EJHC | European Journal of Health Communication

ISSN 2673-5903
ejhc.org

Article *ECREA Aarhus Best Papers Special Issue*

Challenges in Communicating Public Health Data

An Analysis of Italian Regions' Social Media Use During the Covid-19 Pandemic on Facebook

DOI: 10.47368/ejhc.2024.202
2024, Vol. 5(2) 23-43

Gea Ducci
Department of Communication Sciences, Humanities and International Studies, University of Urbino Carlo Bo, Italy

Alessandro Lovari
Department of Political and Social Sciences, University of Cagliari, Italy

Nicola Righetti
Department of Communication, University of Vienna, Austria

Abstract

During the Covid-19 pandemic, public sector organizations have rapidly increased their use of social media platforms to directly communicate with citizens regarding various aspects of the crisis. Given the critical importance of epidemiological data during this period, this study conducts a quantitative analysis of the official Facebook channels of the 20 Italian regions in the initial emergency phase to explore the role of data communication. It employs computational methods for automated classification of the prevailing types of data communication on Facebook posts and a random-intercept negative binomial model to analyze their different impact on engagement. The findings reveal that the most common types of posts incorporate data within the message, either alone or accompanied by a link to the official website. Infographics are also commonly used. Furthermore, the most comprehensive posts, featuring data, a link to a website, and an infographic, had the highest positive impact on engagement. Overall, the study highlights a significant diversity in the way of communicating epidemiologic data, potentially leading to disparities among Italian citizens in receiving information from institutions about the spread of the virus. This poses substantial challenges for public health communication directed at citizens and the relationships between the national and local levels.

Corresponding author:
Gea Ducci: gea.ducci@uniurb.it

Keywords
Health data communication, public sector communication, social media, Covid-19 pandemic, engagement, Italy.

The Covid-19 pandemic has prompted a renewed focus on the strategic role of public sector communication in addressing the emergency (OECD, 2021). Alongside the health crisis, a communication crisis has also emerged due to the proliferation of an infodemic and the increase in information disorders, which pose a significant risk of spreading disinformation (Levandowski et al. 2020). Amidst this environment of information disorder, national and global health authorities, scientists, medical professionals, and other health experts in many countries have maintained high levels of trust even a year after the pandemic began (Ducci et al., 2022; Nielsen et al., 2021; Kessler et al., 2022). Therefore, strategies that foster coordination among institutions at the local, national, and supranational levels are crucial to ensure unambiguous and visible institutional voices through the implementation of integrated and inclusive communication approaches.

The ability of institutions to project a unified voice significantly impacts their credibility and the likelihood of citizens sharing and accepting the provided guidelines (Coombs, 2020), thus nurturing a significant degree of trust. The consistency and coherence of institutional communication should be reflected in all implemented communication activities, particularly in the management of communication within digital environments and social media (Ducci & Lovari, 2021; OECD, 2021).

Many scholars have highlighted that social media has presented new challenges for crisis communication, epidemics, and health-related issues (Guidry et al., 2017, 2023; Jin & Austin, 2022; Moorhead et al., 2013; Sastry & Lovari, 2017). The domestication of social media by public institutions seems to have reached a certain level of maturity and institutionalization (Mergel & Bretscheinder, 2013). However, public health institutions have often been slow to respond to health concerns on social media due to a lack of vision, competencies, and/or skills to strategically engage and communicate with citizens (Guidry et al., 2017, 2023; Tirkkonen & Luoma-aho, 2011). The use of social media by institutions during the pandemic has been accelerated and shown both advantages and disadvantages in different countries (La Rocca et al., 2023; Vaagan et al., 2021). Their use has enabled rapid information dissemination to citizens, addressing various communication needs (Lovari, Bowen, & D'Ambrosi, 2020), including the periodic communication of emergency rules, data communication on virus spread, scientific discoveries, and medical practices, counteracting misinformation, adopting effective strategies to counteract fake news, such as the debunking activities (Chan et al., 2017; Lewandowsky et al., 2020; Memenga et al., 2022); promoting preventive behaviors, providing information about health services, addressing psychological and social crises, and empowering citizens (Lovari, Ducci, & Righetti, 2021).

There are also disadvantages and risks that include increased criticism of institutional activities, reliance on private platforms in the public sector, and the potential influence of platform logic on public-sector communication and healthcare (Locatelli & Lovari, 2021; Lovari & Ducci, 2021). Additionally, institutions may exacerbate inequalities due to the digital divide and the ambiguity of information shared online (Coombs, 2020; Lovari & Righetti, 2020; van Dijck, 2020).

As the Covid-19 pandemic has prompted public sector organizations, including healthcare-related entities, to enhance their digital communication skills and improve public health communication, it has also highlighted significant differences in how institutions communicate with citizens. These disparities reflect the varying speeds of digitization among organizations and the lack of homogeneous communication cultures within the public sector.

In the pandemic context, one of the novel factors has been the extensive use of epidemiological data related to the spread of the virus, the number of patients infected and cured. Data became central to communication strategies especially in the early phase of the Covid-19 pandemic, prompting supranational organizations (i.e., WHO), governments and health organizations at different levels to devote attention to this type of communication format in their digital and social media communication.

Data communication in crisis and emergency situations related to health issues in the public sector is a topic still partially explored by scholars. In this context, this article aims to investigate the use of data for public health communication during a pandemic addressing this gap in health communication studies.

The Importance of Data Communication in Health

One of the aspects that has characterized pandemic communication and have been the subject of intense public debate is the communication of epidemiological data regarding the spread of the virus, which has been practiced consistently by all institutions (international, national, local) since the beginning of the outbreak (February-March 2020).

There are several aspects to consider when communicating data concerning health and healthcare. On the one hand, an important issue is intrinsic to the data as the object of communication. That is, there are issues related to the criteria for acquiring/producing data and all aspects of this process, including the reliability (what kind of data, how they are collected, the criteria by which they are acquired and classified, and labeled, thus also the criteria for signifying the data themselves) (Mira & Massarenti, 2020).

If these issues concern the object of communication, from a more strictly communicative point of view, the manner of communication is fundamental (Watzlawich et al., 1967), which pertains to how data (once acquired and "classified" by the scientific community) are made visible, disseminated, represented, and contextualized. According to Parrott (2009), data and numbers represents one of the two strategies and practices adopted to talk about health, together with stories and narratives. Data and statistics are mainly used to respond to specific questions (like "how many", "how common is it", and "how often") to allow us to compare with some standards and to predict health risks (Lipkus, 2007). This type of information is neutral and impersonal; the challenge lies in understanding what numbers mean and their implications for citizens' lives (Parrott, 2009).

Communicating health-related data with citizens and the lay publics is complicated and challenging due to low levels of statistical, scientific, and health literacy, people's existing beliefs and biases, and the presence of cognitive limitations in the audiences (Nelson et al., 2012; Woloshin et al., 2008). At the same time, publishing and sharing data are important to inform the population about healthcare-related issues and emerging health phenomena and to provide the rationale behind health recommendations and scientific understanding of diseases and access to healthcare services.

Scholars have pointed out that communicating scientific and health data to citizens is strategic to making decisions about immunization policies and treatments and to making transparent community exposure to potential disease outbreaks. Moreover, data and numbers can provide evidence to justify policy interventions or facilitate informed decision-making (Nelson et al., 2009; Parvanta et al., 2011; Rothman & Kiviniemi, 1999). Additionally, from another perspective, sometimes people prefer not to be exposed, or avoid, health data that contradicts their beliefs, attitudes, or behaviors, as they do with written messages, favoring partisanship and polarizations among individuals (Bennett & Iyengar, 2008).

Furthermore, data-based decisions are central to guiding medical practice. They also involve politicians and public managers in the case of health issues that can impact the general population or a specific territory.

The Covid-19 pandemic immediately showed the importance of global, national, and regional data and communication. In the face of an unknown and threatening virus that heavily affected citizens' health, public health organizations, and healthcare services, data and numbers were among the few elements that could orientate citizens, facilitate their knowledge, and allow comparisons among countries and regions. This task was particularly challenged by the specificities of contemporary communication ecologies, characterized by the pervasive role of digital platforms, the growing development of infodemic (WHO, 2020), and the politicization of the virus in many countries (Lilleker et al., 2021; Lovari, 2020).

In this context of infodemic disorder (La Rocca et al., 2023), WHO, Ministries of Health, and public health organizations started to frequently inform the population about epidemiological data (e.g., via online press conferences) and to post them on their social media channels. It is clear that, in social media, it is important to pay attention to visual data communication since an effective visualization of numbers (e.g., using infographics, emoticons, social cards, videos, etc.) can foster the understanding of patterns in data across health, public policy and science, as well as effectively give evidence about statistics in the population (Guidry et al., 2020, 2021, 2023). At the same time, an ineffective design by practitioners and limited graphical literacy (inside the organizations) can impact online users' understanding of data (Franconeri et al., 2021).

In general, visual communication is a key skill needed by communicators to convey important public health information (Anderson et al., 2019). In particular, scholarly literature highlighted an increasing presence and effectiveness of infographics in the social media channels managed by public health organizations (Guidry et al., 2020; Lovari & Righetti, 2020), and their use for scientific communication for Covid-19 topics, producing a positive impact in terms of public engagement and perceptions on social media (Lee et al., 2022; Solito & Materassi, 2021).

Considering the strategic role that public health communication, including through social media platforms, has played during the pandemic, and specifically the importance of data communication on Covid-19, this article focuses on Italy. It explores the communication efforts of the regions on Facebook during the initial phase of the pandemic. Italy is an interesting case study as it was the first European country to confront the swift and devastating spread of the Covid-19 pandemic in the early months of 2020.

Six Types of Data Communication by Italian Regions on Facebook. A Quantitative Study: Aim and Research Questions

In Italy, the use of social media in public sector communication has experienced a significant acceleration during the pandemic (Lovari & Righetti, 2020; Ducci & Lovari, 2021), at both the national and local levels. This is particularly notable due to the highly decentralized and regionally focused nature of the country's public health system (Lovari, Ducci, & Righetti, 2021). The organization of health structures and services is influenced by local health systems in each region. Although national health rules take precedence during emergencies, regional differences in the management of health communication activities (Sendra et al., 2021) are still evident across various levels, including regional authorities, councils, health departments, local health authorities, hospitals, research centers, and care centers.

Given the regions' central role in communicating with citizens during the pandemic, examining their communication on official social media pages is crucial for understanding how communication functions within the Italian public health system. The lack of homogeneity is also evident in how regions have communicated various aspects of the pandemic, especially regarding data related to the spread of the virus.

A previous analysis (Lovari, Ducci, & Righetti, 2021) revealed that 78.8% of public health communication on Facebook predominated with Covid-19 topics, with particular emphasis on epidemiological data (24.6%) and the national and regional regulations addressing the emergency (19.5%). However, posts related to other topics, such as fake news and citizens' behaviors to prevent contagion, generated the highest levels of engagement, even though the Italian regions dedicated fewer posts to these themes on the platform. Notably, there were variations among regions in the communication approaches employed for fake news and epidemiological data.

Simultaneously, the two topics that received the most specific posts from the regions – epidemiological data and emergency regulations – yielded the lowest levels of engagement (likes, reactions, comments, and shares), with a relatively uniform pattern across all 20 regions.

Building upon the aforementioned research findings and the substantial production of posts on epidemiological data by Italian regions, a subsequent qualitative-exploratory study (Ducci, 2021) was conducted to understand the different ways these institutions communicated these data. The term "epidemiological data" generally encompassed the number of deaths, the number of infected people (symptomatic/asymptomatic individuals), patients hospitalized in semi-intensive and intensive care units, and swabs performed.

According to the study, it is clear that regions in Italy have adopted the communication practices of the Italian government and the Ministry of Health by providing daily news bulletins on the spread of the virus or significant changes in data related to affected individuals. Although the approach to communicating on Facebook varies by region, some recurring criteria can be identified.

In this study, Ducci (2021) identified six principal communication types (see Tables 1 and 2 for examples):

Type 1: This type involves presenting the data as a list within the post's text, followed by an infographic that illustrates virus's spread trend.

Type 2: Posts under this category do not contain any data but instead provide a link to the regional website, where detailed information and/or downloadable PDF files can be found.

Type 3: In this type, data are listed in the post's text and are accompanied by a link to the regional website for further information.

Type 4: These posts offer a detailed description of the data in the text, paired with an infographic (ranging from simple to sophisticated). A link to the regional website is also provided at the end of the post.

Type 5: Posts of this type are fully self-contained, featuring both a comprehensive description of the data and an infographic that neatly represents the data, without any external references or links to websites.

Type 6: This final type is characterized by posts where all information is exclusively conveyed through an infographic, with no additional text.

Table 3 summarizes the six types of posts identified in Ducci 2021, based on the following features: presence or absence of data in the text of the post (in the form of a list of data, or data inserted in a more or less articulated text); presence or absence of a more or less complex infographic, in which the data are variously represented; presence or absence of a link to the institution's official website in which the updated data appears.

Given the high level of attention to epidemiological data and based on the findings of previous studies (Ducci, 2021; Lovari, Ducci, & Righetti, 2021), we conducted a quantitative analysis to determine the extent to which Italian regions have adopted the different types of communication for Covid-19 related data on Facebook posts, as well as the levels of engagement they have generated.

Based on this context, the research questions are:

> *RQ1:* What are the prevailing types of Facebook posts adopted by regions to communicate data related to the Covid-19 pandemic?
>
> *RQ2:* Are different types of posts associated with different levels of engagement?

Method

In the initial phase of the pandemic, a mixed-methods study examined the public health communication strategies deployed on the official Facebook pages of Italy's 20 regions. This investigation focused on the quantity and nature of Facebook posts, particularly those discussing Covid-19 topics, as well as the level of engagement they received (Lovari, Ducci, & Righetti, 2021). The analysis encompassed 7,880 Facebook posts published by the official pages of these regions from January 30, 2020, the date the first Covid-19 case was confirmed in Italy, to May 3, 2020, which marked the end of the nationwide lockdown and the commencement of the so-called "Phase 2" of the pandemic.

The decision to analyze communication on Facebook is based on several reasons: first, Facebook is the social network site most used by Italians, with a balanced distribution among the age groups of the population (We are Social, 2022); second, it is the social network site most adopted and institutionalized by public sector organizations and in particular those in the health sector (Lovari, Ducci, & Righetti, 2021; Sendra et al., 2021); and third, the availability of reliable data from the platform.

Data from posts, including elements such as the text of the messages, publication dates, engagement metrics, and content types, were collected using CrowdTangle, a social analytics tool owned by Facebook (CrowdTangle Team, 2020). Posts pertaining to Covid-19 were coded by two researchers, who classified them according to thematic categories derived from prior

studies on public health communication (Cioni & Lovari, 2014) and guidelines on pandemic communication provided by the Italian Ministry of Health (Lovari, 2020). This coding process achieved good inter-coder reliability. For a comprehensive description of the data collection and coding procedures, we refer readers to Lovari, Ducci, & Righetti (2021).

Table 1. Examples of Facebook Posts for Data Communication Types n. 1-3.

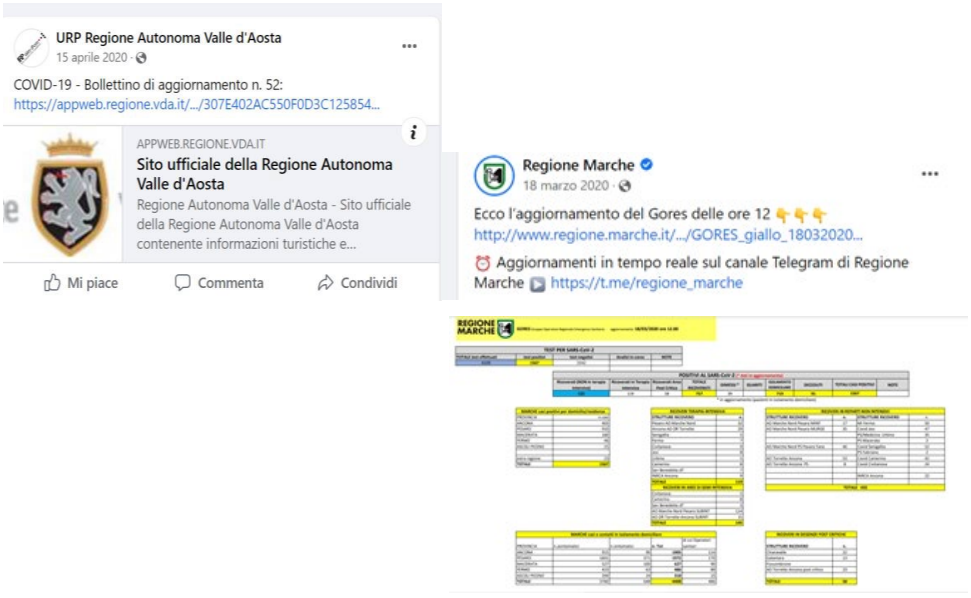
Types	**Examples of Facebook Posts**
Type 1	
Type 2	
Type 3	

Table 2. Examples of Facebook Posts for Data Communication Types n. 4-6.

Types	Examples of Facebook Posts
Type 4	
Type 5	  
Type 6	

Table 3. Post Types Based on the Main Features Identified During the Qualitative Analyses

	Data in the Message		Infographic	Link to the Website
	Data List	Data Description		
Type 1	X			
Type 2				X
Type 3	X			X
Type 4		X	X	X
Type 5		X	X	
Type 6			X	

In the present study, we extracted posts specifically concerned with epidemiological data related to Covid-19 from the original dataset, resulting in a subset of 1,527 posts, which constitutes 19% of the complete dataset. Based on the qualitative analysis of Italian regions' communication about Covid-19 data on Facebook (Ducci, 2021), we operationalized the types of social media posts to measure their prevalence, distribution across regions, and engagement. Specifically, we categorized the types of posts using a combination of three primary elements: the presence of an infographic, the inclusion of a link to an institutional website offering detailed epidemiological information, and the usage of specific epidemiological terminology within the post's text. This third criterion is intended to capture posts that include textual references to quantitative data in the text of the posts. We will now describe the method used to identify these elements, as informed by the annotated dataset detailed in Lovari, Ducci, & Righetti (2021).

We identified posts including infographics if they were related to epidemiological data and were coded as containing infographic images or information related to data and percentages.

The procedure for identifying links to regional or provincial sites consisted of several steps. In the first step, links were extracted using information provided by CrowdTangle or by extracting links from the text of posts. Shortened links (e.g., using services such as bit.ly) were restored to their canonical form. Domains were then extracted from the links and checked to see if they included the terms "region" or "province", or, in the case of Tuscany, the domain of the Regional Council Information Agency. This method exploits the fact that official regional and provincial websites mention the name of the region or province (e.g., "https://www.regione.lombardia.it", "https://www.ufficiostampa.provincia.tn.it", or "http://pti.regione.sicilia.it"). Regarding links to the region's websites, it should be noted that links themselves do not include much information on the actual content of the website pages to which they refer. However, it is reasonable to assume that the content of these website pages is related to epidemiological information. This assumption was validated by manually verifying a random 5% sample of the links to official websites, confirming their relation to epidemiological data. We could not verify the content of only four of the sampled URLs, as the linked pages were no longer online.

The incorporation of epidemiological data within the post represents the third and final primary component of social media posts focused on data communication. For this purpose, a list of terms related to data communication was compiled starting from a series of words typically used in Covid-19 data communication (e.g., "positives", "negatives", "tests", "deaths", "recoveries", etc.) and then identifying additional similar terms used in the texts. The

similarity between words was calculated using the "word embeddings" technique (Mikolov et al., 2013). Word embeddings are "a recently popular representation of language in modern natural language processing. The embeddings learn a vector for each word in a low-dimensional space that predicts the kinds of words that are nearby" (De Marchi & Stewart, 2020, p. 304). The list of the most similar terms extracted in this way was manually verified and applied to the messages to automatically categorize them. The final list included the following words along with their grammatical variations: "cases", "registered", "dead", "death", "infected", "percentage", "trend", "share", "deceased", "passed away", "test", "positive", "negative", "variation", "exceed", "decrease", "high", "rise", "counted", "charts", "hospitalized", "sick", "healed", "discharged", "discharged", "confirmed", "ascertained", and "total". Overall, the described procedure allowed us to identify and quantify the prevalence of different types of posts in the communication of Italian institutions about Covid-19 and answer the first research question.

Using the presence/absence of each of the three elements (infographic, link, message on epidemiological data), we then defined different types of posts, similar to Ducci's (2021) typology, as defined in Table 3. We validated the approach by extracting a stratified sample, segmented by type of post, which comprised about 10% ($N = 162$) of the total. We then manually verified the accuracy of the classification. The results were satisfactory, with an 88% correct identification rate for post types across various categories. The worst-performing categories achieved a 73% accuracy rate, while the best-performing category achieved a 96% rate of correctly classified posts. More specifically, the identification rates for each category are as follows: Data only in the message (Type 1): 96%; Only link to website (Type 2): 73%; Data in the message and link to the website (Type 3): 94%; Data in the message with link to the website and infographic (Type 4): 94%; Data in the message and infographic (Type 5): 88%; Only infographic (Type 6); 84%; Others: 86%. Validation was conducted by a single coder. The process entailed confirming the presence or absence of straightforward, objectively identifiable elements, such as the existence or non-existence of a link or image in a post. Therefore, the usual context of potential interpretive ambiguity, which necessitates the use of multiple coders and the computation of intercoder agreement, did not apply.

We used this data to answer both research questions. With references to the second research question about the level of engagement of the different types of posts, we fitted a negative-binomial mixed model to predict engagement with the identified types of posting. We defined post engagement as the sum of the total interactions on a post (likes, shares, comments, and other Facebook interactions, emojis representing love, wow, haha, sad, and angry). Social media engagement is a count variable with skewed distribution, meaning that there are few posts totaling a high number of interactions, and most have a low level of engagement. We modeled this variable using a negative binomial distribution, a common choice for skewed social media data (e.g., Eberl et al., 2020). We included the total number of likes on the different Facebook pages as a control variable in the model. Page likes were included in the model to account for the fact that the number of interactions is likely correlated to the number of followers of a page, with more followed pages being more likely to reach more people and receive a higher level of interaction on their posts. The page likes variable was measured using the variable likes at posting provided by CrowdTangle, which measures the number of users that liked a page when posting a message. As this measure is often incomplete and includes missing values due to the limitations in the CrowdTangle system, we imputed the missing observations using the average number of likes of the page across the entire observation period.

We also accounted for potential heterogeneity in the levels of engagement in the different geographical macro-areas by including their effect using a random intercept model. The random intercept accounts for possible systematic variation in posts' engagement that might be due to factors correlated to the three geographical areas. Including such an effect permits a more accurate assessment of the specific effect of the types of messages on posts' engagement while controlling for possible systematic geographical variations. All the analyses were performed using the R software (R Core Team, 2022), and the random intercept negative binomial regression model was fitted using the function *glmer.nb* of the library *lme4* (Douglas et al., 2015).

Results

The analysis, conducted at the macro-area level (North, Central, South Italy), focused on posts about epidemiological data related to the Covid-19 pandemic (n = 1,527). Figure 1 shows the proportions of types in the three macro-areas, and Table 4 reports the count of posts for each type and the relative frequencies by each area.

In the dataset, the most common types are those that feature data prominently in the message. The data only in the message (Type 1) and data in the message and link to the website (Type 3) categories each constitute 22% of the total posts, almost half when combined (44%). Following these are the data in the message and infographic (Type 5) and only link to website (Type 2) categories, representing 15% and 13%, respectively. Posts containing infographics – either in combination with data and a website link (Type 4 or Type 5) or solely as the primary content (Type 6) – are relatively less common, comprising 8% (Type 4), 15% (Type 5), and 9% (Type 6) of the dataset, but they still form a significant portion (32%) of the content. The other category makes up 11% of the posts and includes heterogeneous content types not captured by the six primary categories, such as videos, posts linking to the Health Ministry, and posts sharing screenshots of bulletins with Covid-19 case counts.

Figure 1. The Six Types of Posts About Epidemiological Data in the Italian Regions (Total Count)

Table 4. Count of Post Types by Macro-Area and Their Relative Frequencies

Type of Post	Northern Italy *n* = 346	Central Italy *n* = 455	Southern Italy *n* = 726	Total *N* = 1,527
Data only in the message (Type 1)	114 (33%)	113 (25%)	111 (15%)	338 (22%)
Only link to website (Type 2)	41 (12%)	68 (15%)	92 (13%)	201 (13%)
Data in the message and link to the website (Type 3)	127 (37%)	201 (44%)	9 (1.2%)	337 (22%)
Data in the message with link to the website and infographic (Type 4)	2 (0.6%)	3 (0.7%)	120 (17%)	125 (8%)
Data in the message and infographic (Type 5)	41 (12%)	8 (1.8%)	177 (24%)	226 (15%)
Only infographic (Type 6)	1 (0.3%)	4 (0.9%)	132 (18%)	137 (9%)
Other	20 (5.8%)	58 (13%)	85 (12%)	163 (11%)

Note. The total number of posts about epidemiological data by type and by area is also indicated.

Examining the distribution of post types across the three geographical areas—Northern Italy, Central Italy, and Southern Italy—reveals notable heterogeneity in how Covid-19 data is communicated. In Northern and Central Italy, the focus is predominantly on posts that include data either in the message itself or with a link to a website. Specifically, data only in the message (Type 1) and data in the message and link to the website (Type 3) make up 28% and 41% of the posts in these areas, respectively. Contrastingly, Southern Italy presents a different communication pattern. Infographics play a significant role here, with data in the message with link to the website and infographic (Type 4), data in the message and infographic (Type 5), and only infographic (Type 6), together accounting for nearly 59% of the posts of the area. These observations underscore the varied approaches to Covid-19 communication across Italy and hint at regional preferences, whether it be for data-heavy posts in the North and Central areas or more visual, infographic-oriented posts in the South.

We employed a negative binomial mixed-effects model to investigate the impact of various post types and the number of likes at posting on engagement levels. The model incorporated random intercepts for different areas and successfully explained a substantial proportion of the variability in engagement (marginal R^2 = .626, conditional R^2 = .670).

The reference category for the post type comparison is only infographic. This type of post was chosen because of the attention that the scholarly literature has placed on the topic regarding the increasing presence and effectiveness of infographics in the social media of public health organizations (Guidry et al., 2017, 2020; Andersen et al., 2019; Lovari & Righetti, 2020; Lee et al., 2022). The model expresses coefficients as Incidence Rate Ratios (IRR). An IRR of 1 means no engagement difference compared to the reference, while an IRR greater than 1 or less than 1 indicates higher or lower engagement, respectively. The magnitude of the IRR shows the extent of this difference.

Within this model, posts featuring data only in the message (Type 1) exhibited an engagement that was, on average, approximately 1.64 times higher than the only infographic

reference category (IRR = 1.64, 95% CI [1.31-2.04], $p < .001$). Posts that contain only a link to a website (Type 2) had an engagement that was, on average, approximately 0.49 times (51% less) that of the reference category (IRR = 0.49, 95% CI [0.39-0.61], $p < .001$). The most comprehensive posts, featuring data, a link to a website, and an infographic (Type 4) had the highest positive impact on engagement, approximately 3.98 times higher than the reference category (IRR = 3.98, 95% CI [3.10-5.13], $p < .001$). A one-standard-deviation increase in likes at posting also led to an engagement rate approximately 2.31 times higher (IRR = 2.31, 95% CI [2.15-2.49], $p < .001$). Other types of posts did not exhibit a significant difference in average engagement compared to the reference category. Regarding the random effects, posts from Northern, Central, and Southern Italy show engagement levels that are respectively 52% higher, 18% lower, and 20% lower than the overall average engagement.

Table 5. Regression Model Predicting Engagement Based on the Type of Post (Negative-Binomial Mixed Model)

Predictors	Engagement		
	Incidence Rate Ratios	CI	*p*
(Intercept)	259.26	175.32 – 383.39	<.001
Data only in the message (Type 1) (vs only infographic - Type 6)	1.64***	1.31 – 2.04	<.001
Only link to website (Type 2) (vs only infographic - Type 6)	0.49***	0.39 – 0.61	<.001
Data in the message and link to the website (Type 3) (vs Only Infographic - Type 6)	0.85	0.67 – 1.08	.194
Data in the message with link to the website and infographic (Type 4) (vs only infographic - Type 6)	3.98***	3.10 – 5.13	<.001
Data in the message and infographic (Type 5) (vs only infographic - Type 6)	1.07	0.85 – 1.34	.585
Other	0.86	0.67 – 1.10	.235
Likes at posting	2.31***	2.15 – 2.49	<.001
Random effects			
σ^2		.69	
$\tau_{00\ area}$		.09	
ICC		.12	
N_{area}		3	
Observations		1527	
Marginal R^2 / Conditional R^2		.626 / .670	

*$p < .05$, **$p < .01$, ***$p < .001$.

Discussion

Principal Findings

According to the first research question (RQ1), the results revealed varying patterns of data communication on Facebook among Italian regions during the initial months of the Covid-19 pandemic. In terms of the engagement impact of these communication practices (RQ2), it is observed that some types achieve greater engagement and in what way across areas of the country. The most comprehensive posts, featuring data, a link to a website, and an infographic, had the highest positive impact on engagement.

Diverse Data Communication Practices in Italian Public Health

The differences in the use and quantity of the six types of data communication among Italian regions can be attributed to the diverse communication culture and digital practices within Italian public health organizations, as highlighted in previous studies (Cioni & Lovari, 2014; Ducci, 2021; Vaagan et al., 2021). This variety of approaches and practices demonstrates the absence of a coordinated government-level policy for using social media in the health context. Additionally, there is a lack of centralized guidance from the Ministry of Health on effective data communication during crises and emergencies (Solito & Materassi, 2021). These variations may stem from the sudden onset of the Covid-19 pandemic. This rapid development hindered the formulation of comprehensive guidelines. Consequently, this situation reveals an unexplored area. There was limited use of data communication in health messages previously published by Italian public health institutions (Cioni & Lovari, 2014). Additionally, the outdated national pandemic emergency communication plan lacked a dedicated data communication section and was not updated before the pandemic unfolded. Additionally, the results emphasize significant investment by regions of Southern Italy in visually communicating data, with the widespread use of infographics as a key method to make health messages visible and shareable online, even in emergency situations. The pandemic presented a challenge to communicators in an area that lacked institutionalization, leading many regions to improve the quality of their data communication during the pandemic through a combination of learning by doing and imitating successful practices at the national or international level. However, there were instances where data, statistics, and numbers were ineffectively shared through ambiguous social cards, lacking visual graphics or descriptive text. These factors can be related to the different limited graphical literacy of communication practitioners among the regions as reported also in other health related contexts (Franconeri et al., 2021).

Moreover, this empirical evidence stimulate a reflection on the skills of public health communicators who manage content and interactions with citizens in digital environments.

Therefore, there is a need to strengthen digital/graphical communication competences, including data communication, across all Italian regions. This process can facilitate fine-tuning with digital publics and promote greater alignment between citizens' needs and public health organizations' requirements, thus avoiding new inequalities and increasing trust in institutions.

Factors Influencing Facebook Engagement with Public Health Posts in Italy

Regarding the generally low engagement of posts on epidemiological data on Covid-19 that we observed (Lovari, Ducci, & Righetti, 2021), several factors may be at play. For instance, this could be related to data avoidance by a portion of Facebook fans or to the phenomenon of

Covid-19 message fatigue (Guan et al., 2022) experienced during the national lockdown, characterized by fear and uncertainty among the Italian population. We may be witnessing a process of Covid-19 data fatigue, especially in a country that has been publishing and sharing data on traditional and social media every single day for two years following a decision by the national government.

When considering, in greater detail, the varying engagement rates of different types of posts on epidemiological data, one possible explanation could be attributed to the algorithmic factors of the digital platform. It is conceivable that the Facebook algorithm tends to discourage links to external websites (preserving the principle of the so-called "walled garden"). In fact, the type of post with the most negative effect on engagement is the one that shares only a link to an external website. Conversely, posts that include text in the message, along with multi-modal posts featuring text, infographics, and external links, were found to positively impact engagement, consistent with previous public health studies (Guidry et al., 2020).

The presence of text may indicate better communication care and higher quality in delivering public health messages by regional communicators that make them more useful and effective for online users (Guidry et al., 2020, 2023), thereby generating more engagement than simply adding an image or a shared link. Similarly, multi-modal posts (text plus infographic and link) may be favored by the platform in terms of visibility on users' timelines, as richer content is assumed to hold greater value for users and it is thus prioritized by the platform's algorithm.

So, our study confirms the importance of the use of infographics in conveying epidemiological data, alongside careful curation of each post on the official Facebook pages.

Limitations and Resulting Tasks for Future Research

Furthermore, it should be noted that the classification method we used for this study was a methodological simplification compared to the nuanced typology identified by Ducci (2021). Specifically, we captured only references to epidemiological data in the text of the message, regardless of how these references can appear in the text - such as in list form, bullet points, structured text, or detailed descriptions. Despite this simplification, our approach still facilitates easier quantitative detection of this type of content. Future studies will seek to bridge this gap by developing methodologies that more closely match these nuanced types.

Moreover, the study employs computational techniques that are crucial for decoding the overwhelming volume of social media communications in pandemic times. However, the analysis is based on a restricted dataset and should be broadened to include other timeframes.

Additionally, there is potential for refining the classification of data-related posts and further enhancing the computational methods used to identify them on a large scale. For example, it should be noted that our approach to post classification did not distinguish the nuances of style with which the regions treated the data in the text of the post (e.g., whether in the form of a simple list or more elaborate description and commentary), for which a more refined manual classification and, subsequently, the possible training of a machine learning algorithm is necessary. Furthermore, the results cannot be generalized to other social media platforms, as each platform has distinct relational and communicative affordances, influencing message posting, types of engagement, and algorithm operations for organic post visibility. These developments may be the subject of future work.

Despite these limitations, the paper highlights an area of data communication on social media that remains inadequately investigated and understood in current literature. It emphasizes the need for further investigation and provides valuable insights that pave the way for future advancements and developments.

Future research could enhance the classification system we adopted by further refining the automated message classification methodology. This could involve using a supervised machine learning approach, particularly leveraging Large Language Models. Additionally, it would be beneficial to integrate an analysis of the number of views each post category receives, as opposed to focusing solely on engagement value. Views represent a different metric of a post's popularity compared to engagement and are particularly relevant in emphasizing the dissemination of messages, which is crucial in pandemic communication. Regrettably, researchers have not typically had access to such data. However, the prospect of conducting such visualization-based analyses may soon be realized with the data that META intends to provide researchers through the new Meta Content Library and API.

Conclusions

Within the theoretical framework of the public sector communication, this article explored the role of data communication related to the Covid-19 lockdown. It did so through an analysis conducted on the official Facebook channels of Italy's 20 regions, enabling the examination of various types of data communication and their corresponding engagement rates.

In conclusion, the diversity in the communication of epidemiological data among regions may have resulted in differences in how Italian citizens were informed by public institutions about the spread of the virus. This can pose a challenge for public health communication in the relationship between the national level and the local territories, both in ordinary and emergency situations.

Greater homogeneity in communication approaches and increased synergy at the inter-institutional level are particularly desirable in the context of health data management, considering the increasing centrality of this aspect in health communication in general (Parrott 2009). This homogeneity should extend to the adoption of techniques, formats and languages characteristic of social media environments.

It is also important to emphasize the need for uniform competencies among health communication professionals, focusing on accuracy and completeness in social media data communication, which goes beyond just infographics. This includes creating comprehensive posts that combine text, imagery, and other relevant elements to engage citizens more effectively.

Ethical Approval

The social media data collected and analyzed in this project is publicly accessible and lawfully obtained through the official CrowdTangle API, provided by Meta to researchers within a partnership program. This program ensures strictly controlled access and compliance with legal, ethical, and privacy standards.

Funding

The research presented in the article has not received funding.

Conflict of Interest

The authors have no conflicts of interest to declare.

References

Anderson, O. S., Kardia, S., Gupta, K., & August. E. (2019). Are we teaching our students visual communication? Evaluation of writing assignments in public health. *Journal of Visual Communication in Medicine, 43*(1), 62-65. https://doi.org/10.1080/17453054.2019.1698943

CrowdTangle Team (2023). *CrowdTangle*. Facebook.

Bennett, W. L., & Iyengar, S. (2008). A new era of minimal effects? The changing foundations of political communication. *Journal of Communication, 58*(4), 707–731. https://doi.org/10.1111/j.1460-2466.2008.00410.x

Chan, M. S., Jones, C. R., Hall Jamieson, K., & Albarracín, D. (2017). Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. *Psychological Science*, 28(11), 1531–1546. https://doi.org/10.1177/0956797617714579

Cioni, E., & Lovari, A. (2014). Social media for health communication: Implementation issues and challenges for Italian public health authorities. In M. Househ, E. Borycki, & A. Kushniruk (Eds.), *Social media and mobile technologies for healthcare* (pp. 237–263). IGI Global.

Coombs, W.T. (2020). Conceptualizing crisis communication. In R. L. Heath & H. D. O'Hair (Eds.), *Handbook of risk and crisis communication* (pp. 99-118). Routledge. http://dx.doi.org/10.4324/9781003070726-6

De Marchi, S., & Stewart, B. (2020). Wrestling with complexity in computational social science: Theory, estimation and representation. In L. Curini & R. Franzese (Eds.), *The SAGE handbook of research methods in political science and international relations* (pp. 289-310). SAGE Publications.

Ducci, G. (2021). Alcuni nodi della comunicazione pubblica digitale di fronte alla pandemia. Le regioni italiane su Facebook durante il lockdown. *Mediascapes Journal*, 18(18), 141-160. https://rosa.uniroma1.it/rosa03/mediascapes/article/view/17673.

Ducci, G., Boccia Artieri, G., Corsi, M., & Folena, C. (2022). Information sources, trust and public health sector communication on the COVID-19 pandemic: A study on the Italian citizens' perception. *Revista Española de Comunicaciòn en Salud, 13*(2), 182-199. https://doi.org/10.20318/recs.2022.7039

Ducci, G., & Lovari, A. (2021). The challenges of public sector communication in the face of the pandemic crisis: professional roles, competencies and platformization. *Sociologia della Comunicazione, 61*, 9-19. https://doi.org/10.3280/SC2021-061002

Eberl, J. M., Tolochko, P., Jost, P., Heidenreich, T., & Boomgaarden, H. G. (2020). What's in a post? How sentiment and issue salience affect users' emotional reactions on Facebook. *Journal of Information Technology & Politics*, 17(1), 48-65. https://doi.org/10.1080/19331681.2019.1710318

Franconeri, S. L., Padilla, L. M., Shah, P., Zacks, J. M., & Hullman, J. (2021). The science of visual data communication: What works. *Psychological Science in the Public Interest*, 22(3), 110–161. https://doi.org/10.1177/15291006211051956

Guan, M., Li, Y., Scoles, J. D., & Zhu, Y. (2022). COVID-19 message fatigue: How does it predict preventive behavioral intentions and what types of information are people tired of hearing about? *Health Communication*, 38(8), 1631-1640. https://doi.org/10.1080/10410236.2021.2023385

Guidry, J. P. D., Jin, Y., Orr, C. A., Messner, M., & Meganck, S. (2017). Ebola on Instagram and Twitter: How health organizations address the health crisis in their social media engagement. *Public relations review*, 43(3), 477-486. https://doi.org/10.1016/j.pubrev.2017.04.009

Guidry, J. P. D., Meganck S. L., Lovari, A., Messner, M., Medina-Messner, V., Sherman, S., & Adams J. (2020). Tweeting about #diseases and #publichealth: Communicating global health issues across nations. *Health Communication, 35*(9), 1137-1145. https://doi.org/10.1080/10410236.2019.1620089

Guidry, J. P. D., Meganck, S. L., Perrin, P. B., Messner, M., Lovari, A., & Carlile, K.E., (2021). #Ebola: Tweeting and pinning an epidemic. *Atlantic Journal of Communication, 29*(2), 79-92. https://doi.org/10.1080/15456870.2019.1707202

Guidry, J. P. D., O'Donnell, N. H., Meganck, S. L., Lovari, A., Miller, C. A., Messner, M., Hill, A., Medina-Messner, V., & Carlyle, K. E. (2023). Tweeting a pandemic: Communicating #COVID19 across the globe. *Health Communication*, 28(11), 2377-2386. https://doi.org/10.1080/10410236.2022.2071192

Jin, Y., & Austin, L. (Eds). (2022). *Social media and crisis communication.* Routledge.

Kessler S. H., Cano Pardo, M. S., Jobin, A., & Georgi, F. (2022). How informed are the Swiss about COVID-19 and prevention measures? Results of a survey on information awareness, behaviour, and deficits. *European Journal of Health Communication, 3*(3), 118-142. https://doi.org/10.47368/ejhc.2022.306

La Rocca, G., Carignan, W.E., & Boccia Artieri, G. (Eds) (2023), *COVID-19 and the global crisis of information: case studies from Europe, Canada and Mexico*. Palgrave MacMillan.

Lee, S. H, Pandya, R. K., Hussain, J. S., Lau, RJ., Chambers, E. A. B., Geng, A., Jin, B. X., Zhou, O., Wu, T., Barr, L. & Junop, M. (2022). Perceptions of using infographics for scientific communication on social media for COVID-19 topics: A survey study. *Journal of Visual Communication in Medicine, 45*(2), 105-113. https://doi.org/10.1080/17453054.2021.2020625

Lewandowsky, S., Cook, J., Ecker, U. K. H., Albarracín, D., Amazeen, M. A., Kendeou, P., Lombardi, D., Newman, E. J., Pennycook, G., Porter, E. Rand, D. G., Rapp, D. N., Reifler, J., Roozenbeek, J., Schmid, P., Seifert, C. M., Sinatra, G. M., Swire-Thompson, B., van der Linden, S., Vraga, E. K., Wood, T. J., & Zaragoza, M. S. (2020). *The debunking handbook 2020*. Databrary. http://doi.org/10.17910/b7.1182

Lipkus, I. M. (2007). Numeric, verbal, and visual formats of conveying health risks: Suggested best practices and future recommendations. *Medical Decision Making, 27*, 696–713. https://doi.org/10.1177/0272989X07307271

Lilleker, D, Coman, I. A., Gregor, M., & Novelli, E. (Eds). (2021). *Political communication and COVID-19. Governance and rhetoric in times of crisis*. Routledge.

Locatelli, E., & Lovari, A. (2021). Platformization of healthcare communication: Insights from the early stage of the COVID-19 pandemic in Italy. *Catalan Journal of Communication and Cultural Studies, 13(2),* 249-266.

Lovari, A. (2020). Spreading (dis)trust: COVID-19 misinformation and government intervention in Italy. *Media and Communication, 8*(2), 458–461. https://doi.org/10.17645/mac.v8i2.3219

Lovari, A., D'Ambrosi, L., & Bowen, S. (2020). Re-connecting voices: The (new) strategic role of public sector communication after COVID-19 crisis. *Partecipazione e conflitto*, *13*(2), 970-989. http://dx.doi.org/10.1285/i20356609v13i2p970

Lovari., A., Ducci, G., & Righetti, N. (2021). Responding to fake news. the use of Facebook for public health communication during the COVID-19 pandemic in Italy. In M. Lewis, E. Govender, & K. Holland (Eds.), *Communicating COVID-19: Interdisciplinary Perspectives* (pp. 251-275). Pagrave Macmillan.

Lovari, A., & Righetti, N. (2020). La comunicazione pubblica della salute tra infodemia e fake news. *Mediascapes Journal, 15*, 156-173. https://rosa.uniroma1.it/rosa03/mediascapes/article/view/17079

Memenga, P., Eitze S., Shamsrizi, P., Addo, M. M., & Betsch, C. (2022). Debunking misinformation and communicating critical events in vaccine trials: Experimental evidence on vaccination intentions in SARS-CoV-2 pandemic. *European Journal of Health Communication*, *3*(2) 64-96. https://doi.org/10.47368/ejhc.2022.204

Mergel I. & Bretschneider S. I. (2013). A three stage adoption process for social media use in government. *Public Administration Review*, *73*(3), 390-400. https://doi.org/10.1111/puar.12021

Mira, A., & Massarenti, A. (2020). *La pandemia dei dati*. Mondadori Università, Milano.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In C. J. Burges, L. Bottou, M. Welling, Z. Ghahramani, & K. Q. Weinberger (Eds.), *Advances in neural information processing systems*. https://doi.org/10.48550/arXiv.1310.4546

Moorhead, S. A., Hazlett, D. E., Harrison, L., Carroll, J. K., Irwin, A., & Hoving, C. (2013). A new dimension of health care: Systematic review of the uses, benefits, and limitations of social media for health communication. *Journal of Medical Internet Research, 15*, Article e85. https://doi.org/10.2196/jmir.1933

Nelson, D. E., Spieker, M. R., & Hesse, B. W. (2012). Communicating health data. In J. Merrick (Ed.), *Public health yearbook*. Nova Science Publisher.

Nelson, D. E., Hesse, B., & Croyle, R. (2009). *Making data talk: Communicating public health findings to the public policy makers, and the press*. Oxford University Press.

Nielsen, R. K., Schulz, A., & Fletcher, L. (2021, May). *An ongoing infodemic: How people in eight countries access news and information about coronavirus a year into the pandemic.* Reuters Institute. https://reutersinstitute.politics.ox.ac.uk/ongoing-infodemic-how-people-eight-countries-access-news-and-information-about-coronavirus-year

OECD (2021). *OECD Report on public communication: The global context and the way forward*. OECD Publishing. https://doi.org/10.1787/22f8031c-en.

Parrott, R. (2009). *Talking about health*. Wiley-Blackwell.

Parvanta, C., Nelson, D. E., Paravanta, S., & Harner, R. N. (2011). *Essential of public health communication*. Jones Barlett.

R Core Team (2022). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org/

Rothman, A. J., & Kiviniemi, M. T. (1999). Treating people with information: An analysis and review of approaches to communicating health risk information. *Journal of the National Cancer Institute Monographs, 25*, 44-51. https://doi.org/10.1093/oxfordjournals.jncimonographs.a024207

Sastry, S., & Lovari, A. (2017). Communicating the ontological narrative of Ebola: An emerging disease in the time of "epidemic 2.0". *Health Communication*, *32*(3), 329-338. https://doi.org/10.1080/10410236.2016.1138380

Sendra, A., Farré, J., Lovari, A., & Lombi, L. (2021). Post-pandemic routes in the context of Latin countries: The impact of COVID-19 in Italy and Spain. In D. Lupton & K. Willis (Eds.), *The COVID-19 crisis: Social perspectives* (pp. 156–167). Routledge.

Solito, L. & Materassi, L. (2021). If a picture is not worth a thousand words: Digital infographics use during the COVID-19 pandemic crisis. *Sociologia della Comunicazione, 61*, 52-70. https://doi.org/10.3280/SC2021-061005

Tirkkonen, P., & Luoma-aho, V. (2011). Online authority communication during an epidemic: A Finnish example. *Public Relations Review*, 37(2), 172–174. https://doi.org/10.1016/j.pubrev.2011.01.004

Vaagan, R. W., Torkola, S., Sendra, A, Farré, J., Lovari, A. (2021). Critical analysis of the digitization of healthcare communication in the EU: A Comparison of Italy, Finland, Norway, and Spain. *International Journal of Communication. 15*, 1718–1740.

Van Dijck, J. (2020). Governing digital societies: Private platforms, public values. *Computer, Law, & Security Review, 36*, Article 10537. http://dx.doi.org/10.1016/j.clsr.2019.105377

Watzlawick, P., Beavin, J. H., & Jackson, D. D. (1967). *Pragmatics of human communication*. W. W. Norton & Company.

We Are Social. (2022). *Global digital report.* https://wearesocial.com/it/blog/2022/01/digital-2022-i-dati-globali/.

WHO. (2020). *Risk communication and community engagement readiness and response to coronavirus disease.* World Health Organization. https://apps.who.int/iris/bitstream/handle/10665/331513/WHO-2019-nCoV-RCCE-2020.2-eng.pdf.

Woloshin, S., Swartz, L. M., & Welch, H. G. (2008). *Know your chances: Understanding health statistics*. University of California Press.

Author Contributions

Conceptualisation (main idea, theory): Gea Ducci & Alessandro Lovari
Funding acquisition: –
Project administration: Gea Ducci
Methodology (design, operationalisation): Nicola Righetti
Data collection: Nicola Righetti
Data analysis: Gea Ducci, Alessandro Lovari, & Nicola Righetti
Writing – original draft: Gea Ducci, Alessandro Lovari, & Nicola Righetti
Writing – review & editing: Gea Ducci & Nicola Righetti

Author Biographies

Gea Ducci is Associate Professor in Sociology of Culture and Communication at the Department of Communication Sciences, Humanities and International Studies at University of Urbino Carlo Bo (Italy). She is the coordinator of the master's degree course in Communication and Advertising for Organizations, a member of the PhD board in Humanistic Studies and of the Research Lab on Advanced Communication at the same university. She is elected member of the board of Italian Sociology Association. Her research focuses on health communication, public sector communication, social issues communication and digital media.

Alessandro Lovari is Associate Professor of Sociology of Communication at the University of Cagliari (Italy). Lovari's research focuses on public sector communication, public relations, and health communication, investigating Internet and social media impact on organizations' practices and citizens' behaviours. He was visiting research scholar at Purdue University, University of Cincinnati, University of South Carolina, Virginia Commonwealth University (USA). He is author of more than 90 publications in books, encyclopaedias, and international and national journals.

Nicola Righetti is a researcher and lecturer in the field of digital, statistical, and computational methods applied to social research at the Department of Communication of the University of Vienna, where he is also an associate researcher at the computational communication lab of the same university. His research primarily focuses on the study of digital communication related to phenomena of political and social relevance.