CARMA 2026

8th Int. Conf. on Advanced Research Methods and Analytics
1-3 July 2026, Valencia, Spain
DOI: https://doi.org/10.4995/CARMA2026.2026.21969

Captured Voices: Organized Crime Infiltration and Municipal Communication

Nicola Caravaggio, Giuliano Resce, Agapito Emanuele Santangelo
Department of Economics, University of Molise, Italy.

How to cite: Caravaggio, N.; Resce, G.; Santangelo, A. E. (2026). Captured Voices: Organized Crime Infiltration and Municipal Communication In: 8th International Conference on Advanced Research Methods and Analytics (CARMA 2026). Valencia, 1-3 July 2026. https://doi.org/10.4995/CARMA2026.2026.21969

Abstract

This paper examines whether Italian municipalities subjected to mafia-related administrative dissolution differ in their social-media communication patterns during periods of presumed infiltration. We combine three data sources: (i) the official register of 368 municipal dissolution events (1991–2025); (ii) the national registry of local administrators (Anagrafe degli Amministratori Locali), from which we reconstruct each council's mandate period to define the presumed infiltration window (mandate start, dissolution date); and (iii) a corpus of 2,287,279 Facebook and Instagram posts by 5,335 Italian municipalities. Applying BERTopic to the full corpus and comparing topic shares across temporally defined infiltration and control periods, we find that infiltrated municipalities over-represent topics related to water infrastructure, civic ceremonies, and public-order measures, while non-infiltrated municipalities show higher prevalence of cultural programming and participatory content.

Keywords: *mafia infiltration; BERTopic; local government; social media; administrative dissolution; Italian municipalities.*

1. Introduction

The relationship between organised crime and local government has attracted growing attention in economics. A robust body of evidence documents that mafia presence depresses firm performance (Mirenda et al., 2022), distorts the allocation of public transfers (Barone & Narciso, 2015), and reduces long-run economic growth at both the municipal (Fenizia & Saggio, 2024) and regional levels (Pinotti, 2015). The Italian institution of scioglimento per infiltrazione mafiosa, the compulsory dissolution of municipal councils found to be conditioned by organised crime (Legislative Decree 267/2000, Art. 143), provides a unique administrative record of confirmed infiltration. Applied 368 times between 1991 and 2025, the mechanism is concentrated in southern regions as: Calabria, Campania, Sicilia, and Puglia.

Article in Press

This work is licensed under a Creative Commons License CC BY-NC-SA 4.0
Editorial Universitat Politècnica de València

Yet the communicative dimension of infiltrated administrations remains unexplored. If mafia conditioning shapes the priorities and rhetoric of local governments, as the literature on resource misallocation and institutional capture suggests (Alesina et al., 2019; Acemoglu et al., 2020), this may be detectable in how municipalities communicate with citizens on social media. In line with the general trends (Resce, Maynard, 2028; Carneiro et al. 2022ab), Social network like Facebook and Instagram have become primary channels of Italian local government communication, generating millions of publicly available posts whose thematic structure can be systematically analysed at scale.

We make three contributions: (i) reconstruct temporally precise infiltration windows by combining the dissolution register with the Anagrafe degli Amministratori Locali; (ii) apply BERTopic (Grootendorst, 2022) to over 2.2 million Italian municipal posts; and (iii) provide the first large-scale textual characterisation of how infiltrated municipalities communicate.

2. Data and Methods

2.1. Three-Source Data Architecture

Step 1 — **Dissolution register**. The official register on dissolution of municipal councils, compiled from Ministry of the Interior (MINT, 2025a), records 368 dissolution events (1991–2025). Each entry contains the municipality name, province, region, and dissolution date, the formal date on which the elected council was removed and an extraordinary commission appointed. We treat this date as the terminal point of the presumed infiltration period.

Step 2 — **Anagrafe degli Amministratori Locali**. The national registry of local administrators records mandate start and end dates for all elected mayors and councilors (MINT, 2025b). For each dissolved municipality we identify the most recent council mandate preceding the dissolution date. The presumed infiltration window is defined as the interval (mandate start date, dissolution date). This approach follows Fenizia & Saggio (2024), who show that mafia conditioning of local administrations typically begins at or shortly after council formation. Posts published within this window by a subsequently dissolved municipality are classified as infiltrated; all other posts, including pre-mandate posts from the same municipality, constitute the control group.

Step 3 — **Social-media corpus**. We scraped the official Facebook and Instagram accounts of Italian municipalities, yielding 2,287,279 posts from 5,335 unique municipalities after deduplication and quality filtering. Post metadata includes the publication date and municipality code, enabling precise temporal alignment with the mandate windows reconstructed in Step 2. Municipality names were harmonised through lower-casing, removal of the prefix “comune di”, and punctuation stripping. After normalisation, 52 dissolved municipalities were matched, yielding 13,517 allegedly “infiltrated posts”.

Article in Press

2.2. Topic Modelling with BERTopic

BERTopic (Grootendorst, 2022) combines sentence-transformer embeddings, UMAP dimensionality reduction, HDBSCAN clustering (McInnes et al., 2017), and class-based TFIDF (c-TF-IDF) to produce interpretable topic representations. Noise posts (topic_id = −1) are excluded and each topic is characterised by four representative c-TF-IDF terms (e.g., "*idrica · acqua · erogazione · idrico*"). For each municipality–period we compute a topic share vector (proportion of posts assigned to each topic).

2.3. Comparison Strategy

For each topic we compute the mean share across infiltrated posts, the mean share across control posts, and the raw difference (infiltrated − control) in percentage points. A positive difference indicates over-representation in infiltrated administrations. We complement the topic analysis with word-frequency visualisations (word clouds) from cleaned post text, both in aggregate and disaggregated by the four regions with the highest dissolution counts. The identification assumption is that within-municipality differences in topic shares between the infiltration window and surrounding periods reflect the communicative consequences of mafia conditioning rather than time-invariant local characteristics.

3. Results

3.1. Sample Overview

Figure 1 summarises the sample. Infiltrated posts account for 13,517 observations from 52 unique municipalities, against posts when municipalities are non-infiltrated. Panel C shows two dissolution waves: an initial peak in the early 1990s following legislative introduction, and a second sustained wave from 2010 to 2022 (consistent with Fenizia & Saggio, 2024).

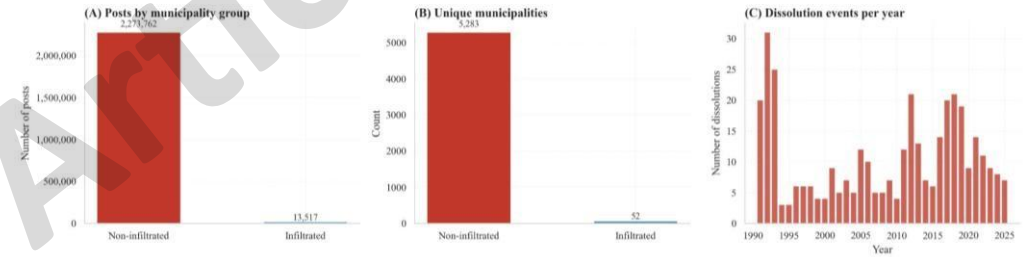


Figure 1. Sample overview. Panel A: posts by group; Panel B: unique municipalities; Panel C: dissolution events per year (1991–2025).

Article in Press

Table 1 disaggregates coverage by region. Calabria records the most dissolution events (127), followed by Campania (111) and Sicilia (90). Sample coverage ranges from 10.2 % (Calabria) to 38.5 % (Puglia), reflecting the limited digital presence of smaller dissolved municipalities.

Table 1. Coverage of dissolved municipalities in the social-media sample, by region.

Region	Mun. in sample	Total dissolutions	Coverage (%)
Calabria	13	127	10.2
Campania	12	111	10.8
Sicilia	15	90	16.7
Puglia	10	26	38.5
Lazio	1	5	20.0
Piemonte	0	3	0.0
Basilicata	0	2	0.0
Emilia-Romagna	1	1	100.0
Liguria	0	1	0.0
Lombardia	0	1	0.0
Valle d'Aosta	0	1	0.0

3.2. Lexical Content: Word Cloud Analysis

We begin with a lexical portrait of infiltrated municipalities (Carneiro et al. 2025). Figure 2 presents the aggregate word cloud for all posts published during presumed infiltration windows. Dominant terms, *città, territorio, comunità, cittadinanza*, reflect a communication style centred on territorial identity and civic belonging. Secondary terms protezione and collaborazione anticipate the systematic differences that the BERTopic analysis will quantify. The prominence of protezione civile-adjacent vocabulary is notable: civil protection expenditure is a documented mafia rent-seeking target in southern Italy (Pinotti, 2015).

Article in Press



Figure 2. Most frequent terms in infiltration-window posts (all regions). Word size proportional to frequency after Italian stopword removal.

Figure 3 disaggregates by region. Calabria is dominated by gioia alongside cittadinanza and territorio. Campania features seismic emergency terminology (magnitudo, sismico, protezione). Sicily shows legislative references (articolo, leggi) and toponyms (Pantelleria, Bagheria). Puglia is driven by Manfredonia (city name) and civic associations and territorial projects terms.



Figure 3. Most frequent terms in infiltration-window posts, by region (top 4 regions by dissolution events). Word size proportional to frequency after Italian stopword removal.

3.3. Topic Differences During Infiltration Windows

The lexical portrait in 3.2 is confirmed and quantified by the BERTopic analysis. Figure 4 displays the 20 topics with the largest absolute mean share difference between infiltrated and control posts. Red bars indicate over-representation during infiltration windows; blue bars indicate higher prevalence in the control group.

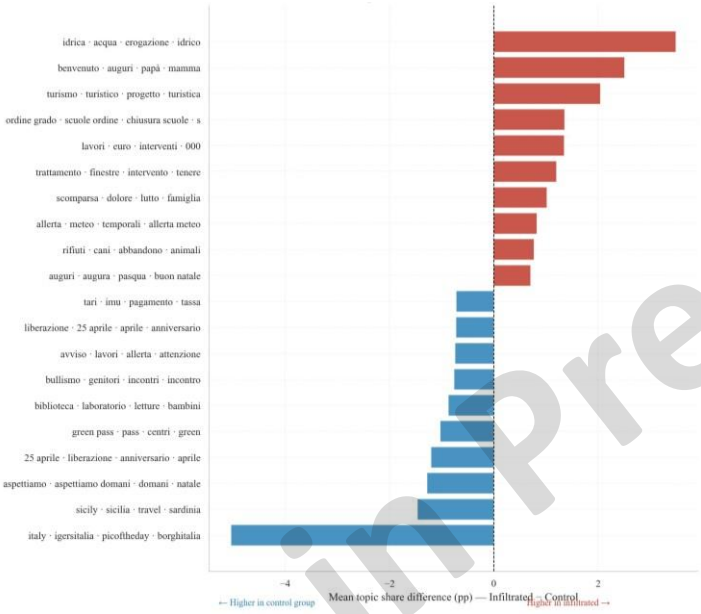


Figure 4. Topics with the largest mean share difference: infiltrated vs. control (percentage points, infiltrated − control).

The largest positive difference belongs to "idrica · acqua · erogazione · idrico" (water supply, ≈3.2 pp). Further over-represented topics include civic celebrations (+2.6 pp), tourism promotion (+2.2 pp), school-order communications (+1.7 pp), and infrastructure works (+1.6 pp), domains in which public procurement is a key channel of organised-crime rent extraction (Pinotti, 2015; Alesina et al., 2019).

Conversely, the control group shows markedly higher shares of Instagram territorial promotion ("italy · igersitalia · picoftheday · borghitalia", −5.4 pp), cultural programming, participatory initiatives, and Liberation Day commemorations.

4. Discussion and Conclusions

This paper provides a first large-scale textual characterisation of Italian municipal communication during periods of presumed mafia infiltration. Combining the dissolution

Article in Press

register, the *Anagrafe degli Amministratori Locali*, and over 2 million social-media posts, we reconstruct temporally precise infiltration windows and compare BERTopic topic distributions.

Three main findings emerge.

First, the lexical portrait of infiltrated municipalities centres on territorial-identity and civicbelonging vocabulary (*città, territorio, comunità*) alongside emergency-adjacent terms such as *protezione and collaborazione*, pointing to a communication style dominated by public-order priorities. Second, the BERTopic analysis quantifies these patterns: infiltrated municipalities over-represent water infrastructure, public-works spending, and civic-ceremony content, precisely the domains the economics literature identifies as primary mafia rent-extraction channels (Barone & Narciso, 2015; Pinotti, 2015; Mirenda et al., 2022), while non-infiltrated municipalities invest more in cultural and participatory communication, consistent with Acemoglu et al. (2020) on institutional quality and local public-good provision. Third, regional word clouds reveal spatial patterns in the communication.

Several caveats apply. The infiltrated group is small (52 municipalities, 0.6 % of posts) and the 16 % matching rate introduces selection bias towards larger, digitally active communes. The analysis does not control for municipality size, political affiliation, or socioeconomic conditions, and the full-mandate-window identification is an approximation. Future work should employ propensity-score matching, event-study designs centred on dissolution dates, and individual administrator career tracking. The BERTopic pipeline and the region-shaped word cloud extension offer scalable tools for public-sector text analytics, opening new empirical avenues at the intersection of computational social science and the economics of organised crime.

References

Acemoglu, D., De Feo, G., & De Luca, G. D. (2020). Weak states: Causes and consequences of the Sicilian Mafia. *'Review of Economic Studies', 87(2), 537–581. https://doi.org/10.1093/restud/rdz009*

Alesina, A., Piccolo, S., & Pinotti, P. (2019). Organized crime, violence, and politics. *'Review of Economic Studies', 86(2), 457–499. https://doi.org/10.1093/restud/rdy036*

Bandiera, O. (2003). Land reform, the market for protection, and the origins of the Sicilian Mafia: Theory and evidence. *'Journal of Law, Economics, and Organization', 19(1), 218–244. https://doi.org/10.1093/jleo/19.1.218*

Barone, G., & Narciso, G. (2015). Organized crime and business subsidies: Where does the money go? *'Journal of Urban Economics', 86, 98–110. https://doi.org/10.1016/j.jue.2015.01.004*

Carneiro, B., Resce, G., & Sapkota, T. B. (2022a). Digital artifacts reveal development and diffusion of climate research. *Scientific Reports*, 12(1), 14146.

Article in Press

Carneiro, B., Resce, G., Läderach, P., Schapendonk, F., & Pacillo, G. (2022b). What is the importance of climate research? An innovative web-based approach to assess the influence and reach of climate research programs. *Environmental Science & Policy*, 133, 115-126.

Carneiro, B., Resce, G., Caravaggio, N., Santangelo, A. E., Ruscica, G., Tucci, G., ... & Coffey, K. (2025). Text mining and machine learning reveal global determinants of food insecurity. *Scientific Reports*, 15(1), 36709.

Fenizia, A., & Saggio, R. (2024). Organized crime and economic growth: Evidence from municipalities infiltrated by the Mafia. *'American Economic Review', 114(7), 2171–2200. https://doi.org/10.1257/aer.20221687*

Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv:2203.05794. https://doi.org/10.48550/arXiv.2203.05794*

McInnes, L., Healy, J., & Astels, S. (2017). hdbscan: Hierarchical density based clustering. *'Journal of Open Source Software', 2(11), 205. https://doi.org/10.21105/joss.00205*

Mirenda, L., Mocetti, S., & Rizzica, L. (2022). The economic effects of Mafia: Firm level evidence. *'American Economic Review', 112(8), 2748–2773. https://doi.org/10.1257/aer.20201015*

Mocetti, S. (2024). Organized crime in Italy: An economic analysis. *'Italian Economic Journal'. https://doi.org/10.1007/s40797-024-00280-w*

MINT 2025a, Italian Ministry of Interior, "List of City Council Dismissals unrelated to Mafia Infiltration," accessed on February 8th, 2025, 2025.

MINT 2025b*,* Italian Ministry of Interior, "Italian Local politicians 1986-2025," available at http://doi.org/10.3886/E197843V1, 2025.

Pinotti, P. (2015). The economic costs of organized crime: Evidence from Southern Italy. *'Economic Journal', 125(586), F203–F232. https://doi.org/10.1111/ecoj.12235*

Resce, G., & Maynard, D. (2018). What matters most to people around the world? Retrieving Better Life Index priorities on Twitter. *Technological Forecasting and Social Change*, 137, 61-75.

Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using Siamese BERT-networks. *In 'Proceedings of EMNLP-IJCNLP 2019' (pp. 3982–3992). https://doi.org/10.18653/v1/D19-1410*

Article in Press