

“It’s Not Hate; It’s Common Sense”: Polycentric Normalization of Religious Hate in Spanish Digital Media

SERGIO ARCE-GARCÍA¹

ELÍAS SAID-HUNG

ÓSCAR DE GREGORIO-VICENTE

Universidad Internacional de La Rioja, Spain

ANA MARGARIDA BARRETO

Universidade Nova de Lisboa, Portugal

This study explores how low-intensity religious hate becomes normalized on Spanish social media, shaped by multiple power centers (editorial policies, algorithms, and social norms), especially during major news events. An analysis of 53,787 comments on articles from 6 ideologically diverse outlets (March–April 2025) found 3.4% expressed religious hate, mainly anti-Muslim, followed by anti-Catholic and anti-Semitic discourse. Using multiple correspondence analyses and semantic clustering, significant links emerged among hate type, outlet ideology, and comment intensity. The results also revealed connections among geopolitical conflicts, migration crises, and peaks in hostility, showing how such dynamics normalize exclusionary narratives. The study calls for greater editorial transparency, event-sensitive moderation, and stronger cooperation between platforms and newsrooms to reduce algorithm-driven amplification of hate.

Keywords: hate speech, governance, digital media ecosystems, algorithmic amplification, affective polarization, symbolic violence, religious intolerance

Sergio Arce-García: sergio.arce@unir.net

Elías Said-Hung: elias.said@unir.net

Óscar De Gregorio-Vicente: oscar.degregorio@unir.net

Ana Margarida Barreto: ambarreto@fcsh.unl.pt

Date submitted: 2026-01-13

¹This research was supported by a grant from the Fundación Pluralismo y Convivencia [Pluralism and Coexistence Foundation] (Project Code: PC-24-0050), Ministry of the Presidency of Spain. The funding body played no role in the design of the study; the collection, analysis, or interpretation of data; or in writing the manuscript.

Copyright © 2026 (Sergio Arce-García, Elías Said-Hung, Óscar De Gregorio-Vicente, and Ana Margarida Barreto). Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at <https://ijoc.org>.

Social media platforms exacerbate hostility toward minority groups, including women, immigrants, lesbian, gay, bisexual, transgender/transsexual, queer, and more (LGBTQ+) individuals, and religious communities, which constitutes a significant issue in digital communication. Notably, over 80% of hate speech remains unremoved, indicating a systemic failure (Alorainy et al., 2022; Bergmanis-Korāts & Haiduchyk, 2024; Carvalho et al., 2023).

Religious hate speech means messages inciting prejudice against people for their beliefs (Kiper, 2023). In this study, religious hostility is conceptualized as a continuum ranging from subtle forms of discursive aggression to explicit threats. These authors distinguish four types: (1) incivility, (2) malicious speech, (3) direct insults, and (4) explicit threats. Each level erodes social cohesion differently (Hartikainen, 2025).

Incivility refers to statements that delegitimize or belittle a religious group through disrespectful or generalizing language, without targeting specific individuals. Malicious speech involves accusations attributing negative intentions, moral inferiority, or collective guilt. The key difference lies in intention: Incivility expresses contempt without imputing hostile motivations, while malicious speech attributes guilt or collective dangerousness. Both differ from direct insults (explicit, degrading language targeting a group or its members) and explicit threats (statements advocating or threatening physical violence), which occupy the highest hostility levels. Explicit threats are widely recognized as unacceptable; direct insults constitute overt aggression, but do not necessarily imply physical harm.

By contrast, incivility and malicious speech operate through discursive ambiguity. Incivility, as the lowest hostility level, may fall within freedom of expression and often appears as cultural criticism or opinion. Malicious speech goes further by attributing negative intentions or collective blame while avoiding explicit aggression. Because these lower-intensity forms can be framed as legitimate commentary, they often circulate without triggering moderation mechanisms, allowing aggressors to deny hostile intent, while affected groups still experience harm. This ambiguity facilitates the gradual normalization of hostile discourse: Repeated exposure to veiled prejudice lowers tolerance thresholds without triggering institutional alarms (Davani et al., 2023; Rajan & Venkatraman, 2021; Uyheng & Carley, 2021). Over time, such expressions can become incorporated into everyday conversational "common sense," particularly in highly polarized digital media environments.

Existing literature focuses on identifying hostility patterns on digital platforms (Alorainy et al., 2022; Bergmanis-Korāts & Haiduchyk, 2024; Carvalho et al., 2023) or examining hate narratives directed at religious groups, often without considering media context (Becker et al., 2022; Wells, 2023). Relatively little empirical research has examined how editorial stance and platform dynamics interact to shape the visibility, intensity, and narrative structure of religious hate in digital public spheres—a gap particularly relevant in polarized media contexts like Spain (Lilleker & Pérez Escolar, 2023). While research indicates that digital media create environments with varying levels of normative friction toward hostile expressions (Arce-García et al., 2024; Labiano Juangarcía et al., 2023; Minooie et al., 2023), empirical studies systematically examining associations among editorial orientation, types of religious hate, and discourse intensity remain limited.

This study investigates how polycentric governance (encompassing X's algorithms, media editorial policies, and social validation dynamics) normalizes low-intensity religious animosity during windows opened by emotionally impactful events. Rather than establishing causal relationships, the study adopts an exploratory multivariate approach to characterize associations among editorial contexts, discursive patterns, and temporal dynamics of religious hostility in Spain's digital media ecosystem.

The study delineates three dimensions: (1) a conceptual dimension, which characterizes religious hatred as a continuum of intensities undermining social cohesion (Hartikainen, 2025; Kiper, 2023; Montero-Díaz et al., 2025); (2) a contextual dimension, which identifies factors facilitating its normalization in digital ecosystems such as echo chambers, algorithms prioritizing engagement, and social validation among ideologically similar individuals (Granovetter, 1973; Uyheng & Carley, 2021; Weinberg et al., 2025) operationalized through associative pattern analysis; and (3) a media dimension, which conceptualizes news media as intermediaries with varying capacities to propagate or mitigate hate based on their ideological stance (Alonso & Palomo, 2025; Minoie et al., 2023).

The study answers three specific research questions:

- RQ1: How do editorial lines shape the perceived legitimacy of different types of religious hatred in user comments?*
- RQ2: How do the volume and intensity of religious hatred vary during normative windows of opportunity generated by emotionally impactful events (migration crises, geopolitical conflicts, terrorist attacks), and what evidence do these variations provide regarding the temporary flexibilization of normative thresholds?*
- RQ3: Which narrative structures characterize each type of religious hatred, and how do these structures allow us to interpret the differentiated ways in which religious hatred is articulated and expressed in the context of high-impact events?*

By combining multivariate association analysis, semantic clustering, and temporal event analysis on a large corpus of user comments, this study contributes empirically to the literature on online hate by demonstrating how media environments, event dynamics, and narrative structures jointly shape the normalization of low-intensity religious hostility in a polarized digital public sphere.

The Spanish context shows distinct characteristics that make it a particularly relevant case: Institutional ambiguity between religious freedom and Catholicism's privileged status in the 1978 Constitution creates a normative gray area.

Normalization and the Continuum of Religious Hostility

Social media platforms facilitate hate dissemination through four structural mechanisms: (1) echo chambers reinforcing selective exposure to hostile content (Popa-Wyatt, 2023), (2) algorithmic prioritization of emotional engagement over factual accuracy (Uyheng & Carley, 2021), (3) the exploitation of anonymity

and weak social ties to amplify messages without normative constraints (Gramigna, 2022; Granovetter, 1973), and (4) populist and informal communication styles enabling unmoderated expression (Muddiman, 2017).

These dynamics are further reinforced by polarized media environments. Research on the Spanish media system shows that news consumption increasingly reflects citizens' ideological preferences, particularly during periods of political conflict. In such contexts, selective exposure to ideologically aligned media outlets can create segmented interpretive communities in which narratives about social issues (including religion and migration) are framed through partisan lenses. This process contributes to affective polarization, whereby political identities shape perceptions of social groups and intensify antagonistic attitudes (Cuéllar Rivero, 2024; Iyengar et al., 2012).

These dynamics worsen when platforms exhibit systemic deficiencies, including inconsistent moderation policies, narrow conceptions of harm that address only extreme cases, and reactive responses (Tarvin & Stanfill, 2022). These platforms function as mechanisms of symbolic governance, exerting distributed normative power that creates hierarchies of visibility without public deliberation (Gillespie, 2018). This is evident in pluralistic religious contexts, where legitimate anticlerical criticism coexists with minority stigmatization, facilitating hate normalization (Kennedy et al., 2022; Wells, 2023).

Low-intensity expressions like incivility and malicious speech occupy a gray area between freedom of expression and hate speech, creating normative ambiguity that enables normalization (Sue et al., 2007). This yields three effects: (1) reduced perception of transgression through multiple interpretations, such as narratives framing Islam as misaligned with national principles, blurring legitimate critique and hostility, a hallmark of microaggressions (Sue et al., 2007; Uyheng & Carley, 2021); (2) complicated moderation, as systems designed for explicit threats fail with low-intensity expressions, leaving 72% circulating unimpeded versus 3% of extreme forms; and (3) prejudice accumulating without institutional alarm, gradually integrating hostility into conversational background noise.

The normalization of low-intensity hate manifests across four dimensions: (1) Cognitive (Bandura, 1999; Kalmoe, 2014) exposure to veiled prejudice desensitizes individuals, transforming hostility into a "valid opinion"; (2) structural (Uyheng et al., 2022) lack of institutional sanction conveys implicit tolerance, authorizing hostile expressions and creating escalation cycles; (3) discursive (Bourdieu, 1991; Hall, 1997) media narratives legitimize hostility as "valid criticism," enabling hate while maintaining self-image as responsible citizens (symbolic violence); and (4) temporal (Sue et al., 2007) unsanctioned prejudice naturalizes into common sense, generating psychological impact equivalent to explicit hate. This process threatens social cohesion, as exposure to covert hostility facilitates escalation toward extreme intolerance (Wiedlitzka et al., 2021).

This process occurs in a mediated social order (Couldry & Hepp, 2018), where news media refract algorithmic dynamics through editorial decisions and moderation policies, shaping how religious conflicts are interpreted and debated in public discourse (Hall, 1997). Normalization operates as symbolic violence: discursive practices that transform hostility into a "debatable" discourse through repetition. This

normalization particularly affects Muslim and Jewish communities, frequently associated with violent conflicts perpetrated by far-right groups (Arce-García et al., 2024; Becker et al., 2022).

These dynamics unfold within a highly polarized media environment. Spain's media system has been characterized as a case of polarized pluralism, where news consumption increasingly reflects ideological and partisan preferences. Recent research shows that during polarized electoral cycles, political attitudes become stronger predictors of media consumption, reinforcing selective exposure to aligned outlets (Valera-Ordaz et al., 2025). In such contexts, audiences cluster around politically compatible sources, creating segmented interpretive communities in which narratives about social and cultural conflicts (including religious issues) circulate with limited cross-ideological contestation.

The dynamics are driven by social validation among individuals with similar ideological views (Walther, 2022), which intensifies prejudices. Algorithmic architectures emphasizing emotional virality exacerbate these dynamics (Uyheng & Carley, 2021), enabling expression through distinct narratives tailored to groups: institutional critiques (anti-Christian), identity representations (anti-Muslim), and security frameworks (anti-Semitic; Davani et al., 2023; Weinberg et al., 2025).

Polycentric Governance of Normalizing Amplification

The normalization of religious hate in digital media can be understood as a polycentric governance process in which the following multiple actors shape the visibility and acceptability of hostile discourse: (1) algorithmic platforms that prioritize emotional engagement, amplifying comments through reactions and creating feedback loops where hostility generates visibility and further hostility (Granovetter, 1973; Uyheng & Carley, 2021); (2) news outlets that shape editorial agendas through coverage decisions and narrative frameworks while applying variable moderation policies that convey institutional tolerance and structure "acceptable norms" (Alonso & Palomo, 2025; Arce-García et al., 2024; Minoie et al., 2023); and (3) users who internalize altered norms through social validation from like-minded individuals, perceiving concentrated support in segmented communities as evidence of acceptability (Walther, 2022).

The actors engage in a four-phase cycle: First, high-impact events like migration crises, terrorist attacks, and polarizing pronouncements generate uncertainty and prompt identity-based explanations. Subsequently, selective media coverage with ideologically differentiated framing activates specific mental schemas. This is followed by user reactions that show hostility, aligned with the media frame, which algorithms amplify due to engagement. Finally, circular reinforcement occurs, as amplification signals relevance, lack of moderation signals tolerance, and users intensify hostility for validation.

The "normative window of opportunity" draws from event-centered approaches, highlighting how disruptive events temporarily reshape social norms (Sewell, 1996, pp. 262–270), analogous to the "policy windows" in agenda-setting theory (Kingdon, 1984). We operationalize it as a three-day period surrounding a high-impact event, consistent with research showing reactions concentrate within the first 48–72 hours (Uyheng & Carley, 2021). This window requires three conditions: emotionally impactful events prompting

identity explanations, intensive media coverage activating identity narratives, and a temporary absence of moderation signaling institutional tolerance.

All three conditions must concur; the absence of any one prevents normative relaxation. Although norms tend to revert to baseline after three days (Uyheng & Carley, 2021), the ratchet effect suggests that successive windows produce gradual normalization, progressively elevating the tolerance threshold and transforming what were once exceptional expressions into a residual “new normal” (Kalmoe, 2014).

The Spanish Context of Multifaceted Governance Around Religious Hatred

Spain’s strong anticlerical traditions have historically framed criticism of the Catholic Church within secularization debates (Ledesma, 2001). VOX’s emergence has consolidated anti-Muslim narratives framing Islam as incompatible with national identity (Corral et al., 2023), while growing polarization has fueled religious hostility, particularly around the Israel–Gaza conflict. These dynamics unfold in a society of nominal Catholicism and growing pluralism, where religious discourse intersects with ideological polarization (Cabrera & González, 2022; Observatorio del Pluralismo Religioso, 2025). In 2024, Spain recorded 1,955 hate crimes; racism/xenophobia led with 804 cases, while anti-Semitism and aporophobia showed the largest year-on-year increases (Ministerio del Interior, 2024).

Spain’s secularization represents a hybrid model that, distinct from other European contexts, retains Catholic Church privilege (Torres Gutiérrez, 2025), creating a normative gray area. This framework permits anticlerical critique between legitimate inquiry and hostility, while antagonism toward minorities uses Catholic tradition as an exclusionary national identity symbol.

Within this framework, polycentric governance is delineated by (1) the tension between Catholicism and secularization, which normalizes anticlerical criticism (Herranz-de-Rafael & Fernández Prados, 2024); (2) xenophobic narratives from the far right, which link national identity with anti-Muslim discourse (Pérez Escolar et al., 2025); and (3) migration crises that serve as catalysts, creating “normative windows of opportunity” for such animosity (Pons Raga, 2024, p. 10).

In this context, expressions of religious hostility often intersect with partisan identities, legitimized as “defense” through selective media coverage, and media consumption patterns. Within Spain’s “polarized pluralism” media system (Carratalá & Valera-Ordaz, 2020; Humanes & Valera-Ordaz, 2022), selective exposure to ideologically aligned media may reinforce identity-based interpretations of religious issues, shaping how hostility toward different religious groups is articulated in public discourse.

Methodology

This study is a longitudinal, exploratory analysis of social media content, examining religious hate speech patterns in user comments on X from six Spanish media outlets with diverse ideological orientations during March–April 2025. Using multivariate methods, including multiple correspondence analysis, semantic

clustering via UMAP-HDBSCAN,² and time-series analysis, the research investigates associations among media outlet stance, religious hate types (anti-Catholic, anti-Muslim, anti-Semitic), discourse intensity, and impactful external events. The objective is not to establish causality between editorial variables and hostility patterns, but to characterize their associations in the Spanish context.

Objectives and Hypotheses

This study aims to analyze the characteristics and patterns of religious hate speech in responses to news content from Spanish media outlets on X. To this end, the following specific objectives have been established:

- SO1: To determine the prevalence, distribution of religious hate (anti-Catholic, anti-Muslim, anti-Semitic), and intensity levels in relation to news content published by media outlets on X.*
- SO2: To characterize the patterns of association among media outlets, types of religious hate, and intensity levels using multiple correspondence analysis, evaluating both the statistical significance (χ^2 test) and the stability of the graphical representation (explained variance, quality of representation $\cos^2$) of these associations.*
- SO3: Identify and characterize the narrative frameworks structuring anti-Christian hatred (internal institutional criticism) versus anti-Muslim/anti-Semitic hatred (geopolitical-identity conflicts), assessing whether these distinctions reflect different modes of articulating religious hostility depending on the target group.*
- SO4: Analyze the temporal relationship between high-impact news events and peaks in the volume and intensity of religious hate speech.*

The hypotheses that guide the development of this study are:

- H1: Editorial lines generate differentiated regimes of acceptability for religious hatred, which is reflected in systematic associations between the medium and the type of hostility.*
- H2: The normalization of low-intensity religious hatred is a systemic process that operates across the board, regardless of the editorial line, indicating that the lack of consistent moderation is a structural characteristic of the digital ecosystem, not a differentiated editorial decision.*
- H3: High-impact events produce temporal variations in the intensity of hate, revealing periods of relaxation in normative expression thresholds.*

² Uniform manifold approximation and projection (UMAP) for dimensionality reduction with hierarchical density-based spatial clustering of applications with noise (HDBSCAN) for identifying clusters.

H4: The semantic and narrative structure of religious hate varies significantly depending on the target group.

Study Population and Sample

Data collection was conducted in March–April 2025 on X using the Basic API, analyzing responses to news content from six Spanish media outlets.³ This methodology recognizes that comment sections on social media are significant hotspots for toxicity, disinformation, and offensive language, with implications for inclusive debate and media credibility (Frischlich et al., 2019; Madhyastha et al., 2023). The March–April 2025 period was selected for its analytical relevance, coinciding with a post-electoral cycle marked by high polarization (following the 2023 elections). It encompasses high-impact events such as the DANA storm in Valencia (March 1), protests during *masclètàs* (March 1, in an emotionally charged post-DANA context), the migration crisis with the *cayuco* tragedy (March 13; 153 messages), Israel bombings (March 19), and the Popular Party (PP) video (March 6–7).

Of 53,787 messages collected, 19,121 contained hate speech across general, political, misogynistic, sexual, xenophobic, and religious categories (Montero-Díaz et al., 2025; Said-Hung et al., 2024). The analysis focused on religious hate messages targeting anti-Christian, anti-Muslim, and anti-Semitic sentiments.

Platform X was selected based on two factors: (1) its position as a leading social media platform for news consumption in Spain, with significant news outlet presence (Reuters Institute, 2023); and (2) its architecture supporting real-time interactions and short-text format, facilitating contextualized hate analysis (Pérez et al., 2022; Walther, 2024).

The selection of media outlets applied three criteria: (1) activity on X, (2) national news coverage, and (3) public ideological perception as reported by the Center for Sociological Research (2023). Six outlets were selected across three ideological blocs: progressive (El País, n.d.; Público, n.d.), centrist (20Minutos, n.d.; El Confidencial, n.d.), and right-wing (ABC, n.d.; OKDiario, n.d.). This distribution characterized patterns between editorial stance and types of religious hate, maximizing editorial variation without attempting to cover the entire Spanish media ecosystem. The classification draws on CIS Study 3421 (2023), in which El País (n.d.) and Público (n.d.) score 3.2–3.5 on a 1–10 left-right scale, ABC (n.d.) and OKDiario (n.d.) score 7.1–8.2, and El Confidencial (n.d.) and 20Minutos (n.d.) score 4.8–5.3, reflecting centrist appeal across mixed ideological audiences. This public perception-based approach ensures representativeness without self-referential bias.

Coding and Reliability Validation Procedure

Hate speech was labeled following the methodology developed by Said-Hung et al. (2024), which establishes rigorous protocols for classifying hostile expressions. Specifically, two researchers were trained over four weeks in March 2025 for this task. The training included (a) an introduction to religious hatred

³Supplementary files about methodology: <http://bit.ly/4pFS4CX>.

typology—anti-Catholic, anti-Muslim, and anti-Semitic sentiments, as outlined by Montero-Díaz et al. (2025); (b) instruction in recognizing message intensities, rated 1–4, where 1 represents uncivil messages, 2 denotes malicious messages, 3 indicates insulting messages, and 4 signifies threatening messages, through examining examples; and (c) calibration between coders through joint coding of 200 messages, with discussions to resolve discrepancies.

This study uses “anti-Catholic,” “anti-Muslim,” and “anti-Semitic” terminology for three reasons. First, “anti-Semitism” is standard in international hate speech literature (Becker et al., 2022; International Holocaust Remembrance Alliance (IHRA), 2016; Weinberg et al., 2025), denoting hostility toward Jews as an ethnoreligious group, rather than “anti-Judaism,” which refers to religious-theological rejection. Second, “anti-Muslim” better characterizes hostility directed at Muslims as a group, avoiding psychopathological implications of “Islamophobia” (Contreras Mazarío, 2016; Halliday, 1999), which may conflict with our framework of symbolic violence. Third, empirical analysis supports this terminology: Anti-Semitic narratives are structured around geopolitical conflicts (Israel-Palestine), while anti-Muslim narratives focus on identity and security stereotypes, both distinct from institutional anti-Catholic critiques. This nomenclature aligns with narrative structures observed in the analyzed corpus.

- Acceptability thresholds: A maximum tolerated error of 7% was established for both categorization dimensions.
- Quantitative error (percentage of messages categorized differently by coders): $\leq 7\%$ qualitative error (suitability of labeling according to type and intensity of hate). This threshold was reached after the initial calibration phase, after which each coder independently labeled their portion of the corpus.
- Academic validation: Two scholars with expertise in hate speech analysis conducted an independent review of labeled messages quantitatively. They randomly selected 10% of the total sample ($n = 1,912$ messages) for qualitative analysis to verify coders’ labeling matched the established typology and intensity criteria. This validation process facilitated identifying and rectifying potential systematic coding biases.
- Agreement statistics: Since it is impossible to calculate Cohen’s kappa on such a large corpus with two independent coders (because of technical limitations of data overload), agreement was assessed using:
 - Percentage of direct agreement in validation subsample: 91% (95% CI: 89–93%)
 - Krippendorff’s alpha for ordinal data (intensity of hate): 0.84 (95% CI: 0.79–0.89)
- Disagreement resolution protocol: When initial discrepancies arose among coders (7% of initial corpus), this protocol was implemented: (1) joint review of the message, (2) discussion of applied criteria, (3) consensus by majority vote among coders, and (4) re-execution of academic validation. This methodology ensures that classification of 635 religious hate messages, which forms the basis of empirical analysis, is reproducible and validated through standardized content analysis procedures.

Techniques and Data Processing

The analysis encompasses three dimensions: (1) the conceptual dimension, characterizing religious hatred as a continuum of intensities that undermine social cohesion (Hartikainen, 2025; Kiper, 2023; Montero-Díaz et al., 2025); (2) the contextual dimension, identifying dynamics of social validation, selective exposure to media frameworks, and algorithmic architectures through association patterns (Granovetter, 1973; Uyheng & Carley, 2021; Weinberg et al., 2025); and (3) the media-related dimension, conceptualizing media as intermediaries influencing hatred's dissemination or containment through editorial decisions and comment management (Alonso & Palomo, 2025).

For SO1 and SO2, multiple correspondence analysis (MCA) depicted relationships among categorical variables (medium, type of hate, intensity) within a two-dimensional framework. To operationalize H1, three proxies were used: (1) co-occurrence between medium and type, assessed through MCA and χ^2 test; (2) differentiated semantic structure analyzed using UMAP-HDBSCAN; and (3) differentiated distribution of intensities by medium, as an indirect indicator of editorial tolerance.

MCA visualizes associations among categorical variables (media outlet, hate type, intensity) in two-dimensional space, accounting for 95.1% of variance (D1⁴: 78.5%, D2: 16.6%; $\cos^2 > 0.70$). This enables identification of editorial acceptability regimes without implying causality. Uniform manifold approximation and projection for dimension reduction (UMAP)-HDBSCAN then clusters semantic narratives: Multilingual embeddings (384 dimensions) reduce to 2D via UMAP, with density-based clustering (HDBSCAN; separability > 0.60) distinguishing anti-Catholic institutional critiques from anti-Muslim/anti-Semitic geopolitical conflicts. These methods complement descriptives, triangulating evidence on polycentric governance.

Messages were converted into 384-dimensional vectors using multilingual embeddings, projected via UMAP (refer to technical details in Supplementary Material A.1), and analyzed through density clustering (HDBSCAN; see A.2). This methodology identifies distinct narratives, distinguishing between anti-Christian discourse (institutional critique) and anti-Muslim/anti-Semitic discourse (geopolitical-identity conflicts), enabling assessment of their associations with editorial lines and intensities.

For SO4, ± 3 -day windows were delineated around events and validated through multivariate sensitivity analysis spanning 1–7 days. This operationalization determines whether peaks in volume and intensity are associated with high-impact events, such as migration crises, geopolitical conflicts, and terrorist attacks. It also assesses if low intensities are prevalent, indicating normalization.

Robustness was assessed using standardized criteria (refer to metrics in Supplementary Material A.3): (1) AMS: variance explained ($> 90\%$), cosine square (> 0.70), chi-square test ($\alpha = 0.05$); (2)

⁴ Variance (D1): The first component captures the direction of maximum variance in the original data. Variance (D2): The second component is orthogonal to D1 and captures the remaining maximum variance. $\cos^2$ measures how well an original variable is represented in a principal component—specifically, the square of the cosine of the angle between the variable and the component axis. Values close to 1 indicate that most of the variable's information resides in that component.

semantic analysis: qualitative-quantitative triangulation through keyword inspection per cluster, intercluster separability (> 0.60); (3) temporal analysis: validation of ± 3 -day window using multivariate sensitivity. The integration of MCA, semantic clustering, and temporal analysis forms a robust protocol that links the three levels of polycentric governance, with evidence triangulation strengthening exploratory conclusions without assuming causal determinism. Development used Python 3.11.6 (see libraries and versions in Supplementary Material A.4).⁵

Results

This study analyzes, using multivariate techniques, the representation of religious hate in Spanish media. The results are organized according to the proposed specific objectives (SO1–SO5) to ensure clarity and thematic coherence, avoiding redundancies.

Prevalence and Distribution of Religious Hate

To address objective SO1 and provide a preliminary answer to Q1, the distribution of 640 religious hate messages identified in March–April 2025 was analyzed by medium of origin, type of hate, and intensity level. Tables 1 and 2 present this descriptive characterization and provide the empirical basis for subsequent multivariate analyses examining the associations among these variables.

Table 1. Religious Hatred Detected According to the Media Type of Hatred (% Row/Column Totals).

News Media	Anti-Catholic	Anti-Muslim	Anti-Semitic	Total
20Minutos (n.d.)	35.7% (25)	28.6% (20)	15.70% (11)	8.8% (56)
ABC (n.d.)	26.2% (44)	32.1% (54)	9.5% (16)	18.0% (114)
El País (n.d.)	20.6% (65)	25.6% (81)	22.5% (71)	34.2% (217)
El Confidencial (n.d.)	36.7% (11)	36.7% (11)	10.0% (3)	3.9% (25)
OKDiario (n.d.)	23.4% (22)	42.6% (40)	9.6% (9)	11.1% (71)
Público (n.d.)	21.8% (45)	25.1% (52)	24.6% (51)	24.1% (148)
Total	33.1% (212)	40.4% (258)	25.2% (161)	100% (640)

Table 2. Religious Hatred Detected According to Medium and Intensity (% Row/Column Totals).

News Media	1 (Uncivil Messages)	2 (Malicious Messages)	3 (Insulting Messages)	4 (Threatening Messages)	Total
20Minutos (n.d.)	39.3% (22)	44.6% (25)	14.3% (8)	1.8% (1)	8.8% (56)
ABC (n.d.)	32.2% (56)	23.6% (41)	8.0% (14)	2.3% (4)	17.8% (115)
El País (n.d.)	24.8% (79)	28.3% (90)	14.8% (47)	0.9% (3)	34.1% (219)
El Confidencial (n.d.)	36.7% (11)	33.3% (10)	13.3% (4)	0.0% (0)	3.9% (25)
OKDiario (n.d.)	26.5% (26)	33.7% (33)	10.2% (10)	2.0% (2)	11.1% (71)

⁵Methodology supplementary files: <http://bit.ly/4pFS4CX>.

Público (n.d.)	23.6% (53)	25.0% (56)	17.4% (39)	0.4% (1)	23.3% (149)
Total	38.6% (247)	39.8% (255)	16.9% (122)	1.7% (11)	100% (635)

Religious hate distribution (Tables 1–2) shows El País (n.d.) accounted for 34.0% of cases (219), Público (n.d.) 23.1% (149), ABC (n.d.) 17.6% (115), OKDiario (n.d.) 11.0% (71), 20Minutos (n.d.) 8.7% (56), and El Confidencial (n.d.) 3.9% (25). This concentration within progressive outlets reflects higher comment volume, not greater proportions of hate.

Anti-Muslim sentiment predominated with 258 instances (42.5%), followed by anti-Catholic with 212 cases (30.6%) and anti-Semitic with 161 cases (26.9%), reflecting Spain's religious intolerance patterns. Intensity levels 1 and 2 comprised 78% (247 at level 1 [38.4%], 255 at level 2 [39.6%]); level 3 represented 18.9% (122 cases), and level 4 was residual (11 cases, 1.7%).

Semantic Structure and Differentiated Discursive Approaches

To answer Q3 and comply with SO3, semantic analysis was implemented using transformer embeddings, UMAP reduction, and HDBSCAN clustering. Figures 1 and 2 reveal differentiated narrative frameworks according to the target group.

Anti-Christian messages are characterized by institutional criticism using terms such as "church," "pope," "Vatican," "priests," "pedophiles," and "money," with accusations of corruption and abuse within the ecclesiastical hierarchy. The tone is predominantly anticlerical and derogatory toward clerical institutions and figures.

Anti-Semitic messages focus on "Israel," "genocide/genocidal maniac," "Zionist/Zionism," "Palestinians," "war," and "USA," linking Jews to Israel and accusing them of genocide. The rhetoric is belligerent and conspiratorial, reinforcing negative stereotypes.

Anti-Muslim messages center on "Islam," "Moors/Moor," "terrorists," "women," "Spain/Europe," and "Hamas/Hamas," with stereotypes of violence and concerns about migration and security. The tone is xenophobic and security-oriented, reflecting negative perceptions of the Muslim community.

This analysis substantiates the theory of polycentric governance by illustrating that religious animosity is expressed through distinct narrative frameworks: Anti-Christian discourses function as internal institutional critiques, whereas anti-Muslim and anti-Semitic narratives manifest as external geopolitical-identity conflicts.

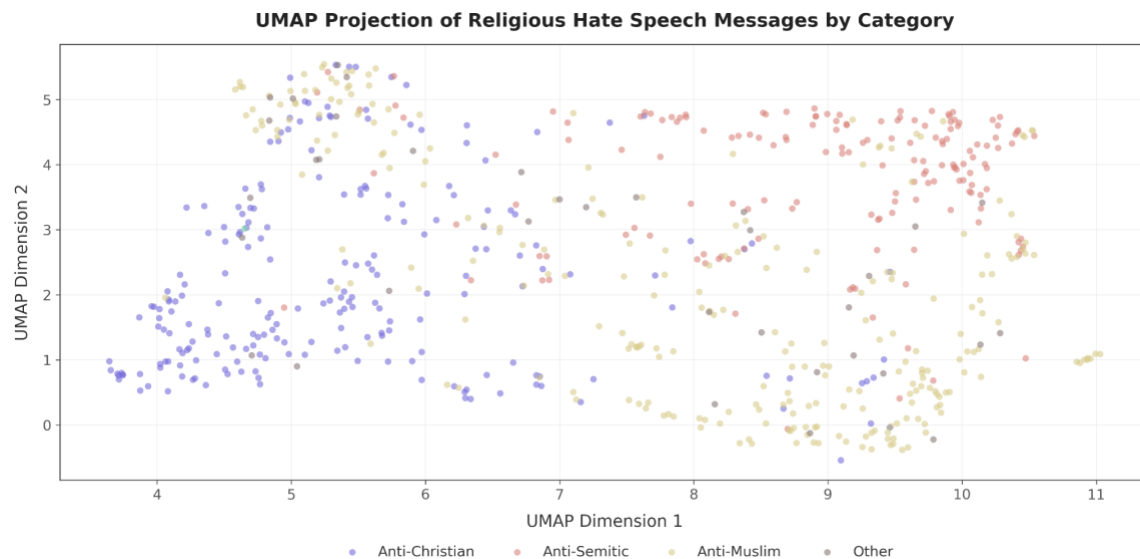


Figure 1. UMAP projection in two dimensions around the objective and focus of the messages.

Figure 1 shows semantic bifurcation: anti-Christian discourse (left, terms: “church,” “pope”) versus anti-Semitic/anti-Muslim discourse (right, terms: “Israel,” “terrorists”), reflecting institutional critique versus geopolitical conflict. The Y-axis (discourse type) separates geopolitical approaches at the top, addressing conflicts with inflammatory language, from religious/cultural approaches at the bottom, which delegitimize religious practices and promote intolerance.

The yellow cluster shows anti-Muslim sentiment and integrates Islam within a geopolitical context, while the purple cluster, marked by anti-Christian sentiment, emphasizes institutional critique. Figure 2 categorizes these patterns: Cluster 3 (purple, $n = 113$, $X = 4.70$, $Y = 1.43$) is the largest, focusing on Catholic institutions with terms “church,” “pope,” and “Catholic,” indicating rejection of these institutions. Cluster 1 (purple, $n = 22$, $X = 4.66$, $Y = 3.19$) engages with abstract religious discourse and geopolitical themes. This analysis shows religious animosity patterns: Anti-Christian sentiment appears as institutional critique, while anti-Semitic and anti-Muslim sentiments manifest as geopolitical-identity conflicts.

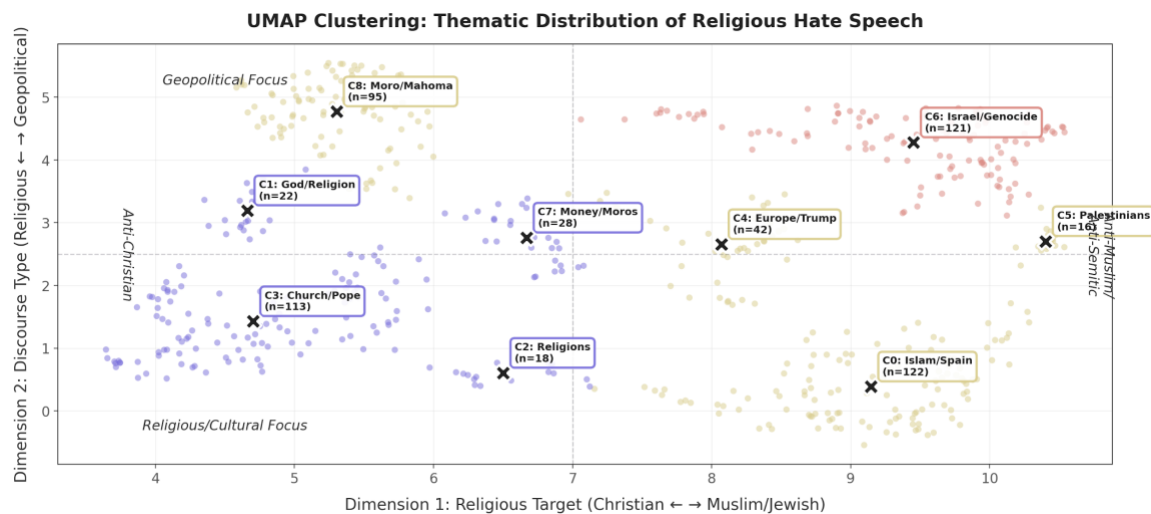


Figure 2. UMAP projection of the clusters in two dimensions according to the objective and type of discourse.

Cluster 3 (purple, $n = 113$, $X = -4.70$, $Y = -1.43$) dominates anti-Catholic discourse with institutional critique keywords ("church" frequency = 0.42, "pope" 0.31, "pedophiles" 0.18), rejecting corrupt clerical authority ("money," "Vatican"), aligning with Spain's left anticlerical reform tradition, not ethnic targeting. In contrast, clusters 0/6 ($n = 122/121$) split anti-Muslim/anti-Semitic narratives: Cluster 0 (yellow, $X = 9.15$, $Y = 0.39$) stresses cultural clash ("Islam," "Moors," "Spain"; separability = 0.68), cluster 6 geopolitical violence ("Israel," "genocide," "Hamas"). Intercluster separability exceeding 0.60 confirms robust structural divergence, fully supporting H4 on differentiated narrative frameworks.

The lower-right quadrant illustrates anti-Muslim discourse with a religious emphasis. Cluster 0 ($n = 122$, $X = 9.15$, $Y = 0.39$) is the largest, focusing on "Islam," "Muslims," and "Spain," thereby reflecting cultural perceptions of Islam within Spanish society. In contrast, cluster 2 ($n = 18$, $X = 6.50$, $Y = 0.61$), which is smaller, serves as a comparative analysis of religions situated in the transitional zone between anti-Christian and anti-Muslim discourse.

The upper-right quadrant delineates the geopolitical anti-Muslim/anti-Semitic discourse group. Specifically, cluster 6 ($n = 121$) is characterized by anti-Semitic discourse concerning Israel and genocide, employing terms such as "Israel," "genocide," and "Palestinians." Cluster 5 is centered on Israeli-Palestinian conflict, while cluster 4 links Islam to European politics using terms like "Europe," "Trump," and "terrorists." The upper-left quadrant integrates a multifaceted geopolitical strategy: Cluster 8 employs specific terminology aimed at Muslims in an intermediate context, while cluster 7 amalgamates economic and religious narratives.

The patterns reveal a structured logic: The X-axis delineates anti-Christian (left) from anti-Muslim/anti-Semitic (right) discourses, while the Y-axis differentiates religious/cultural (lower) from

geopolitical (upper) narratives. Anti-Christian discourse is thus positioned as internal institutional critique and anti-Muslim/anti-Semitic discourse as geopolitical-identity narratives. Media focusing on security issues are predicted to amplify anti-Muslim hate, while outlets with institutional critique traditions are expected to amplify anti-Catholicism. MCA corroborates these predictions, indicating that associations between media and discourse types reflect polycentric governance rather than random occurrence.

Association Among Media Outlet, Type of Hate, and Intensity

To address Q1 and achieve SO2, a two-dimensional MCA was conducted, projecting relationships among categorical variables (media outlet, type of hate, intensity) onto a reduced factor space. This analysis determines if systematic associations exist between a media outlet's editorial line (perceived ideological orientation: progressive, center, right) and the types and intensities of religious hate in its comment sections. Figures 3 and 4 illustrate these associations as biplots, where spatial proximity signifies statistically significant co-occurrences, validated by the χ^2 test.

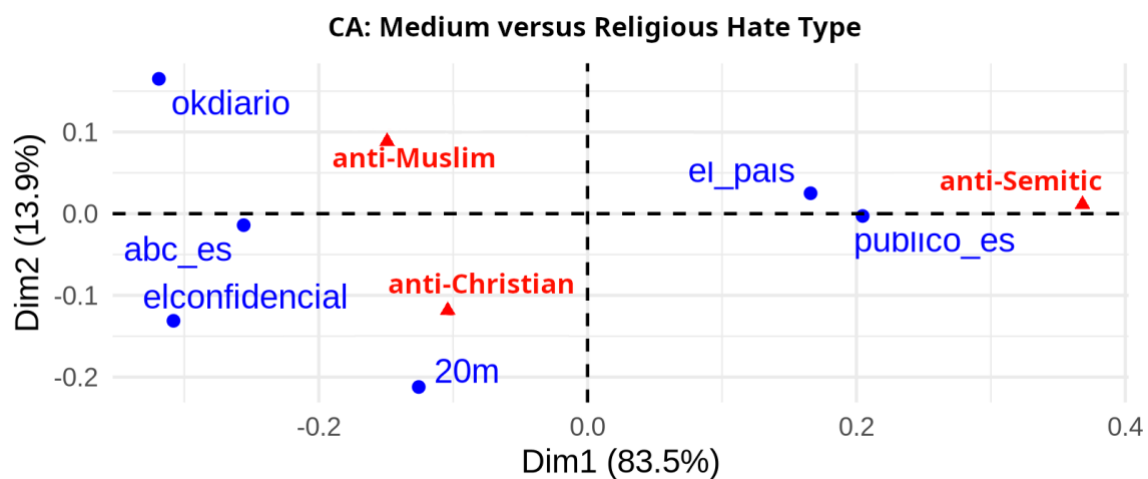


Figure 3. Two-dimensional diagram of media outlets according to the type of religious hate.

The MCA analysis of intensity distributions reveals an additional governance aspect: identifying which hate types are amplified in media outlets and the levels of hostility they exhibit. This detail helps determine whether editorial permissiveness is consistent across types or if outlets show varying tolerance for low- versus high-intensity expressions.

Statistical proximity between media and hate forms, while not implying causality, indicates systematic co-occurrences influenced by news agendas, audiences, or social network dynamics. These findings suggest differentiated protocols based on type and intensity: Media outlets linked to "insult" (3) need language filters and educational counternarratives; proximity to "threat" (4) requires stringent

measures, including authority reporting. Counterdiscourse campaigns can be advantageous by focusing on structured types and media outlets where presence is pronounced.

The correspondence analysis between medium and type of religious hate yields a significant representation, accounting for 97.4% of the variance explained. The study shows high $\cos^2$ values (means ≈ 0.815 , intensities ≈ 0.946) and a χ^2 test ($p \approx 0.0016$), indicating associations are not due to chance. This statistical reliability ensures the two-dimensional positions reflect the underlying relationships, supporting the map's interpretation with confidence. This is crucial for identifying associations and informing editorial, moderation, or monitoring decisions.

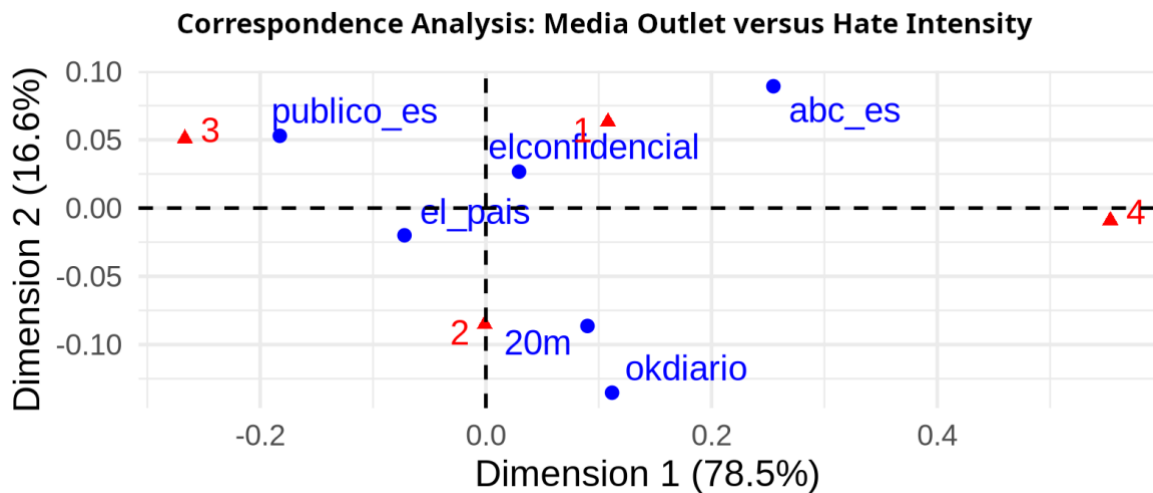


Figure 4. Two-dimensional diagram of mean versus intensity of hatred.

MCA on media outlets, types of hate speech, and intensity accounts for 95.1% of the variance, with the first dimension explaining 78.5% and the second 16.6%. The intensity categories "insult" (3) and "threat" (4) most significantly structure these dimensions, while "incivic" (1) and "malicious" (2) also influence them. Media outlets ABC and Público are predominant in the first dimension, OKDiario is associated with the second dimension, and El Confidencial is linked with the third dimension.

Media outlets' proximity to vectors like "threat" (4) suggests heightened intensity, while proximity to "insult" (3) or "incivic" (1) indicates verbal aggression. From an operational perspective, media outlets in these categories should review their content, headlines, and comment dynamics. This may involve revising style guides, implementing moderation, and temporarily closing comments on polarizing content.

The graphical reliability is high, with 95.1% of variance explained and $\cos^2$ values for media and intensities of 0.815 and 0.946, ensuring accurate representation. The χ^2 test ($\chi^2 = 19.03$, $p \approx 0.212$) is not statistically significant because of the low frequency of categories such as anti-evangelical, which limits prediction robustness (random forest precision = 0.41, 95% CI ≈ 0.34 – 0.48). While the medium elucidates

anti-Muslim and anti-Catholic patterns, it remains inadequate for predicting anti-evangelical and anti-Semitic patterns.

Figures serve as instruments for exploration and monitoring, but do not enable definitive assertions. Static association patterns characterize governance at equilibrium; however, significant emotional events temporarily lower normative thresholds. The temporal analysis examines how surges in hate volume and intensity correlate with external events, creating normative windows of opportunity, when acceptability thresholds are relaxed.

Temporal Dynamics and Promotion With Events

To address Q2 and satisfy OE4, time-series analysis was performed using ± 3 -day windows surrounding each event. This approach aims to determine whether emotionally impactful events lead to surges in religious hostility, indicating relaxed discursive acceptability thresholds, as posited by the normative window of opportunity theory. Figure 5 shows significant correlations between news events and spikes in religious animosity, establishing a connection among media coverage, digital discourse, and religious intolerance.

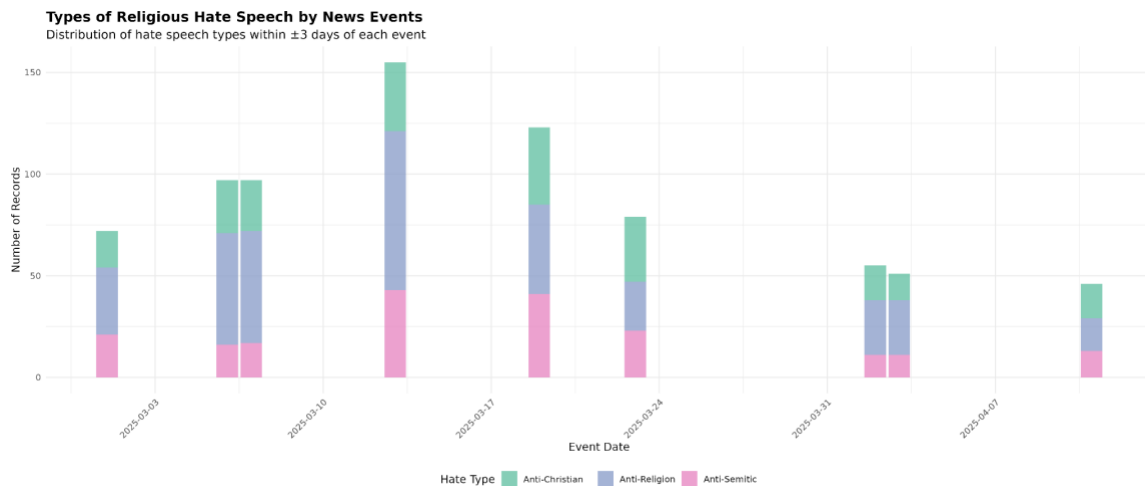


Figure 5. Association of messages with relevant news throughout the analyzed period.

Event selection criteria for Figure 5 involve emotional salience (media coverage $> 10\times$ baseline), theoretical fit (identity activation), and empirical peaks (volume $+20\%$, intensity $+0.3$). DANA/masclatà protests (March 1) link via Catholic rituals amid disaster, spiking anti-Catholic (intensity 2.38); cayuco (March 13) triggers anti-Muslim migration frames (153 messages); Israel bombings/Cruces (19) blend ritual/geopolitics for anti-Semitic surge. Non-inclusion of minor events (e.g., routine politics) due to $< 5\%$ variance explained.

Figure 5 presents the main findings on temporal dynamics:

- High volume of religious hatred: Notable concentrations are observed around specific events. The bombing in Israel, coinciding with the Cremà in Valencia (March 19), generated 117 posts within ± 3 days; the cayuco tragedy (March 13) reached 153 posts; a PP video (March 6–7) accumulated 100–102 posts.
- High average intensity: The protest during the mascletà in the context of the DANA storm in Valencia (March 1) presented the highest average intensity (2.38), followed by the Cremà plus bombings in Israel (2.09) and the PP video (1.84), suggesting that the emotional charge affects both volume and tone.
- Differentiated patterns by type of hatred: General hatred predominates in migration contexts; anti-Semitism intensifies in conflicts between Israel and Gaza; anti-Christianity predominates in local political events. International events (armed conflicts) trigger anti-Semitism; migrant tragedies provoke a general antireligious discourse, linked to cultural rejection or xenophobia; local political events (e.g., the PP video) generate anti-Christianity, reflecting ideological tensions or criticism of religious leaders.

Discussion

The findings offer systematic empirical evidence supporting the four hypotheses, illuminating patterned associations that characterize polycentric governance within Spain's polarized digital media ecosystem. MCA (Figures 3–4; 95.1% variance explained, $\cos^2 = 0.815$) reveals notable spatial proximity between conservative outlets (ABC, OKDiario) and anti-Muslim hate, contrasting with more dispersed patterns among progressive media (El País, Público). This provides partial support for H1 on differentiated editorial acceptability regimes. While visual associations are robust, statistical significance remains marginal ($\chi^2 = 19.03$, $p = 0.212$) and predictive power limited (random forest precision = 0.41, 95% CI [0.34–0.48]), likely because of confounding factors such as audience self-selection, varying comment volumes, and unmeasured moderation practices. These patterns suggest (without establishing causality) that outlet ideologies co-occur with specific hostility types, structuring normative friction in comment sections.

Building on this, H2 finds strong confirmation in the uniform dominance of low-intensity hate (levels 1–2: 78.4% overall; Table 2), spanning outlets from 84% in 20Minutos to 53% in El País. This cross-ideological prevalence underscores systemic normalization rather than targeted editorial bias, aligning with prior evidence that 72% of low-intensity expressions evade moderation because of their normative ambiguity (Sue et al., 2007). Such persistence transforms veiled prejudice into conversational “common sense,” eroding cohesion without triggering overt sanctions.

Temporal dynamics further validate H3: Figure 5 documents clear spikes in volume and intensity (+0.3–0.4 points) within three-day normative windows around high-impact events (e.g., cayuco migration tragedy with 153 messages; DANA/mascletà protests: intensity 2.38), followed by reversion to baseline. These surges empirically demonstrate event-driven relaxation of discursive thresholds, with cumulative ratchet effects potentially elevating long-term tolerance (Kalmoe, 2014). The specificity of spikes to identity-salient triggers (migration for anti-Muslim, rituals for anti-Catholic) reinforces the theory's conditional logic.

H4 emerges most robustly from UMAP-HDBSCAN clustering (intercluster separability > 0.60), which delineates anti-Catholic discourse as institutional critique (cluster 3, $n = 113$: “church” frequency = 0.42, “pope” 0.31, “pedophiles” 0.18; $X = -4.70$, $Y = -1.43$) from anti-Muslim/anti-Semitic geopolitical frames (cluster 0 $n = 122$: “Islam/Moors/Spain”; cluster 6 $n = 121$: “Israel/genocide/Hamas”). This structural divergence confirms distinct narrative scaffolds enabling differentiated hostility articulation, with transitional clusters (e.g., cluster 2) hinting at hybrid amplification risks.

The convergence of H1–H4 validates the polycentric governance model. Three interdependent actors mutually enhance normative influence through reinforcing cycles: algorithmic platforms amplifying emotionally engaging content regardless of validity (Uyheng & Carley, 2021), news outlets whose variable moderation policies signal institutional tolerance (Alonso & Palomo, 2025; Minooie et al., 2023), and users internalizing altered norms through peer validation (Walther, 2022). These levels function interdependently—algorithms require selective editorial coverage to activate cognitive schemas; media sustain poor moderation, believing hostility is commercially beneficial; users perceive hostility as acceptable only by observing institutional tolerance. Each amplifies others, creating cycles distinct from isolated algorithmic effects. Operationally, this constitutes symbolic violence (Bourdieu, 1991; Hall, 1997)—domination through narratives rendering hostility “common sense.” The 72% of hate circulating at low-moderate intensity exemplifies symbolic violence: functioning unintentionally, exploiting ambiguity enabling hostility denial, producing psychological impact equivalent to explicit expressions, and incrementally undermining social cohesion.

Conclusion

This study presents multivariate quantitative evidence indicating that religious animosity normalization in Spain functions as a polycentric process. This process is governed by three interdependent entities: algorithmic architectures, news media editorial decisions, and social validation dynamics among ideologically segmented users. Contrary to views that emphasize algorithms or attribute hostility to user characteristics, the data reveal that news media serve as critical intermediaries, possessing autonomous decision-making capacity and structuring differentiated tolerance regimes.

Despite limitations in the unproven causality of the analyzed data, the indirect operationalization of moderation through the perceived ideological orientation of media outlets, and temporal constraints on data access on platforms such as X, the findings remain valid within their specific scope. This validity stems from the research design and analytical techniques employed. Moreover, the study identifies documented patterns within the analyzed context, which contribute to redefining our conceptualization of the normalization of online hate, thereby challenging three previous theoretical simplifications:

Redefining the role of the media: Previous research has focused on algorithmic amplification (Uyheng & Carley, 2021) or user characteristics (Bandura, 1999). This study shows that news media act as autonomous agents with decision-making abilities, creating distinct regimes of representation. These regimes systematically amplify specific hatreds while restraining others. The empirical inclusion of the media dimension is the central theoretical contribution of this study.

Operationalization of symbolic violence: Bourdieu (1991) and Hall (1997) conceptualized symbolic violence as domination through internalized frameworks that legitimize hierarchies. This study reveals that 72% of hate at low to moderate intensity constitutes symbolic violence. This is because it (a) functions without explicit intent, (b) exploits ambiguity that permits denial of hostility, (c) produces psychological impact comparable with overt expressions, and (d) undermines social cohesion.

Conceptual shift regarding normative responsibility: Current regulatory frameworks assume harm is primarily linked to explicit content. This study shows that normalizing low-intensity behaviors is equally detrimental and requires shifting from editorial responsibility for intent to institutional responsibility for effect, regardless of stated intentions.

These findings have immediate implications. The governance of online religious hostility cannot be reduced to algorithmic decisions or editorial responsibility; it requires systemic consideration of interactions among technology, media, and citizens in polarized contexts. While H1 (editorial differentiation) needs further validation, H2 (systemic normalization) is robust and has clear implications: The structural mechanisms of Spain's digital ecosystem permit 72% of low-intensity religious hate to circulate in normalized form, with consequences for social cohesion that are insufficiently theorized in prior literature.

Several limitations temper generalizability. First, the sample (640 religious hate messages from 53,787 comments) derives from six outlets on X, potentially overlooking platform-specific dynamics (e.g., Instagram's visuals) or underrepresented voices (e.g., non-news sites). Second, exploratory MCA/UMAP-HDBSCAN characterizes associations, not causality—editorial effects may reflect audience self-selection or unmeasured moderation. Third, the March–April 2025 window captures post-electoral polarization, but limits long-term trends; religious hate evolution (e.g., post-2025 events) remains unexplored. Inference scope is thus context-bound to Spain's hybrid secularism.

Future research should incorporate editorial policies, algorithmic dynamics, audience demographics, and electoral cycles to enhance causal rigor. Concurrently, data-driven interventions across regulation, platform design, and editorial practices are imperative, as current frameworks address only fractional issues.

References

20Minutos [@20m]. (n.d.). *Posts* [X profile]. X. Retrieved May 15, 2026, from <https://x.com/20m>

ABC [@abc_es]. (n.d.). *Posts* [X profile]. X. Retrieved May 15, 2026, from https://x.com/abc_es

Alorainy, W., Burnap, P., Liu, H., Williams, M., & Giommoni, L. (2022). Disrupting networks of hate: Characterising hateful networks and removing critical nodes. *Social Network Analysis and Mining*, 12(1), 1–22. <https://doi.org/10.1007/s13278-021-00818-z>

<https://doi.org/10.65476/4vd67z56>

- Arce-García, S., Said-Hung, E., & Montero-Díaz, J. (2024). Unmasking coordinated hate: Analysing hate speech on Spanish digital news media. *New Media & Society*, 27(10), 5848–5868. <https://doi.org/10.1177/14614448241259715>
- Bandura, A. (1999). A social cognitive theory of personality. In L. Pervin & O. John (Eds.), *Handbook of personality: Theory and research* (pp. 154–196). New York, NY: Guilford Press.
- Becker, J. J., Ascone, L., & Troschke, H. (2022). Antisemitic comments on Facebook pages of leading British, French, and German media outlets. *Humanities & Social Sciences Communications*, 9(1), 339. <https://doi.org/10.1057/s41599-022-01337-8>
- Bergmanis-Korāts, G., & Haiduchyk, T. (2024). *Social media manipulation for sale. 2024 experiment on platform capabilities to detect and counter inauthentic social media engagement*. Nato Strategic Communications Centre of Excellence. <https://stratcomcoe.org/pdfjs/?file=/publications/download/Social-Media-Manipulation-2024-DIGITAL-1eb50.pdf>
- Bourdieu, P. (1991). *Language and symbolic power*. Cambridge, UK: Polity Press.
- Cabrera, L., & Rosa-González, F. (2022). Creyentes religiosos en España por clases sociales: Resultados de 268.261 personas: 2013 a 2022 [Religious believers in Spain by social class: Results from 268,261 people, 2013 to 2022.]. *International and Multidisciplinary Journal of Social Sciences*, 11(2), 87–117. <https://doi.org/10.17583/rimcis.10708>
- Carratalá, A., & Valera-Ordaz, L. (2020). Polarized pluralism facing new politics and online columnism. *Studies on Journalistic Messages*, 26(26), 71–83. <https://doi.org/10.5209/ESMP.67287>
- Carvalho, P., Caled, D., Silva, C. G., Batista, F., & Ribeiro, R. M. (2023). The expression of hate speech against Afro-descendant, Roma, and LGBTQ+ communities in YouTube comments. *Journal of Language Aggression and Conflict*, 2(2), 171–206. <https://doi.org/10.1075/jlac.00085.car>
- Center for Sociological Research. (2023). *Study on social media audiences*. <https://www.cis.es/documents/d/guest/es3421mar-pdf>
- Centro de Investigaciones Sociológicas (2023). *Estudio sobre audiencias de medios de comunicación social (I)* (Estudio nº 3421) [Study on social media audiences (I) (Study No. 3421)]. <https://www.cis.es/es/estudios/estudio-sobre-audiencias-de-medios-de-comunicacion-social-i->
- Contreras Mazarío, J. M. (2016). Islamophobia and religious hatred in Spain. *Secularism and freedoms: Legal Writings*, 16, 75–110. https://www.boe.es/biblioteca_juridica/anuarios_derecho/abrir_pdf.php?id=ANU-J-2016-10007500110

- Corral, A., Fernández, C., & Prieto-Andrés, A. (2023). Islamophobia and far-right parties in Spain: The 'Vox' discourse on Twitter. In A. Mekki-Berrada & L. d'Haenens (Eds.), *Islamophobia as a form of radicalisation: Perspectives on media, academia and socio-political scapes from Europe and Canada* (pp. 161–176). Leuven University Press. <https://doi.org/10.2307/j.ctv34h08d2.14>
- Couldry, N., & Hepp, A. (2018). *The mediated construction of reality*. Hoboken, NJ: John Wiley & Sons.
- Cuéllar Rivero, R. (2024). The role of media in contexts of affective polarization: A systematic literature review. *Revista Española De Ciencia Política*, 64, 179–201. <https://doi.org/10.21308/recp.64.07>
- Davani, A., Atari, M., Kennedy, B. F., & Dehghani, M. (2023). Hate speech classifiers learn normative social stereotypes. *Transactions of the Association for Computational Linguistics*, 11, 300–319. https://doi.org/10.1162/tacl_a_00550
- El Confidencial [@elconfidencial]. (n.d.). *Posts* [X profile]. X. Retrieved May 15, 2026, from <https://x.com/elconfidencial>
- El País [@el_pais]. (n.d.). *Posts* [X profile]. X. Retrieved May 15, 2026, from https://x.com/el_pais
- Frischlich, L., Boberg, S., & Quandt, T. (2019). Comment sections as targets of dark participation? Journalists' evaluation and moderation of deviant user comments. *Journalism Studies*, 20(14), 2014–2033. <https://doi.org/10.1080/1461670x.2018.1556320>
- Gillespie, T. (2018). *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media*. New Haven, CT: Yale University Press.
- Gramigna, R. (2022). Inside Facebook's semiosphere. How social media influence digital hate and fuel cyber-polarization. *Social Semiotics*, 32(5), 606–633. <https://doi.org/10.1080/10350330.2022.2157171>
- Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6), 1360–1380. <https://doi.org/10.1086/225469>
- Hall, S. (1997). The spectacle of the other. In S. Hall, S. Nixon, & J. Evans (Eds.), *The spectacle of the "other"* (3rd ed., pp. 157–234). London, UK: SAGE Publications. <https://doi.org/10.4135/9781036232825.n4>
- Halliday, F. (1999). "Islamophobia" reconsidered. *Ethnic and Racial Studies*, 22(5), 892–902. <https://doi.org/10.1080/014198799329305>
- Hartikainen, E. I. (2025). Religious intolerance and the semiotics of secular indeterminacy in Brazil. *Signs and Society*, 13(1), 78–101. <https://doi.org/10.1017/sas.2024.4>
- <https://doi.org/10.65476/4vd67z56>

- Herranz-de-Rafael, G., & Fernández Prados, J. S. (2024). Religious beliefs and socialization: An empirical study on the transformation of religiosity in Spain from 1998 to 2018. *Religions*, 15(7), 848. <https://doi.org/10.3390/rel15070848>
- Humanes, M. L., & Valera-Ordaz, L. (2022). Partisanship, ideology, and selective exposure: A longitudinal analysis of media consumption in Spain (2008–2019). *Media and Communication*, 11(2), 113–126. <https://doi.org/10.17645/mac.v11i2.6280>
- International Holocaust Remembrance Alliance. (2016). *Working definition of antisemitism*. <https://www.holocaustremembrance.com/resources/working-definitions-charters/working-definition-antisemitism>
- Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, not ideology: A social identity perspective on polarization. *Public Opinion Quarterly*, 76(3), 405–431. <https://doi.org/10.1093/poq/nfs038>
- Kalmoe, N. P. (2014). *With passion and principle: Trade-offs and strategy in political argument*. Chicago, IL: University Chicago Press.
- Kennedy, B. F., Atari, M., Mostafazadeh Davani, A., Yeh, L., Omrani, A., Kim, Y., Coombs, K., Havaladar, S., Portillo-Wightman, G., Gonzalez, E., Hoover, J., Azatian, A., Hussain, A., Lara, A., Cardenas, G., Omary, A., Park, C., Wang, X., Wijaya, C., . . . Dehghani, M. (2022). Introducing the gab hate corpus: Defining and applying hate-based rhetoric to social media posts at scale. *Language Resources and Evaluation*, 56(1), 79–108. <https://doi.org/10.1007/s10579-021-09569-x>
- Kingdon, J. W. (1984). *Agendas, alternatives, and public policies*. Boston, MA: Little, Brown.
- Kiper, J. (2023). Religious hate propaganda: Dangerous accusations and the meaning of religious persecution in light of the cognitive science of religion. *Religions*, 14(2), 185. <https://doi.org/10.3390/rel14020185>
- Labiano Juangarcía, R., Azurmendi Adarraga, A., & Novoa Jaso, M. F. (2023). The constructive role of the media in hate speech controversies: The Valtònyc and Hasél cases. *Journalism Practice*, 18(2), 452–489. <https://doi.org/10.1080/17512786.2023.2242818>
- Ledesma, M. P. (2001). Studies on anticlericalism in Contemporary Spain. *International Review of Social*, 46(2), 227–255. <http://doi.org/10.1017/S0020859001000128>
- Lilleker, D. G., & Pérez-Escobar, M. (2023). Demonising migrants in contexts of extremism: Analysis of hate speech in UK and Spain. *Politics and Governance*, 11(2), 127–137. <https://doi.org/10.17645/pag.v11i2.6302>

Madhyastha, P., Founta, A.-M., & Specia, L. (2023). A study towards contextual understanding of toxicity in online conversations. *Natural Language Engineering*, 29(6), 1538–1560. <https://doi.org/10.1017/s1351324923000414>

Ministerio del Interior. (2025). *Informe sobre la evolución de los delitos e incidentes de odio en España 2024 [Report on the evolution of hate crimes and hate incidents in Spain 2024]*. <https://acortar.link/qSyHlq>

Minooie, M., Taylor, J. B., & Vargo, C. J. (2023). Agendamelding and COVID-19: The dance of horizontal and vertical media in a pandemic. *Frontiers in Political Science*, 5, 1021855. <https://doi.org/10.3389/fpos.2023.1021855>

Montero-Díaz, J., Said-Hung, E., & Moreno-Delgado, A. (2025). Assessment of hate in Spanish media: Typologies, intensities, and digital scenarios. *Mediterranean Journal of Communication*, 16(2), e28666. <https://doi.org/10.14198/MEDCOM.28666>

Muddiman, A., & Stroud, N. J. (2017). News values, cognitive biases, and partisan incivility in comment sections. *Journal of Communication*, 67(4), 586–609. <https://doi.org/10.1111/jcom.12312>

Observatorio del Pluralismo Religioso en España. (2025). *Informe de resultados BREC 2025 [BREC 2025 — Report of results.]*. <https://acortar.link/5ttwv3>

OKDiario [@okdiario]. (n.d.). *Posts [X profile]*. X. Retrieved May 15, 2026, from <https://x.com/okdiario>

Pérez, J. M., Luque, F., Zayat, D., Kondratzky, M., Moro, A., Serrati, P. S., Zajac, J., Vilaplana de Miguel, P., Debandi, N., Gravano, A., & Cotik, V. (2022). Assessing the impact of contextual information in hate speech detection. *IEEE Access*, 11, 30575–30590. <https://doi.org/10.1109/ACCESS.2023.3258973>

Pérez Escolar, M., Morejón Llamas, N., & Alcaide-Pulido, P. (2025). Populist rhetoric and hate speech: Analyzing xenophobic narratives in Vox's 2023 election campaign. *Politics and Governance*, 13, 9346. <https://doi.org/10.17645/pag.9346>

Pons Raga, M. (2024). Dissident blood: Neo-Santeria in Barcelona and the refusal of sacrifices. *Religions*, 15(11), 1373. <https://doi.org/10.3390/rel15111373>

Popa-Wyatt, M. (2023). Online hate: Is hate an infectious disease? Is social media a promoter? *Journal of Applied Philosophy*, 40(5), 788–812. <https://doi.org/10.1111/japp.12679>

Público [@publico_es]. (n.d.). *Posts [X profile]*. X. Retrieved May 15, 2026, from https://x.com/publico_es

<https://doi.org/10.65476/4vd67z56>

- Rajan, B., & Venkatraman, S. (2021). Insta-hate: An exploration of Islamophobia and right-wing nationalism on Instagram amidst the COVID-19 pandemic in India. *Journal of Arab & Muslim Media Research*, 14(1), 71–91. https://doi.org/10.1386/JAMMR_00025_1
- Reuters Institute. (2023). *Digital News report 2023*. <https://reutersinstitute.politics.ox.ac.uk/es/digital-news-report/2023/espana>
- Said-Hung, E., Montero-Díaz, J., Ruíz-Iniesta, A., Blanco-Valencia, X., Pérez, D., De Gregorio-Vicente, O., & Cubillas-Mercado, J. (2024). *El odio en los medios informativos digitales en España, 2023* [Hate speech in Spain's digital news media, 2023]. Hatemedia Project. <https://doi.org/10.5281/zenodo.13893173>
- Sewell, W. H. (1996). Three temporalities: Toward an eventful sociology. In T. J. McDonald (Ed.), *The historic turn in the human sciences* (pp. 245–280). Ann Arbor, MI: The University of Michigan Press.
- Sue, D. W., Capodilupo, C. M., Torino, G. C., Bucceri, J. M., Holder, A. M., Nadal, K. L., & Esquilin, M. (2007). Racial microaggressions in everyday life: Implications for clinical practice. *American Psychologist*, 62(4), 271–286. <https://doi.org/10.1037/0003-066x.62.4.271>
- Tarvin, E., & Stanfill, M. (2022). "YouTube's predator problem": Platform moderation as governance-washing, and user resistance. *Convergence*, 28(3), 822–837. <https://doi.org/10.1177/13548565211066490>
- Torres Gutiérrez, A. (2025). Asymmetries, lights, and shadows of the legal situation of religious minorities in Spain. *Religions*, 16(9), 1144. <https://doi.org/10.3390/rel16091144>
- Uyheng, J., & Carley, K. M. (2021). Characterizing network dynamics of online hate communities around the COVID-19 pandemic. *Applied Network Science*, 6(1), 1–21. <https://doi.org/10.1007/S41109-021-00362-X>
- Uyheng, J., Bellutta, D., & Carley, K. M. (2022). Bots amplify and redirect hate speech in online discourse about racism during the COVID-19 pandemic. *Social Media + Society*, 8(3), 1–14. <https://doi.org/10.1177/20563051221104749>
- Walther, J. B. (2022). Social media and online hate. *Current Opinion in Psychology*, 45, 101298. <https://doi.org/10.1016/j.copsyc.2021.12.010>
- Walther, J. B. (2024). The effects of social approval signals on the production of online hate: A theoretical explication. *Communication Research*. Advance online publication. <https://doi.org/10.1177/00936502241278944>

Weinberg, D. B., Levy, M. D., Edwards, A., Kopstein, J., Frey, D., Antonaros, P., Baci, N., Cohen, N. D., Fernandez, J., & Ni, Y. (2025). Hidden in plain sight: Antisemitic content in QAnon subreddits. *PLoS One*, 20(3), e0318988. <https://doi.org/10.1371/journal.pone.0318988>

Wells, T. (2023). Bad adopters or bad proponents of technology? Facebook and the violence against Muslims in Myanmar. *Third World Quarterly*, 45(8), 1309–1324. <https://doi.org/10.1080/01436597.2023.2285808>

Wiedlitzka, S., Prati, G., Brown, R., Smith, J., & Walters, M. (2021). Hate in word and deed: The temporal association between online and offline Islamophobia. *Journal of Quantitative Criminology*, 39(2023), 75–96. <https://doi.org/10.1007/S10940-021-09530-9>