understanding of the impact of OD in paving the way for the data economy (Peña-López 2017). In particular, the report evaluates portal maturity by classifying[4] the countries into four different clusters: beginners, followers, fast trackers, and trendsetters. Italy belongs to the trendsetter group (top maturity quadrant), together with Ireland, Spain, France and Cyprus.

---

[1] https://obamawhitehouse.archives.gov/sites/default/files/microsites/ogi-progress-report-american-people.pdf (Retrieved 11 November /2017).

[2] https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52010DC0245R(01)andfrom=EN (Retrieved 11 November 2017).

[3] https://www.gov.uk/government/publications/open-data-charter (Retrieved 11 November 2017).

[4] The ranking is built on four dimensions. (1) The open data policy is a dimension focusing on the presence at the national level of specific policies on open data, licencing norms, and the extent of coordination at national level to provide guidelines and guidance to national, local and regional administrations, and set up coordinated approaches towards data publication. (2) The open data portal looks at the development and level of sophistication of national portals to feature available open data, and comprises considerations around functionality, usage (user analytics), variety of data available on the portal, and the approach to ensuring the portal's sustainability. (3) Open data impacts analyses existing methodologies and strategic awareness at the national level towards capturing and measuring impact, and zooms in on the impact of open data at country level on four dimensions: political, social, environmental and economic. (4) Open data quality explores the extent to which national portals have a systematic and automated approach to harvesting, the currency and reliability of featured data, and the compliance level in terms of the metadata standard DCAT-AP (Cecconi and Radu 2018, p. 8).

In this framework, the paper aims to analyse the characteristics of "Italian OD governance structure" in term of actors and existing relations between them, and the "OD governance" as a process. For what concern the analysis of the OD governance structure the research aims are: to determine who are the main actors in the Italian OD infrastructure; understand the relations between the social actors involved in the OD network; and analyse the main conversation topics online on Twitter.

In a second stage of the research, through a participatory research approach, we try to better understand the features of the OD Governance. This is done in term of procedures, roles, and finally the pros and cons of use/reuse of OD in Italy through the opinions of key actors.

This article is organised as follows. Firstly, we outline the theoretical foundations of open government and the civic hacking culture; secondly, we illustrate the research design and the methodology used to define the main actors in the Italian OD system, to select the expert hackers, and to realise the online focus group; and thirdly, we present the results of the research concluding on the pros and cons of the OD system in Italy.

1.1 Open data and civic hacking: background and innovation

Castells ([2001](#)) describes the internet culture as a culture consisting of a technocratic confidence in progress made by the human race by the means of technology. This is applied by hacker communities, who thrive on open and freely accessible technological creativity deriving from virtual networks that aims at reinventing society and manifesting itself in new economy mechanisms created by profit-motivated entrepreneurs.

The development of web spaces and the cyberculture (Turkle [1997](#); Castells [2001](#)), and the power of big data and of web-mediated communication, constitute the new frontier of network communication (Coyle and Vaughn [2008](#)).

Another interesting interpretation, by Pierre Lévy ([1994](#), [2002](#)), sees the development of ICT (Information and Communications Technology) as a determiner in the emergence of a new public sphere which has, in turn, given rise to new demands for interactivity.

According to Levy's paradigm (2008), the following three fundamental principles characterise current users of the web:

1. Universality—the need to navigate within systems that are functional and easily accessible to all.
2. Transparency—the desire for all sources of information to be totally accessible.
3. Inclusivity—the desire to be treated as a participant by the administration and to have unrestricted access to sources of information, which, in turn, should conform to the principles of universality and transparency.

The digitalised public sphere, as defined by Levy, does in fact make it necessary to construct an interactive relationship with institutions in the form of a network.

In this context, public administration needs to carry out an extensive process of innovation in order to interact with, and meet the needs of, a new social, technical and economic dimension, in order to bring about improvements to institutional activities (Scavo and Shi [2000](#); Gascò [2003](#)).

Castells' treatment of the informational welfare paradigm describes innovation in terms of the connection between the welfare state and network society (Castells and Himanen [2002](#)). The process of informational innovation for the welfare state involves a virtuous

circle made up of a dynamic network in which the three dimensions of the network (the technological, the social and the economic) are incorporated into the traditional institutional system.

Additionally, the innovation process for the welfare state involves the application of information technology to the purposes of social wellbeing and a structural renewal of the welfare state by means of a more dynamic organisation of the network. An innovation of this kind can help to give rise to greater productivity in public services, leading to a reduction in their cost and therefore relieving the fiscal pressures on the welfare state.

Castells' (2001) treatment of this paradigm can therefore become a practical proposition only if a network is constructed within which the academic system, the business system, and the institutional system are all interconnected. It is therefore important that public measures should exist to add value to the business world, both by an increase in economic productivity and by creating a system of public and private interdependence.

In this framework we can place the development of open governance systems and civic data hacking. Indeed, civic hacking can be defined as a form of data activism, requesting, digesting, contributing to, modelling and contesting data (Schrock 2015).

1.2 The Italian OD infrastructure

In Italy, the development of OD is regulated by Article 9 of Decreto Crescita 2.0 (Law 221/2012), which lays down the following specifications.

a. Free access data format—in other words, a data format that is publicly available with full documentation and is neutral where the technology necessary for use of the data themselves is concerned.
b. Open access data. Data are to have the following characteristics: (1) They are to be available in accordance with the terms of an authorisation that permits their use, even for commercial purposes, in an unbundled format and by anyone. (2) They are to be accessible via information and communication technology, including public and private telematic networks, in free access formats as defined above. (3) They are to be suitable for automatic use by computer programs and equipped with the relevant metadata. (4) They are to be made available free of charge via information and communication technology, including public and private telematic networks, or else at a nominal charge to cover their reproduction and circulation.

The text of Article 9 also gives great importance to the key role of OD as a means of ensuring transparency of administration, stating that activities designed to enable telematic access and the reutilisation of data derived from public administration departments fall within the parameters for evaluation of management performance.

The key words 'evaluation' and 'transparency' acquire a double meaning when used in the context of OD. On one (ontological) level, free access data originate as a means of ensuring an institution's transparency and of offering to stakeholders an opportunity to carry out an evaluation of its administrative performance. On another (methodological) level, OD become a means of evaluating the transparency of the public administration: the typology of the data that have been made available, and the manner in which they are made available, offer the user an opportunity to assess the level of transparency achieved by the public administration.

To this purpose, AgID[5] has published a vade mecum or guidebook which sets out guidelines whereby data quality can be assessed by determining how coherent, comprehensive and up to date they are, in order to promote the predominance of increasing availability of data; the acquisition of data of unknown quality; the processing of frequently inadequate information; a focus on the quality requirements of data; a reduction in the dispersal of data among various owners and users; an increase in the amount of reusable data (by eliminating semantic ambiguity); the co-existence of legacy systems with open systems; a reduction in duplication and a commitment of resources; improvements to processes that cause errors; an estimate of the cost of a low quality product; and a gradual elimination of paper-based data capture models.

In their article 'Semantic Web', Berners-Lee et al. (2001) classify the quality of OD according to reference parameters on a value scale from one to five stars[6]: (a) one star relates to OD that are freely available on the web; (b) two stars relate to structured data that are legible using a software program such as Excel; (c) three stars relate to OD structured to be accessible without the use of proprietary software such as CSV format; (d) four stars relate to data which, besides having all the characteristics given above, are distributed in RDF format, SPARQL format, or other formats defined by W3C[7]; and (e) five stars relate to linked data (Berners-Lee 1999, 2009), or data which, besides containing all the characteristics given above, are interlinked.

The authorisations that enable these data to be accessed also play a significant role in this context. In Italy, the authorisations granted (IODL) are regulated by FormezPA and are classified as IODL 1.0 (October 2011) and IODL 2.0 (March 2012). These two classes of authorisation have the same validity and give the same rights—the right to create derived work and to use their validity for commercial purposes and compatible authorisations—but they differ in the stipulations they impose for the provision of data: IODL 1.0, unlike IODL 2.0, obliges the authorisation holder to publish and to share derived work with holders of the same authorisation or of a compatible authorisation.

Legislative Decree No. 82 (7 March 2005)[8] set the Italian public administration on the road to digitalisation. The text of this decree set out both the right of access and the means of access for businesses and the public to services of public utility via the web.

In practice, though, the process of informational innovation for Italian institutions is as detailed in Decreto Crescita (Growth Decree) 2.0,[9] a document which sets out the evolution of Italian digital administration in the following terms: (1) the creation of an Agency for Digital Italy; (2) the establishment of a digital agenda; (3) the introduction of new electronic media capable of ensuring certified communication between one government department and another, and between those departments and specific users; (4) regulation of OD; (5) digitalisation of traditionally paper-based documents used for institutional communication (medical prescriptions, sick notes) and of academic textbooks; (6) abolition of the digital divide; (7) promotion of electronic currency in order to facilitate payments, while at the same time ensuring security and enabling traceability of economic transaction[10]; (8)

---

[5] Agency for a Digital Italy.

[6] https://www.w3.org/DesignIssues/LinkedData.html (Retrieved 10 January 2018).

[7] W3C refers to the World Wide Web Consortium, a non-governmental organisation which sets out to define communication standards and protocols in order to maximise potential use of the web.

[8] Digital administration code.

[9] D.lgs.179/2012, enacted into Law 221 on 17 December 2012.

[10] Printed statement from Ministry for Economic Development, 27 January 2014.

encouragement of strategic research and innovation projects; (9) creation of a technical committee of smart communities; and (10) specification of the formal requirements for the creation of start-ups (innovative businesses).

1.3 Digital citizenship and civic hacking

Examination of various points of this Decree makes it clear that it was D.lgs.82/2005 that laid the juridical foundations for the establishment of a digital public administration capable of interconnecting a variety of social and institutional actors into a single informational network, complete with up-to-date technological support.

It should be emphasised that this document aimed to create a virtuous circle by establishing a connection between public and cultural institutions on the web, and is therefore not to be regarded as merely an updated version of the earlier digital administration code.

Growth Decree 2.0 can be described, in a nutshell, as the ideal type of informational welfare defined by Castells and Himanen (2002), since it recognises that it is essential for the various systems—economic, business, institutional and research—to be interconnected. Another advantage of the Decree is that it takes into account the skills of the participants: the connection between big data, OD and the free circulation of knowledge via the internet gives rise to forms of hackerism and provides a basis for a redefinition of the whole concept of active citizenship.

The FormezPA guidebook *Misurazione della qualità dei dati* (How to Assess Data Quality) states that the preliminary for successful assessment of the total quality of data is the creation of an interactive system in which there is constant interaction between the institutions and the general public. In the words of the guidebook the citizen, who is conversant with the real situations affecting him has an active part to play in a cooperative model designed to improve the quality of data. It occurs more and more frequently that an institution requests co-operation from citizens in order to verify complex information designed to achieve intersectoral global quality.[11]

The ideal type of digital citizen in this context is therefore an individual who is capable of being simultaneously both a user of public administration services and a potential source of informational redesign of welfare resources.

It must be emphasised that the 'bottom-up' response relates not merely to the straightforward evaluation of a service but to a collective intelligence (Lévy 1994; 2002) which manifests itself in the shape of self-organised collectives making use of the latest technology to create parastatal and para-administrative forms of civic activism.

Castells (2001) describes this process, seen as the creation of a 'bottom-up' power source capable of cooperating with the institutional system, as 'civic hackerism'. The creation of the informational welfare paradigm and of its associated processes is achieved as a result of the combination of public activity, business activity, and what he terms the 'hacker ethic'.

The concept of the hacker ethic relates to a situation where users (i.e. citizens) are actively involved in the development of informational welfare, thereby becoming simultaneously users of the services provided and creators of feedback designed to improve the system.

---

[11] https://www.agid.gov.it/sites/default/files/documenti_indirizzo/iso_25024_agid_misurazione_della_quali ta_dei_dati.pdf (Retrieved 14 December 2017).

The term 'hacker' is widely applied to IT experts who can create and manage software and can make use of sources to devise complex computer systems. They are described by Castells (2001) and Himanen (2001) as fundamental elements in the establishment of web cultures, since the hacker culture is evidently a value-based culture possessing a strong community spirit which manifests itself in the creation of informal institutions.

The hacker culture is really a collective of subcultures, independent of the web, retaining an awareness of their origins, their values and significant shared experiences. It has its own legends, heroes, villains, popular rhymes, games, taboos and dreams. For this reason, hackers, considered as a group, are extremely creative people who are distinguished, in part, by their rejection of 'normal' values and working habits: their culture, although less than fifty years old, possesses unusually rich and well-informed traditions.[12]

Hacking activities extend into various IT and social areas. In the context of digital citizenship, the birth of the open government philosophy gave rise to the civic hacking movement.

Civic hackers are groups of IT experts who make use of OD from public administrations in order to call to account and improve the services offered by those administrations. Civic hackerism can therefore be seen as aiming to create a 'hinge' between the general public and the relevant institutions. Civic hacking can broadly be described as a form of alternative/activist media that employ or modify the communication artifacts, practices, and social arrangements of new information and communication technologies to challenge or alter dominant, expected, or accepted ways of doing society, culture, and politics (Shrock 2015, p. 2).

Ben Campbell describes the activities of civic hackers as "citizens developing their own applications which give people simple, tangible benefits in the civic and community aspects of their lives".[13]

We can find several case studies in recent literature dealing with the relationship between open government strategies and civic hacking in the US (Frecks 2014; Carr and Lassiter 2017; Choi and Tausczik 2017), in UK urban data initiatives (Thakuriah et al. 2016), in the innovation experiences of the implementation of technological strategies for the solution of social problems in Mexico (Tena-Espinoza-De-Los-Monteros 2016), in the evaluation of potentials and limits of crowdsourcing software as tools for civic participation in Russia and France (Ermoshina 2016), and others. All these studies recognise the potential role of civic hacking groups in increasing participation and at the same time creating strategies of OD reuse.

1.4 The research: stages, materials and methods

The theoretical panorama outlined in the previous sections delineates a reticular and multi-structural system of data and social actors. The advantage of such a network is that it creates a virtuous circle of information that can mainly develop into a reinforcement

---

[12] The hackers Eric S. Raymond and Raphael Finkel, anxious to outline the principles that regulate the hacker system, have written The Jargon File, which sets out all the principles that regulate the hacker system, including the code of conduct. The Jargon File can be considered for all practical purposes as the hackers' code of ethics. The Jargon File can be accessed at https://catb.org/jargon/html/ (Retrieved 30 January 2018).

[13] The complete document can be accessed at https://scumways.com/civichackingtalk/ (Retrieved 4 February 2018).

and transformation of institutional structures, and the stimulation of social and economic growth, and, potentially, increased participation.

Given that the essential aim of Growth Decree 2.0 is to create an interactive network system in which public administration, the general public and the economic system are linked together, the research questions are: (a) Does this network exist? Who are the main actors and what are the interactions among them? (b) What are the main conversation topics online? (c) What are the real opportunities and what are the shortcomings relative to the use of OD by the social actors involved?

In order to answer to these questions, the research aims: (1) to map the OD network, identifying the key actors of the Italian OD system; (2) to map thematic topics emerging form the Twitter content analysis; and (3) to organise an online focus group with the OD opinion leaders to explore issues surrounding the pros and cons of Italian ODG and OD reuse.

To achieve these objectives, the research is structured in three stages.

In the first stage, we analyse the civic hacker community active on Google groups and on Twitter. To detect the main opinion leader in the field of OD, it was useful to develop an index of potential opinion leaders (POL) on Twitter.

In the second step, we sought to identify the main topics of OD matter on Twitter using text mining and social network analysis (SNA) (Jung and Park 2015; Smith 2015; Jalali and Park 2017; Wasserman and Faust, 1994).

There is a large scientific literature of applications of SNA to Twitter's environment. The social media structure lends itself to a network analysis method because it is composed of three-dimensional relationships: (1) asymmetry between users (friends and followers); (2) asymmetry between users and concepts (@mention); and (3) symmetry between concepts and concepts (#hashtag). These three dimensions are related to the entire Twitter-sphere (Dijck, 2011).

Finally, the contents of Twitter are very short (each Tweet a maximum of 280 characters) and can be analysed easily using SNA, a collection of research techniques that treat a word as a node in a network and a semantic relationship (in terms of the concurrence) between words as a social relationship connecting those words (Chung and Park 2010; Jung and Park 2015).

Through SNA, it is possible point out the semantic network of environmental OD in the Twittersphere. Moreover, using a modularity algorithm, it is possible to detect some semantic subsets. Segmentation of the dataset points out some cliques that map the most important levels of OD discussion in the Twitter space.

Therefore, in the third stage of the research, considering the results of the first and second steps, we promote an online focus group with the key OD actors. The aim of the focus group is to synthesise, with a participatory approach, the pros and cons of the Italian OD system.

2 Results

2.1 Italian OD governance: civic hacking groups and Italian opinion leaders on the OD system

In the first step of the research, the analysis of digital trails within the Google groups brought to light 670 discussion chats on the OD issue, and 50 Italian groups participating

in the conversations. The distinguishing feature of these networks is that they are connected to local reuse projects.

Of these fifteen communities (the Italian groups cited in the preceding paragraph), there are three (SpaghettiOpenData, OpenPuglia.org and Opendatasicilia) which stand out for their recency and their large volume of posts.

SpaghettiOpenData is a community of civic hackers who use this forum to participate in, and comment on, the subject of OD. This is an extensive community that has been active since 2010 and currently numbers 1,300 members.

SpaghettiOpenData is the most significant of all the networks examined, not only for its quantity of interactions and the frequency of its interventions, but also because it can therefore be seen as occupying a central position, and functioning as a hub, in the network of the groups.

OpenPuglia.org is a network comprising 123 civic activist users; the purpose of the group is to promote cooperative projects for the reuse of OD within new informational portals.

Opendatasicilia is a network comprising 116 users whose purpose is to publicise and propagate open government culture and Open Data procedure in our region, and to open a participative public discussion in our cities. Interestingly, not only do members of this group also belong to SpaghettiOpenData, but those who belong to both groups are the most active members in either group.

Analysis of the Google groups has revealed the existence of virtual communities wholly composed of civic activists in free cooperation with institutions on the basis of a new direction for civic information. The existence of these bottom-up networks is a phenomenon directly attributable to the informational welfare paradigm (Castells and Himanen 2002), where the hacker ethic plays a crucial role.

The selection and identification of OD opinion leaders in Italy was carried out by analysing the digital footprints of Twitter social media users in relation to the OD theme.

The Twitter data mining process was executed through the Social Grabber[14] online platform, through which it was possible to download data in streaming API mode and in relation to specific keywords. The download phase made it possible to set up a dataset consisting of 1,905 tweets relating to the hashtags #opendata, #trasparenza (transparency), #partecipazione (participation) and #opengov in the period from 14 January 2019 to 21 January 2019. Subsequently, the dataset was processed using tools programmed in Python, with which it was possible to extract two datasets related to: (1) a spreadsheet containing all the tweets, users, the number of retweets and retweets with text; and (2) a dataset containing a list of all the users who produced content, the number of posts produced in the observed time, the number of followers and friends, the number of user retweets, and, finally, the number of mentions.

To identify influential users, a potential opinion leader POL index was developed. The index is composed of the following indicators (Fig. 1): the number of followers, the number of posts, the number of mentions and the number of retweets of the content.

The index is obtained by normalising the indicators, rescaling them between 0 and 1. The standardisation process was realised with the min–max elementary indicators, $Z = (x - \min(x))/(\max(x) - \min(x))$, aggregated with the arithmetic mean.

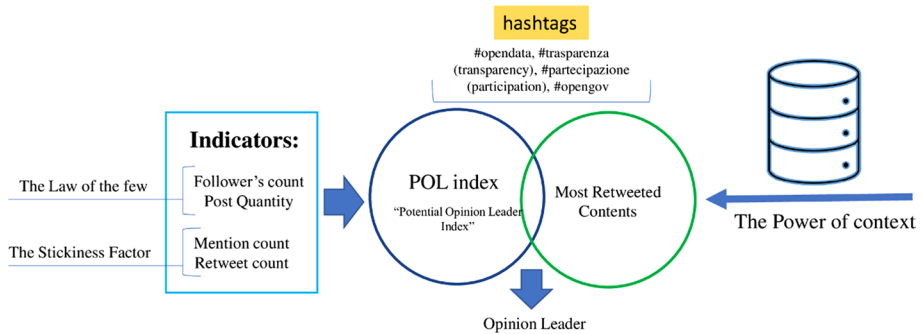---

[14] www.emilianodelgobbo.com/socialgrabber.

Fig. 1 The process of identification of opinion leaders within the six hashtag topics

Table 1 The most significant users

User	POL index	Value of most retweeted content
OL1	35,426,692	288
OL2	21,488,046	77
OL3	2,004,909	45
0L4	17,924,044	41
OL5	16,502,673	33
OL6	16,260,209	23
OL7	15,527,291	23

The choice of indicators is based on the assumption of The Million Follower Fallacy (Cha et al. 2010): the number of followers of a user alone does not represent an element able to highlight an influencer, since certain users have many followers only by virtue of their social role and not in relation to actual influence. Several empirical researches elaborate indices considering followers, retweets and mentions (Leavitt et al. 2009; Marchetti, cited in Bentivegna 2014; Choi 2015).

Here, according to the paradigms of Katz and Lazarsfeld (1955) and Gladwell (2006), the chosen indicators, useful to describe the degree of popularity of the user, are: (1) his media production (follower count, post quantity), (2) his position in the Twitter network (mention count, retweet count).

As we can see in the Fig. 1, the obtained index allowed us to find opinion leaders on this issue.

Then, we considered what content was actually quoted by the public, obtaining a higher retweet value (the power of context). From the results, a sample of users was selected from both lists.

The analysis carried out a sample of users representing the opinion leaders on some specific hashtag topics: #opendata, #trasparenza (transparency), #partecipazione (participation) and #opengov (Table 1). Six of those influencers agreed to participate in the focus group in the third research stage.

Fig. 2 Word cloud of most frequent terms

Table 2 Top words by frequency (more than 300)

Rank	Top word	Frequency
1	Open Data	2498
2	Lombardia	586
3	Civic Hacking	462
4	AgID Gov	420
5	Rousseau	389
6	Digital	322
7	Transparency	318
8	Spaghetti	302

2.2 Conversational topics: the semantic social network analysis

Between 15 July 2019 and 21 July 2019, using NodeXL software (Smith et al. 2010), we downloaded the tweets of the influential actors detected. A new dataset was created by merging the dataset with the tweets previously used to determine the influencers and the tweets of the OD opinion leaders.

The database was filtered by deleting retweets that contained additional text and posts unrelated to the OD theme. Then, through the R package RcmdrPlugin.temis (Bouchet-Valat and Bastin 2013), text mining was carried out.

The dataset consisted of 176,293 terms (excluding stop words). Automatic analysis of the text highlighted a series of recurring terms and the related co-occurrences.

Through the analysis of frequent words, it is possible to identify a set of 52 terms that characterise the OD semantic sphere (Fig. 2).

In the set of the most frequent words, eight terms have higher frequency (more than 300): open data, Lombardia, civic hacking, AgID Gov, Rousseau, digital, transparency, Spaghetti (Table 2).
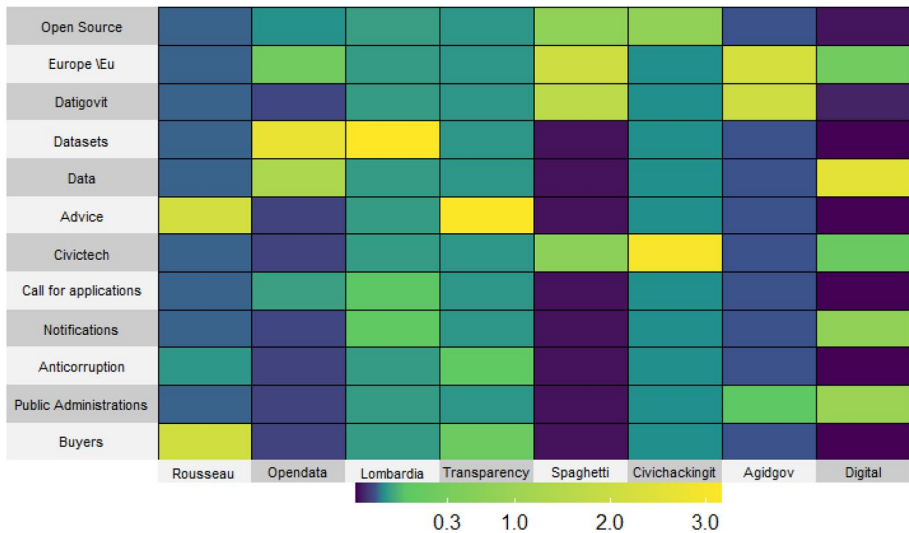
Fig. 3 Heatmap graph of the co-occurrence of the eight most frequent terms with the most relevant words

The term 'Lombardia' is one of the most frequent because the Lombardy region represents a virtuous case of organisation and management of regional OD datasets. The terms 'civic hacking' and 'Spaghetti' concern the civic hackers' movement SpaghettiOpenData, while 'AgID Gov' is an abbreviation used to define the digital government agenda. 'Digital' refers to the whole sphere of data and platforms that characterise the ecosystem of OD, and 'transparency' refers to what we can consider one of the main goals of OD strategy—to guarantee a transparent communication system between the public administration and its stakeholders.

The eight terms shown in Table 2 are particularly important as keywords of the OD world.

Considering these keywords, and through a co-occurrence analysis, we structured the semantic area that characterises the research object.

The first result concerns the existence of some words, with a particularly significant content, which co-occur several times with the most frequent terms (MFT) (Fig. 2).

The heatmap (Fig. 3) represents the links in the co-occurrence of the MFT with the selected words, considering: (1) relevance to the topic and (2) relevance to links with more than one MTF.

The matrix can be considered as a first map of the OD semantic universe.

From this first analysis, three sub-semantic categories emerge. The first is characterised by the link between MFT 'Rousseau'[15] and 'transparency' with 'buyers', 'anticorruption' and 'advice', whereas the second is composed of the terms 'Lombardia' and 'open data' that co-occur with 'call for application', 'notification'[16] and 'dataset'. The third core

---

[15] Rousseau is the technological platform, taking its name from the philosopher Jean-Jacques Rousseau, created by Casaleggio Associati and used by the Italian Five Star Movement for the implementation of direct democracy.

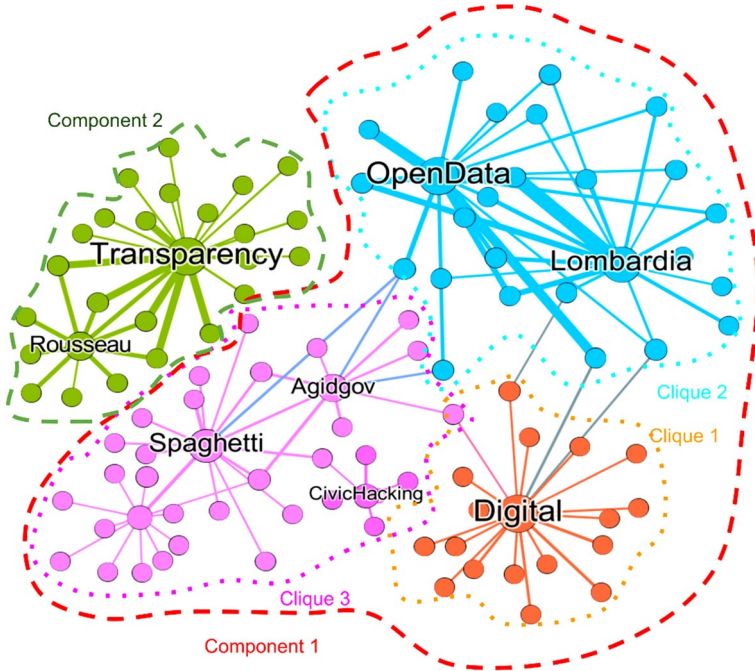[16] The term 'notification' also co-occurs with MTF Digital.

Fig. 4 The semantic network

category comprises the co-occurrence links between MFT 'open data', 'Spaghetti', 'AgID Gov' (Agenda Digitale di Governo) and 'digital' with the words 'public administration', 'Datigovit',[17] 'Europe', 'open source' and 'civitec' (civic technologies).

The third stage of the text mining process aims to explore deeply the links between the keywords of the OD conversational structure in Twittersphere. The semantic analysis has been extended to 25 most meaningful words for every eight MFT.

A new symmetric adjacency matrix (weighted in the edges) is built, starting from previous co-occurrence analysis (Bullinaria and Levy 2007). The semantic SNA is done by Gephi software (Bastian 2009) and the graph display using the ForceAtlas2 layout algorithm (Jacomy et al. 2014) (Fig. 4).

The semantic network of co-occurrences appears as an indirect network composed of 92 nodes and 113 links,[18] and has two distinct components[19] (Tarjan 1972): the first consists of four connected modules and the second of a single module (Lambiotte et al. 2015). In the first component, the nodes with the highest degrees are 'Lombardia', 'open data', 'digital', 'Spaghetti' and 'civic hacking'. The stronger connections are Lombardia-datasets, open data-datasets, EU-open data, open data-data, AgID Gov-SpaghettiOpenData, SpaghettiOpenData-Team Digitale, Carige transparency, consultancy transparency.

---

[17] Datigovit indicates a database managed by Agenzia Per L'Italia Digitale, available at https://www.dati.gov.it/

[18] The supplementary materials list the graph nodes divided by cliques.

[19] Identification of the components was analysed with the modularity algorithm (Blondel et al. 2008; Lambiotte et al. 2015).

Table 3 The four cliques and the nodes that compose them

Sub-cliques	Words	Semantic area
Digitale (Digital)	Digitale (digital); confini (borders); Agenda; amministrazione (administration); civico (civic); codice (code); convegno (conference); cybersecurity; Europa (Europe); garanteprivacy (privacy guarantor); GDPR; open by default; organizzato (organised); PA; protezione (protection); Amministrazioni Pubbliche (public administrations)	Institutional administration, privacy, conference, security
Lombardia–open data	Arese; bandi (call for applications); bilancio (statements); classificazione (classification); comune (common); contributi (contributions); convenzioni (conventions); corrispettivi (fees); CRL (Certificate Revocation List); dataset	Virtuous and interactive administration
	Dlgs (Decreto Legislativo); entrate (revenue); art. (article); Lombardia; EU; dati (data); open data; gara (public notice); capitale (capital city);	
	data; denunce (complaints); finanziario (financial); avvisi (notifications)	
SpaghettiOpenData–AgID Gov	AgID Gov, Codicecup, dataportal, datigovit, difensore, istat, mitgov, openantani, alimentare, Team Digitale, Spaghetti, controlli, dettaglio, francia, industriaitasl, merge (evento), ministero, ogp (open government), opendatasicilia, opengovitaly, openstreetmap, raduno, sanitari, sod (SpaghettiOpenData), wcarrara (wiki carrara), wikimediaitalia, panettieri, ristoranti civic hacking; civitech; datonatale, newsletter, spunti, open source	Civic hacking groups, datasets, production sectors, semantic area
Trasparenza (Transparency)	Rousseau; acquirenti (buyers); alternata (alternate); Carige (Carige bank); Casaleggio; consulenza (advice); conto (bill)	M5S transparency culture
	parlamentare (parliamentary); agricoltura (agriculture); anticorruzione (anticorruption); auspici (auspices); banca (bank); bancadati (database)	
	buon andamento (good performance)	
	cominciamo (let's start); conoscere (recognise); dasapere (to know); difeso (defended); ente (authority) ERSAP (Regional Agency for Agricultural Development)	

In the second component, the major nodes with the highest degrees are 'transparency' and 'Rousseau', while the highest co-occurrences can be observed for the term pairs transparency-bill, transparency-Rousseau, and transparency-bank (Table 3).

2.3 The cliques

In the first component, we can distinguish four cliques, named on the bases of the most frequent words: (1) digital; (2) Lombardia-open data; (3) SpaghettiOpenData-digital (Table 2).

'Digital' is an ego network, or a network where a node is linked with all other nodes (Everett and Borgatti 2005) and shares nodes exclusively with Lombardia-open data. Lombardia-open data presents a sub-network composed of two nodes with a higher degree ('Lombardia' and 'open data') and a strong edge between the words 'Lombardia', 'open data', 'dataset', 'EU' and 'data'. This second clique shares nodes with 'digital' and 'SpaghettiOpenData-Team Digitale'.

SpaghettiOpenData-Team Digitale is made of three nodes with larger connections—SpaghettiOpenData, AgID Gov and civic hacking—and a higher co-occurrence link between the SpaghettiOpenData and Team Digitale nodes. This clique shares nodes with Lombardia-open data.

Finally, 'transparency' cannot be considered a clique since it is the only nodule of the second graph component. It consists of two nodes with a higher degree ('transparency' and 'Rousseau') and a stronger edge between the nodes transparency-Rousseau, Transparency-Carige, Transparency-account, Transparency-bank, Transparency-consulting, Rousseau-Carige and Rousseau-consulting.

From the semantic analysis, it can be seen how the discussion on open data in Italy is divided into three levels: (1) institutional (digital team), (2) territorial governments (Lombardia), and (3) hacker groups in collaboration with Team Digitale. They coincide with the governance level, and the different integrated semantic spheres (Table 3).

At the macro level, which we call institutional, the topics mainly concern the processes that are currently the basis of the OD system, including the digital strategy of the public administration—agID, agID Gov, European, administration, PA (public administration)—and the problem of protection and use of data: GDPR, privacy, cybersecurity, open by default.

At the meso level, which we call administrative, the topics refer mainly to the Lombardia region's OD system which, besides being an example of good practice in the Italian OD environment, turns out to be a very active user in the Twitter context. Here, the keywords are related to the processes of digital PA (tenders, budget, capital, municipality, tender, contributions, agreements, fees, date, dataset, data, complaints, legislative decree, revenue, financial).

Finally, at the micro level, the emerging topics concern the concrete use and reuse of OD in formal and informal teams: Team Digitale and SpaghettiOpenData. Team Digitale is a commissioner's organisational structure appointed by the Council of Ministers on 16 September 2016 whose aim is to promote digitisation in public administrations. SpaghettiOpenData, as previously illustrated, is an Italian collective of civic hackers working on reuse projects. The keywords that fall under this specific topic are composed of four sub-dimensions: OD sources (AgID Gov, Codicecup, dataportal, datigovit, defender, istat, mitgov, opengovitaly ogp), reuse projects (openantani, spaghetti, industriaitasl, opendatasicilia,

openstreetmap, sanitary, sod, wcarrara), meetings and gatherings (meeting), and informal satire (datonatale, openantani).

2.4 The online focus group

The third stage of the research aims at defining the pros and cons of the Italian OD infrastructure with the help of an online focus group. Starting from six influencers, we arrived at twelve participants through snowball sampling. Among them we had civic hackers, public service employees, private sector employees and bloggers.

The arguments for the focalised interview were organised on the basis of the results of the text mining and co-occurrences analysis (second stage of the research) and integrated with topics not emerging from the online content, but from the analysis of literature and reports.

The issues discussed in the focus group were:

1. The relationship between democracy, participation and the OD infrastructure.
2. Usability, reuse and advantages offered by OD.
3. The effects of OD on public administration performance.
4. Weaknesses and shortcomings of the Italian system.
5. The roles of active citizenship and civic hacking.
6. OD as an opportunity for accountability and policy evaluation.

Considering the opinions expressed by the expert members of the focus group, we can better identify the pros and cons of OD in Italy.

Using a qualitative analysis of the online conversation, we can focus on four main topics: the nature of the OD infrastructure in Italy, considered mainly as a public service; the potential impact of this service on democracy (participation and transparency); the management of the informational infrastructure (governance and reuse problems);and the opportunity for accountability and policy evaluation.

To start with, it is important to emphasise that there is a general agreement that OD should be regarded as an open government tool. In the words of our experts:

> Open data is, in the true sense of the words, a public service, using modern technology to address the classic requirements of democracy, […] to assist citizens, businesses, and public administration itself to take decisions based upon data and thereby to solve a specific problem or meet a complex informational need. (Focus group participant).

OD is considered to be a true informational infrastructure in the service of the general public and of democracy, and as such capable of initiating social development not only in the field of public administration (increased effectiveness, transparency and efficiency, and greater relevance of policy systems and services), but also for businesses, civil society organisations and individual citizens:

> One of the things that public administration finds most difficult to understand is that a genuine appreciation of the informational assets possessed by the general public could reduce administrative expenses by facilitating and accelerating the exchange of information between public administrations, thereby rendering the whole system

more efficient, which in the final analysis would benefit the associated market economy. (Focus group participant).

There is currently a heated national and international debate on the subject of economic growth.[15] The capacity of OD to generate wealth is evidently one of its strengths, an unperceived but certainly relevant benefit. In the words of an expert participant in a focalised interview, the OD service.

> by the reuse of public informational assets, activates that famous economic lever (employment growth, new business models, new supply and production chains, new services), and this is something which in Italy we have great difficulty bringing about.

'Reuse' is a key word which relates to the advantages but also to the weaknesses and indeed dangers of the OD system. Regarding the opportunities offered, the word refers both to the possibilities for economic growth mentioned above and, even more, to its societal effects. The use of OD in a diverse variety of fields (the environment, the infrastructure, services, census data, etc.) gives rise to a series of relevant questions, such as what regulations and authorisations should govern the use of open data by third parties; how to create and manage satisfactory and efficacious governance systems; how to generate services and products to be spent on the market by non-profit companies and organisations; how individual citizens should participate, in cases where data is both legible and comprehensible; or how cooperation should be organised between volunteer organisations or by individuals who act as mediators between institutions and citizens. These questions are especially applicable to civic hacking.

Data reuse therefore becomes a central issue both when data are used by public administration departments themselves and when they are reused by businesses or members of the general public. The result of such a system of governance, even if currently expressed solely in terms of recommendations of good collaborative practice,[16] could amount to a revolution in the relationship between the citizen and public administration.

It would seem that OD could trigger new positive dynamics able to restore trust in political institutions. Focusing on open government, we can imagine a kind of democratic revolution based on new relations between the citizen, the state and the market. The reference here is to Castells and Himanen's (2002) informational welfare paradigm.

As one of the experts involved in the focus group stated, the citizen has the opportunity to play his own rights as he becomes an active participant in and monitor of public administration procedures.

ICT in general and OD in particular are capable of creating a more direct link between political and administrative bodies and civil society (businesses and the third sector), making a strong contribution to the organisational redirection of the bureaucratic apparatus towards openness to the general public, a virtuous system of exchange between one administration and another, and between an administration and a citizen.

The OD service represents a means of going beyond the idea of open government, where flows of information and participatory communication exert a positive effect on political and administrative decision-making.

For all these reasons, OD is evidently a public service with great potential impact in terms of economic growth, knowledge and awareness, support to decision-making, transparency and policy evaluation, and creation of governance systems.

However, one must take into account weaknesses in the Italian OD system that could impede its development. For the sake of simplicity, we can summarise some critical

Table 4 Pros and cons of Italian OD environment

Pros	Cons
Democracy and participation. OD could trigger new positive dynamics able to restore trust in political institutions	Technological weaknesses—i.e. delays in the establishment of a standardised IT infrastructure whose data are uniform, useable and worthwhile
Efficiency and digitalisation in the PA. OD as a public service can stimulate enhancement of the ICT culture in the PA and increase its efficiency	Insufficient ICT culture within public administration, especially among senior management and decision-makers
Transparency and accountability. OD could amount to a revolution in the relationship between the citizen and public administration	Insufficient financial resources for updating obsolete IT systems
Governance exists in the Italian OD system. It is clear that it has an important role and links active citizenship and civic hacking with the institutions	Regulatory weaknesses, in privacy, in authorisations for use, and in safeguarding authors' rights to reuse data
Reuse projects. OD in Italy follow a development model based on a collaborative process which is typical of the community of developers[a]	Reuse. Weaknesses in the transformation of data into accessible and useable information

[a]https://developers.italia.it/en/datigov/

areas: a) technological weaknesses—i.e. delays in the establishment of a standardised IT infrastructure whose data are uniform, useable and worthwhile; b) insufficient ICT culture within public administration, especially among senior management and decision-makers; c) insufficient financial resources for updating obsolete IT systems; d) regulatory weaknesses, in privacy, in authorisation for use, and in safeguarding authors' rights for the reuse of data; e) weaknesses in the transformation of data into accessible and useable information; and f) links between OD, transparency and evaluation of the public administration.

This last issue requires a more detailed explanation. During the interviews, some participant observers maintained that to associate the OD service with public administration's need for transparency and accountability might restrict its development and its potential: OD must be kept separate from transparency. There are datasets that can help citizens with their monitoring, but as a rule, the release of public data is to be considered as being on a par with any other infrastructure, such as road networks or water pipelines. The data are the responsibility of an administrative body, in exactly the same way as the roads are, although this does not mean that they are not for the benefit of all.

The results of the focus group analysis are summarised in the table below and grouped in pros and cons (Table 4).

3 Conclusions

OD should be linked to an open governance strategy in which the government builds an open system that interacts with its environment (see Fig. 5). The aim should be to structure a participatory system able to create multiple benefits: economic, political and social, and operational and technical.

Taking into consideration all the data collected in the three stages of the research, we can conclude that:
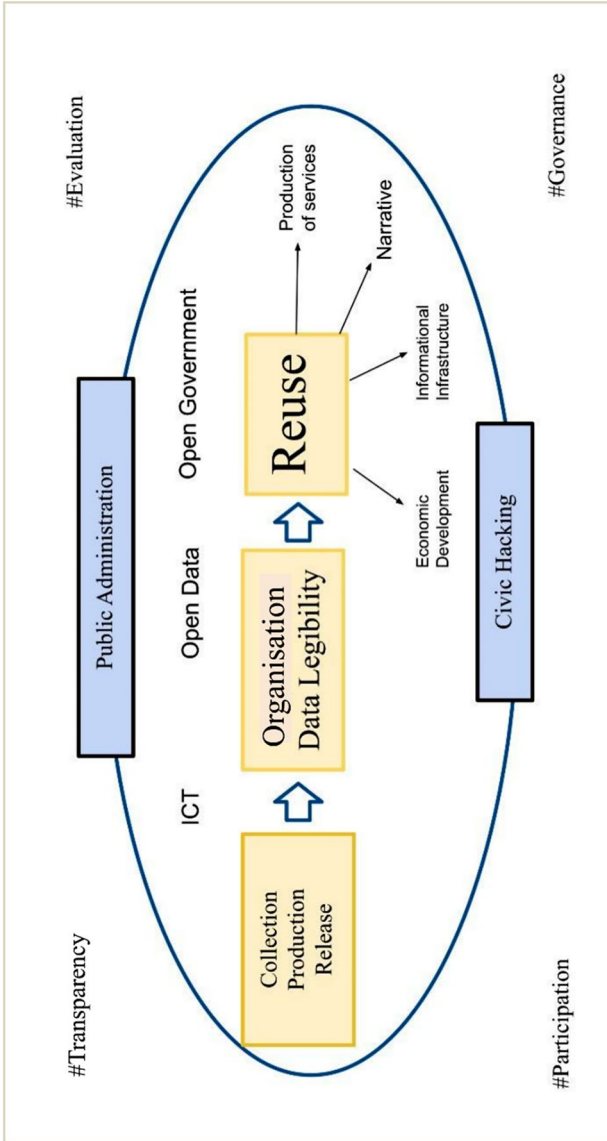
Fig. 5 The open data environment

1. Italian OD governance exists. The semantic analysis shows how discussion of OD in Italy is divided into three levels: institutional, territorial (Lombardia), and the hacker groups in collaboration with the Digital Team. They coincide with the governance level and the different integrated semantic spheres.
2. The topics emerging from the Twittersphere on this issue are transparency, evaluation, governance, participation and anticorruption. These words define values that characterise the Italian OD system. In particular, transparency, participation and anticorruption are linked to the clique associated with the e-democracy project of the Italian Five Star Movement for the implementation of direct democracy (Fig. 4 and Table 3).
3. An important role is played by voluntary organisations. The civic hacking groups collaborate with national and local institutions in the promotion of reuse projects, and training and support for administrations. Moreover, especially in the digital space, OD and ODG are configured as important influencers, sharing information and discussing conference initiatives, news and reuse projects.
4. Despite the leading role of the civic hackers community, it emerges, mainly from the online focus group, that the technological and cultural shortcomings of the Italian bureaucratic model are often resistant to innovations.

Focusing on this last point, we can underline how there are some obstacles to realising the opportunities that OD service can offer. First of all, there is the Italian public administration's cultural delay in the field of communication and information technology infrastructure. In the specific case of OD, the problem is with reuse. Indeed, there are many technical problems (Vetrò et al. 2016) with the quality of the data and the information infrastructure, in particular: data are not in a well-defined accessible format; there is an evident absence of standards; there is a lack of support to make data available; there is inaccurate information; data are obsolete or invalid; and there is a lack of territorial consistency, with the same data held in different forms.

The rigidity and the resistance of the bureaucracies define a lack of faith in the potential of such forms of active participation as civic hacking and reuse. One interviewee stated that "Civic hacking would have a lot to contribute, but trust is required: public administration has as little trust in anything originating from the public as the public has in anything not originating from public administration."

Despite all these problems, the innovative process that can be initiated by means of an IT infrastructure of this kind could lead to a change in the system of governance, and therefore in the democratic system as a whole, in the new millennium. The possibility of using boundless quantities of data could, if made available to the public, set in motion virtuous processes that would undoubtedly include the practice of participation as an essential component of a deliberative democratic process (Ryan and De Stefano 2000).

Moreover, if we consider how the introduction of the General Data Protection Regulation (GDPR Regulation EU 2016/679) has made the use of big data more difficult both in economic matters and in the field of research, open data has, therefore, become a valuable means of access to data, the use of which can also facilitate scientific research initiatives.

References

Bastian, M.: Gephi: An Open Source Software for Exploring and Manipulating Networks. AAAI Publications, Third International AAAI Conference on Weblogs and Social Media (2009).

Bentivegna, S. (ed.): La Politica in 140 Caratteri: Twitter e SpazioPubblico. FrancoAngeli, Milano (2014)

Berners Lee, T.: Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by its Inventor. HarperCollins, New York (1999)

Berners-Lee, T.: Linked data: the great unveiling. Technology Entertainment and Design, Long Beach, CA (4 February 2009). https://www.ted.com/talks/tim_berners_lee_on_the_next_web. Accessed 6 Oct 2019.

Berners-Lee, T., Hendler, J., Lassila, O.: The semantic web. Sci. Am. **284**, 34–43 (2001)

Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre, E.: Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment. 2008(10). Article ID: P10008 (2008).

Bouchet-Valat, M., Bastin, G.: RcmdrPlugin.temis, a graphical integrated text mining solution in R. R Journal. 5(1), 188–196 (2013).

Bullinaria, J.A., Levy, J.P.: Extracting semantic representations from word co-occurrence statistics: a computational study. Behav. Res. Methods **39**, 510–526 (2007)

Carr, S.J., Lassiter, A.: Big data, small apps: premises and products of the civic hackathon. In: Carr, S.J., Lassiter (eds) Seeing Cities Through Big Data, pp. 543–559. Springer, Cham (2017).

Castells, M.: Internet Galaxy. Reflections on the Internet. Oxford University Press, Oxford (2001)

Castells, M., Himanen, P.: The Information Society and the Welfare State: The Finnish Model. Oxford University Press, Oxford (2002)

Cecconi, G., Radu, C.: Open Data Maturity in Europe 2018. https://www.europeandataportal.eu/sites/default/files/edp_landscaping_insight_report_n4_2018.pdf (2018). Accessed 6 Oct 2019

Cha, M., Haddadi, H., Benevenuto, F., & Gummadi, K.P: Measuring user influence in twitter: the million follower fallacy. In fourth international AAAI conference on weblogs and social media. (2010). https://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/viewPaper/1538.

Choi, S.: The two-step flow of communication in Twitter-based public forums. Soc. Sci. Comput. Rev. **33**, 696–711 (2015)

Choi, J., Tausczik, Y.: Characteristics of collaboration in the emerging practice of open data analysis. In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, pp. 835–846. ACM (2017).

Chung J., Park, H. W.: Textual analysis of a political message: The inaugural addresses of two Korean presidents. Soc. Sci. Inf. **49**(2), 215–239 (2010).

Coyle, C.L., Vaughn, H.: Social networking: communication revolution or evolution? Bell Labs Tech. J **13**, 13–17 (2008)

Dijck, J.V.: Tracing Twitter: the rise of a microblogging platform. Int. J. Media Cult. Polit. **7**, 333–348 (2011)

Donker, F.W., Van Loenen, B.: How to assess the success of the open data ecosystem? Int. J. Digital Earth. **10**, 284–306 (2017)

Ermoshina, K.: Is there an app for everything? Potentials and limits of civic hacking. Observatorio (OBS*). 10(ESPECIAL), 116–140 (2016).

Everett, M., Borgatti, S.P.: Ego network betweenness. Soc. Netw. **27**, 31–38 (2005)

Frecks, L.M.: Descriptive information, statistics and little known facts derived from a database of US civic hacking groups. In: Proceedings of the 15th Annual International Conference on Digital Government Research, pp. 324–324. ACM (2014).

Gascò, M.: New technologies and institutional change in public administration. Soc. Sci. Comput. Rev. **21**, 6–14 (2003)

Gerunov, A.: Open data: policy and implementation in Bulgaria. In: Big Data, Knowledge and Control Systems Engineering Conference Sofia, pp. 1–11 (2015).

Gladwell, M.: The Tipping Point: How Little Things Can Make a Big Difference. Little Brown, Boston (2006)

Himanen, P.: The Hacker Ethic and the Spirit of the Information Age. Random House, New York, NY (2001)

Huijboom, N., Van den Broek, T.: Open data: an international comparison of strategies. Eur. J. of ePractice. **12**, 4–16 (2011)

Jacomy, M., Venturini, T., Heymann, S., Bastian, M.: ForceAtlas2: a continuous graph layout algorithm for handy network visualization designed for the Gephi software. PLoS ONE **9**(6), e98679 (2014). https://doi.org/10.1371/journal.pone.0098679

Jalali, S.M.J., Park, H.W.: Conversations about open data on Twitter. Int. J. Contents. **13**(1), 31–37 (2017)

Janssen, M., Charalabidis, Y., Zuiderwijk, A.: Benefits, adoption barriers and myths of open data and open government. Inf. Syst. Manag. **29**, 258–268 (2012)

Jetzek, T.: The value of open government data. GeoforumPerspect. **12**, 47–56 (2013)

Jung, K., Park, H.W.: A semantic (TRIZ) network analysis of South Korea's 'open public data' policy. Government Inf. Q. **32**(3), 353–358 (2015)

Katz, E., Lazarsfeld, P.: Personal Influence: The Part Played by People in the Flow of Mass Communication. The Free Press, Glencoe (1955)

Kundra, V.: Digital fuel of the 21st century: Innovation through open data and the network effect. Joan Shorenstein Center on the Press, Politics and Public Policy. https://shorensteincenter.org/wp-content/uploads/2012/03/d70_kundra.pdf (2012). Accessed 8 June 2019

Lambiotte, R., Delvenne, J.-C., Barahona, M.: Laplacian dynamics and multiscale modular structure in networks. IEEE Trans. Netw. Sci. Eng. (2015).

Leavitt, A., Burchard, E., Fisher, D., Gilbert, S.: The influentials: new approaches for analyzing influence on Twitter. Web Ecol. Project. **4**, 1–18 (2009)

Lévy, P.: L'Intelligence Collective. Pour uneAnthropologie du Cyberespace. La Découverte, Paris (1994)

Lévy, P.: Cyberdémocratie. Essai de PhilosophiePolitique. Editions Jacob, Paris (2002)

Peña-López, I.: Open data maturity in Europe 2017. https://www.capgemini.com/consulting/wp-content/uploads/sites/30/2017/11/open-data-maturity-europe-2017 (2017). Accessed 6 Oct 2019

Ryan, K.E., De Stefano, L. (eds.): Evaluation as a Democratic Process: Promoting Inclusion, Dialogue, and Deliberation: New Directions for Evaluation. Jossey-Bass, San Francisco (2000)

Scavo, C., Shi, Y.: Public administration: The role of information technology in the reinventing government paradigm. Normative predicates and practical challenges. Soc. Sci. Comput. Rev. **18**, 166–178 (2000). https://doi.org/10.1177/089443930001800206

Schrock, A.R.: Civic hacking as data activism and advocacy: a history from publicity to open government data. New Media Soc. **18**, 581–599 (2015)

Smith, M.: Catalyzing social media scholarship with open tools and data. J. Contemp. Eastern Asia. **14**(2), 87–96 (2015)

Smith, M., Ceni, A., Milic-Frayling, N., Shneiderman, B., Mendes Rodrigues, E., Leskovec, J., Dunne, C.: NodeXL: A free and open network overview, discovery and exploration add-in for Excel 2007/2010/2013/2016. Social Media Research Foundation (2010). https://www.smrfoundation.org. Accessed 16 Oct 2019

Tarjan, R.: Depth-first search and linear graph algorithms. SIAM J. Comput. **1**, 146–160 (1972)

Tena-Espinoza-De-Los-Monteros, M.A.: Civic innovation as a response to social problems. In: Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality-TEEM '16 (2016). Doi:https://doi.org/10.1145/3012430.3012555

Thakuriah, P., Dirks, L., Keita, Y.: Digital infomediaries and civic hacking in emerging urban data initiatives. In: Thakuriah, P., Tilahun, N., Zellner, M. (eds.) Seeing Cities Through Big Data: Research Methods and Applications in Urban Informatics. Series: Springer Geography, pp. 189–207. Springer, NY (2016). doi:https://doi.org/10.1007/978-3-319-40902-3_11

Turkle, S.: Life on the Screen: Identity in the Age of the Internet. Simon and Schuster Inc, New York (1997)

Ubaldi, B.: Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives. OECD Working Papers on Public Governance, No. 22, OECD Publishing (2013).

Vetrò, A., Canova, L., Torchiano, M., Minotas, C.O., Iemma, R., Morando, F.: Open data quality measurement framework: Definition and application to open government data. Government Inf. Q. **33**, 325–337 (2016)

Wasserman, S., Faust, K.: Social Network Analysis: Methods and Applications (Vol. 8). Cambridge University Press, Cambridge (1994).

Yu, H., Robinson, D.G.: The new ambiguity of open government. UCLA L. Rev. Discourse. **59**, 178 (2011)