Accepted manuscript

Creating Open Government Data ecosystems: network relations among governments, user communities, NGOs and the media

Luigi Reggi [1] & Sharon S. Dawes [2,3]

[1] *Agency for Territorial Cohesion, Italian Presidency of the Council, Rome, Italy.*
[2] *Rockefeller College of Public Affairs and Policy, University at Albany, SUNY, USA*
[3] *Center for Technology in Government, University at Albany, SUNY, USA*

Corresponding author:
Luigi Reggi
Agency for Territorial Cohesion - Italian Presidency of the Council
Via Sicilia 162, 00198 Rome, Italy.
luigi.reggi.esp@agenziacoesione.gov.it

Abstract

Open Government Data (OGD) ecosystems are composed of public, private and non-profit actors playing specific roles related to the availability and use of publicly accessible government information. The literature considers the presence of healthy ecosystems as crucial for effective use of OGD, with positive effects on democracy, policy effectiveness, and economic development. This paper employs the Exponential Random Graph model (ERGM) technique to empirically explore relations among the actors of an OGD ecosystem for public participation in the context of the European Policy in Italy. The models estimate the likelihood of an ecosystem connection between actors as documented online via Twitter, by considering the type of actor - namely government organizations, user communities, NGOs and the media - and their locations. The analysis showed that governmental organizations as data providers and intermediaries play a crucial role in disseminating OGD and facilitating their use by local communities. Government organizations as policy makers were much less active. In addition, NGOs and the media were less disposed than government actors to serve as data intermediaries and less likely than local communities to engage in policy deliberation. These patterns suggest that the nature and level of engagement by various actors may be influenced by their interest in the specific purpose of the ecosystem. Finally, co-location is a powerful predictor of the creation of new connections among actors of all kinds, demonstrating that effective local data use can be enabled and encouraged by national data provision.

1 Introduction

The introduction of Open Government Data (OGD) policies in many countries has been accompanied by high expectations in terms of better government transparency and accountability, public participation, innovation, and economic development (Janssen et al., 2012; Ruijer et al., 2017; Zuiderwijk et al., 2014b). However, a long list of disappointments has followed acknowledgment of several "myths" regarding the power of OGD to automatically create these benefits (Janssen et al., 2012). Scholars now tend to associate the actual fulfilment of these ambitious promises with the development of healthy open data ecosystems of both public and private actors that enable meaningful use of the information released via governmental portals (Dawes et al., 2016; Gupta et al., 2020; Oliveira et al., 2019). In particular, OGD policies with the specific purpose to foster transparency, accountability and public participation need not only proactively reach out to potential users, but they must also create the conditions for effective data use in order to have an impact on both social capital and democratic decision-making (Fung et al., 2013; Ruijer et al., 2017).

Although the study of OGD ecosystems is still in its infancy, several contributions have identified a few "keystone" actors and functions. Actors comprise public, private and non-profit organizations, while typical functions include data provision and data use, as well as an intermediation function between providers and users (Attard et al., 2016; Davies, 2011; Dawes et al., 2016; Martin et al., 2017; Ubaldi, 2013; Zuiderwijk et al., 2014b). To stimulate meaningful participation in public debate and coproduction of policies and services, keystone actors need to create stable relations with each other, allowing information to flow from the public sector to communities of users - either directly or through intermediaries - and then back to policy makers in the form of actionable feedback (Janssen & Helbig, 2016). A "healthy ecosystem" exists when these essential traits are substantially developed and working through a significant number of active relations among keystone actors (McBride et al., 2020; Reggi & Dawes, 2016).

To date, empirical research on how connections among ecosystem actors are likely to form is limited to a few qualitative case studies, and only a few of these studies investigate what factors have the potential to strengthen or weaken the development of the ecosystem (Dawes et al., 2016; Harrison et al., 2012; Khayyat & Bannister, 2017; McBride et al., 2019; Slobodova & Becker, 2020). This paper offers a contribution through a quantitative analysis that applies social network analysis methods and estimates the likelihood of a connection to form between actors of two different types. It is a first attempt to systematically explore the creation of linkages between different actors in a specific ecosystem, and the factors that may influence this process.

This paper addresses the following research questions: 1) What types of actors are more likely to develop connections with other actors in an OGD ecosystem? In particular, what are the strongest (or weakest) connections among the different types of actors? 2) Do actors' locations influence the probability of developing connections?

To respond to these questions, a dataset was constructed including all relevant actors and relations in an OGD ecosystem focused on government spending in Italy and aimed at stimulating public participation in European development policies. This ecosystem, called "A Scuola di OpenCoesione" (ASOC), is a country-wide OGD program of the Italian government that engages actors at different levels of government, as well as schools, non-governmental organizations (NGOs), civil society organizations (CSOs), the media, and community organizations to use OGD to assess and discuss funded projects. Using the Exponential Random Graph model (ERGM) technique, the analysis showed that governmental organizations as data providers and intermediaries play a crucial role in disseminating OGD and facilitating their use by local communities. Government organizations as policy makers are much less active. In addition, NGOs and the media are less disposed than government actors to serve as data intermediaries and less likely than local communities to engage in policy deliberation. Finally, co-location is a powerful predictor of the creation of new connections among actors of all kinds demonstrating that local communities can make effective use of data collected and released by national-level providers.

The paper is structured as follows. The next section identifies key actors and roles from the recent literature on OGD ecosystems. The subsequent section deals with the factors potentially affecting the likelihood of a new ecosystem relationship to form. We then describe the case of ASOC. Section 4 presents the methods used in the analysis, followed by the main results in sections 5 and 6. The final section draws some conclusions, discusses the limitations of the study, and offers avenues for future research.

2 Actors, roles and relations in OGD ecosystems

The ecosystem metaphor is often used in the literature to "evoke an image of biological ecologies with their complex dynamics and diverse species" which co-evolve over time (Nardi & O'Day, 1999, p. 50). According to Harrison et al. (2012), ecosystems are intended to represent "interdependent social systems of actors, organizations, material infrastructures, and symbolic resources that must be created in technology-enabled, information-intensive social systems" (p. 900). OGD ecosystems derive from the dynamic interplay of both social and technical elements, which include individuals, formal and informal organizations with different interests and goals, as well as technological infrastructures such as the data, the technology employed to produce and disseminate them, and the tools that are used to create value from them (Davies, 2011; Dawes et al., 2016; Zuiderwijk et al., 2014b).

Some authors envision multiple OGD ecosystems that can be simultaneously stimulated around specific types of government data, or the practices of particular government agencies as they interact with relevant stakeholders and user communities (Gupta et al., 2020; Harrison et al., 2012; Lassinantti et al., 2019; Oliveira et al., 2019). The presence of multiple ecosystems is also introduced when considering the participation of different levels of government, from federal to local (Kassen, 2018).

This section considers the extant literature that deals specifically with OGD ecosystems that have the goal of fostering public participation in decisionmaking. First, it considers main actors and their roles (subsection 2.1). Second, evidence from work in governance network theory is analyzed to explore issues about the nature of ecosystem relations (subsection 2.2).

2.1 Keystone actors and their roles in OGD ecosystems for public participation

Identifying key types of actors and their roles is crucial to understanding the dynamics of these ecosystems (Helbig et al., 2012; Zuiderwijk et al., 2014b). The literature identifies different types of actors that can be defined as the "keystone species" of an OGD ecosystem, i.e. different actors that enable vital functions in the ecosystem, either as creators of information, services, or tools, or as mediators between different actors (Van Schalkwyk et al., 2015). These essential roles can co-evolve over time (Harrison et al., 2012).

Table 1 presents a set of "keystone actors" in an ecosystem whose goal is public participation, based on our review of relevant literature. Information on the main level of government at which the actors are active is also provided. Each type of actor usually plays multiple roles in the ecosystem. Roles are based on the main functional components of the OGD ecosystem for public participation, including policy making, data provision, intermediation, and data use (Dawes et al., 2016; Ruijer et al., 2021). Other actors such as businesses or universities are cited as well in the literature, but their roles are less salient when considering OGD ecosystems for public participation.

Table 1 – Keystone actors and roles in selected OGD ecosystem frameworks

Keystone actor type	Main roles	References
Governments (political leaders and public agencies)	*Policy maker:* Define OGD overall strategies, access rules and accompanying policies (e.g. participation mechanisms). It orchestrates the activity of other actors. It is responsible for engagement quality, legitimacy and data quality. Use the feedback from OGD users in policy programming, implementation, or evaluation. Use OGD from the same or other public agencies.	Harrison et al. (2012), Dawes et al. (2016), Reggi and Dawes (2016), Ruijer et al. (2018), Martin et al. (2017)
	OGD user: Use OGD from the same or other public agencies at different levels of government	Harrison et al. (2012), Reggi and Dawes (2016), Ruijer et al. (2018)
	OGD provider: Publish OGD. Perform proactive actions to stimulate meaningful use and ecosystem creation.	Davies (2011), Harrison et al. (2012), Ubaldi (2013), Zuiderwijk et al. (2014b), Dawes et al. (2016), Reggi and Dawes (2016), Attard et al. (2016), Martin et al. (2017), Kassen (2018)
Media, Civic media	*OGD user:* Use OGD and translate them into interesting stories.	Ubaldi (2013), Zuiderwijk et al. (2014b), Reggi and Dawes (2016), Martin et al. (2017), Attard et al. (2016), Lassinantti et al. (2019), Zuckerman (2014)
	OGD intermediary: Provide additional data from own investigations. Augment raw OGD thanks to the inputs from the public. Influence policy decisions by exerting external pressure.	
NGOs, CSOs, Civic Tech communities, informal advocacy groups	*OGD user:* Uses OGD to create new analyses or services.	Davies (2011), Harrison et al. (2012), Dawes et al. (2016), Attard et al. (2016), Ruijer et al. (2018), Lassinantti et al. (2019)
	OGD intermediary: Build capacities at the community level and create a culture that appreciates the relevance of the data. Aggregate citizen-generated data to collect input on a policy issue or for a public service. Repackage and republish OGD in more usable form for other actors.	Ubaldi (2013), Zuiderwijk et al. (2014b), Reggi and Dawes (2016), Attard et al. (2016), Martin et al. (2017)
Local communities, individuals	*OGD user:* Analyze the data, contextualize it, use the data to make decisions. Use the data for public participation purposes. Create user-generated data based on OGD.	Harrison et al. (2012), Ubaldi (2013), Zuiderwijk et al. (2014b), Attard et al. (2016), Kassen (2018), Lassinantti et al. (2019)
	OGD beneficiary: Use products or services created thanks to the availability of OGD.	Harrison et al. (2012), Dawes et al. (2016)

The first type of actor that can be defined as a "keystone" actor based on the literature is Government. Actors belonging to this type are identified as the "central actors, taking the initiative within networked systems organized to achieve specific goals related to innovation and good government" (Harrison et al., 2012). They comprise individual public agencies at different levels of government, which often interact in order to develop public policies or implement policies and services. Acting as policy makers, political leaders and administrators are responsible for multiple phases of the policy cycle from agenda setting to policy evaluation. Not only they are proactive users of OGD (Yang & Wu, 2021), but they also interact with public and private organizations responsible for policy implementation on the ground. Governments are supposed to "take the input of citizens seriously" (Pateman, 1970), ensuring that participation is sustained and has a genuine impact on public decisions. Several studies highlight the opportunity for governments to integrate that input in the form of user-generated data and information based on the availability of OGD and produced by external actors such as NGOs, civic communities and the media (Janssen & Helbig, 2016).

Governments also define specific OGD strategies though their decisions about resources, rules and regulations. They often play the role of "orchestrator" to "ensure consistency among tasks and to oversee whether the various stakeholders work in concert to contribute to meaningful engagement", as well as taking responsibility for a sufficient "quality of engagement, legitimacy of the process and the usability of the data and information" (Janssen & Helbig, 2016).

Two other keystone actors are the media (including civic media) and NGOs (including Civil Society Organizations, civic technology communities) which are often associated with the role of intermediaries or "infomediaries" between the government and OGD users (Janssen & Zuiderwijk, 2014). They play the role of information brokers in the network (Fleming &

Waguespack, 2007; Wasserman & Faust, 1994). As advanced users of OGD, these formal and informal organizations provide value in data interpretation, aggregation, and analysis. These actors also include data activists (Baack, 2015), civic media (Zuckerman, 2014) and advocacy groups (Dawes et al., 2016). In particular, media outlets and civic platforms are expected to exert pressure over policy decisions by fostering a public debate based on the data and the evidence collected, while NGOs and other advocacy groups can develop applications and services that broaden the original scope of OGD provision and aggregate the input from users to enable new feedback flows that loop back to the data providers and other decision makers (Harrison et al., 2012).

Local communities of users comprising individuals, and informal groups are a fourth key type of actor. They are often the final beneficiaries of the products and services developed by businesses, the media, or the civic technology community. They can use the data and analytical products to make better decisions themselves (Fung et al., 2007), and to stimulate an "informed citizenry" able to make arguments that enable deeper contributions to policy deliberations and useful feedback on the workings of government (Janssen & Helbig, 2016). They can also play a more active role as prosumers (Attard et al., 2016), that is by generating new data and information that is used as a basis for public participation processes (Meijer & Potjer, 2018).

2.2 The relations among keystone actors

The relations among the actors in an ecosystem represent a second fundamental aspect. From the point of view of governance network theory, these relations are considered flows of information from one actor to another. In the case of OGD ecosystems, OGD itself is information generated within the public sector and then shared with external actors. Similarly, other transactions based on this information can be used to develop new services, analyses, or news articles or to create new information that actors – especially the government – can use to improve public policies and services, or to make better decisions.

Based on the literature, in Figure 1 we developed a representation of the four keystone actor types and their main relations (Attard et al., 2016; Davies, 2011; Dawes et al., 2016; Harrison et al., 2012; Kassen, 2018; Lassinantti et al., 2019; Martin et al., 2017; Ubaldi, 2013; Zuiderwijk et al., 2014b). All connections are two-way information exchanges between actors of different types. The content of information exchanges depends on the source and the target of the connection. Table 2 summarizes the main content of information exchange among these actors and roles.

Figure 1 - Keystone actors, roles and relations in OGD ecosystems for public participation

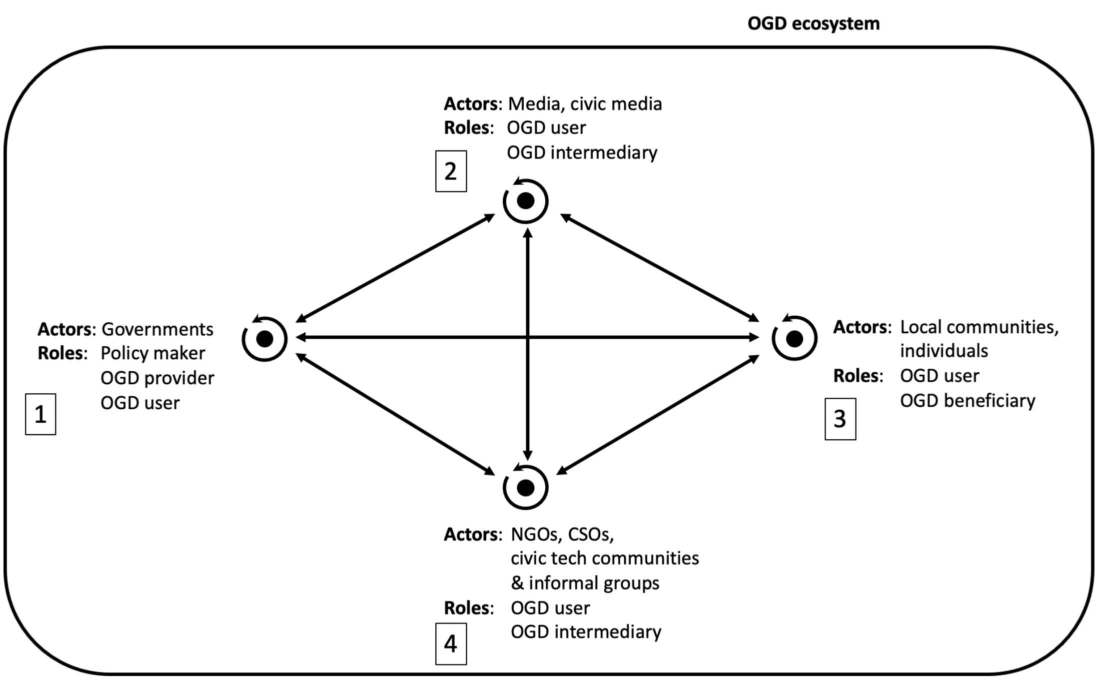


Table 2 - Content of information exchange among actors

Main relations	Description	References
1 → 3	Government stimulates the use of OGD by providing raw and aggregated data and tools for analysis to different user targets. OGD providers offer support to users to ensure sustainable and meaningful data use.	Davies (2011), Harrison et al. (2012), Ubaldi (2013), Martin et al. (2017), Dawes et al. (2016), Kassen (2018), Khayyat and Bannister (2017)
1 → 2		
1 → 4		
2 → 1	Influence policy decisions by exerting external pressure.	Zuiderwijk et al. (2014b), Reggi and Dawes (2016), Martin et al. (2017), Attard et al. (2016), Lassinantti et al. (2019)
2 → 3	Provide interpreted data and stories.	
4 → 1	Influence policy decisions though active engagement and advocacy. Provide feedback to OGD providers on data quality. Advocate for better OGD policies and strategies	Harrison et al. (2012), Dawes et al. (2016), Reggi and Dawes (2016)
4 → 3	Provide interpreted data, analyses, and visualizations	Ubaldi (2013), Dawes et al. (2016)
3 → 1	Provide additional data and feedback on policy making that can be used for public participation purposes. Provide feedback to OGD providers on data quality and request additional data.	Dawes et al. (2016), Davies (2011), Harrison et al. (2012), Zuiderwijk et al. (2014b), Reggi and Dawes (2016), Khayyat and Bannister (2017)
3 → 2	Provide additional data and feedback on policy making that can be aggregated for public participation purposes or for influencing policy making	Zuiderwijk et al. (2014b), Ubaldi (2013), Reggi and Dawes (2016)
3 → 4		
2 ↔ 4	Distribute the information of some successful open data cases or indicate that there are needs for using open data in some specific areas.	Dawes et al. (2016), Ubaldi (2013)

3 Factors affecting the creation of relations among actors

The literature also highlights several factors that can facilitate or hinder the creation of connections among different actors. Some of these factors are related solely to the characteristics of each actor, while other factors consider two or more actors sharing common features. We focus on two factors, namely the type of actor and the actor's location to develop initial hypotheses for the subsequent analysis.

3.1 Type of actor

A first factor is related to the nature of each actor and its role in the ecosystem. Based on the literature above, we can hypothesize that some actors are more likely to develop connections than others. In particular, individual actors classified as government actors play a pivotal role not only as OGD providers and ecosystem initiators, but also as orchestrators of joint initiatives and stimulators of OGD use (Gupta et al., 2020; Martin et al., 2017; Slobodova & Becker, 2020). The number of connections involving this type of actor is therefore expected to be high, compared to other types of actors, at least at the initial stage of ecosystem development.

Other actors are also expected to develop connections, but they generally face several barriers. For users, potentially significant benefits of OGD use for public participation justify the creation of connections. However, these benefits are also accompanied by legal, political, social, institutional, economic, operational and technical barriers (Janssen et al., 2012; Zuiderwijk et al., 2014a) These can include extensive use of technical jargon in policy documents and data descriptions, and lack of interpretive tools and technical knowledge among users. For example, civic communities that lack specific training may not have statistical, policy, or legal skills to make sense of the data in order to use them for analysis or participation (Gascó-Hernández et al., 2018). These problems can prevent interested parties from developing connections with OGD providers for data use, with policy makers and intermediaries for debating policy issues based on the data, or for channeling their feedback to decision-makers. Therefore:

H1.1: *Connections are more likely to develop between government actors and other actors*

In addition to government actors, we can expect a high number of connections generated by other actors playing the role of data intermediary. In an ecosystem focused on accountability and participation, these actors often serve as information brokers by making data more intelligible for other user groups and by advocating for action themselves (McGee et al., 2018). Their mediating role also implies a wide range of potential connections with actors in diverse domains including advocacy, civic technology, and public policy (Harrison et al., 2012). Therefore:

H1.2: *Connections are more likely to develop between actors whose roles include OGD intermediary and other actors*

3.2 Location

The second factor is related to the physical location or co-location of ecosystem actors. Information exchanges through technology are especially frequent when the actors are located within networks of "particular and localized relationships" (Nardi & O'Day, 1999). This idea is linked with the concept of geographic proximity in network theory, which is expected to improve network effectiveness (Jones et al., 1997), reduce conflicts and engender trust among actors (Polzer et al., 2006). Sharing the same geographic location represents a homophily feature in social network analysis. According to this principle, two nodes of a network are more likely to develop a connection when they share common characteristics (McPherson et al., 2001; Wasserman & Faust, 1994). The importance of location is also consistent with several findings in the field of collaborative governance and coproduction of public services. In particular, collaboration seems to work better at the local level, where regional, municipal and other local institutions can have an advantage in applying this approach. For example, Scott and Thomas (2017) demonstrated that actors engaged in the Collaborative Governance Regimes of a regional environmental policy were likely to report an increase in financial, human and technical resources compared to other actors. Accordingly:

H2: *Sharing the same geographic location will increase the likelihood of a connection*

We now turn to a specific case in which these hypotheses can be tested with empirical data.

4 The case of "At the School of Open Cohesion" (ASOC)

In 2012, the Italian government launched OpenCoesione, an OGD portal publishing detailed information on every project financed by its Cohesion Policy. Currently, nearly 1.5 million projects worth 148.1 billion Euros are tracked on the portal, reporting data on the source of funding, financial progress, actors responsible for policy programming and implementation, project location, and timing (Reggi & Dawes, 2016). These data are updated every two months. This high level of granularity allows user communities to locate and follow the progress of individual projects financed in their neighborhoods, thus creating opportunities for new information and input to be discussed with local actors and policy makers. However, this information is not enough to know whether a project is delivering its promises, since no data are published about underlying policy decisions, specific objectives and, more importantly, results.

Initially, the OpenCoesione portal was mainly known by people and organizations directly involved in EU Cohesion Policy, such as the managing authorities of the EU funds and a limited number of journalists, policy evaluators and researchers. No specific dissemination activities were carried out targeting the broader audience of policy beneficiaries and citizens (Lo Russo, 2016). In order to address this issue, OpenCoesione staff at the Presidency of the Council of Ministers in partnership with the Ministry of Education and the European Commission launched an initiative called "At the School of Open Cohesion" (ASOC)[1]. This initiative offers interested high schools in Italy a Massive Open Online Course (MOOC) through which students can acquire basic data analysis skills, understand key processes behind local development policy, and identify organizations responsible for specific projects (Ciociola & Reggi, 2015). The students select one

[1] www.ascuoladiopencoesione.it

relevant project on OpenCoesione.gov.it that is being implemented in their city, such as transport infrastructure, an urban renovation project, or investments to preserve cultural heritage. The students conduct a set of structured civic monitoring activities on the project's history, objectives, policy rationales, administrative processes, and results, using tools and methods from Monithon[2], a civil society initiative engaging NGOs and local communities in assessing the effectiveness of EU-funded projects. Methods include field visits, interviews with experts and administrators, and a Strength, Weakness, Opportunity, and Threat (SWOT) analysis (Gascó-Hernández et al., 2018). The results generally include an analysis of the status of the project, its progress and effectiveness, and suggestions for improvement or problem solving. Students then connect with NGOs, the media, transparency activists, political representatives, and public managers to present their results or ask for support. Finally, this evidence is discussed in public events involving local policy makers where students and local communities ask questions and suggest improvements (Ciociola & Reggi, 2015; Reggi & Dawes, 2016).

Based on the existing literature on this case (Ciociola & Reggi, 2015; Reggi & Dawes, 2016) we identified the specific types of actors and roles included in Table 3. Compared to the theoretical representation of keystone actors types, we found that not all the roles in Figure 1 and Table 2 were relevant for this case. In particular, NGOs and the media appear to act only in the role of data intermediaries, as they hardly ever make direct use of the data, while local communities (the students and their teachers) are typically proactive users. We also found that some government organizations, namely the Europe Direct Centers (EDIC) played an additional role of OGD intermediary.

[2] www.monithon.it

Table 3 – Case-specific types of actors, their roles and characteristics in the ASOC network

	Type of actor	Main role	Description	Level	Specific role in the program
A	Government	OGD provider	OpenCoesione.gov.it and supporting staff	National	Provide access to data and data interpretation. Stimulate use/re-use of OGD. Facilitate relations with national-level administration
B	Government	Policy maker	Political leaders or public managers / employees at National, regional and municipal agencies.	Mainly local	Discuss / consider / integrate feedback from the bottom-up in the policy making
C	Government	OGD intermediary	"EU Direct" centers (EDICs), a network of information and documentation centers, and speakers in all EU countries[34].	Local	Support data users. Facilitate relations with regional and local policy makers
D	Local communities	OGD user	High school students and their teachers participating in the ASOC program.	Local	Data analysis. Civic monitoring. Develop suggestions for policy makers. Organize accountability forums with administrators
E	NGOs	OGD intermediary	Civil Society Organizations, open data activists and NGOs	Local	Support the schools. Interested in accountability and influencing policy decisions
F	Media	OGD intermediary	National and local newspapers, TVs, web magazines	Mainly local	Foster evidence-based debate

From 2015 to 2020, more than 26,000 students, supported by 3,000 teachers, analyzed about 800 publicly funded projects worth almost 8 billion Euros, stimulating public deliberation by conducting fieldwork and organizing public events[5]. While ASOC activities mainly take place off-line, such as in-class data analysis or interviews with public authorities, students are required to document them all using Twitter, which is employed in this paper as the main source of information to identify and select relevant actors and connections.

---

[3] https://europa.eu/european-union/contact/meet-us_en

[5] https://opencoesione.gov.it/it/pillole/data-card-ASOC/

5 Research methods

The analysis makes use of Twitter data to estimate the likelihood of a connection among actors in the ASOC ecosystem. First, a qualitative analysis was applied to the content of Twitter interactions to verify the presence of relevant connections. Second, three Exponential Random Graph Models (ERGMs) - an advanced application of Social Network Analysis - were developed to test the hypotheses we formulated in relation to our research questions.

5.1 Dependent and independent variables

Qualitative data from Twitter were used to identify actors and connections in the ASOC ecosystem. The analysis initially considered five years of data from 2016-2020 consisting of 35,218 tweets and 2,649 twitter accounts that used the hashtags *#ASOC1516, #ASOC1617, #ASOC1718, #ASOC1819, and #ASOC1920.* Each hashtag represents one annual edition of the ASOC program. The data were imported in R through the "rtweet" package using Twitter's APIs.

The dependent variable is a network connection between two individual actors. A connection is formed between two actors when a shared activity is documented in the selected tweets by looking at the content of the Twitter conversation. In particular, a set of potential connections for each actor (i.e., each Twitter account) was initially identified based on the presence of other accounts mentioned in the same tweet. Both organizations and individuals are considered in order to include both formal and informal exchanges (Mergel & Bretschneider, 2013). In this initial selection, accounts were removed if they never used the hashtags associated with the program (e.g., accounts that were mentioned by some of the actors but that never actually participated, such as @GretaThunberg), or if they were bots or fake accounts. The remaining potential connections were then manually coded based on the content of each tweet. Connections were included in the analysis if the tweet explicitly mentioned a shared activity between two or more accounts, namely OGD analysis, in-class activities, fieldwork, joint participation in events, or dissemination of findings. The coding of tweets was validated by accessing other public sources (e.g., students' websites, other social media, official website of Italy's "Open data day" event). The final selection included 455 actors and 1,177 connections. Two independent variables were constructed from the edge lists based on systematic coding of each Twitter account.

Type of actor is a categorical variable that classifies actors into groups based on their type and role as they emerged from case analysis (see Table 3, from type A to F). A residual category "other actor" (G) includes universities, research institutes, businesses, policy professionals, and individual citizens. The coding of Twitter accounts for classification was based on available information included in the relevant bio, tweets, and links to websites. The result of this coding was then systematically compared with available information from the ASOC program (i.e., lists of students' accounts and participating organizations). These validity checks were performed by two members of the ASOC staff.

The *location* variable was attributed to each actor based on the information available on the relevant Twitter biography and the content of the tweets posted. One of the 21 Italian regions (subnational areas) was attributed to each node. Actors operating in all regions such as Ministries or national NGOs were categorized as "national level" actors.

In addition, the *number of followers* was employed as a control variable. This variable can be considered a proxy for the strength of an actor's online presence, which is expected to directly affect the number of connections generated (Hofer & Aubert, 2013).

5.2 Exponential Random Graph Modeling

As the final step of the empirical analysis, three curved ERGMs were developed to estimate the likelihood that two actors would develop a connection, given the type of actors involved, their location, and the overall network structure. ERGM is a social network analysis technique for estimating the probability of a connection to form between two actors in a network, based the characteristics of both the individual actors and the patterns of ties comprising the whole network (Harris, 2013). ERGM uses a "familiar logistic regression-like statistical form" (Harris, 2013, pp. 5-6), so the interpretation of the coefficients is similar to that in a logit model. The advantage of using ERGM is that it allows for statistically sound inference in the presence of highly interdependent network relationships, given that the characteristics of the whole network are taken into account (Hunter et al., 2008). Therefore, compared to simple logistic regression, ERGM provides better fit (Cranmer et al., 2016).

In this analysis, the categorical variable "type of actor" is used to estimate the likelihood of a connection between two actors. A method called "nodal attribute mixing" in the ERGM package in R allows for estimating this probability considering all possible pairs of actor types (e.g. an actor classified as "government" with an actor classified as "NGO"). The location variable is employed as a homophily term. Homophily indicates that two individuals are more likely to develop connections if they have some characteristics in common (McPherson et al., 2001). In this case, the location factor is operationalized as co-location, indicating that two individual actors have the same geographic region in common. Finally, three variables are used to control for the characteristics of the whole network, namely the number of connections ("Edges"), the Geometrically Weighted Degree ("GWDegree"), which represents the sum of degree counts with geometrically decreasing weights, and the Geometrically Weighted Edgewise Shared Partnerships

("GWESP"), which considers the tendency for tied nodes to have multiple shared partners (Hunter, 2007; Hunter & Handcock, 2012). These variables are usually called "structural terms" since they consider the "structure" of the whole network.

We employ the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) as estimators of the models' prediction error, so lower values of AIC and BIC are associated with better fit.

6 Results

6.1 Descriptive statistics

Table 4 shows how the selected actors are distributed by type and location. The types of actors included in the table reflect both their main characteristics as keystone actors – governments, NGOs, the media and the OGD users – and their main roles as they emerged in the ASOC case. The OGD provider type of actor includes 14 Twitter accounts at the national level, comprising the official OpenCoesione and ASOC organizations, as well as individual staff members. Other types include policy makers at the national and regional levels, NGOs, the media, and local communities of OGD users (i.e., the students and their instructors).

Table 4 - Number of actors in the ASOC ecosystem, by type and location

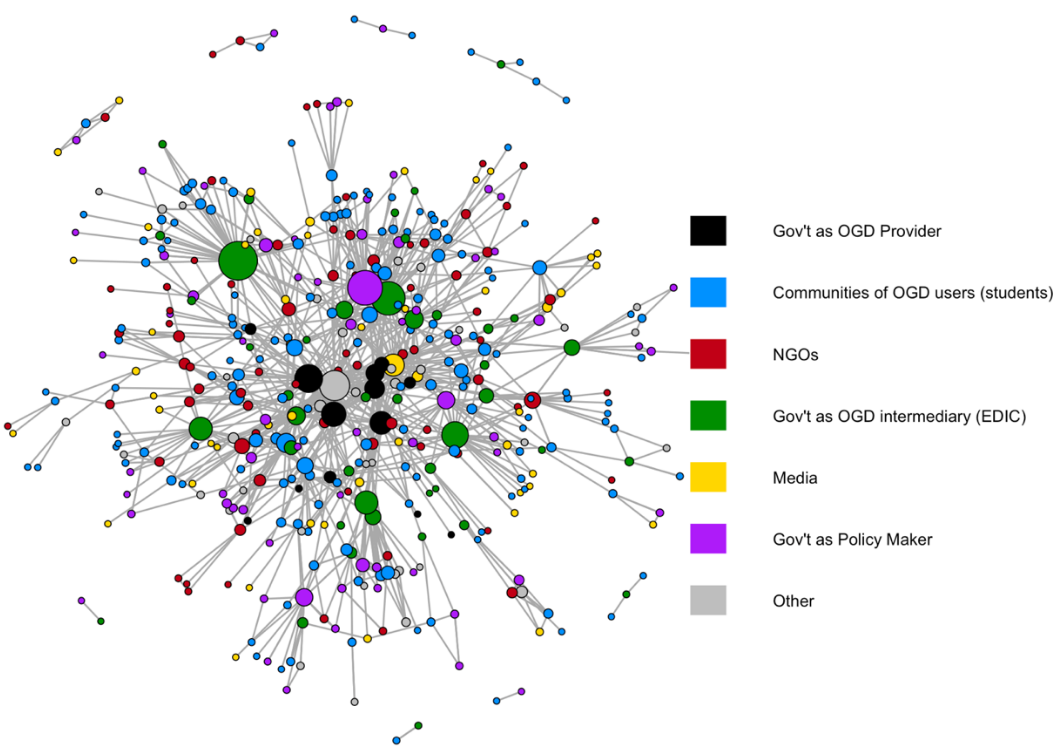
	A - Gov't as OGD provider	B - Gov't as Policy Maker	C- Gov't as intermediary (EDIC)	D- Local com-munities (Students)	E - NGOs	F- Media	G - Other
National level	14	5	4		19	20	7
Abruzzo		3	2	6	1	2	
Basilicata		3	2	2	3	1	1
Calabria		6	6	30	1	4	2
Campania		5	6	27	6	9	4
Emilia-			1	2	2	3	1
Friuli-		1					
Lazio		7	1	7		5	3
Liguria						1	
Lombardia		6	4	16	3	1	2
Marche			1	1			
Molise		2	2	7	4	2	1
Piemonte		5	3	8		1	7
Puglia		4	2	13	1	8	2
Sardegna		6	3	17	3	5	3
Sicilia		7	2	18	4	7	1
Toscana		7	3	8	2	1	3
Trentino						1	
Umbria		1	1	3	3	3	1
Veneto						1	
Total	*14*	*68*	*43*	*165*	*52*	*75*	*38*

Descriptive statistics regarding location of actors show that they mainly operate at the local level. The exception is the data provider, which is a national portal. All types of actors (except the students and their teachers) also include organizations at the central level, such as the Ministry of Education (policy maker) or ForumPA (a media organization based in Rome).

Figure 2 and 3 show how these actors are connected and were created using R. Each actor is represented by a colored circle based on the actor type. The size of the circle is proportional to the number of connections created. In all, 455 actors have developed 1177 connections. The graphs show the presence of a strong core component and a number of peripheral areas connected to the center.

The colors of the circles represented in Figure 2 help distinguish between the case-specific types of actors, which tend to be complementary and mutually interdependent when forming relations. In particular, the OGD provider (ASOC staff) is at the very core of the network, showing moderate to high levels of degree centrality. The EDICs show a major role as information hubs at regional and local levels, connecting different types of users from the center to the periphery of the network. Policy makers, NGOs and the media hold less important central positions. The grey circle very close to the center ("Other" category) represents the Italian national institute of statistics (ISTAT), which is directly involved in the ASOC program for helping students analyze statistical data.

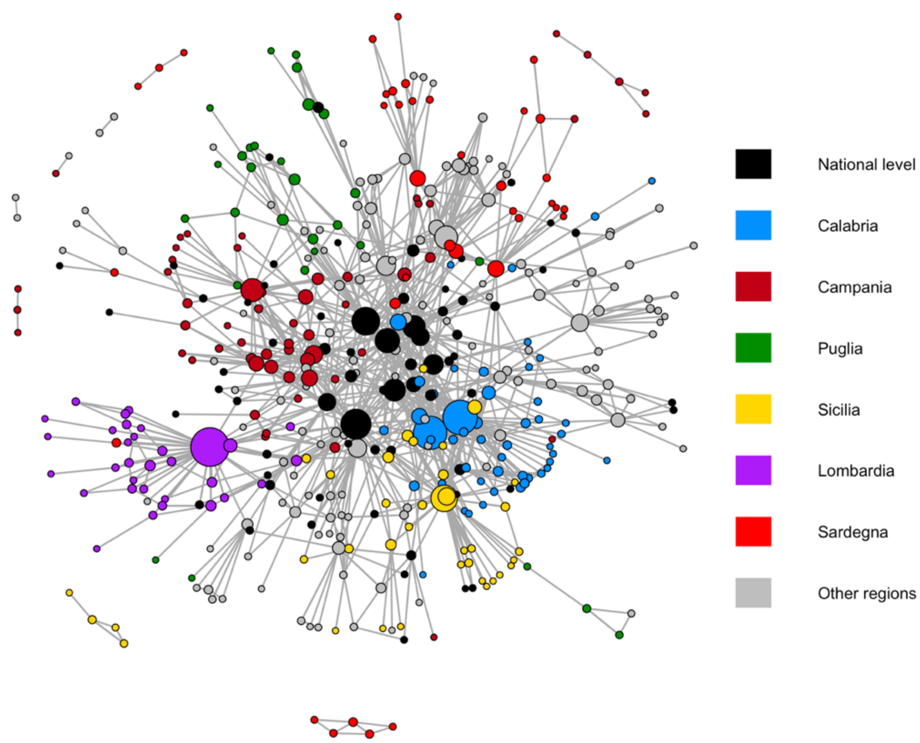
Figure 2 – ASOC Network: Type of actor



Note: Each circle represents an actor in the network (i.e., a Twitter account). Edges between two actors are formed when a shared activity is found by analyzing the Twitter exchange. The size of the circles is proportional to the number of connections of each actor.

Circles in Figure 3 show the location of each actor. Actors at the national level (e.g., OGD provider and other Ministries) hold central positions, connecting with a wide array of actors from different geographical areas. Local actors tend to form clusters that are usually developed around the regional EDICs (e.g., Lombardy, Sicily, Calabria), or the Regional Managing Authorities of EU funds (e.g., Campania, Calabria, Sardinia). The four regions receiving the most funding (namely Sicily, Campania, Calabria, and Puglia) are also the most represented in terms of active actors in the network. Nodes belonging to each region tend to form clusters of contiguous nodes in the represented space.

Figure 3 - ASOC Network: Location



Note: Each circle represents an actor in the network (i.e., a Twitter account). Edges between two actors are formed when a shared activity is found by analyzing the Twitter exchange. The size of the circles is proportional to the number of connections of each actor.

6.2 Modeling

Three ERGMs are estimated and presented in Table 5. Model 1 focuses on the “type of actor” variable, in which every individual actor was classified (see Table 3). The models estimate coefficients based on the following question: “what is the likelihood that a connection will form between two individual actors if the first actor is classified as ‘type x’ and the second as ‘type y’”? In other words, we calculate the odds of a connection between two actors based on their respective types -- all possible pairs of types of actors are considered. Odds ratios are reported in Table 5, as in logistic regression. Model 2 introduces the location variable as a pair of actors having the same region in common. Model 3 adds the “structural terms” to account for the structure of the whole network, as well as the number of followers as an additional control. The values for both the AIC and BIC decrease from Model 1 to 3, indicating that Model 3 is the best-fitting model.

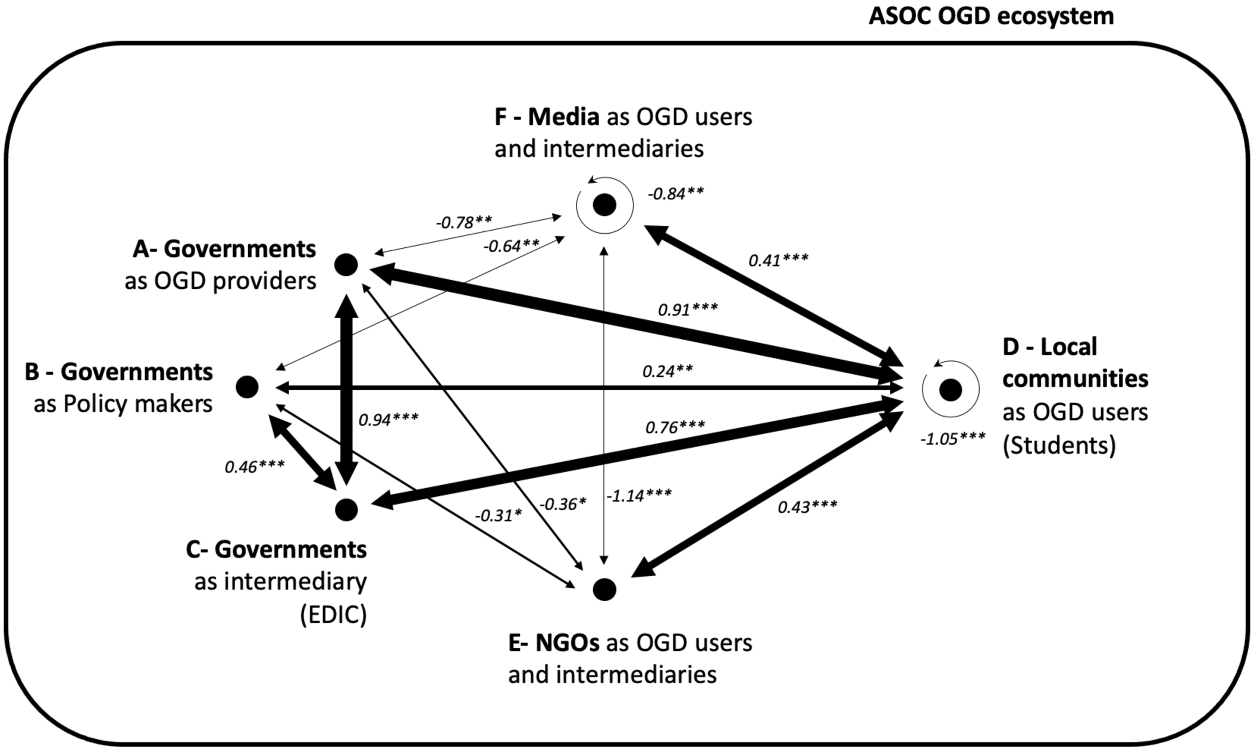
Table 5 – ERGM coefficients (odds ratios)

	Model 1	Model 2	Model 3
Type of actor			
EDIC (C) ↔ EDIC (C)	0.71 (0.25)***	0.91 (0.27)***	-0.12 (0.29)
EDIC (C) ↔ Media (F)	0.10 (0.22)	-0.04 (0.23)	-0.17 (0.22)
Media (F) ↔ Media (F)	-0.72 (0.41)*	-1.55 (0.42)***	-0.84 (0.41)**
EDIC (C) ↔ NGO (E)	0.40 (0.17)**	0.41 (0.18)**	-0.05 (0.16)
Media (F) ↔ NGO (E)	-0.95 (0.27)***	-1.61 (0.29)***	-1.15 (0.27)***
NGO (E) ↔ NGO (E)	0.24 (0.19)	-0.17 (0.2)	-0.05 (0.16)
EDIC (C) ↔ OGD provider (A)	1.75 (0.20)***	1.82 (0.22)***	0.94 (0.21)***
Media (F) ↔ OGD provider (A)	0.28 (0.34)	-0.96 (0.35)***	-0.79 (0.31)**
NGO (E) ↔ OGD provider (A)	0.91 (0.22)***	-0.02 (0.23)	-0.36 (0.19)*
OGD provider (A) ↔ OGD provider (A)	3.12 (0.28)***	1.31 (0.3)***	0.29 (0.25)
EDIC (C) ↔ Policy Maker (B)	0.67 (0.16)***	0.65 (0.17)***	0.47 (0.16)***
Media (F) ↔ Policy Maker (B)	-0.79 (0.27)***	-0.98 (0.28)***	-0.64 (0.27)**
NGO (E) ↔ Policy Maker (B)	-0.33 (0.19)*	-0.46 (0.2)**	-0.32 (0.18)*
OGD provider (A) ↔ Policy Maker (B)	0.46 (0.28)*	0.37 (0.29)	-0.03 (0.26)
Policy Maker (B) ↔ Policy Maker (B)	-0.01 (0.23)	-0.16 (0.24)	0.15 (0.23)
EDIC (C) ↔ Community (D)	0.74 (0.11)***	0.97 (0.13)***	0.76 (0.11)***
Media (F) ↔ Community (D)	-0.16 (0.14)	0.17 (0.15)	0.41 (0.14)***
NGO (E) ↔ Community (D)	0.01 (0.12)	0.32 (0.13)**	0.44 (0.11)***
OGD provider (A) ↔ Community (D)	0.26 (0.20)	1.28 (0.22)***	0.91 (0.19)***
Policy Maker (B) ↔ Community (D)	-0.033 (0.12)	0.08 (0.13)	0.24 (0.12)**
Community (D) ↔ Community (D)	-1.10 (0.17)***	-0.84 (0.19)***	-1.06 (0.18)***
Homophily term			
Same location		3.05 (0.07)***	1.81 (0.06)***
Controls			
Edges	-4.66 (0.08)***	-6.6 (0.19)***	-7.09 (0.17)***
GWDegree (**α**= 0.075)			3.14 (0.46)***
GWESP (**α**= 0.495)			1.69 (0.07)***
No. of followers (log)		0.08 (0.01)***	0.03 (0.01)***
Fit			
AIC	10967	9070	7926
BIC	11177	9300	8194

*** p< 0.01; ** p<0.05; * p<0.10. Standard errors in parenthesis. Baseline for Type of Actor: Other actors. Variables have variance inflation factors of 2.2 or lower, which indicates there is no or very little multicollinearity.

The statistically significant coefficients of Model 3 (odds ratios) are also reported in Figure 4, which shows the relations among the types of actors included in the case. The thickness of the arrows is proportional to the value of the coefficient.

Figure 4 – Statistically significant coefficients (odds ratios) from Model 3 (Type of actor)



*** p< 0.01; ** p<0.05; * p<0.10. Baseline: "Other" Type category.

In ERGM, the coefficients can be interpreted as in logistic regression. In this case, the connection between data provider and data users is one of the strongest. For example, the odds that a government organization acting as an OGD provider will develop a connection with a user community (i.e., students) are 2.5 [1.3-3.7, p<0.01] times higher than a connection to the baseline "other" category. However, the connection between the user community and government as policy maker, while positive, is much weaker (0.24, p>0.01), suggesting that the use of OGD to give feedback to policymakers is less common. Negative coefficients are produced when the likelihood of a connection between two types of actors is lower compared to the baseline “other” actor.

Strong relations are also present between EDICs and OGD providers (0.94, $p>0.01$) and between the EDICs and the students (coefficient is 0.76, $p>0.01$). The EDICs also appear to be more capable of establishing robust connections with policy makers (0.46, $p>0.01$).

OGD users in the case have fairly good probabilities to connect with both the media and NGOs (odds are 51% and 53% higher than the baseline, respectively), but then neither the media nor the NGOs are likely to develop relations with government (coefficients are always negative and statistically significant). A relation between media and NGOs is particularly unlikely.

Policy makers, apart from the EDICs, seem to have good probabilities to connect directly with students (mostly during shared events and interviews). The odds of a connection of this kind are 24% higher than the baseline.

Table 5 also includes the coefficient of the location term, which is a very strong predictor of the likelihood of a connection. In fact, the odds of forming a relation between two actors that share the same location are 6.11 [5.75-7.14, $p<0.01$] times higher than between two actors that operate in different regions or at different levels of government.

The control variable *Number of followers* is positive and statistically significant, as expected. Also, the three structural terms - *Edges*, *GWDegree,* and *GWESP* - are all statistically significant with high coefficients, thus confirming their important role as controls in this analysis. These additional variables reflect the key characteristics of the whole network and help improve the coefficients' estimates as well as the overall fit (Cranmer et al., 2016).

7 Discussion

This study is a first effort to analyze the formation and structure of a specific OGD ecosystem with both qualitative and quantitative data about the relations within the network. In doing so, the analysis reveals some useful ways to describe and potentially evaluate the robustness of an OGD ecosystem in order to identify strengths and weaknesses and consider potential variations and improvements.

The analysis of the ASOC OGD ecosystem confirmed the presence of all "keystone actors" identified in the literature. However, a few expected roles for some of the actors were not found, namely the OGD user role for the media and NGOs, and the beneficiary role for the local communities, since the local communities all use OGD directly from the source. By contrast, in the literature, the beneficiary role derives from the presence of a service or application developed by an intermediary. In addition, governmental organizations emerged from the case as the most important intermediary, contrary to the literature which tends to assign this role to NGOs and the media. In the ASOC case, this role was played by the local centers of the European Commission (EDICs) which acted as the intermediary between the students and the policy makers.

With reference to the first research question, the first hypothesis (H1.1) on the prominent role of the government was largely confirmed. While government organizations combined, were shown to play crucial central roles, not all roles were equally prominent. The OGD provider mainly established a direct relation with data users. The EDICs, by comparison, had higher probabilities to develop multiple connections with the provider, the users, and the policy makers. Among the government actors, policy makers were least connected with other actors. Taken together, these findings confirm that the national government is a crucial stimulator of OGD use (Koppenjan & Klijn, 2004; Martin et al., 2017) but its influence is not consistent across different roles.

The intermediary function was also confirmed to be important (H1.2), but only when governments played this role. The facilitating role of the EDICs was highlighted as strong and statistically significant, while NGOs and the media, which are typically expected to play an important intermediary role, showed only moderate probabilities to develop connections with OGD users. In the case, the EDICs, a country-wide network of regional and local governments acting as EU representatives, receive an annual budget from the European Commission to help students make sense of OGD, and then interact with policy makers and other relevant actors by organizing meetings and public events. Feldman and Khademian (2007) identification of the role of government as "information broker" is consistent with these findings.

By contrast, while NGOs and the media connected with data users, they were not likely to connect with policy makers to stimulate a debate or to engage with authorities responsible for policy implementation. This lack of connection may limit the effectiveness of the whole ecosystem in generating meaningful public debate and participation. It may be that NGOs and investigative or data journalists suffer from capacity and financial sustainability issues (Davies, 2011; Dawes et al., 2016; Harrison et al., 2012; Reggi & Dawes, 2016; Ubaldi, 2013), which makes this type of actor at the same time crucial but under-resourced and fragile. However, it is also possible that NGOs and the media were simply not interested in the projects the students chose to analyze. If this is the case, we could expect that ecosystems formed around different purposes could lead to different actors becoming involved and possibly playing different roles that would affect the network of relationships and outcomes.

Regarding the second research question, Hypothesis 2.1 about location as a factor influencing the probability of a connection was confirmed, as the coefficient in the model is both very high and statistically significant. Sharing the same geographic location seems to have a strong effect on the likelihood that two actors will develop a connection. This result implies that while OGD are disseminated at the national level through a single open data portal, most of the action takes place at the local level. This evidence from the case confirms the existence of multiple ecosystems that stem from the availability of the same data (Harrison et al., 2012). Heimstädt et al. (2014) assert that "ecosystems have to be understood as being nested, with municipal and regional ecosystems at the micro level, national ecosystems on the meso level and one global ecosystem at the macro level. All these ecosystems mainly deal with data from their own level, but they may intersect with the other levels at times" (p. 131). The case demonstrates that such nested ecosystems co-exist using data about local projects that is collected and released at the national level. The case also confirms that OGD users can share information effectively at the local level. For example, physical proximity of the students with local NGOs, media and the EDIC centers allowed the organization of meetings and face-to-face interactions, facilitating the sharing of knowledge and expertise (Gil-Garcia et al., 2016). At this level, user communities likely have greater interest, and a clearer picture of their needs and their possible contributions to improve policymaking (Gilman, 2016).

8 Conclusions

This paper applied social network analysis methods to study the formation and structure of an OGD ecosystem focused on the use of European development funding in Italy. It used Twitter data to identify and select relevant actors and their connections within the A Scuola di

OpenCoesione (ASOC) open data program. The application of these methods was useful to highlight the nature and intensity of the interactions among different types of actors in the ecosystem, showing potential strengths and weaknesses of current OGD ecosystems. The results could be particularly beneficial for academics to model the components and the dynamics of ecosystems, as well as for public agencies when designing OGD programs.

The analysis of this case confirmed that governmental organizations are central actors in open government data ecosystems, as they are represented in the literature. They played strong roles as data providers and intermediaries and a weaker role as policy makers. NGOs and the media were likely to create connections only with data users and seldom with policy makers, and therefore they did not play an intermediary role as expected from the literature. In this specific case, the voices the of OGD user group composed of students, teachers and their communities did reach policy makers. However, policy makers themselves did not play strong roles beyond simply receiving feedback from users either directly or from users assisted by other government organizations acting as intermediaries. Thus, the overall purpose of the ASOC network to enable and encourage civic engagement and public participation in policy making was only partly achieved.

These results suggest several theoretical and practical implications. First, while conceptual work on OGD ecosystems envisions NGOs and the media as the most likely intermediaries linking data users to policy makers for purposes of public participation and citizen feedback, the case indicates that the intermediary *role* is crucial, but it can effectively be played by different actors. In the studied case, this role was played by specialized government organizations created for the purpose of encouraging data use and user feedback to policy makers. Direct government actions to strengthen the intermediary function could be useful to compensate for the weakness of usual intermediaries, especially in the phase of ecosystem creation.

Second, the development of a robust ecosystem that actively engages all the likely keystone actors seems to depend at least in part on its purpose. ASOC was established partly to teach students about data analysis and civic monitoring and in this it seems to have succeeded. However, for an ecosystem to also succeed in influencing public policy making, it seems important that its purpose be tied more directly to that goal. For example, had the students focused on analyzing projects that were already highly debated in the local context they might have stimulated traditional and civic media to develop independent investigations more likely to engage policy makers.

Third, the analysis confirmed that co-location has a strong influence on the formation of network connections, as shown by the existence of multiple ecosystems at the subnational level, connected to the broader national network. One of the main lessons from this case is that effective local data use does not necessarily depend on local data provision. The case shows it is possible to enable and encourage local data use by releasing local-level data from a national-level source. In the ASOC case, local data use was further supported by consistent, proactive strategies that created networks of interested local actors and supported them through government organizations whose purpose is to make connections between users and policy makers.

An increased commitment to create stable policy mechanisms for citizen engagement and policy co-production by national and local government might be an incentive for NGOs to devote more time and effort to participating (Khayyat & Bannister, 2017). Policy mechanisms may include the creation of formal venues for debating the suggestions from NGOs and civic communities about the effectiveness of government spending. For example, in the specific context of the EU policy, the Managing Authorities of EU funds could promote the inclusiveness of the

Monitoring Committees, making them more open to the input from civil society. In addition, governments could address the chronic sustainability issue of non-governmental intermediaries by giving financial incentives for independent, participatory evaluation of public policies by NGOs and others. One example is again from the EU context, where the European Commission has recently funded several local projects to promote civic monitoring of EU projects based on OGD and to discuss the results with the public administrations involved[6].

Finally, based on the limitations of this study, we outline a few ideas for future research. First, this analysis focused on only one ecosystem and one particular type of government data, with limited external validity. Consequently, the nature and intensity of interactions will likely vary in other contexts. Second, Twitter data do not reflect all kinds of information exchanges among the actors. In particular, the media and the policy makers may be less interested in using the specific program hashtags, while policy actors may choose not to use Twitter as frequently as the other actors (Panagiotopoulos & De Widt, 2016). Third, this analysis considered actors and connections across five different years, consolidated in one dataset, thus not accounting for their evolution over time. Fourth, the available data allowed us to consider location as the only factor that influenced the creation of connections.

To address these limitations, the same method could be replicated for other instances in the future, such as other OGD programs in different countries or focusing on different public issues or different types of data or different data providers. For example, a replication could be useful to

[6] This initiative is named "Support for citizen engagement in the implementation of cohesion policy". Eligible recipients of the funding are European NGOs. https://ec.europa.eu/regional_policy/en/newsroom/news/2020/03/18-03-2020-engaging-citizens-in-cohesion-policy-new-call-for-proposals-for-civil-society-organisations-published.

find out whether NGOs and the media are more (or less) involved in other ecosystem information exchanges, and why. A comparison between OGD ecosystems in different contexts would be especially useful. For example, the Italian case could be compared with similar cases at the national level in countries with different administrative traditions and socio-economic characteristics. Replicating this method in the EU would be facilitated by the presence of a common set of mandatory data that all Member States are required to publish in their websites, as well as by the recent development of several ASOC instances in different countries[7]. Such studies would further refine our conceptual understanding of OGD ecosystems that operate in different contexts. To facilitate the use of the same method in other instances of OGD ecosystems and reduce the time-consuming effort, ad-hoc tools could be developed, guiding the users through the different qualitative and quantitative phases of the process.

This kind of analysis could be complemented by qualitative and quantitative data from surveys, interviews with key actors, public sources, and internal government documents to address questions of how and why ecosystems develop and operate, and what might be the critical success factors that shape their ability to persist and achieve relevant goals. Qualitative studies could be especially helpful in further exploring the motivations for policy makers, NGOs and the media to engage in ecosystem activities, their main barriers and challenges, and the potential incentive structure or set of rules that might be created to promote further engagement by all types of actors. Other studies could investigate additional factors that might explain the likelihood of actors making a connection, such as the characteristics, skills and experiences of the actors, or the setting,

[7] https://ec.europa.eu/regional_policy/en/policy/communication/inform-network/asoc

purpose and other characteristics of the projects or policies of interest to them. In addition, a temporal analysis (e.g., temporal ERGM) could be employed to account for the evolution of an ecosystem network over time.

9 Bibliography

Attard, J., Orlandi, F., & Auer, S. (2016). Data driven governments: creating value through open government data *Transactions on Large-Scale Data-and Knowledge-Centered Systems XXVII* (pp. 84-110): Springer.

Baack, S. (2015). Datafication and empowerment: How the open data movement re-articulates notions of democracy, participation, and journalism. *Big Data & Society, 2*(2), 2053951715594634.

Ciociola, C., & Reggi, L. (2015). A Scuola di OpenCoesione: From open data to civic engagement. In J. Atenas & L. Havemann (Eds.), *Open Data as Open Educational Resources* (pp. 26-37). London: Open Knowledge, Open Education Working Group.

Cranmer, S. J., Leifeld, P., McClurg, S. D., & Rolfe, M. (2016). Navigating the Range of Statistical Tools for Inferential Network Analysis. *American Journal of Political Science*.

Davies, T. (2011). Open data: Infrastructures and ecosystems. *Open Data Research*.

Dawes, S. S., Vidiasova, L., & Parkhimovich, O. (2016). Planning and designing open government data programs: An ecosystem approach. *Government Information Quarterly, 33*(1), 15-27.

Feldman, M. S., & Khademian, A. M. (2007). The role of the public manager in inclusion: Creating communities of participation. *Governance, 20*(2), 305-324.

Fleming, L., & Waguespack, D. M. (2007). Brokerage, boundary spanning, and leadership in open innovation communities. *Organization science, 18*(2), 165-180.

Fung, A., Graham, M., & Weil, D. (2007). *Full disclosure: The perils and promise of transparency*: Cambridge University Press.

Fung, A., Russon Gilman, H., & Shkabatur, J. (2013). Six models for the internet+ politics. *International Studies Review, 15*(1), 30-47.

Gascó-Hernández, M., Martin, E. G., Reggi, L., Pyo, S., & Luna-Reyes, L. F. (2018). Promoting the use of open government data: Cases of training and engagement. *Government Information Quarterly, 35*(2), 233-242. doi:https://doi.org/10.1016/j.giq.2018.01.003

Gil-Garcia, J. R., Pardo, T. A., & Sayogo, D. S. (2016). From bureaucratic machines to inter-organizational networks: Characterizing the response to the World Trade Center crisis. *Transforming Government: People, Process and Policy, 10*(4), 568-590.

Gilman, H. R. (2016). *Democracy Reinvented: Participatory Budgeting and Civic Innovation in America*: Brookings Institution Press.

Gupta, A., Panagiotopoulos, P., & Bowen, F. (2020). An orchestration approach to smart city data ecosystems. *Technological Forecasting and Social Change, 153*, 119929.

Harris, J. K. (2013). *An introduction to exponential random graph modeling* (Vol. 173): Sage Publications.

Harrison, T. M., Pardo, T. A., & Cook, M. (2012). Creating Open Government Ecosystems: A Research and Development Agenda. *Future Internet, 4*(4), 900-928. doi:10.3390/fi4040900

Heimstädt, M., Saunderson, F., & Heath, T. (2014). From toddler to teen: growth of an open data ecosystem. *JeDEM-eJournal of eDemocracy and Open Government, 6*(2), 123-135.

Helbig, N., Cresswell, A. M., Burke, G. B., & Luna-Reyes, L. (2012). *The dynamics of opening government data*. Retrieved from https://ctg.albany.edu/media/pubs/pdfs/opendata.pdf

Hofer, M., & Aubert, V. (2013). Perceived bridging and bonding social capital on Twitter: Differentiating between followers and followees. *Computers in Human Behavior, 29*(6), 2134-2142.

Hunter, D. R. (2007). Curved exponential family models for social networks. *Social networks, 29*(2), 216-230.

Hunter, D. R., & Handcock, M. S. (2012). Inference in curved exponential family models for networks. *Journal of Computational and Graphical Statistics*.

Hunter, D. R., Handcock, M. S., Butts, C. T., Goodreau, S. M., & Morris, M. (2008). Ergm: A package to fit, simulate and diagnose exponential-family models for networks. *Journal of Statistical Software, 24*(3).

Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information Systems Management, 29*(4), 258-268.

Janssen, M., & Helbig, N. (2016). Innovating and changing the policy-cycle: Policy-makers be prepared! *Government Information Quarterly*.

Janssen, M., & Zuiderwijk, A. (2014). Infomediary business models for connecting open data providers and users. *Social Science Computer Review, 32*(5), 694-711.

Jones, C., Hesterly, W. S., & Borgatti, S. P. (1997). A general theory of network governance: Exchange conditions and social mechanisms. *Academy of Management Review, 22*(4), 911-945.

Kassen, M. (2018). Open data and its institutional ecosystems: A comparative cross-jurisdictional analysis of open data platforms. *Canadian Public Administration, 61*(1), 109-129.

Khayyat, M., & Bannister, F. (2017). Towards a model for facilitating and enabling co-creation using open government data. *Information Polity, 22*(4), 211-231.

Koppenjan, J., & Klijn, E.-H. (2004). *Managing uncertainties in networks: Public private controversies*: Routledge.

Lassinantti, J., Ståhlbröst, A., & Runardotter, M. (2019). Relevant social groups for open data use and engagement. *Government Information Quarterly, 36*(1), 98-111.

Lo Russo, M. (2016). Who is using open data on cohesion policy, how and why? *Collana Materiali UVAL*(34).

Martin, S., Turki, S., & Renault, S. (2017). *Open Data Ecosystems.* Paper presented at the International Conference on Electronic Government and the Information Systems Perspective.

McBride, K., Aavik, G., Toots, M., Kalvet, T., & Krimmer, R. (2019). How does open government data driven co-creation occur? Six factors and a 'perfect storm'; insights from Chicago's food inspection forecasting model. *Government Information Quarterly, 36*(1), 88-97.

McBride, K., Olesk, M., Kütt, A., & Shysh, D. (2020). Systemic change, open data ecosystem performance improvements, and empirical insights from Estonia: A country-level action research study. *Information Polity*(Preprint), 1-26.

McGee, R., Edwards, D., Hudson, H., Anderson, C., & Feruglio, F. (2018). *Appropriating technology for accountability: messages from Making All Voices Count*. Retrieved from https://opendocs.ids.ac.uk/opendocs/bitstream/handle/123456789/12837/MAVC_CIPGv2_Final_online.pdf

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology, 27*(1), 415-444.

Meijer, A., & Potjer, S. (2018). Citizen-generated open data: An explorative analysis of 25 cases. *Government Information Quarterly, 35*(4), 613-621.

Mergel, I., & Bretschneider, S. I. (2013). A three-stage adoption process for social media use in government. *Public Administration Review, 73*(3), 390-400.

Nardi, B. A., & O'Day, V. (1999). *Information ecologies: Using technology with heart*: Mit Press.

Oliveira, M. I., Barros Lima, G. d. F., & Farias Lóscio, B. (2019). Investigations into Data Ecosystems: a systematic mapping study. *Knowledge and Information Systems*.

Panagiotopoulos, P., & De Widt, D. (2016). Digital Networks in Public Administration: The Case of #Localgov. In Scholl et al. (Ed.), *Electronic Government: 15th IFIP WG 8.5 International Conference, EGOV 2016, Guimarães, Portugal, September 5-8, 2016, Proceedings* (pp. 277-286). Cham: Springer International Publishing.

Pateman, C. (1970). *Participation and democratic theory*: Cambridge University Press.

Polzer, J. T., Crisp, C. B., Jarvenpaa, S. L., & Kim, J. W. (2006). Extending the faultline model to geographically dispersed teams: How colocated subgroups can impair group functioning. *Academy of Management Journal, 49*(4), 679-692.

Reggi, L., & Dawes, S. (2016). Open Government Data Ecosystems: Linking Transparency for Innovation with Transparency for Participation and Accountability. In Scholl et al. (Ed.), *Electronic Government: 15th IFIP WG 8.5 International Conference, EGOV 2016, Guimarães, Portugal, September 5-8, 2016, Proceedings* (pp. 74-86). Cham: Springer International Publishing.

Ruijer, E., Dingelstad, J., & Meijer, A. (2021). Studying complex systems through design interventions probing open government data ecosystems in the Netherlands. *Public management review*, 1-21.

Ruijer, E., Grimmelikhuijsen, S., & Meijer, A. (2017). Open data for democracy: Developing a theoretical framework for open data use. *Government Information Quarterly, 34*(1), 45-52.

Ruijer, E., Grimmelikhuijsen, S., van den Berg, J., & Meijer, A. (2018). Open data work: understanding open data usage from a practice lens. *International Review of Administrative Sciences*, 0020852317753068.

Scott, T. A., & Thomas, C. W. (2017). Winners and losers in the ecology of games: Network position, connectivity, and the benefits of collaborative governance regimes. *Journal of Public Administration Research and Theory, 27*(4), 647-660.

Slobodova, O., & Becker, S. (2020). Zooming into the ecosystem: agency and politics around Open Data Platforms in Lyon and Berlin. *Frontiers in Sustainable Cities, 2*, 20.

Ubaldi, B. (2013). *Open Government Data: Towards an Empirical Analysis of OpenGovernment Data Initiatives*. OECD.

Van Schalkwyk, F., Willmers, M., & McNaughton, M. (2015). Viscous open data: The roles of intermediaries in an open data ecosystem. *Information Technology for Development*, 1-16.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications* (Vol. 8): Cambridge university press.

Yang, T.-M., & Wu, Y.-J. (2021). Looking for datasets to open: An exploration of government officials' information behaviors in open data policy implementation. *Government Information Quarterly*, 101574.

Zuckerman, E. (2014). New media, new civics? *Policy & Internet, 6*(2), 151-168.

Zuiderwijk, A., Helbig, N., Gil-García, J. R., & Janssen, M. (2014a). Special Issue on Innovation through Open Data: Guest Editors' Introduction. *Journal of Theoretical and Applied Electronic Commerce Research, 9*, i-xiii.

Zuiderwijk, A., Janssen, M., & Davis, C. (2014b). Innovation with open data: Essential elements of open data ecosystems. *Information Polity, 19*(1, 2), 17-33.