societies

MDPI

Article

Assessing Climate Hazard Resilience Through AI-Based Analysis of Online Data: Empirical Evidence from Galicia

Dmitry Erokhin * and Nadejda Komendantova

International Institute for Applied Systems Analysis, 2361 Laxenburg, Austria; komendan@iiasa.ac.at
* Correspondence: erokhin@iiasa.ac.at

Abstract

Climate hazards increasingly unfold as information crises alongside physical impacts, producing rapid shifts in what people search for and discuss online. This case study demonstrates how AI-supported analysis of online data can complement conventional disaster intelligence by providing a scalable social sensing layer for climate hazard resilience in Galicia. It integrates Google Trends as a proxy for changing public attention and information demand, and YouTube videos and comment threads to capture public sensemaking and resilience-relevant signals. Monthly Google Trends series were used for eight hazards, with floods showing the highest mean interest, followed by wildfires and heatwaves. For the three highest-salience hazards, the study analyzed YouTube comments using gpt-5-mini to extract sentiment, emotions, topics, institutional trust cues, collective efficacy cues, calls to action, impacts, vulnerable groups, and coping actions. The corpus included 184 heatwave comments, 20,427 wildfire comments, and 4882 flood comments. Across hazards, discourse is predominantly negative but differs in structure. Heatwave threads skew toward mockery and normalization, wildfire threads center on anger, governance and low institutional trust, and flood threads combine solidarity with demands for localized warnings and guidance. The study translates comment-level signals into traceable policy recommendations emphasizing actionable risk communication, early warning and response capacity, and trust-building practices. The study concludes with an operational pipeline concept for continuous monitoring and dashboard-based decision support, while emphasizing limitations related to Google Trends sampling and normalization, platform and API biases, and model-mediated uncertainty.

Keywords: climate resilience; crisis informatics; social sensing; Google Trends; YouTube analysis; large language models; sentiment analysis; risk communication; early warning systems; Galicia; Spain

Check for updates

Academic Editor: Chun Kai Leung

Received: 14 February 2026
Revised: 21 May 2026
Accepted: 11 June 2026
Published: 12 June 2026

1. Introduction

Climate hazards increasingly unfold as datafied events. Alongside physical impacts, they generate rapid shifts in what people search for, watch, share, and debate online. These digital traces matter for resilience because they reflect how communities perceive evolving threats, seek protective guidance, interpret official warnings, and coordinate informal support. Crisis informatics research has long argued that emergencies are also information crises, in which official response systems and public communication become tightly coupled through networked media [1,2]. Social media and other online platforms can therefore serve as a complementary observational layer for disaster risk reduction, enabling faster situational awareness and more responsive risk governance, particularly when used as an input to decision-support workflows rather than as a substitute for operational monitoring [3,4].

A key promise of AI-supported online data analysis is the ability to convert high-volume, heterogeneous digital content into structured indicators that are interpretable for preparedness, response, and recovery. Reviews of digitally enabled disaster management describe a growing ecosystem of methods that extract public needs, emotional responses, misinformation dynamics, and evolving concerns from online data, with potential to strengthen decision-making when integrated with institutional capacities and validated against ground truth [5,6]. Sensemaking studies further show that online discourse during extreme events is not merely noise but a social process through which people interpret risk, negotiate uncertainty, and coordinate action, which are exactly the processes that resilience policies seek to support [7].

Within this digital layer, two sources are particularly relevant for measuring risk salience and for social listening in climate hazards. First, internet search data can provide a near-real-time proxy for information needs and perceived exposure. The infodemiology and infoveillance framework formalized the use of online information patterns such as search and communication behavior to infer population-level concerns and to inform public policy [8,9]. More recent work illustrates that search behavior can track hazard salience and self-protective responses. For example, research on wildfire smoke uses Google search query data to measure awareness and protective behavior in relation to exposure [10]. Related analyses show that Google Trends can capture temporal spikes during catastrophic events and recurring seasonal cycles, supporting the idea that search dynamics reflect collective attention and perceived threat [11,12]. In the context of climate hazards, this makes search data a plausible proxy for perceived exposure and concern, as well as a scalable complement to surveys, especially when interpreted alongside objective hazard indicators and media and event timelines.

Second, social and video platforms provide richer qualitative context on what people think is happening, what they blame or trust, what help they request, and what adaptation options they consider feasible. Systematic and narrative reviews in disaster risk management emphasize that social media analytics can reveal public sentiment, urgent needs, and reactions to risk reduction measures. These insights can inform more targeted communication, resource allocation, and recovery planning [13,14]. YouTube is particularly important in this landscape because it functions simultaneously as a search engine, a news and explanatory media environment, and a discussion space, often where risk communication content and public reactions coexist.

Galicia (northwest Spain) provides a salient setting for demonstrating such an approach. The region faces compound climate-related risks across coastal, hydrological, and fire regimes. Research documents coastal flood vulnerability along the Galician coast and rías (deep, finger-like coastal inlets and estuaries), including risk to beaches and urbanized shorelines under sea-level rise and storm conditions [15,16]. Research also highlights wildfire risk and its interaction with social vulnerability at the municipal scale, underscoring that impacts are uneven and mediated by demographic and socioeconomic conditions [17,18]. Drought dynamics and future drought occurrence in Galicia have likewise been analyzed under climate change, with implications for water resources and health [19,20]. These hazards shape not only physical exposure but also risk perception, trust in institutions, and the demand for timely guidance, which are the dimensions that online search and discourse signals can help observe.

Despite the growing body of work, three gaps remain insufficiently addressed. First, crisis informatics and disaster social media studies have extensively examined online communication during emergencies, but fewer studies integrate search interest data with video platform comment discourse in a single empirical framework that links public attention, sensemaking, and resilience-relevant policy needs. Second, existing applications of Google Trends and social media analytics often focus on single hazards or descriptive public interest patterns, while less attention has been paid to comparing how different climate hazards generate distinct information problems, trust dynamics, emotional responses, and actionability gaps within the same regional context. Third, research on climate resilience in Galicia has documented important physical and social vulnerabilities related to floods, wildfires, drought, and coastal risks [15–20], but there remains limited evidence on how residents and online publics perceive, discuss, contest, and translate these risks into expectations for institutions and protective action.

This study addresses these gaps by combining Google Trends and YouTube discourse analysis for multiple hazards in Galicia, using AI-assisted annotation to transform unstructured online data into interpretable indicators of hazard salience, public concern, trust, collective efficacy, vulnerability awareness, coping behavior, and policy demand. Accordingly, the study is guided by the following overarching research question: how can AI-supported analysis of online search behavior and platform discourse be used to assess the social-information dimensions of climate hazard resilience in Galicia and translate them into actionable policy insights?

Three subsidiary research questions structure the empirical analysis. RQ1 asks which climate hazards generate the most salient online search interest signals in Galicia during 2022–2025 and what information needs are reflected in related Google Trends queries. RQ2 asks how YouTube discussions of highest-salience hazards differ in sentiment, emotions, topics, institutional trust, collective efficacy, vulnerability awareness, coping actions, and calls to action. RQ3 asks how these online signals can be synthesized into traceable policy recommendations and an AI-enabled monitoring pipeline for climate hazard resilience.

2. Methodology

This study combines Google Trends and YouTube data to examine public attention, interpretation, and resilience-related discourse around climate hazards in Galicia, Spain. Google Trends was used as a quantitative proxy for hazard salience and information demand, while YouTube videos and comments provided qualitative evidence of public concerns, expectations, emotions, and perceived resilience needs during climate-related events [8,14,21] (Figure 1).

Google Trends data were extracted on 13 January 2026, using geographic filtering for Galicia, Spain. The analysis covered January 2022 to December 2025, the four most recent complete calendar years available at the time of collection. This period avoids partial-year bias, captures multiple seasonal cycles, and includes high-salience events likely to shape public information demand, including the exceptional European heat summer of 2022, major wildfire episodes in Galicia in mid-2022, and nationally salient flood disasters in late 2024.

Google Trends was queried using topics rather than raw search terms, because topics aggregate semantically related searches across spelling, phrasing, and language variants. This is especially useful in multilingual and morphologically diverse contexts and reduces sensitivity to individual query wording [22]. The hazard set was defined ex ante by triangulating climate hazard categories from the scientific and applied risk literature with hazards reflected in Spanish and regional reporting and then restricting the list to hazards with available Google Trends Topics.

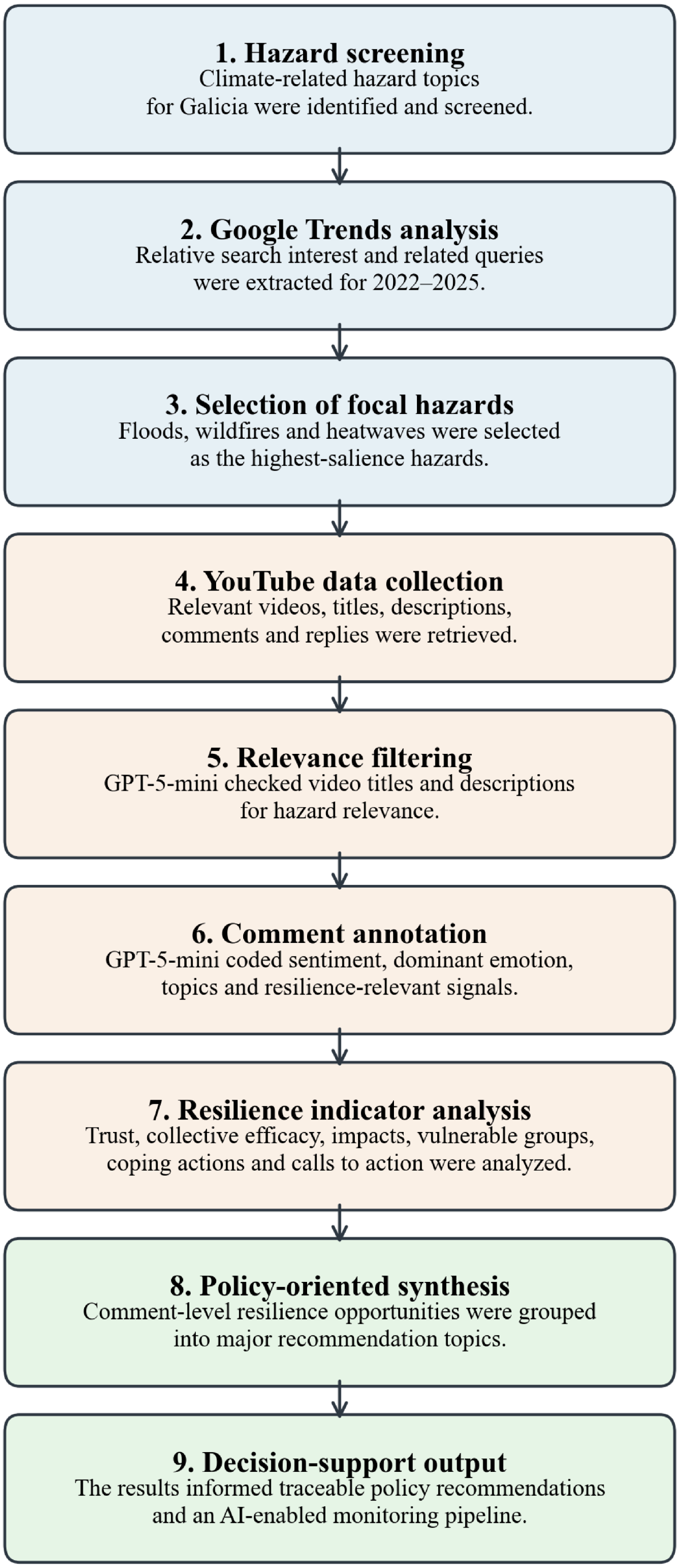

Figure 1. Study methodology.

For each hazard topic, relative search interest and related query metadata were extracted. Relative search interest was interpreted as a signal of public attention and information demand, while related queries provided contextual evidence of co-occurring concerns such as warning seeking, health guidance, protective actions, or recovery needs [22–24]. The three hazards with the highest average search interest in Galicia from 2022 to 2025 were selected for downstream YouTube discourse analysis.

YouTube data were collected on 14 January 2026, using the YouTube Data API v3 [25]. Searches were conducted in Spanish for the three focal hazards identified through Google Trends: "inundación Galicia", "incendios forestales Galicia", and "ola de calor Galicia". The API search endpoint was queried with type = video, order = relevance, maxResults = 50, relevanceLanguage = es, and regionCode = ES. Pagination continued until no further nextPageToken was returned, and duplicate video IDs were removed while preserving the retrieval order.

For each video, metadata were collected using the videos.list endpoint, including video ID, URL, title, description, publication date, channel information, duration, view count, like count, and comment count. Comment threads were retrieved using commentThreads.list with maxResults = 100, textFormat = plainText, and order = relevance. Replies were collected separately through comments.list using the parent comment ID. Videos with disabled or inaccessible comments were excluded from the comment corpus. The final dataset contained 338 videos including 5 heatwave videos, 303 wildfire videos, and 30 flood videos. These yielded 184 heatwave comments, 20,427 wildfire comments, and 4882 flood comments. Unlike the Google Trends data, YouTube retrieval was not restricted to the 2022–2025 period. Therefore, the YouTube dataset is interpreted as a relevance-ranked discourse corpus available at the time of collection.

Large language model annotation was used in two stages. First, gpt-5-mini assessed the relevance of each video title and description to the corresponding hazard topic. The relevance prompt instructed the model to classify each video as hazard-related or not, using strict inclusion and exclusion criteria and relevant Spanish and Galician terminology. Uncertain cases were treated conservatively.

Video relevance filtering prompt, flooding example:

> "You are a strict classifier for whether a YouTube video's title and description are related to the environmental topic of floods/flooding/inundation. Return TRUE if it is about flooding events or impacts, including river overflow/crecida, coastal flooding, flash floods, stormwater or urban flooding, heavy rainfall causing inundation, river level warnings, dam or spillway releases, road closures, evacuations or rescues, emergency response, or official updates. Include Spanish and Galician terms such as: 'inundación', 'inundacións', 'inundacións en Galicia', 'riada', 'crecida', 'desbordamiento', 'desbordamento', 'desbordou', 'chea', 'enchente', 'anegación', 'anegadas', 'alerta', 'aviso', 'Protección Civil', '112 Galicia', 'AEMET', and 'MeteoGalicia'. Return FALSE if the video is unrelated, metaphorical, or about water without flooding, including phrases such as 'a flood of…', or videos about music, sports, politics, gaming, routine rain with no inundation, fire, earthquake, or other hazards without flood context. If uncertain, prefer FALSE with low confidence. The rationale must be no more than 12 words. Return one result per provided key".

Second, YouTube comments were annotated using a fixed JSON schema. For each comment, the model returned structured fields for sentiment, dominant emotion, main topic, resilience signals, and a short resilience opportunity. The resilience-signal fields captured risk perception, impacts mentioned, vulnerable groups, coping actions, adaptation or policy suggestions, information needs, institutional trust, collective efficacy, and calls to action. Fixed JSON output was used to reduce free-form interpretation and improve consistency across hazards.

Comment annotation prompt, common structure:

> "You analyze YouTube comments about [hazard] in Galicia, Spain. The output must be a single valid JSON object and nothing else. Do not include markdown, code fences, explanations, or extra text. Return exactly the following fields: sentiment, as a number from −1.0 to +1.0; emotions, as one dominant emotion label chosen from: amusement, mockery, concern, frustration, anger, fear, sadness, hope, relief, disbelief, neutral, or other; main_topic, as a short 2–6 word label; resilience_signals, including risk_perception, impacts_mentioned, vulnerable_groups_mentioned, coping_actions_mentioned, adaptation_or_policy_suggestions, information_needs, trust_in_institutions, collective_efficacy, and call_to_action; and resilience_opportunity, as one actionable insight of maximum 18 words. For risk_perception, use only low, medium, or high. For impacts_mentioned, choose up to three unique items from: health, injury_drowning, evacuation, property_damage, agriculture, water_contamination, ecosystems, infrastructure, economy, other, or none. For vulnerable_groups_mentioned, choose up to three from: elderly, children, respiratory_patients, outdoor_workers, low_income, tourists, other, or none. For coping_actions_mentioned, choose up to three from: move_to_higher_ground, avoid_travel, evacuation, seek_updates, help_neighbors, donations, sandbags, other, or none. For adaptation_or_policy_suggestions, choose up to three from: drainage_infrastructure, river_management, early_warning, evacuation_plans, emergency_services, land_use_policy, public_safety_messaging, other, or none. For information_needs, choose up to three from: flood_location, rain_forecast, river_levels, road_closures, evacuation_routes, shelter_info, safety_guidance, aid_resources, other, or none. All list fields must always be lists. If no category applies, use ['none'], and if 'none' appears, it must be the only item. For trust_in_institutions and collective_efficacy, use low, medium, high, or unclear; use unclear when the signal cannot be inferred. For call_to_action, use yes only when the comment explicitly urges action; otherwise use no. The resilience_opportunity must contain one actionable insight, with no quotes or newline".

The model was also instructed to propose one action to strengthen resilience for each comment. These comment-level actions were synthesized into policy-recommendation categories through an iterative LLM-assisted coding process. First, the generated actions were reviewed to identify recurring higher-level themes. Second, each action was assigned to one of these themes. The resulting categories were interpreted as policy-relevant clusters while retaining links to the original action statements and source comments, preserving traceability and allowing the evidence behind each recommendation to be inspected.

Climate hazard resilience was operationalized through online indicators capturing the social-information dimensions of resilience. The Google Trends indicator measured relative search interest and related queries as evidence of hazard salience and information demand. YouTube indicators included sentiment, dominant emotion, topic category, institutional trust, collective efficacy, calls to action, impacts mentioned, vulnerable groups mentioned, coping actions, and adaptation or policy suggestions. These indicators provide a reproducible basis for assessing which hazards attract public attention, how risks are interpreted, whether institutions and collective capacity are trusted, which harms and vulnerable groups are recognized, and whether public discourse contains actionable coping or policy signals.

This mixed digital trace design aligns with infodemiology, crisis informatics, risk communication, and disaster-risk-reduction research, which treat search behavior and social media discourse as measurable signals of public attention, information needs, affective response, trust, vulnerability awareness, and actionability [1–9,13,14,22,26–29]. YouTube is treated both as an information source and a discussion venue, consistent with emergency management research recognizing audiovisual platforms as part of the crisis communication ecosystem [14,21,30]. The use of LLMs as text annotators is supported by recent evidence showing that modern models can perform common annotation tasks such as relevance, stance, topic, and frame detection with high accuracy and agreement [31–33].

3. Results

The findings are structured into two main parts: first, an examination of Google Trends signals as a quantitative proxy for information-seeking behavior and perceived risk salience across multiple hazards; second, a qualitative analysis of YouTube discourse to explore public sentiment, concerns, and resilience-related themes for the three highest-salience hazards.

3.1. Google Trends Signals

A comprehensive list of major climate-related hazards was first compiled to determine which hazards are represented as searchable topics in Google Trends. The identified topics included the following:

- “heat wave”;
- “cold wave”;
- “flood”;
- “drought”;
- “water stress”;
- “sea level rise”;
- “ocean acidification”;
- “glacial lake outburst flood”;
- “typhoon”;
- “blizzard”;
- “dust storm”;
- “tornado”;
- “wildfires”;
- “landslide”;
- “avalanche”;
- “permafrost thaw”;
- “coastal erosion”.

Monthly Google Trends time series were then extracted for Galicia for the period 2022–2025 (Figure 2), and the hazards with sufficient search volume and temporal variation were compared to identify the most interpretable public attention signals. Based on this screening, floods, wildfires, and heatwaves were selected as the top three hazards, since their search ecosystems are strongly grounded in locally relevant, action-oriented information needs. Flood-related queries concentrate on real-time monitoring and place-specific impacts (e.g., tracking inundation status and river/reservoir conditions), while also showing amplification during nationally salient flood disasters. Wildfire queries are even more immediate, dominated by “today/now” and map-based tracking, hyperlocal searches (e.g., “incendios Ourense”, “incendios Pontevedra”, “incendio Pontevedra”, “incendios activos Ourense”, “incendios activos Pontevedra”, “incendios activos en Pontevedra hoy”, “Boiro”, and “incendio Boiro”), and spillovers into governance and restriction-related interest during severe seasons. Heatwave queries combine authoritative monitoring with practical planning, with users repeatedly seeking timing/duration (“when will it end/how long will it last"), relying heavily on institutional warning information (e.g., AEMET products), and linking heat to cascading risks such as wildfire.

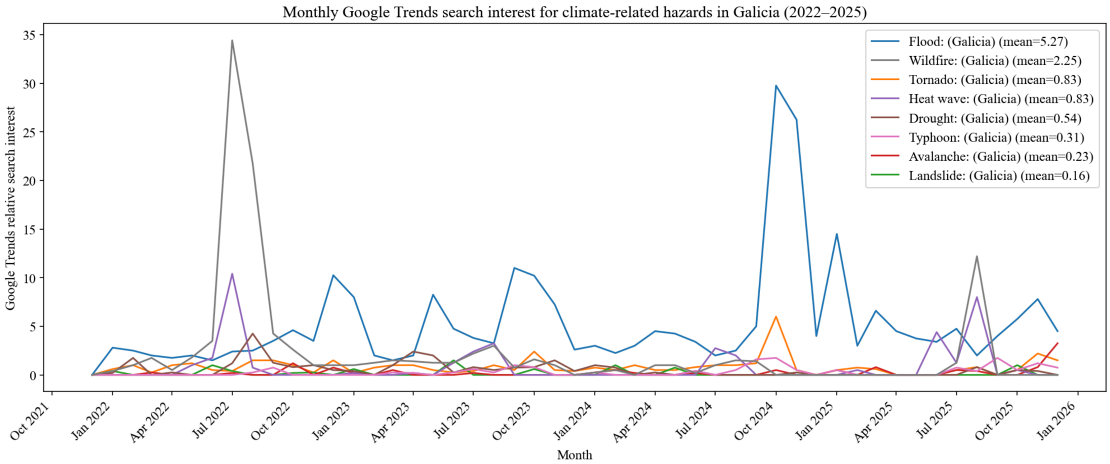

Figure 2. Monthly Google Trends search interest for climate-related hazards in Galicia (2022–2025).

3.2. YouTube Comment Analysis for High-Salience Hazards

To complement the Google Trends evidence with an analysis of public discourse, these three hazards with the clearest and most interpretable search interest signals in Galicia were also used to retrieve and analyze YouTube content and associated comment threads using query pairs:

- "inundación Galicia" (flood Galicia);
- "incendios forestales Galicia" (wildfire Galicia);
- "ola de calor Galicia" (heatwave Galicia).

3.2.1. Heatwaves in Galicia

The YouTube search yielded five videos meeting the study criteria and a total of 184 comments for analysis. The sentiment distribution for heatwave-related comments is skewed toward negativity, with a dense concentration of observations in moderately negative ranges and a smaller positive tail. The histogram (Figure 3) indicates that the modal region lies around mildly to moderately negative sentiments (approximately -0.5 to -0.3), while strongly negative expressions extend toward the lower bound of the scale. A secondary cluster appears near neutral sentiment, and positive sentiment is present but comparatively sparse, extending to moderately positive values. This pattern suggests that heatwave discourse in this sample is more often evaluative and critical than affirmational, but it is not uniformly negative. Rather, it combines a substantial negative affect with a meaningful, though smaller, share of neutral and positive reactions.

Emotion coding helps to explain the negative skew (Figure 4). Mockery is the dominant emotional tone, accounting for roughly 45% of comments, far exceeding any other category. Neutral affect represents the next largest share at approximately 10%, followed by frustration at a similar magnitude. Amusement, concern, and anger each contribute smaller but non-trivial proportions, while hope and relief are also present at modest levels. Fear and disbelief appear less frequently. The prominence of mockery indicates that, within these comment threads, heatwave discussion often serves as a site for sarcasm, derision, or ridicule directed either at other commenters, public figures, media narratives, or the framing of the hazard itself rather than functioning primarily as an exchange of practical risk information.

https://doi.org/10.3390/soc16060188

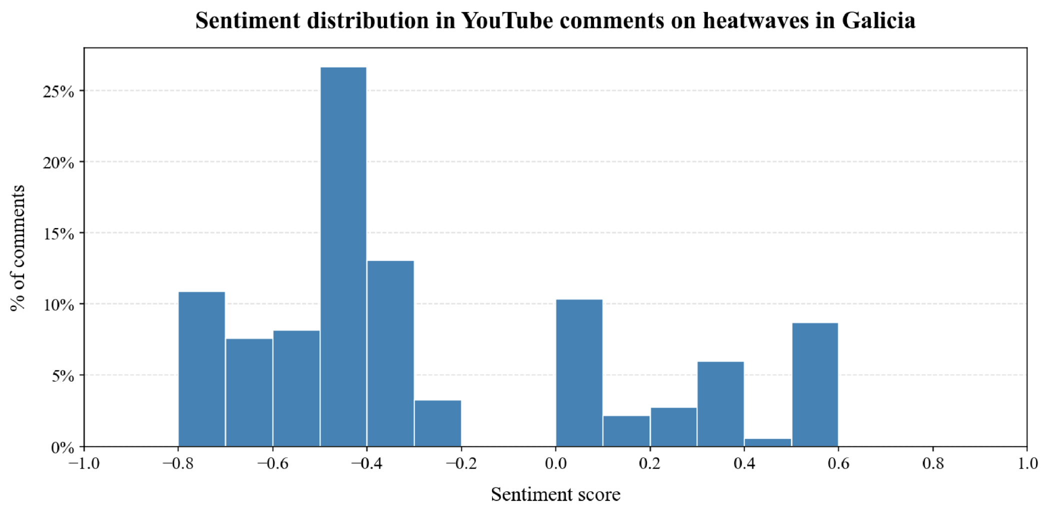

Figure 3. Sentiment distribution of YouTube comments on heatwaves in Galicia.

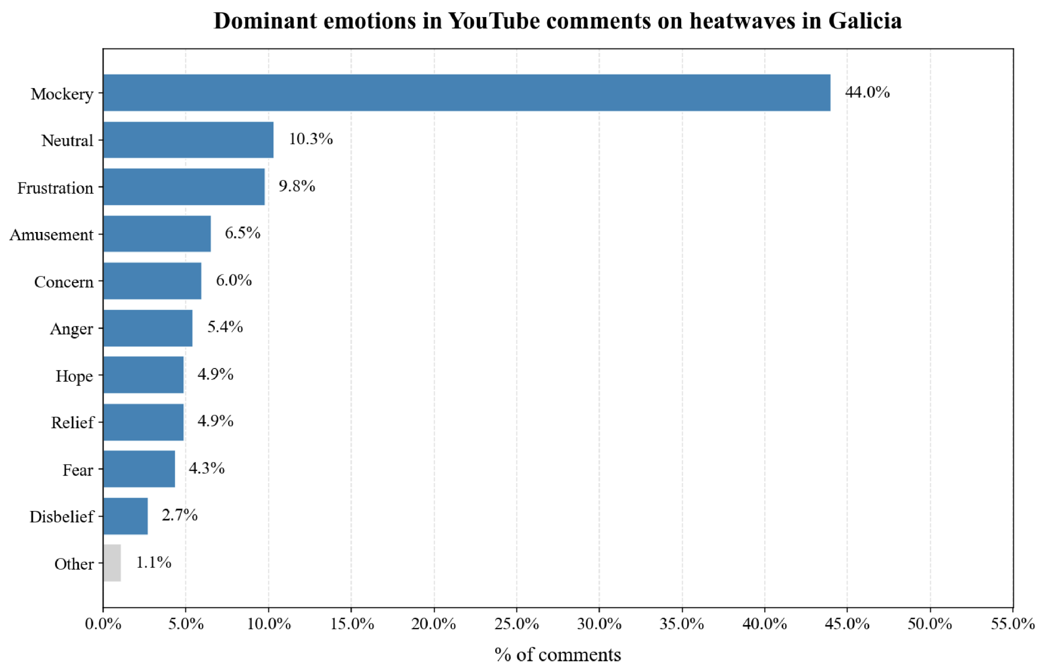

Figure 4. Dominant emotions in YouTube comments on heatwaves in Galicia.

Consistent with this affective profile, the main topic distribution shows that heatwave comments are structured around a blend of everyday normalization, acute risk talk, and contested interpretation (Figure 5). The largest topic category is seasonal and lifestyle framing of heat (approximately 12% of comments), reflecting discourse that treats heat as part of expected summer conditions, lifestyle preferences, or routine seasonal variation. At the same time, heatwave impacts and acute risk forms the second-largest category (around 10%), indicating that a sizable fraction of commenters foreground consequences and hazard salience rather than normalization. The category of interpersonal attacks and ridicule constitutes a similarly large share (close to 10%), mirroring the dominance of mockery in the emotion distribution and indicating that a substantial portion of thread activity is socially adversarial rather than informational. Media and weather reporting, including forecasts (roughly 9%), temperature records and comparisons (also around 9%), together demonstrate that many comments react to reported maxima, comparisons with other places or past summers, and the credibility or framing of forecasts. Travel/place references and regional identity (approximately 8%) further suggest that commenters frequently anchor heat in geographic identity, local pride, and regional comparisons. Climate skepticism, denial, and politicization comprise another notable block (around 7%), indicating that a heatwave serves as a trigger for broader ideological contestation about climate narratives. Adaptation and coping measures and low-content or unclear replies each account for about 6%, while smaller yet persistent categories capture conspiracy narratives, climate change awareness/attribution, policy–economy–social issues, heat trends/escalation, accusations of manipulation or fake news, and pollution/environmental impacts.

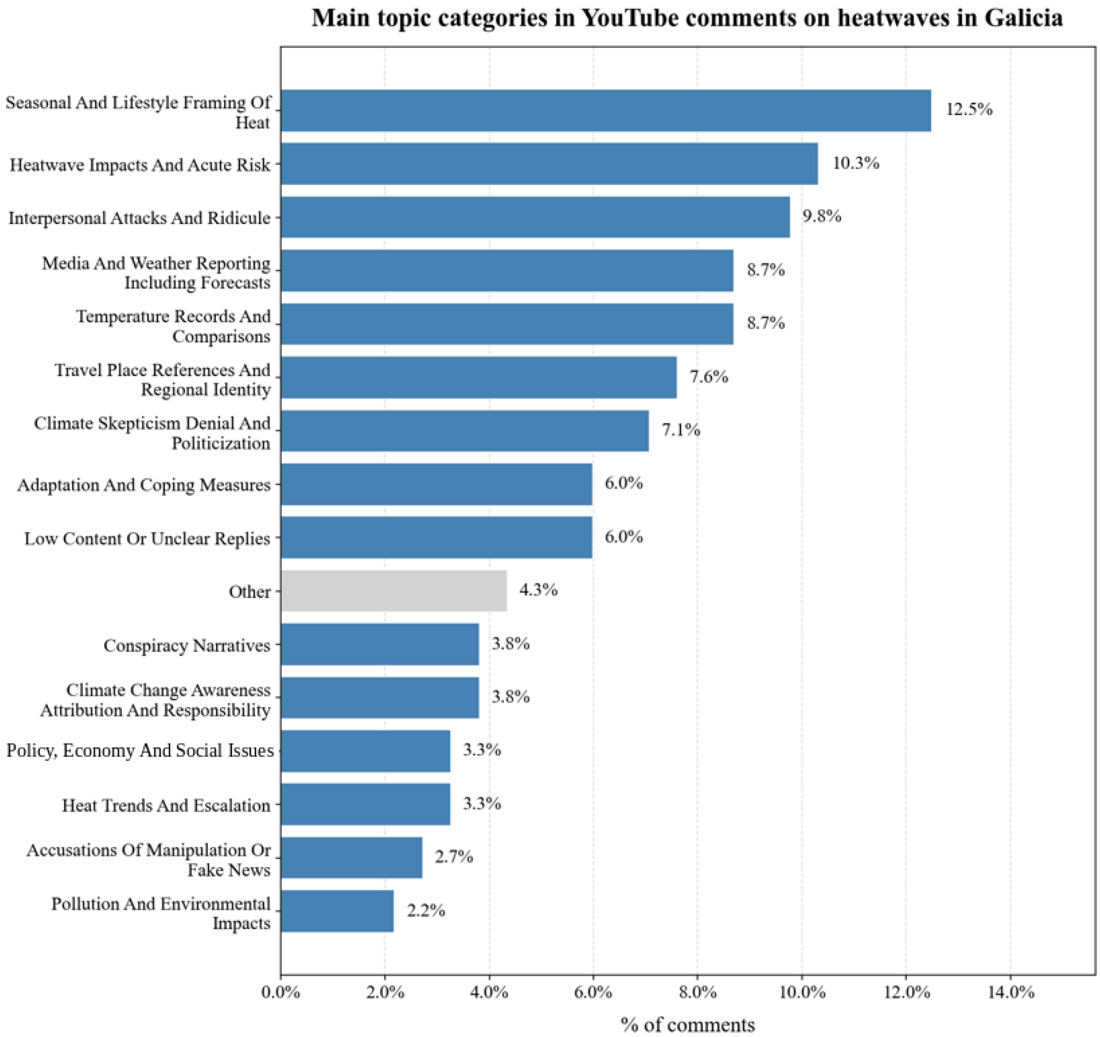

Figure 5. Main topic categories in YouTube comments on heatwaves in Galicia.

Resilience-related signals appear unevenly across the comment set (Figure A1). Explicit trust in institutions is absent from most comments. Nearly 70% contain no interpretable trust cue, while the subset that does address institutional credibility is heavily weighted toward low trust (approximately 28%), with only marginal representation of medium or high trust.

A similar pattern holds for collective efficacy (Figure A2). The overwhelming majority of comments do not express a sense of collective capacity to respond, and when collective efficacy is expressed, it is more often low than medium or high. These distributions suggest that heatwave discussions in this YouTube sample are rarely oriented toward coordinated response or institutional reliance. Instead, they more frequently revolve around commentary, critique, and social positioning.

Calls to action are present but limited, appearing in roughly one-fifth of comments, indicating that direct suggestions (for example, urging protective behavior or advocating for specific actions) are a minority discourse form relative to commentary and debate (Figure A3).

When impacts are mentioned, health is by far the most common category, referenced in approximately 15% of comments, underscoring that the most salient articulated consequence of heat is human well-being (Figure 6). Ecosystem impacts appear at substantially lower levels (around 5%), and work-related impacts are mentioned in roughly 3% of comments, suggesting some recognition of labor exposure and productivity constraints. References to wildfire and water impacts are present but comparatively infrequent (each on the order of a few percent), and explicit mentions of agriculture, energy, and infrastructure impacts are rare (around 1% each). Notably, the low absolute rates for most impact categories indicate that, even in a hazard-focused thread, detailed articulation of sectoral consequences is limited and concentrated on health.

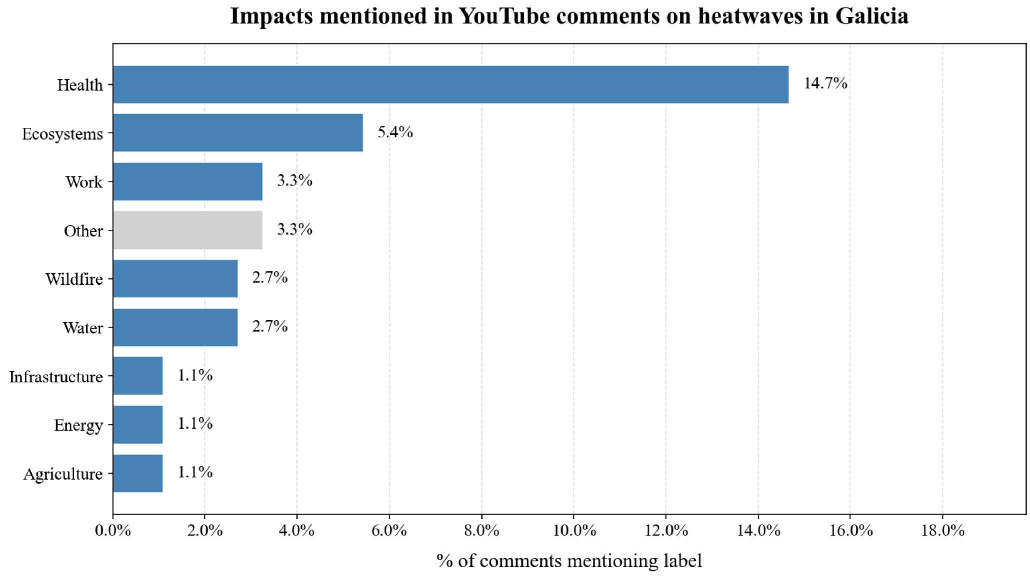

Figure 6. Impacts mentioned in YouTube comments on heatwaves in Galicia.

Mentions of vulnerable groups are also relatively scarce, appearing in small single-digit percentages (Figure 7). The elderly are the most frequently referenced vulnerable group (approximately 3%), followed by outdoor workers at a similar but slightly lower level, while references to children and tourists are uncommon. This pattern implies that vulnerability awareness exists but is not a central organizing feature of comment discourse, despite the well-established differential risks posed by extreme heat.

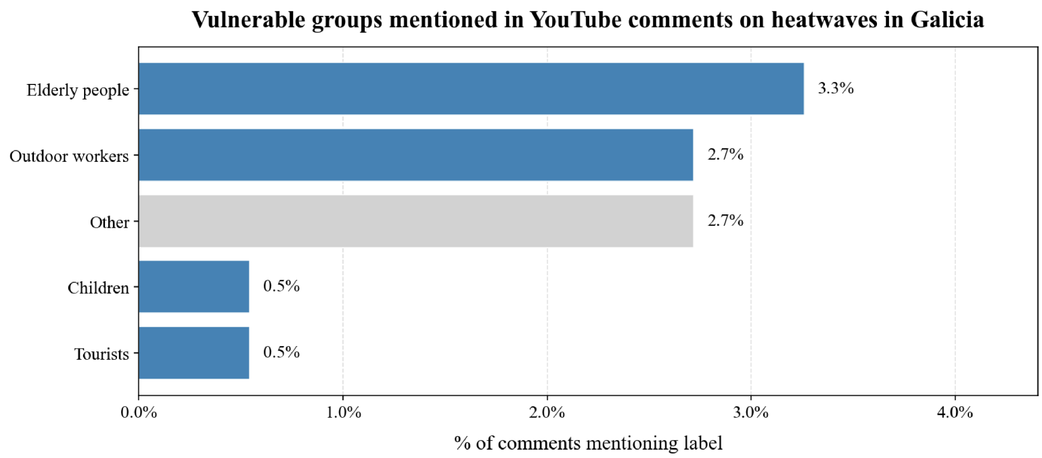

Figure 7. Vulnerable groups mentioned in YouTube comments on heatwaves in Galicia.

Coping actions show the same sparsity (Figure 8). Beyond a general other category that captures heterogeneous and less-standardized responses, explicit references to air conditioning and shade appear only in the low single digits, with mentions of hydration and schedule changes being rarer.

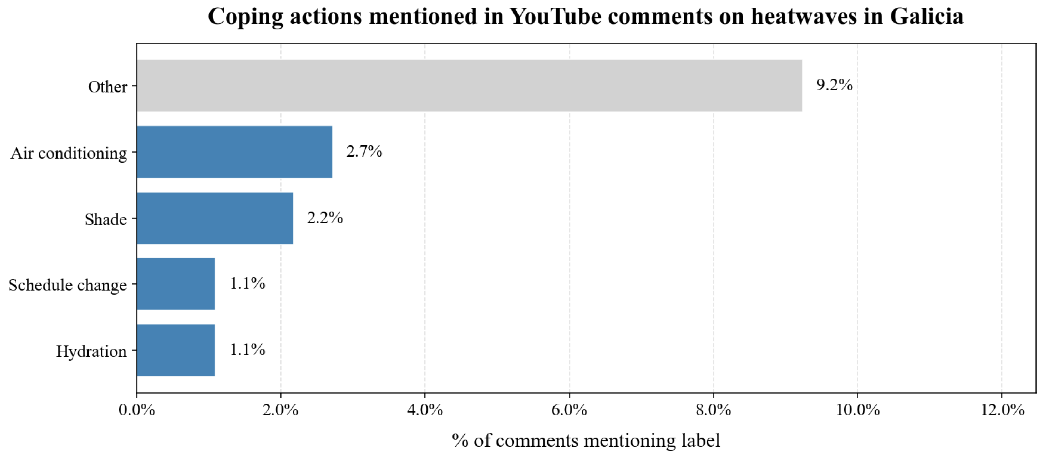

Figure 8. Coping actions mentioned in YouTube comments on heatwaves in Galicia.

Adaptation or policy suggestions are infrequent overall (Figure A4). When present, they appear primarily as heterogeneous proposals grouped under other, with a smaller identifiable share pointing to urban shade interventions. These resilience indicators suggest that the YouTube comment space surrounding heatwaves in Galicia is more strongly characterized by affective and interpretive contestation often expressed through mockery, media critique, and politicized climate talk than by sustained exchange of practical coping guidance or collective adaptation planning, even though health impacts emerge as a clear focal point when concrete consequences are discussed.

3.2.2. Wildfires in Galicia

The YouTube search yielded 303 videos meeting the study criteria and a total of 20,427 comments for analysis. The sentiment distribution for wildfire-related comments is strongly skewed toward negativity (Figure 9). The histogram shows a dense mass of observations in the moderately to strongly negative range, with the modal region concentrated around approximately −0.75 to −0.55. Negative sentiment extends toward the lower bound of the scale, indicating the presence of highly adverse evaluations rather than only mild dissatisfaction. A smaller accumulation appears close to neutral sentiment, while a comparatively thin positive tail extends into moderately positive values. The distribution indicates that wildfire discourse in this sample is dominated by critical and adverse reactions, while still containing a limited but non-negligible share of neutral and affirmative expressions.

Emotion coding helps clarify the intensity and direction of this affective profile (Figure 10). Anger is the dominant emotional tone, accounting for roughly 37% of comments, and substantially exceeding all other categories. Mockery and frustration form the next tier (each around 15–16%), followed by sadness at approximately 8–9%. Hope, concern, and neutral affect appear at lower but still visible levels (roughly 5% each), while disbelief and amusement are present in smaller proportions. Relief and fear occur only rarely. This distribution suggests that wildfire threads function less as low-stakes commentary and more as an arena for grievance, blame, and moral evaluation, with anger and frustration structuring much of the interactional tone.

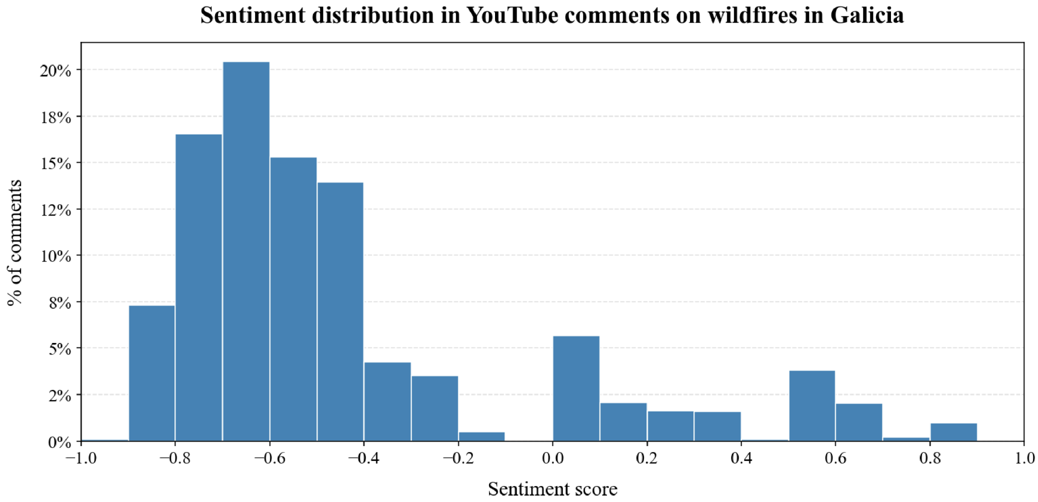

Figure 9. Sentiment distribution of YouTube comments on wildfires in Galicia.

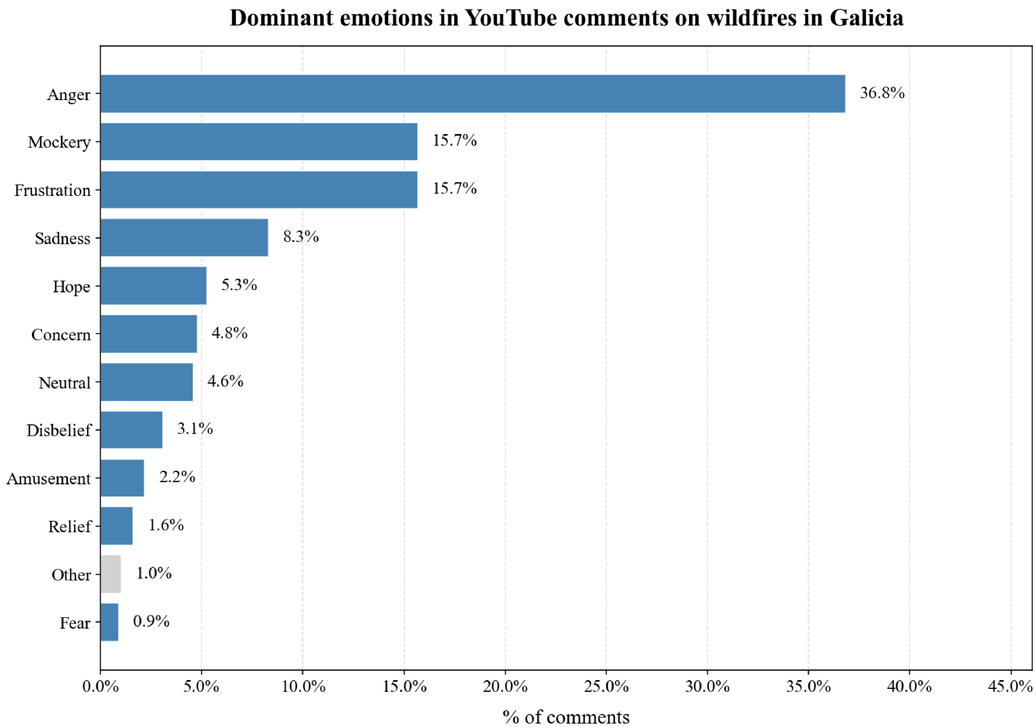

Figure 10. Dominant emotions in YouTube comments on wildfires in Galicia.

Consistent with these emotional patterns, the main topic distribution indicates that wildfire discourse is organized around accountability, attribution, and institutional performance (Figure 11). Politics and governance is the largest topic category, representing roughly 30–31% of comments, which implies that commenters frequently interpret wildfire events through lenses of public responsibility, policy, and institutional decision-making. The category of interpersonal attacks and ridicule comprises a substantial share (approximately 7–8%), indicating that a meaningful portion of thread activity remains socially adversarial rather than informational. Several hazard-specific and response-oriented topics then follow closely, including wildfire general/updates (around 7%), emergency response and firefighters (about 6–7%), and causes, arson, and punishment (about 6%). Smaller but persistent categories include solidarity and support (around 4–5%), smoke and health (roughly 3–4%), misinformation and conspiracy (about 3–4%), and impacts and losses (about 3%). Emotion and sympathy, prayers and rain hopes, and media and information distrust each contribute additional shares in the low single digits. Prevention and land management appears at around 1–2%, while evacuation and safety, humor and sarcasm, and climate and environment occur only marginally. These topical distributions depict a discourse that is strongly politicized and accountability-oriented, while also containing distinct strands focused on operational response, causal attribution, and tangible consequences.

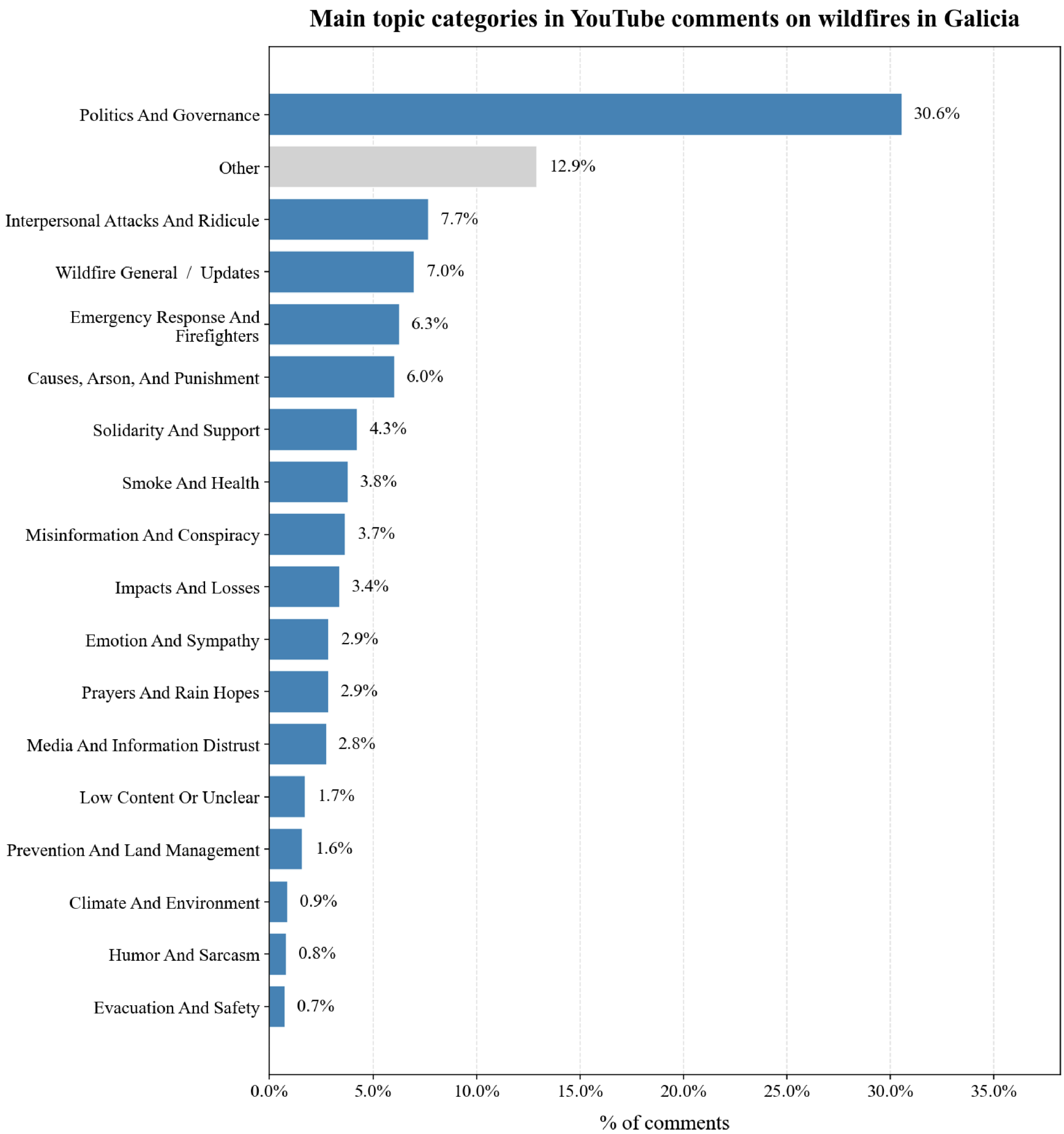

Figure 11. Main topic categories in YouTube comments on wildfires in Galicia.

Resilience-related signals further underscore the prominence of institutional evaluation (Figure A5). Trust in institutions is low in a majority of comments (approximately 54%), while a large minority of comments contain no interpretable trust cue (around 45%). High trust is rare (about 2%), and medium trust is marginal. This pattern suggests that when commenters do engage institutions as objects of evaluation, the prevailing stance is distrustful, with affirmational trust emerging only sporadically.

Collective efficacy is more weakly articulated but still present (Figure A6). About two-thirds of comments are unrelated to collective capacity, indicating that many contributions do not frame wildfire response in terms of shared agency or coordinated action. Among those that do, low collective efficacy is the most common (approximately 25%), while medium and high collective efficacy appear only in small shares (roughly 7% and 4%, respectively). This distribution indicates that expressions of joint capacity to respond exist but are far from dominant and are more often framed pessimistically than optimistically.

Calls to action occur in a minority of comments, but at a visible level (Figure A7). Approximately 22% of comments contain an explicit call to action, while the remainder do not. This suggests that alongside critique and interpretation, the comment space includes a distinct action-oriented register in which commenters urge specific behaviors, interventions, or political responses, though this register is not the primary mode of engagement.

When concrete impacts are mentioned, ecological and material losses are most salient (Figure 12). Ecosystem impacts are referenced most frequently (around 10%), closely followed by property damage (just under 10%). Broader economic consequences are also prominent (around 6%), and health impacts appear in approximately 5% of comments. Infrastructure and agriculture impacts are each mentioned at about 2–3%, while evacuation-related impacts appear at roughly 1–2%. Mentions of water and smoke/air quality appear at comparatively low levels (under 1% each in this coding). This suggests that commenters most often foreground environmental degradation and tangible asset loss, while also registering economic disruption and health burdens to a meaningful extent.

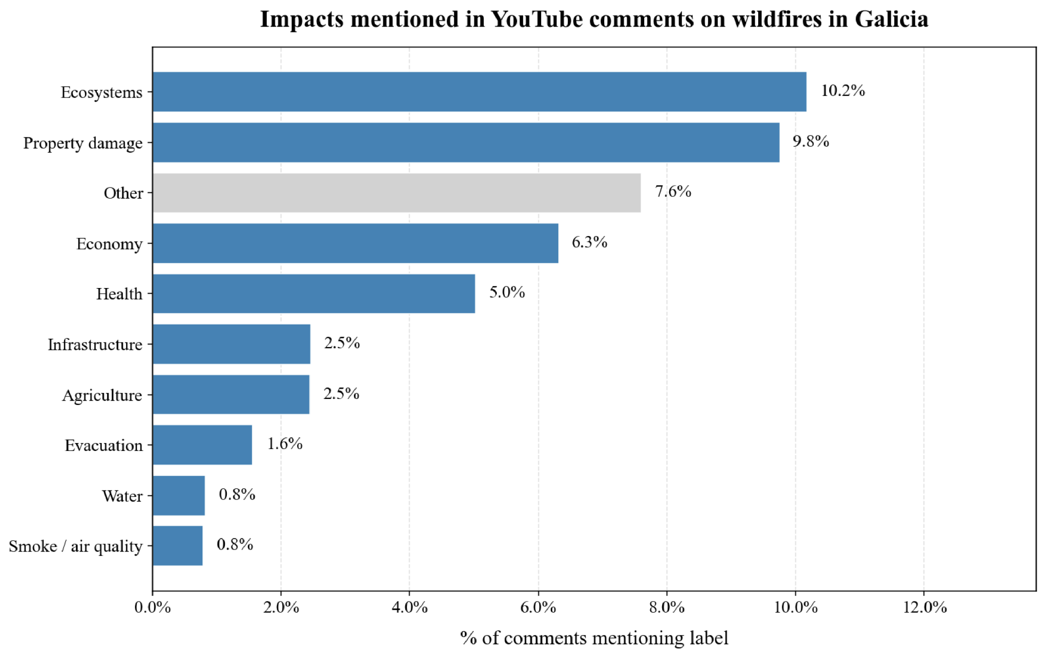

Figure 12. Impacts mentioned in YouTube comments on wildfires in Galicia.

Mentions of vulnerable groups are comparatively limited in absolute terms (Figure 13). Outdoor workers are the most frequently referenced group (around 1.8%), followed by low-income populations (about 1.1%). References to the elderly and animals or pets appear at under 1% each, while children are mentioned at roughly half a percent. Tourists, respiratory patients, and people with chronic illness are only rarely referenced. This distribution indicates that vulnerability awareness is present but not central in the wildfire comment discourse, despite the well-established unevenness of exposure and capacity during wildfire events.

Coping actions are also relatively sparse, with a few categories standing out (Figure 14). Helping neighbors is the most frequently coded coping-oriented behavior (around 2.8%), suggesting that mutual aid is a recognizable, albeit still minority, theme. Political demand and references to firefighting tactics each appear at around 1.4%, and evacuation-related actions at roughly 1%. Requests to seek updates and references to prayers appear in the sub-1% range, and donations are mentioned in a small minority of comments. Other actions, including reporting or denouncing, masking, air filtering, protecting property, fuel management, and controlled burns, are rare in this comment corpus. The pattern suggests that wildfire threads contain some articulation of mutual aid and response preferences, but that detailed behavioral guidance is not a dominant discourse form.

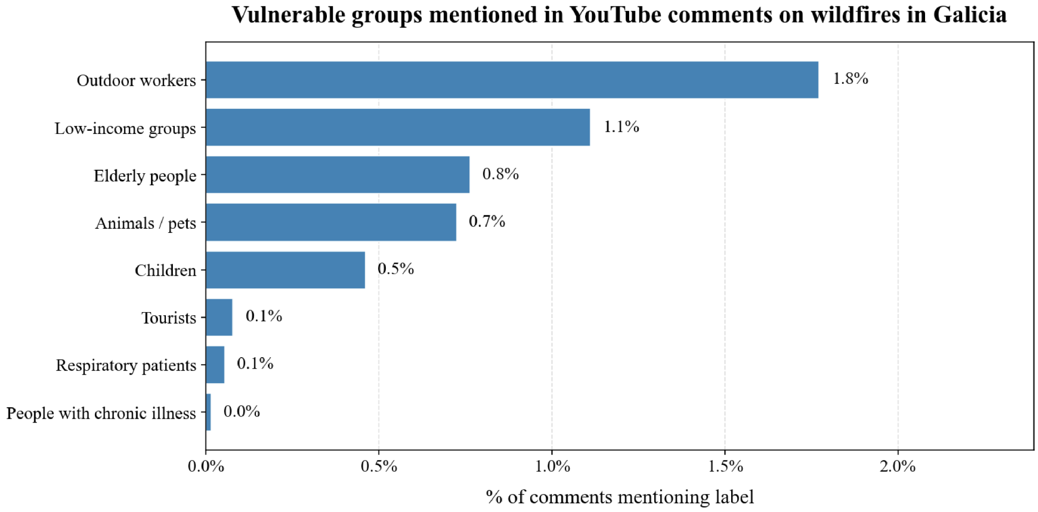

Figure 13. Vulnerable groups mentioned in YouTube comments on wildfires in Galicia.

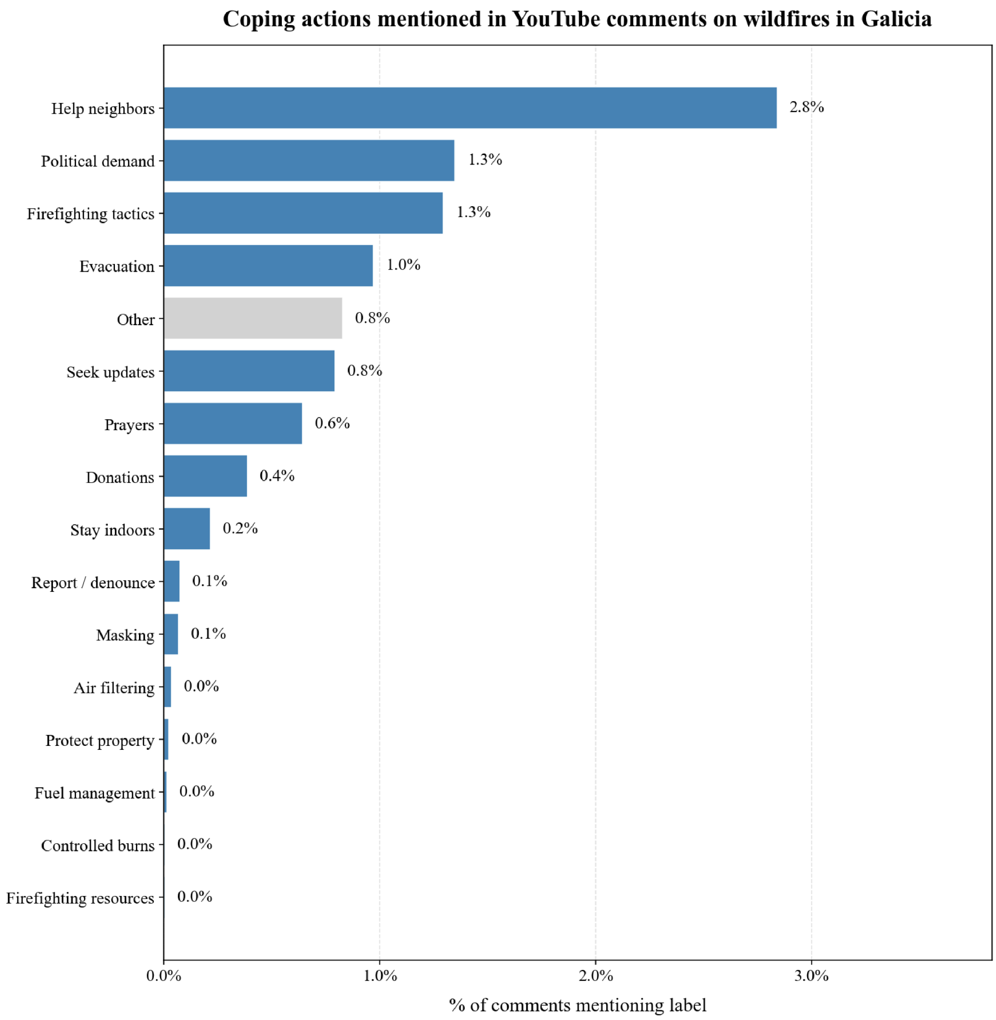

Figure 14. Coping actions mentioned in YouTube comments on wildfires in Galicia.

Adaptation and policy suggestions, when present, are structured primarily around capacity and land-based risk management (Figure 15). The most frequently referenced category concerns firefighting resources (approximately 8%), indicating substantial attention to operational capacity, equipment, staffing, or resourcing. Fuel management is the next most common (around 5%), followed by land-use policy (around 5%). Controlled burns, early warning, and evacuation plans appear at under 1% each, and public health messaging and harsher penalties appear at only a few tenths of a percent. Enforcement or bans and governance or anti-corruption are marginal. These distributions indicate that wildfire discourse is not only evaluative but also policy-relevant. When commenters articulate forward-looking solutions, they most often emphasize response capacity, vegetation and fuel interventions, and land-use governance as levers for risk reduction.

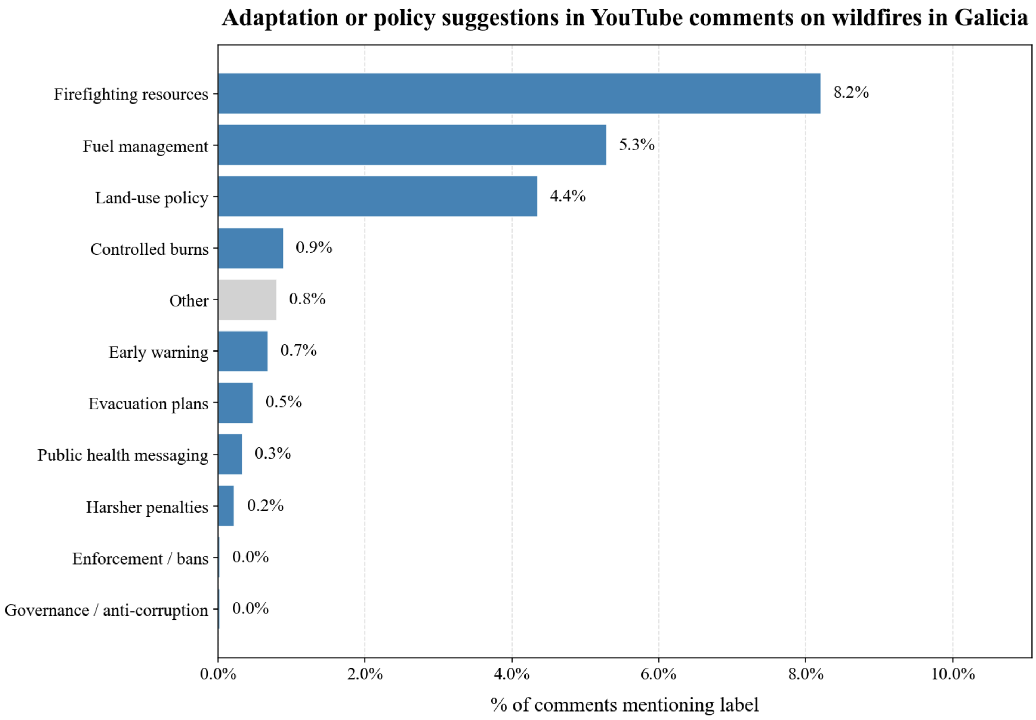

Figure 15. Adaptation and policy suggestions mentioned in YouTube comments on wildfires in Galicia.

3.2.3. Floods in Galicia

The YouTube search yielded 30 videos meeting the study criteria and a total of 4882 comments for analysis. The sentiment distribution for flood-related comments is strongly negative (Figure 16). The histogram shows a dense concentration of observations in the moderately to strongly negative range, with the highest mass centered roughly between -0.75 and -0.55, indicating that critical and adverse evaluations dominate the discourse rather than mild dissatisfaction. A smaller accumulation appears near neutral sentiment, while a thinner positive tail extends into moderately positive values.

Emotion coding clarifies how this negativity is expressed (Figure 17). Anger is the single most common coded emotion (about 17%). Mockery (~16%) and frustration (~14%) are nearly as prominent, while sadness forms another major component (~12%). Concern and hope are both substantial (each roughly 8%), suggesting that many commenters frame floods not only as an object of blame but also as a condition requiring protection, preparedness, or recovery. Neutral affect accounts for about 7%, and amusement for roughly 6%, indicating that low-stakes engagement and joking remain present even in disaster discussion. Disbelief (~4%) and fear/relief (each around 3%) appear less often.

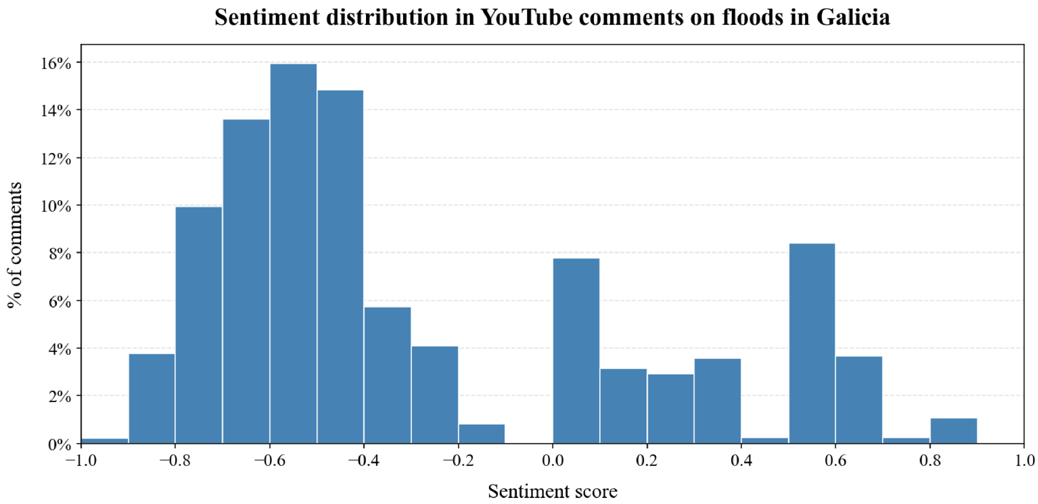

Figure 16. Sentiment distribution of YouTube comments on floods in Galicia.

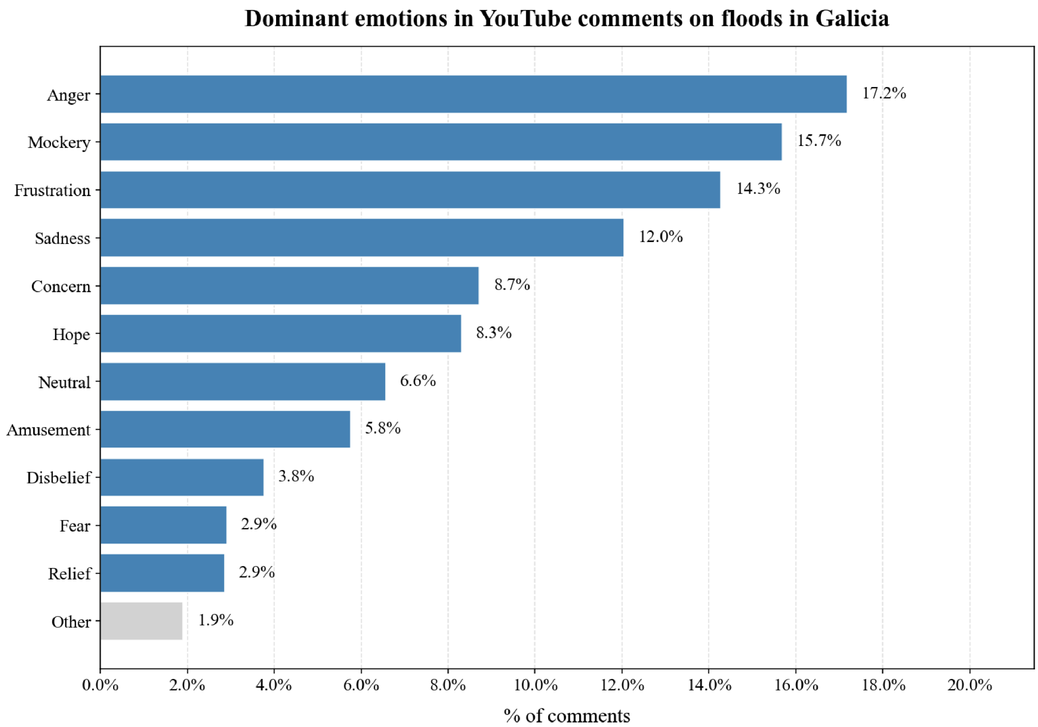

Figure 17. Dominant emotions in YouTube comments on floods in Galicia.

The main topic distribution indicates that flood discourse is organized around a small set of large interpretive frames plus a long tail of minor themes (Figure 18). The other category remains sizable at roughly one quarter of comments, reflecting substantial topical fragmentation beyond the major topics. Among the substantive categories, condolences, sympathy, prayers, and solidarity form one of the largest blocks (about 15%), while politics, government performance, corruption, and accountability constitute another major share (roughly 14%). More event-focused discussion accounts for around 6%. Socially adversarial participation, including personal attacks, insults, and mockery, is also prominent and represents about 6%, and low-content reactions, chatter, greetings, and off-topic comments contribute another ~5%. Emergency-relevant interpretation is visible through warnings, alerts, and emergency management (about 5%), alongside smaller but persistent categories such as thanks/gratitude/praise (~3%), media coverage and censorship/sensationalism (~3%), and hazard-adjacent discussion including coastal waves/sea, storms/tsunami talk and video/footage timestamps or audio/music meta (each around ~2%). Additional categories, including aid logistics, rescue/heroism, casualties, information-seeking, identity/accent/nationality talk, flood cause attribution, and flood impacts, appear in low single digits to ~1% each, indicating a broad but secondary layer of operational, informational, and interpretive commentary.

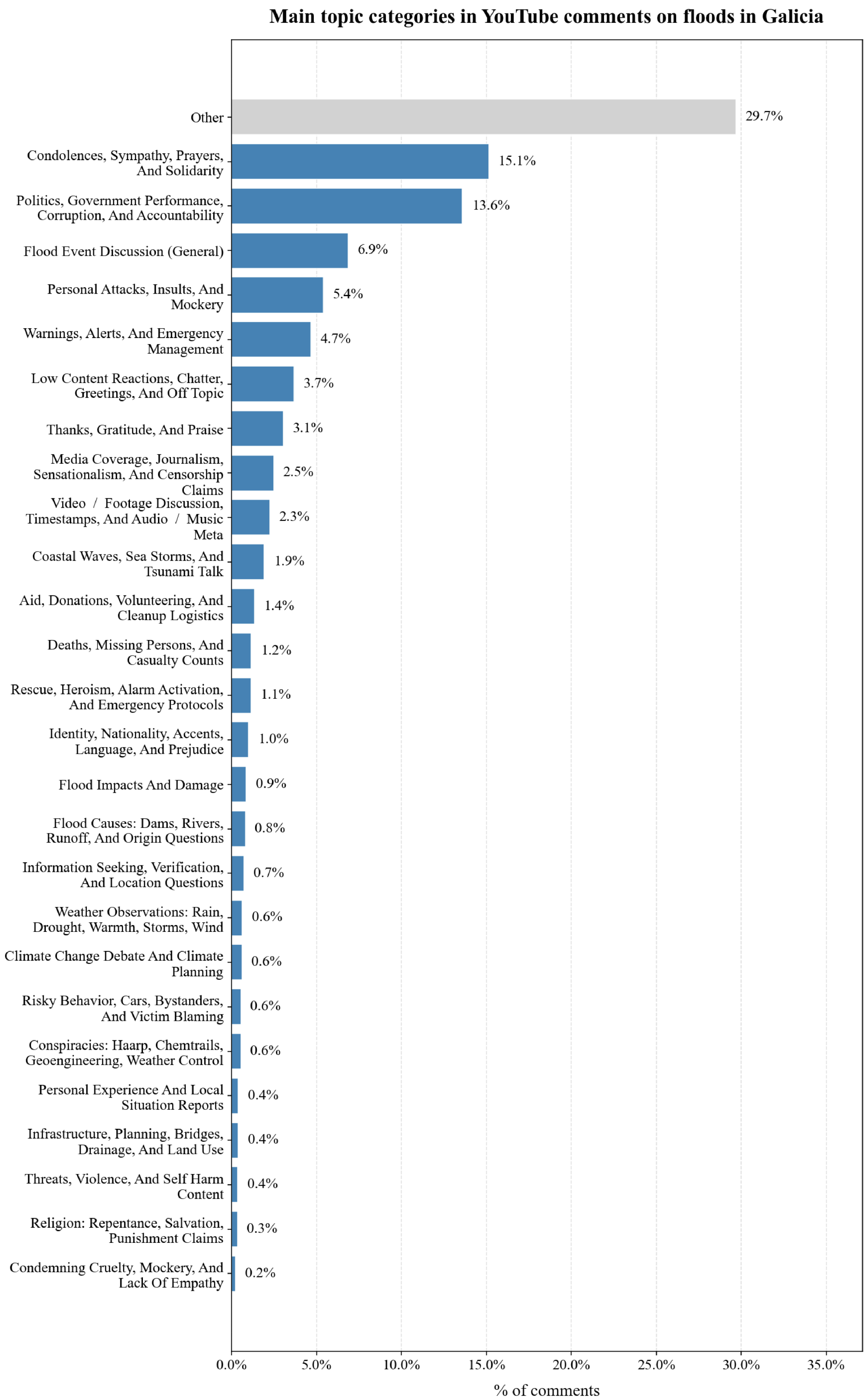

Figure 18. Main topic categories in YouTube comments on floods in Galicia.

Resilience-related signals highlight a strong pattern of institutional ambiguity and skepticism (Figure A8). The majority of comments are unrelated to trust in institutions (about 65%). Where trust is expressed, it is much more often low (~32%) than affirmational. High trust (~2%) and medium trust (~1%) are rare. This suggests that institutional evaluation is present but unevenly articulated.

Collective efficacy is even less directly articulated (Figure A9). Roughly 80% of comments contain no clear collective efficacy cue. Among those that do, low collective efficacy (~13%) is most common, while medium (~6%) and high (~3%) appear only in small minorities. This indicates that while some commenters reference the shared capacity to respond (community readiness, coordination, mutual aid), explicit confidence in collective response is limited and more often pessimistic than optimistic.

Calls to action occur in a minority of comments but are clearly visible (Figure A10). Approximately 18% of comments contain an explicit call to action, while about 82% do not. This suggests that flood threads are primarily evaluative and expressive (emotion, blame, solidarity), with a smaller but meaningful action-oriented register urging specific behaviors, preparedness steps, or political/institutional responses.

When commenters mention concrete impacts, direct harm and material loss dominate (Figure 19). Injury/drowning references are most frequent (about 7.5%), closely followed by property damage (just over 7%). Infrastructure impacts are also substantial (around 4.7%), while evacuation (~3%) and health impacts (~2.3%) appear at moderate levels. Broader economic consequences are referenced less often (~1.5%), and ecosystem impacts appear in under 1% of comments. Mentions of water contamination, agriculture, missing people, and mud/debris-related outcomes are comparatively rare in this coding. Flood impacts are framed most often through immediate life-and-death stakes, damaged property, and disrupted infrastructure.

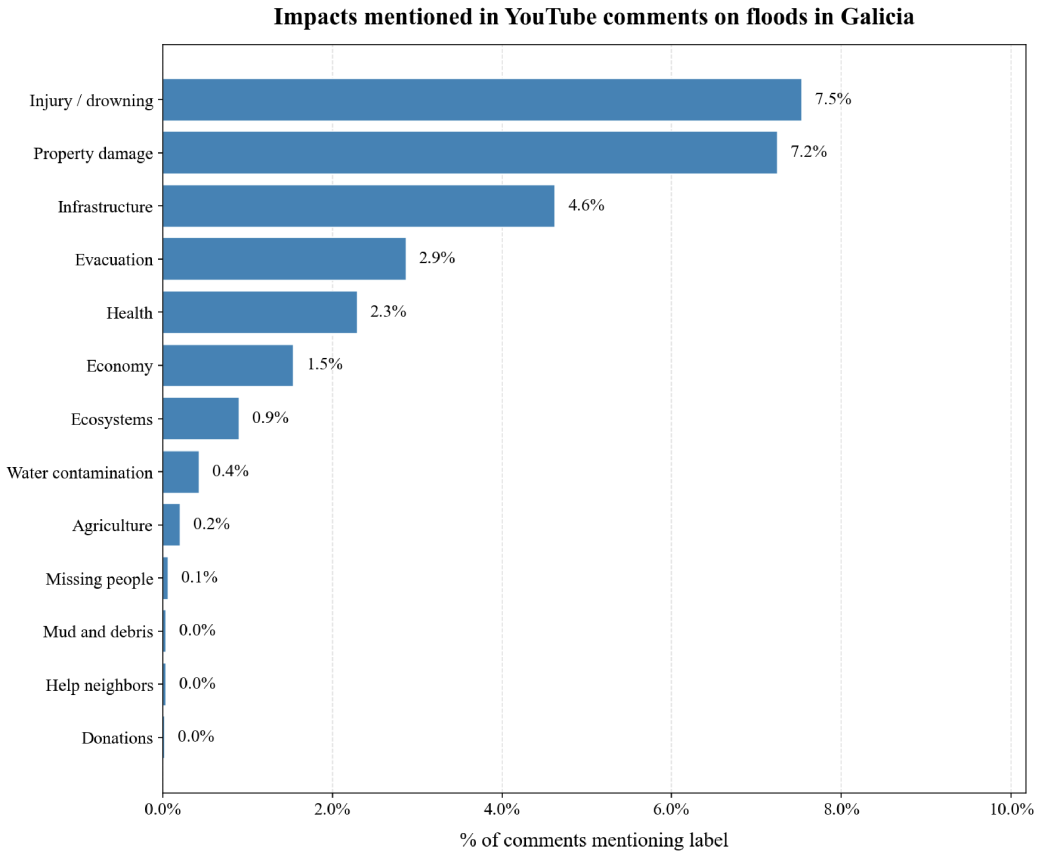

Figure 19. Impacts mentioned in YouTube comments on floods in Galicia.

Mentions of vulnerable groups are comparatively limited (Figure 20). Low-income groups are referenced most often (roughly 0.8%), followed by children (~0.6%) and the elderly (~0.5%). Animals/pets appear at around 0.4%, while outdoor workers, tourists, and disability/mobility are rarely mentioned. This pattern suggests that while vulnerability awareness exists, it is not a central organizing theme of flood comment discourse.

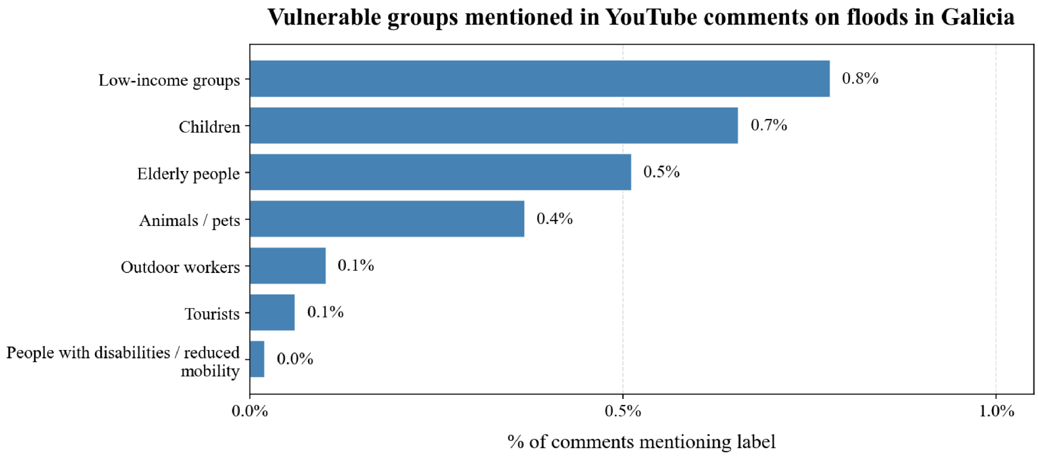

Figure 20. Vulnerable groups mentioned in YouTube comments on floods in Galicia.

Coping actions, when present, foreground mutual aid and immediate protective behaviors (Figure 21). Helping neighbors is the most frequently coded action (~4.4%), followed by donations (~2.4%) and evacuation (~2.1%). Seeking updates is also notable (~1.8%) and moving to higher ground (~1%) and avoiding travel (~0.8%) appear as secondary precautionary behaviors. Cleanup, asking where to donate, and sandbags appear only rarely. This suggests that flood threads contain a recognizable, though minority, practical-action layer that emphasizes mutual aid, evacuation-related behaviors, and information seeking.

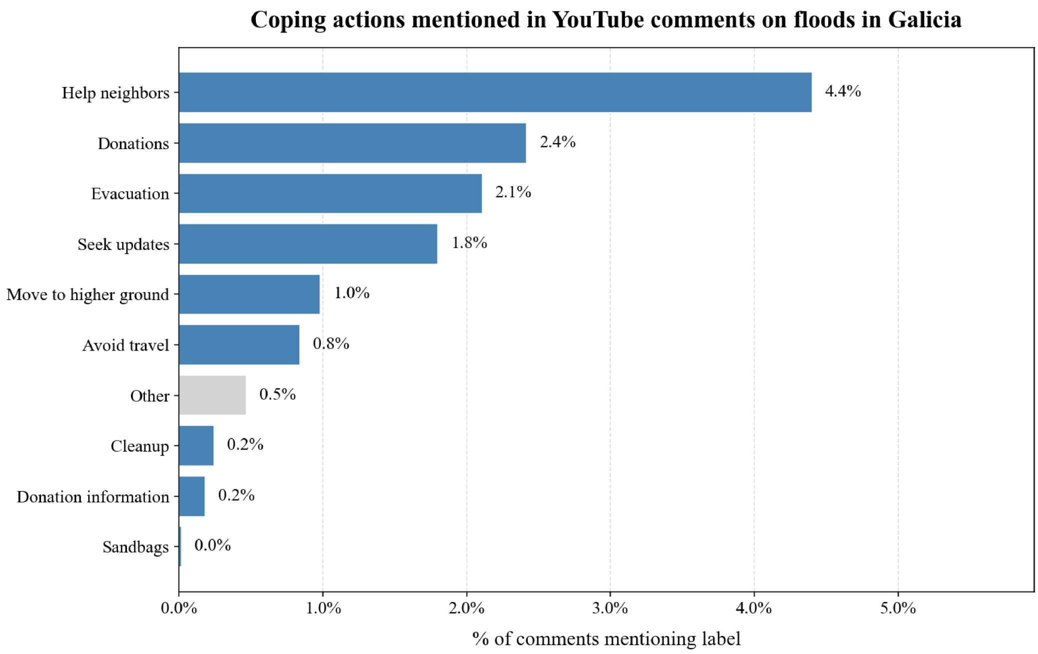

Figure 21. Coping actions mentioned in YouTube comments on floods in Galicia.

Adaptation and policy suggestions, when voiced, are structured primarily around preparedness capacity and risk communication (Figure 22). The most frequently referenced category is early warning (~4.5%), closely followed by emergency service capacity (~4.2%)

and public safety messaging (~2.9%). Land-use policy appears at around 1%, while evacuation plans, drainage infrastructure, and river management appear at around ~0.8% each. Mentions of disaster funds, trash/maintenance, and anti-corruption as explicit adaptation levers are marginal. This distribution suggests that when commenters propose forward-looking measures, they prioritize improving warning systems, strengthening emergency response capacity, and clarifying public safety communication, with structural planning and infrastructure interventions present but less central.

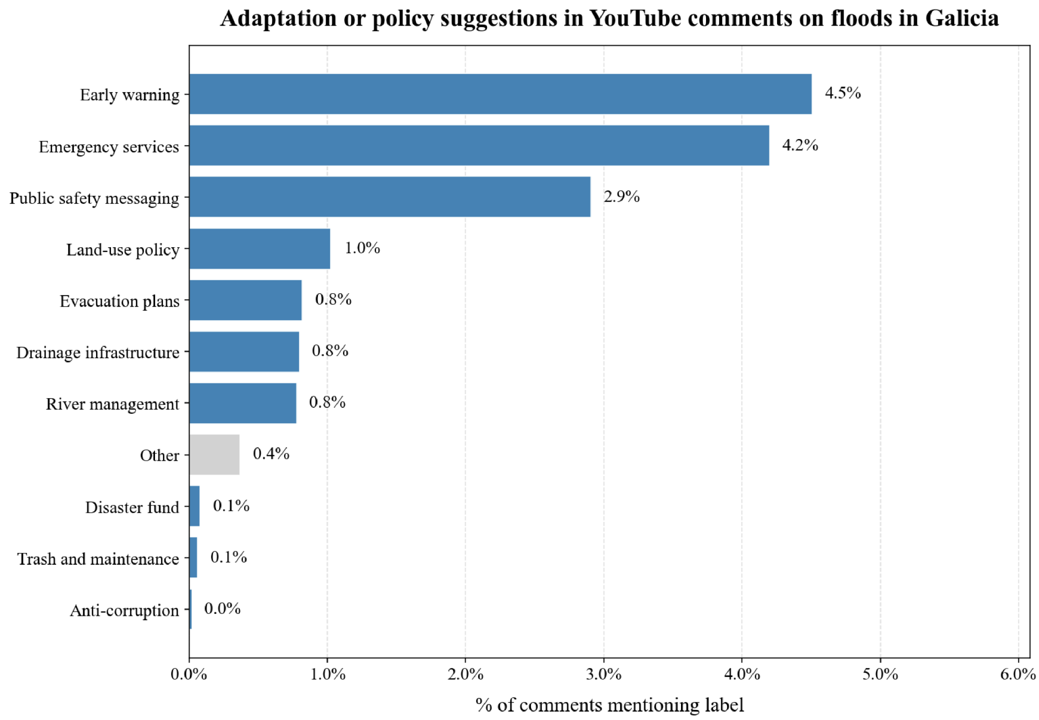

Figure 22. Adaptation and policy suggestions mentioned in YouTube comments on floods in Galicia.

3.3. Data-Driven Policy Recommendations

The qualitative evidence from YouTube comment discourse was translated into policy advice through a structured interpretive pipeline. Comments were first interpreted as signals of resilience-relevant gaps, vulnerabilities, or enabling conditions, and then reformulated as actionable resilience opportunities. These opportunities were grouped within and across hazards to produce higher-level policy recommendations, allowing the analysis to move from dispersed public reactions to a coherent agenda for climate-risk governance (Figure 23).

Across floods, wildfires, and heatwaves, the recommendations converge on the need for real-time, local, and actionable public guidance. Crisis communication should move beyond general messaging and provide spatially specific risk information, time-relevant updates, and clear protective actions, including triggers for evacuation, sheltering, travel avoidance, or seeking assistance. A unified public information system, supported by maps, dashboards, and plain-language advisories, would help households, responders, and institutions work from a common information base.

A second cross-hazard priority is practical preparedness and behavioral risk reduction. The comments repeatedly point to avoidable harms that could be reduced through standardized, scenario-specific guidance before, during, and after events. Such guidance should be accessible, realistic under household and infrastructure constraints, and focused on feasible actions rather than treating preparedness as only an individual responsibility.

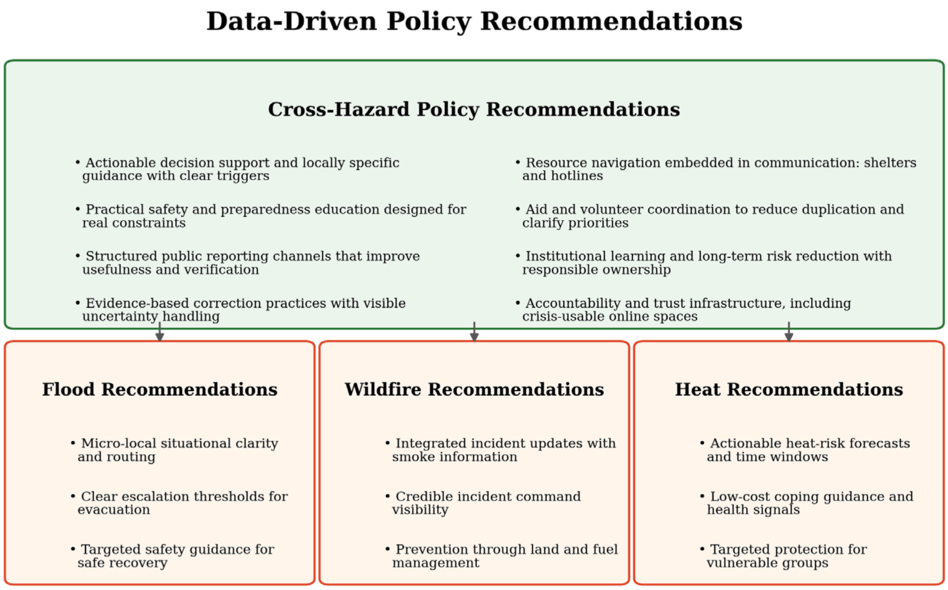

Figure 23. Data-driven policy recommendations.

A third recommendation concerns the use of public observation while limiting rumors, misinformation, and unsafe self-exposure. Local reports and informal updates can improve situational awareness, especially when conditions change rapidly or vary at small spatial scales. However, these inputs require structured reporting channels, verification prompts, and safeguards so that citizen reporting complements official monitoring without amplifying misinformation or encouraging risky behavior.

The analysis also highlights trust infrastructure as a central component of resilience. Misinformation, toxic conflict, and low institutional trust appeared across hazard contexts. Policy responses should therefore emphasize rapid evidence-based corrections, transparent communication about what is known and unknown, and crisis-usable online spaces where verified updates are visible and safety information is not displaced by harassment, political conflict, or rumor cycles.

Public expressions of sympathy, anger, and solidarity also translate into operational needs. Crisis communication should include clear pathways to shelters, hotlines, family reunification services, verified aid channels, psychosocial support, and volunteer or donation coordination. These systems can reduce duplication, fraud, and confusion while strengthening public confidence that assistance is reaching affected people.

Beyond immediate response, the recommendations point to institutional learning and long-term risk reduction. Public-facing after-action reviews, clear ownership of follow-up commitments, transparent performance summaries, and investments in maintenance, prevention, and response capacity can help rebuild trust and reduce repeated impacts.

Hazard-specific recommendations refine this cross-hazard agenda. For floods, the priority is micro-local situational clarity, including localized flood intelligence, road and infrastructure status, evacuation thresholds, safe routing, and practical guidance on floodwater, electrical hazards, contaminated water, and cleanup. For wildfires, the recommendations emphasize transparent incident command, credible post-incident investigations, prevention through land and fuel management, smoke and health guidance, evacuation logistics, and targeted support for vulnerable groups. For heatwaves, the focus is on actionable heat-risk alerts, countering normalization of heat danger, low-cost protective guidance, symptom-based escalation advice, targeted protection for older adults, outdoor workers, and high-exposure neighborhoods, and public cooling or built-environment measures that reduce exposure during peak risk periods.

4. Discussion

Across 2022–2025, the Google Trends layer indicates that climate hazards appear in Galicia's online information environment as episodic attention shocks rather than smooth or continuous cycles. This pattern is consistent with crisis informatics arguments that emergencies are also information crises, in which attention, interpretation, and coordination evolve dynamically through networked media [34].

Floods dominate Galicia's hazard-related search interest, with frequent moderate surges and at least one extremely high peak. This suggests that floods function as a chronic and repeatedly revisited risk in the region's online information needs, consistent with Galicia's hydro-meteorological exposure and recurring storm impacts. It also aligns with infodemiology and infoveillance frameworks, in which search behavior is treated as a proxy for evolving public concern and information demand [8,22]. By contrast, wildfires show a quieter baseline but sharper spikes, reflecting a more seasonal and incident-driven attention pattern. Wildfire-related queries emphasizing now/today, maps, locations, and counts suggest that information seeking is triggered by acute proximity cues such as smoke, evacuations, disruption, and dramatic imagery, mirroring evidence from other contexts that online search behavior can track perceived exposure and protective responses [10]. Heatwaves occupy a more ambiguous position. Search interest rises during specific hot periods but remains below floods and wildfires on average, while YouTube discourse still reveals substantial contestation. This split reinforces an important interpretive point. Search salience measures information demand, not necessarily risk severity. Heat can remain deadly even when it is socially normalized or treated as routine, as shown by large-scale European mortality analyses [35].

The YouTube results show three distinct public meaning-making regimes that correspond to differences in hazard characteristics, perceived immediacy, and governance controversy. Heatwave comment threads are dominated by mockery and lifestyle or seasonal framing, with comparatively low levels of explicit institutional trust, collective efficacy, and concrete coping guidance. This suggests an actionability gap. Many comments recognize heat as a topic of discussion, but fewer translate it into protective behavior or collective adaptation. This pattern is consistent with risk communication models in which familiar, controllable, or socially contested hazards may prompt identity signaling and argument rather than practical response. The Protective Action Decision Model emphasizes that cues, information sources, and perceived protective efficacy shape whether people convert warnings into action [36]. For heatwaves in Galicia, the resilience challenge is therefore not only forecasting accuracy, but also counteracting benign framing through locally grounded, feasible actions such as cooling, hydration, schedule adjustments, symptom-based escalation, and credible public health messaging. This is consistent with guidance emphasizing that heat harms are predictable and preventable when warnings are paired with actionable interventions [37,38].

Wildfire discourse is structurally different. Anger is dominant, politics and governance form the largest topic category, low institutional trust is frequent, and solution-oriented comments concentrate on firefighting resources, fuel management, and land-use policy. This resembles a blame–responsibility–capacity discourse, in which losses are interpreted through questions of institutional competence, accountability, and prevention. The pattern supports crisis informatics findings that online disaster communication is not merely noise but an interpretive and coordination process with implications for trust and compliance [34,39]. It also fits the broader trajectory of wildfire risk under climate change, including evidence that climate-driven fire–weather extremes are increasing in many regions and that Mediterranean fire regimes are sensitive to heat and dryness [40,41].

Flood discourse combines substantial condolences and solidarity with anger, skepticism, and comparatively more discussion of warnings, alerts, and immediate protective behaviors such as evacuation, helping neighbors, and donations. This mixture of high emotion, visible mutual aid, and demand for localized updates matches the dynamics of flood events, where conditions change rapidly and often at micro-spatial scales. Because floods generate visible environmental cues and urgent time pressure, the value of information is highest when it is specific, credible, place-based, and immediately executable [36].

A key contribution of this case study is therefore the demonstration that different hazards generate different public information modes and require different policy levers. For heatwaves, the main resilience gap is risk underestimation and low actionability, pointing to heat–health warning systems, targeted protection for elderly people and outdoor workers, and low-cost coping guidance. For wildfires, the central gap is legitimacy, accountability, and capacity credibility, pointing to transparent incident communication, post-incident reviews, fuel and land management, and integrated smoke–health guidance. For floods, the central gap is micro-local situational clarity, pointing to nowcasting-style public updates, road and infrastructure status, evacuation triggers, and safety guidance for common injury pathways. This hazard-specific differentiation is consistent with climate adaptation and resilient development framing, which stresses context-dependent vulnerability, governance capacity, and risk communication as central determinants of outcomes [42].

Although the empirical analysis focuses on Galicia, the proposed approach is transferable to other regions facing climate-related hazards, provided that it is recalibrated to local hazard profiles, languages, platform-use patterns, institutional arrangements, and validation data. The general logic of combining search interest signals, platform discourse, AI-assisted annotation, and policy-oriented synthesis can be applied in other coastal, fire-prone, flood-prone, or heat-exposed regions. However, cross-regional transfer should not assume that the same keywords, platforms, emotional patterns, trust dynamics, or policy priorities will appear elsewhere. The Galicia case should instead be understood as both an empirical contribution and a methodological demonstration. It provides evidence for Galicia while offering a replicable template for AI-supported social sensing in other climate-risk contexts.

5. Limitations

The findings should be interpreted as indicative signals of online attention, public concern, trust, and actionability rather than as representative measurements of Galicia's population or direct measurements of physical hazard severity. The empirical corpus is highly unbalanced across hazards. The heatwave sample includes five videos and 184 comments, the wildfire sample includes 303 videos and 20,427 comments, and the flood sample includes 30 videos and 4882 comments. As a result, the wildfire findings are likely to be more stable because they draw on a much larger and more diverse corpus, while the heatwave findings should be treated as exploratory and more sensitive to the characteristics of a small number of videos. Future research should address this imbalance through stratified sampling, minimum sample thresholds, repeated data collection, and normalized comparison strategies.

Google Trends also introduces important measurement constraints. Its core output is not an absolute count of searches, but a normalized index derived from sampled query streams. Repeated retrievals may therefore produce variation, and the resulting series may reflect sampling variance, measurement error, and threats to internal validity if reliability and sensitivity checks are not conducted. Recent methodological studies document these sources of instability and highlight frequent misapplications in applied work, showing that Google Trends outputs can be over-interpreted without careful design and diagnostics [43–46]. In this study, Google Trends data are therefore treated as relative attention signals rather than incidence measures. For decision support, they should be interpreted alongside independent data streams such as hazard records, impact data, meteorological observations, and survey evidence.

The YouTube layer is shaped by platform, interface, and API effects. What users see and comment on is influenced by ranking and recommendation systems, while audits indicate that YouTube's Search API may return results that are incomplete, inconsistent across requests, and temporally unstable, limiting replicability in historical samples [47,48]. Related work also documents drift and systematic biases in YouTube's recommendation environment, reinforcing that retrieved content should be understood as a platform-shaped sample rather than a comprehensive archive of public discourse [49]. Videos or comments removed before collection, disabled comment sections, omitted replies, and search instability may shape the final corpus.

The comment data also reflect self-selection and visibility biases. The analysis may overrepresent people who are active on YouTube, willing to comment publicly, emotionally motivated to respond, or exposed to highly ranked and recommended videos. Conversely, people who are less digitally connected, less politically expressive, older, rural, linguistically marginalized, or directly affected but not active online may be underrepresented. Disaster research has long shown that social media participation is uneven across demographic groups and places, and that online visibility does not necessarily correspond to physical exposure, vulnerability, or need [50–52]. Consequently, the findings should not be used to infer demographic representativeness or exact public opinion, but rather to identify recurring themes, concerns, information gaps, and signals for further investigation.

Another linguistic limitation concerns the YouTube search strategy. Although Galicia is a bilingual region and climate hazard discourse may occur in both Spanish and Galician, the YouTube retrieval used Spanish query strings. This decision was made to maximize retrieval volume, improve comparability across hazards, and capture videos circulated through regional and national Spanish-language media. However, it may have underrepresented Galician-language videos, locally oriented channels, and discourse using Galician hazard terminology. As a result, the YouTube corpus should not be interpreted as a comprehensive representation of all climate hazard discourse in Galicia. Future research should implement bilingual or multilingual query expansion, compare Spanish and Galician retrieval results, and assess whether language choice affects the sentiment, topic, trust, and policy demand patterns identified in the analysis.

A further limitation concerns the comment-level nature of the sentiment and emotion analysis. Each comment is assigned an overall sentiment score and dominant emotion label, but this does not identify which specific actor, institution, place, impact, or issue is responsible for that sentiment or emotion. For example, a negative comment may express anger toward public authorities, sadness about victims, distrust of media coverage, or frustration about land management. Aggregate labels cannot fully distinguish these targets. Future research should therefore use aspect-based sentiment and emotion analysis to identify whether negative or positive evaluations refer to emergency response, media communication, evacuation management, land-use policy, institutional credibility, or the hazard itself.

GPT-based annotation and classification add another layer of model-mediated uncertainty. Even with a structured rubric, outputs may be sensitive to prompt wording, category definitions, sarcasm, multilingual expressions, ambiguous comments, and edge cases in interpretation. They may also reflect latent biases embedded in training data or evaluation procedures. Measurement and machine learning guidance emphasizes the need for validation against human-coded benchmarks, robustness testing, and transparent documentation of model decisions and uncertainty [53–55]. Future research should compare AI-coded outputs with human-coded validation samples, test inter-coder agreement, apply alternative models, repeat data collection at different time points, and integrate a qualitative close reading of selected comment threads.

Finally, the spatial precision of the data is limited, which restricts direct operational use for disaster management. Google Trends is geographically aggregated and does not identify the exact location of individual searches. YouTube videos and comments often refer to broad places such as Galicia, a province, or a municipality, and many comments contain no reliable location information. Even when place names appear, they may refer to where a video was filmed, where the commenter lives, where the hazard occurred, or a place mentioned in the news. These meanings cannot always be separated automatically. This spatial imprecision limits the use of the results for emergency dispatch, evacuation planning, or resource allocation. Future work should combine online discourse signals with geocoded official incident data, meteorological observations, remote sensing, emergency-call data, local news, and verified citizen reports to improve spatial validation and operational relevance.

These limitations support a decision-support rather than replacement framing. The pipeline is designed to complement established disaster intelligence, not substitute for it. Responsible deployment therefore requires governance safeguards, traceable audit trails, validation against external benchmarks, and periodic calibration to operational ground truth [29,56].

6. AI-Enabled Resilience Pipeline as an Operational Outcome

The pipeline is a core operational outcome of the study. It translates the Google Trends search-signal layer, YouTube social-listening layer, and AI-assisted annotation into an end-to-end decision-support workflow for climate hazard resilience. As shown in Figure 24, the workflow connects heterogeneous inputs to operational outputs and embeds calibration, validation, and model-updating loops.

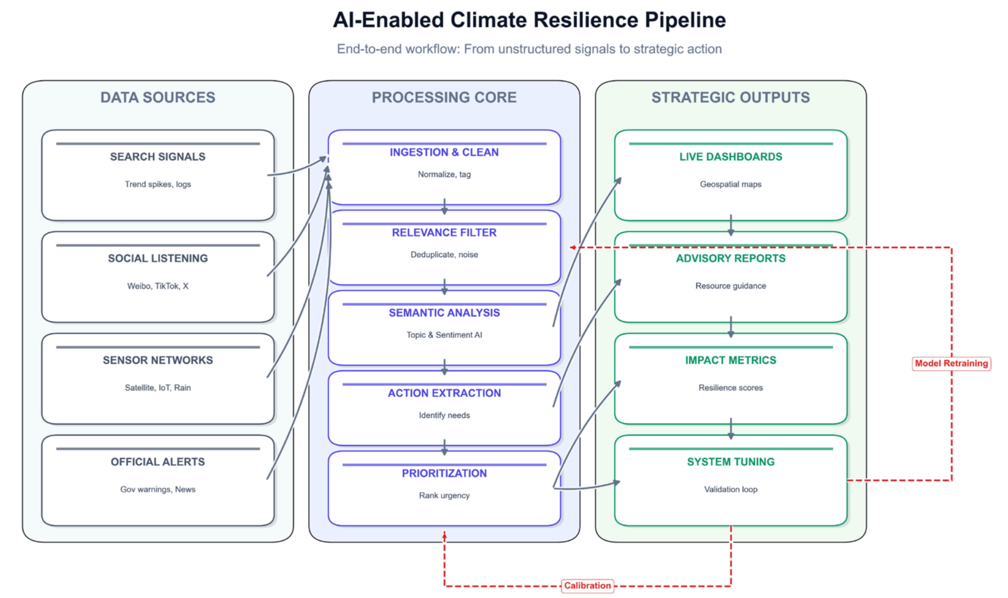

Figure 24. AI-enabled climate resilience pipeline.

The data-source layer combines attention signals, social listening, sensor and satellite evidence, and official alerts. Search data provide rapid, population-scale indicators of shifting relevance [22,57], while platform discourse adds context on perceived risk, contested interpretations, and emergent needs [58]. Sensor networks and remote sensing anchor these signals in observed hazard dynamics and impacts [59,60], and official alerts provide authoritative benchmarks for decision-support use.

The processing core converts raw inputs into structured evidence. Ingestion, cleaning, and relevance filtering support reproducibility and interpretability [61,62]. Semantic analysis summarizes topics, stance, sentiment, and emotion, while action extraction identifies impacts, vulnerable groups, coping suggestions, and other decision-relevant variables. Prioritization then ranks signals by urgency, potential harm, and operational leverage.

The strategic-output layer turns these indicators into live dashboards, advisory reports, and impact metrics. Dashboards support situational awareness and anomaly detection, especially when combined with official alerts and remote sensing layers [59,60]. Advisory reports explain emerging needs and plausible actions, while metrics make trends comparable across hazards and over time.

Finally, the calibration and retraining loops keep the system accountable and adaptive. Outputs should be checked against ground truths such as hazard intensity, verified impacts, and post-event assessments, reflecting AI governance guidance that treats monitoring and documented evaluation as continuous obligations [61,62]. Model retraining addresses changing language, narratives, misinformation, and platform effects, including YouTube API and search limitations that affect representativeness and replicability [63].

7. Conclusions

This study answers the overarching research question by showing that AI-supported analysis of online search behavior and platform discourse can assess the social-information dimensions of climate hazard resilience in Galicia. Google Trends captures public attention and information demand, while YouTube comments reveal how people interpret hazards, express trust or distrust, identify impacts, and call for action. These sources can be transformed into practical indicators for decision support, although they should complement rather than replace official hazard monitoring.

In response to RQ1, the Google Trends analysis shows that floods, wildfires, and heatwaves generated the most salient and interpretable online search interest signals in Galicia during 2022–2025. Floods had the highest mean search interest and frequent surges, reflecting demand for real-time and place-specific information. Wildfire searches were more episodic and incident-driven, often focused on current locations, maps, and emergency updates. Heatwave searches were lower on average but reflected needs for timing, duration, official warnings, and practical protection.

In response to RQ2, YouTube discussions differed clearly across the three hazards. Heatwave discourse was often marked by mockery, normalization, and limited coping guidance, suggesting an actionability gap despite recognition of health risks. Wildfire discourse was more negative and anger-driven, dominated by governance, accountability, low institutional trust, and demands for better firefighting capacity, prevention, and land management. Flood discourse combined grief, solidarity, criticism, and stronger demand for localized warnings, evacuation guidance, road-status updates, and immediate protective information.

In response to RQ3, the online signals can be synthesized into traceable policy recommendations by LLM-assisted semantic grouping and topic assignment of comment-level resilience actions into broader intervention areas. Across hazards, the strongest recommendations are to provide real-time, locally actionable risk communication, develop scenario-specific preparedness guidance, support verified citizen reporting, strengthen trust through transparent communication, and connect solidarity and help offers to verified aid, shelter, hotline, and volunteer systems. Hazard-specific priorities include micro-local flood guidance, transparent wildfire response and prevention, and heat-risk communication targeted at older adults, outdoor workers, and exposed neighborhoods.

Author Contributions: Conceptualization, D.E. and N.K.; methodology, D.E. and N.K.; software, D.E.; validation, D.E. and N.K.; formal analysis, D.E.; investigation, D.E.; resources, N.K.; data curation, D.E.; writing—original draft preparation, D.E.; writing—review and editing, N.K.; visualization, D.E.; supervision, N.K.; project administration, N.K.; funding acquisition, N.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the HORIZON-MISS-2024-CLIMA-01 SystR Project, grant number 101212761.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are publicly available.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

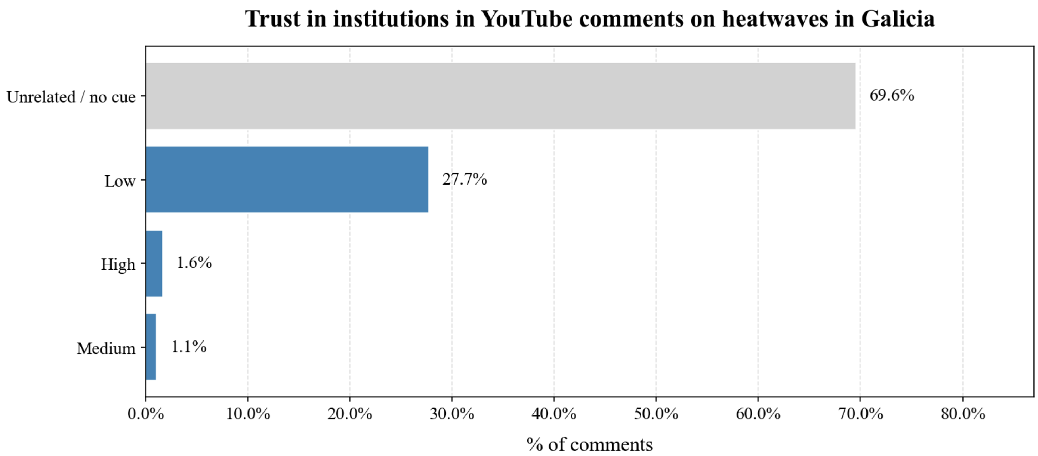

Figure A1. Trust in institutions in YouTube comments on heatwaves in Galicia.

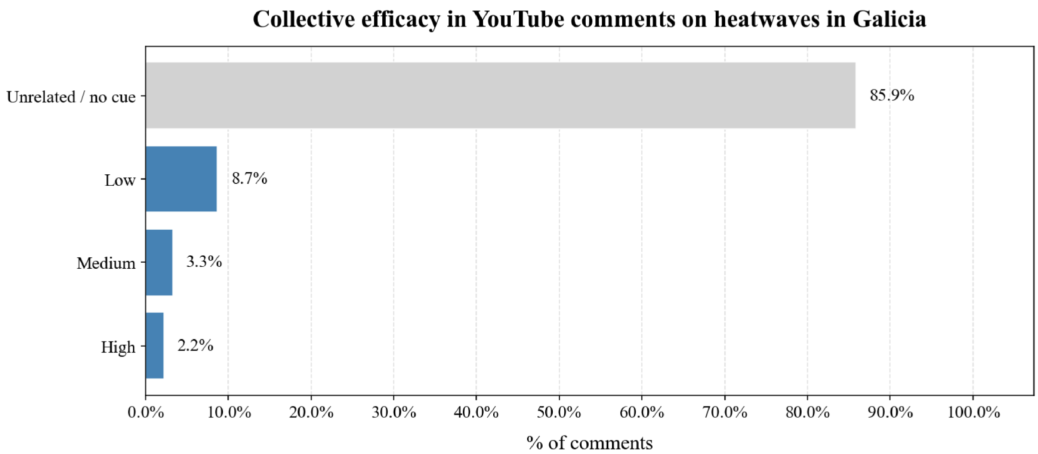

Figure A2. Collective efficacy in YouTube comments on heatwaves in Galicia.

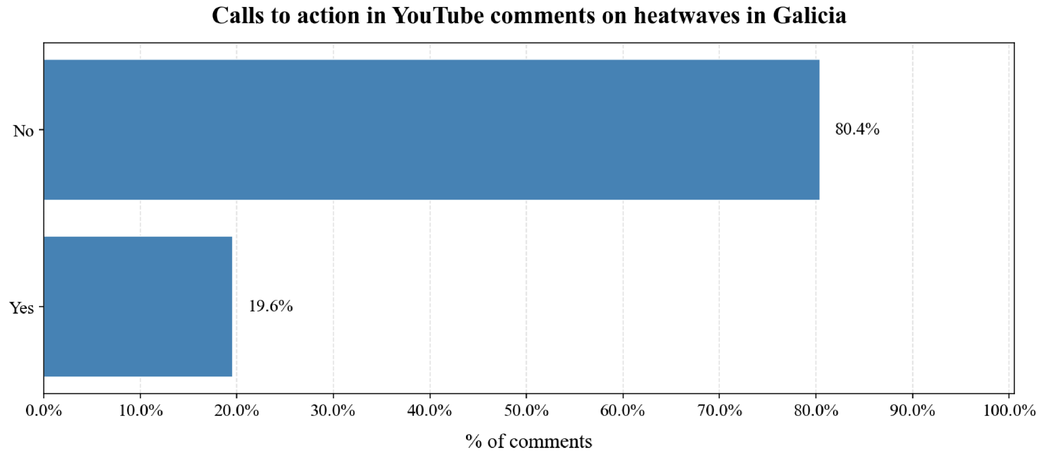

Figure A3. Call to action in YouTube comments on heatwaves in Galicia.

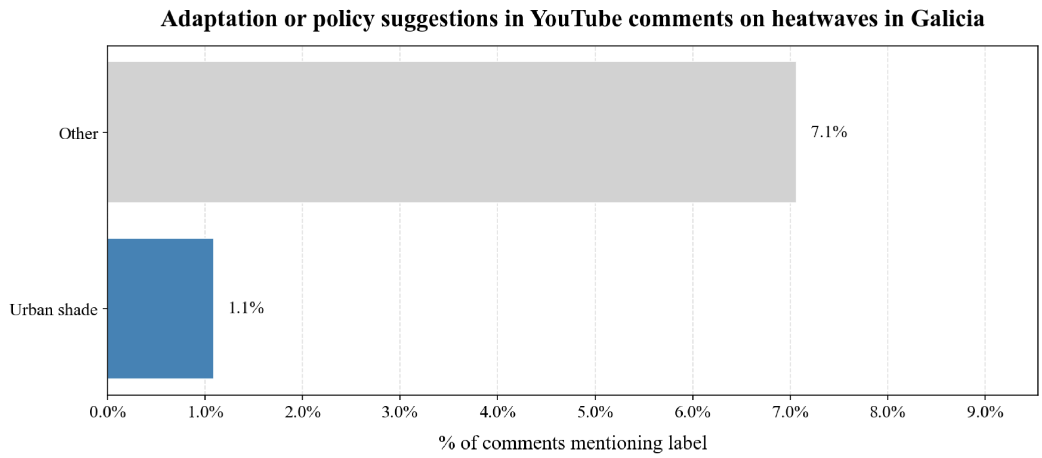

Figure A4. Adaptation and policy suggestions mentioned in YouTube comments on heatwaves in Galicia.

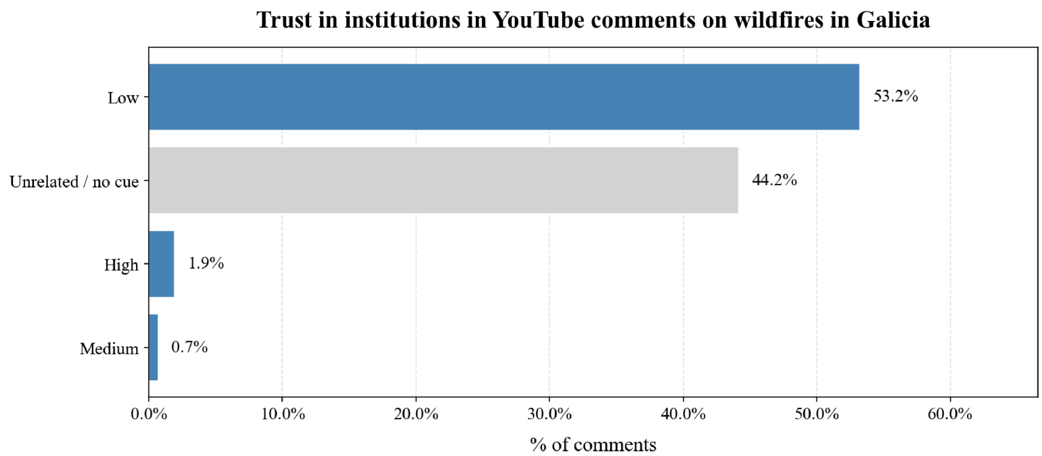

Figure A5. Trust in institutions in YouTube comments on wildfires in Galicia.

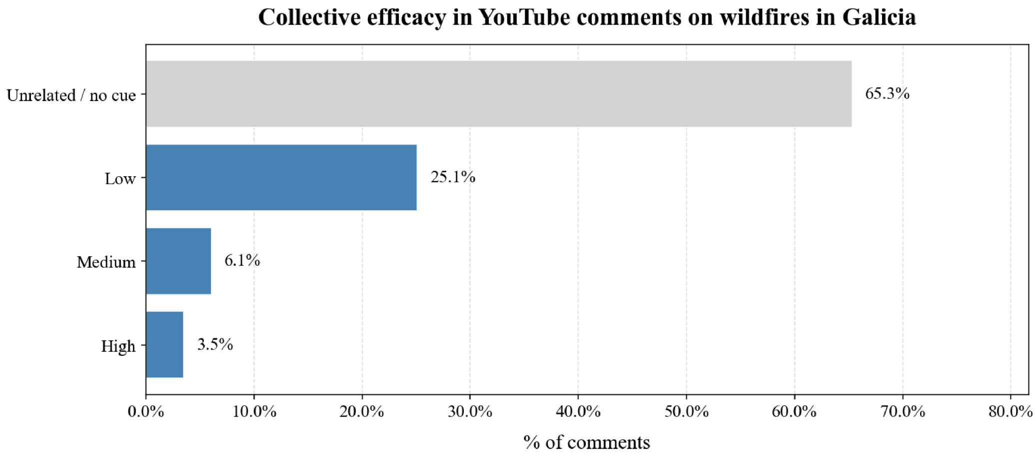

Figure A6. Collective efficacy in YouTube comments on wildfires in Galicia.

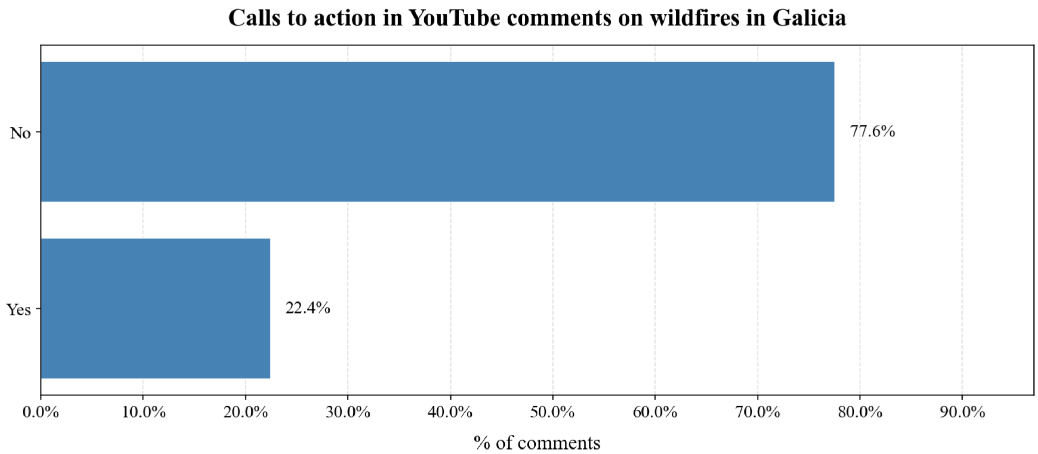

Figure A7. Call to action in YouTube comments on wildfires in Galicia.

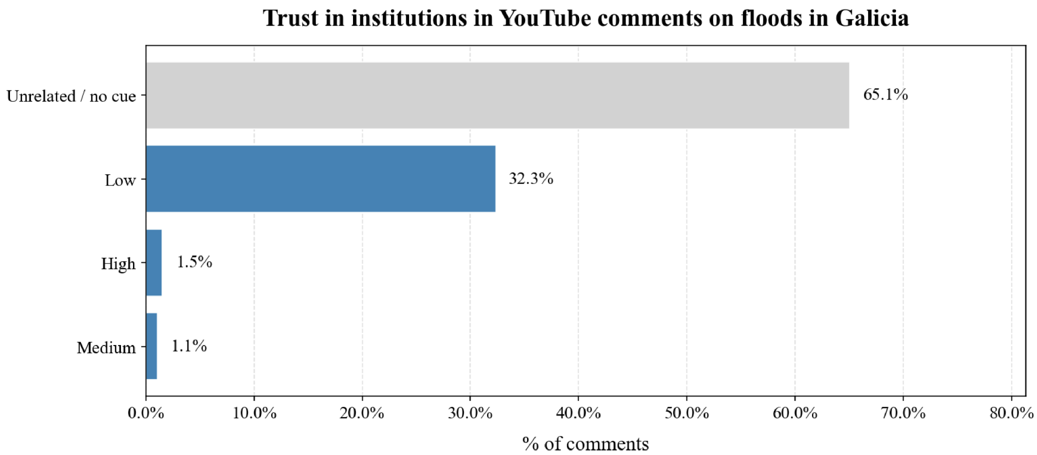

Figure A8. Trust in institutions in YouTube comments on floods in Galicia.

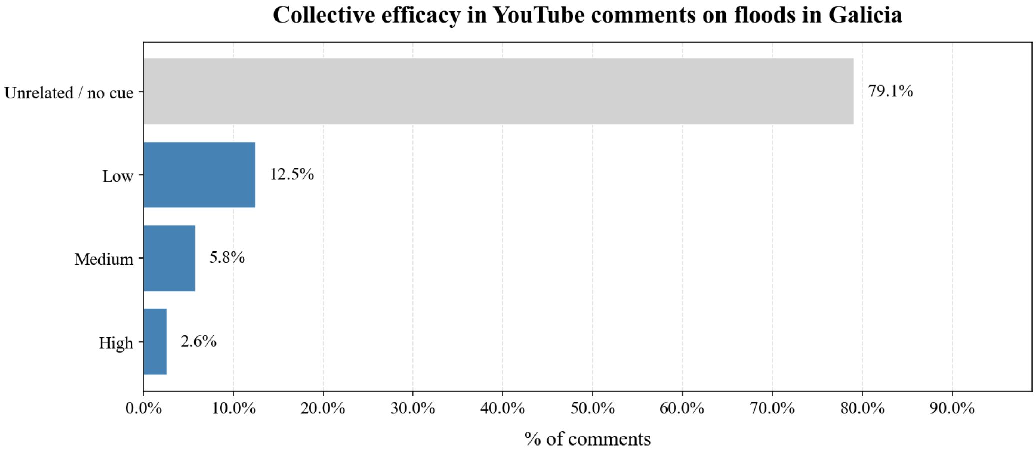

Figure A9. Collective efficacy in YouTube comments on floods in Galicia.

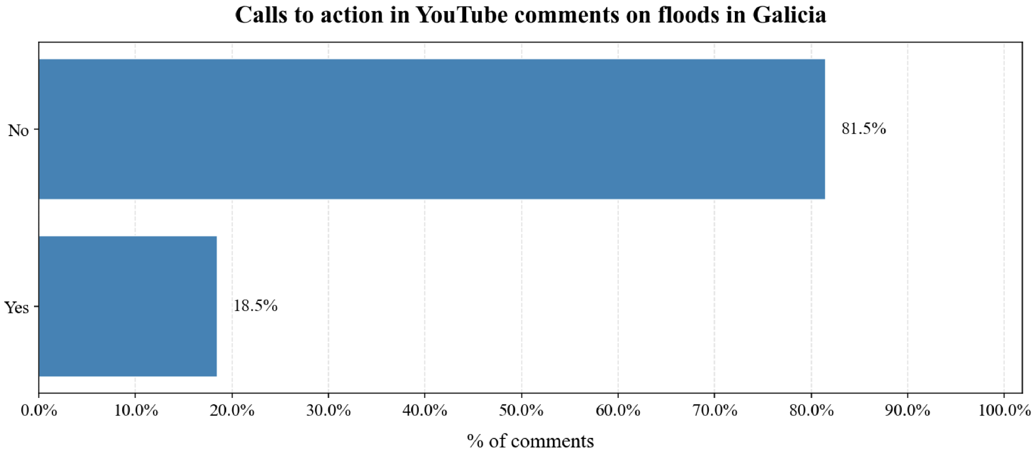

Figure A10. Call to action in YouTube comments on floods in Galicia.

References

1. Palen, L.; Vieweg, S.; Liu, S.B.; Hughes, A.L. Crisis in a networked world: Features of computer-mediated communication in the April 16, 2007, Virginia Tech event. *Soc. Sci. Comput. Rev.* **2009**, *27*, 467–480. [CrossRef]
2. Soden, R.; Palen, L. Informating crisis: Expanding critical perspectives in crisis informatics. *Proc. ACM Hum.-Comput. Interact.* **2018**, *2*, 162. [CrossRef]
3. Erokhin, D.; Komendantova, N. Social media data for disaster risk management and research. *Int. J. Disaster Risk Reduct.* **2024**, *114*, 104980. [CrossRef]
4. Tang, J.; Yang, S.; Wang, W. Social media-based disaster research: Development, trends, and obstacles. *Int. J. Disaster Risk Reduct.* **2021**, *55*, 102095. [CrossRef]
5. Fischer-Preßler, D.; Bonaretti, D.; Bunker, D. Digital transformation in disaster management: A literature review. *J. Strateg. Inf. Syst.* **2024**, *33*, 101865. [CrossRef]
6. Phengsuwan, J.; Shah, T.; Thekkummal, N.B.; Wen, Z.; Sun, R.; Pullarkatt, D.; Thirugnanam, H.; Ramesh, M.V.; Morgan, G.; James, P.; et al. Use of social media data in disaster management: A survey. *Future Internet* **2021**, *13*, 46. [CrossRef]
7. Stieglitz, S.; Bunker, D.; Mirbabaie, M.; Ehnis, C. Sense-making in social media during extreme events. *J. Contingencies Crisis Manag.* **2018**, *26*, 4–15. [CrossRef]
8. Eysenbach, G. Infodemiology and infoveillance: Framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. *J. Med. Internet Res.* **2009**, *11*, e11. [CrossRef]
9. Mavragani, A. Infodemiology and infoveillance: Scoping review. *J. Med. Internet Res.* **2020**, *22*, e16206. [CrossRef] [PubMed]
10. Burke, M.; Heft-Neal, S.; Li, J.; Driscoll, A.; Baylis, P.; Stigler, M.; Weill, J.A.; Burney, J.A.; Wen, J.; Childs, M.L.; et al. Exposures and behavioural responses to wildfire smoke. *Nat. Hum. Behav.* **2022**, *6*, 1351–1361. [CrossRef] [PubMed]

11. Erokhin, D.; Komendantova, N. Analyzing public interest in geohazards using Google Trends data. *Geosciences* **2024**, *14*, 266. [CrossRef]
12. Santín, C.; Moustakas, A.; Doerr, S.H. Searching the flames: Trends in global and regional public interest in wildfires. *Environ. Sci. Policy* **2023**, *146*, 151–161. [CrossRef]
13. Gopal, L.S.; Prabha, R.; Thirugnanam, H.; Ramesh, M.V.; Malamud, B.D. Leveraging social media for managing natural hazard disasters: A critical review of data collection strategies and actionable insights. *EGUsphere* 2024, *preprint*. [CrossRef]
14. Saroj, A.; Pal, S. Use of social media in crisis management: A survey. *Int. J. Disaster Risk Reduct.* **2020**, *48*, 101584. [CrossRef]
15. Nieto, C.E.; Martínez-Graña, A.M.; Encinas, B. Analysis of the risk of coastal flooding due to rising sea levels in Ría of Arosa (Pontevedra, Spain). *Appl. Sci.* **2023**, *13*, 12099. [CrossRef]
16. Toubes, D.R.; Gössling, S.; Hall, C.M.; Scott, D. Vulnerability of coastal beach tourism to flooding: A case study of Galicia, Spain. *Environments* **2017**, *4*, 83. [CrossRef]
17. Chas-Amil, M.L.; Nogueira-Moure, E.; Prestemon, J.P.; Touza, J. Spatial patterns of social vulnerability in relation to wildfire risk and wildland-urban interface presence. *Landsc. Urban Plan.* **2022**, *228*, 104577. [CrossRef]
18. Novo, A.; Fariñas-Álvarez, N.; Martínez-Sánchez, J.; González-Jorge, H.; Fernández-Alonso, J.M.; Lorenzo, H. Mapping forest fire risk—A case study in Galicia (Spain). *Remote Sens.* **2020**, *12*, 3705. [CrossRef]
19. Alvarez, I.; Pereira, H.; Lorenzo, M.N.; Picado, A.; Sousa, M.C.; Taboada, J.J.; Dias, J.M. Drought projections for the NW Iberian Peninsula under climate change. *Clim. Dyn.* **2024**, *62*, 4775–4791. [CrossRef]
20. Salvador, C.; Nieto, R.; Linares, C.; Diaz, J.; Gimeno, L. Effects on daily mortality of droughts in Galicia (NW Spain) from 1983 to 2013. *Sci. Total Environ.* **2019**, *662*, 121–133. [CrossRef] [PubMed]
21. Luna, S.; Pennock, M.J. Social media applications and emergency management: A literature review and research agenda. *Int. J. Disaster Risk Reduct.* **2018**, *28*, 565–577. [CrossRef]
22. Mavragani, A.; Ochoa, G. Google Trends in infodemiology and infoveillance: Methodology framework. *JMIR Public Health Surveill.* **2019**, *5*, e13439. [CrossRef] [PubMed]
23. Basrowi, R.W.; Sundjaya, T.; Pratiwi, D.; Rajab, N.M.; Amanda, R.; Komarudin, H.; Amalia, G. Digital insights into workplace breastfeeding in Indonesia: A Google Trends analysis of barriers and opportunities. *Nutrients* **2025**, *17*, 3433. [CrossRef]
24. Flame, A.C.; Coulson, S.; Lee, J. Analysis of Google Trends data on facial paralysis as a proxy of public interest. *Aust. J. Otolaryngol.* **2025**, *8*, 27. [CrossRef]
25. Deubel, A.; Breuer, J.; Kohne, J.; Mohseni, M.R. *Overview of Working with Data from YouTube*; GESIS Guides to Digital Behavioral Data No. 12; GESIS—Leibniz Institute for the Social Sciences: Mannheim, Germany, 2024. [CrossRef]
26. Deng, Z.; Jiang, H.; Fan, B.; Tang, X. Balancing act: How do local governments use social media in crises? *Int. Public Manag. J.* **2025**, *28*, 425–446. [CrossRef]
27. Mihunov, V.V.; Jafari, N.H.; Wang, K.; Lam, N.S.; Govender, D. Disaster impacts surveillance from social media with topic modeling and feature extraction: Case of Hurricane Harvey. *Int. J. Disaster Risk Sci.* **2022**, *13*, 729–742. [CrossRef]
28. Rudra, K.; Ganguly, N.; Goyal, P.; Ghosh, S. Extracting and summarizing situational information from the Twitter social media during disasters. *ACM Trans. Web* **2018**, *12*, 17. [CrossRef]
29. Zade, H.; Shah, K.; Rangarajan, V.; Kshirsagar, P.; Imran, M.; Starbird, K. From situational awareness to actionability: Towards improving the utility of social media data for crisis response. *Proc. ACM Hum.-Comput. Interact.* **2018**, *2*, 195. [CrossRef]
30. Austin, L.; Fisher Liu, B.; Jin, Y. How audiences seek out crisis information: Exploring the social-mediated crisis communication model. *J. Appl. Commun. Res.* **2012**, *40*, 188–207. [CrossRef]
31. Aguda, T.D.; Siddagangappa, S.; Kochkina, E.; Kaur, S.; Wang, D.; Smiley, C. Large language models as financial data annotators: A study on effectiveness and efficiency. In Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024), Torino, Italy, 20–25 May 2024; pp. 10124–10145. Available online: https://aclanthology.org/2024.lrec-main.885/ (accessed on 20 May 2026).
32. Bhat, S.; Varma, V. Large language models as annotators: A preliminary evaluation for annotating low-resource language content. In Proceedings of the 4th Workshop on Evaluation and Comparison of NLP Systems, Bali, Indonesia, 1–2 November 2023; pp. 100–107. [CrossRef]
33. Gilardi, F.; Alizadeh, M.; Kubli, M. ChatGPT outperforms crowd workers for text-annotation tasks. *Proc. Natl. Acad. Sci. USA* **2023**, *120*, e2305016120. [CrossRef]
34. Palen, L.; Anderson, K.M. Crisis informatics—New data for extraordinary times. *Science* **2016**, *353*, 224–225. [CrossRef]
35. Ballester, J.; Quijal-Zamorano, M.; Méndez Turrubiates, R.F.; Pegenaute, F.; Herrmann, F.R.; Robine, J.M.; Basagaña, X.; Tonne, C.; Antó, J.M.; Achebak, H. Heat-related mortality in Europe during the summer of 2022. *Nat. Med.* **2023**, *29*, 1857–1866. [CrossRef] [PubMed]
36. Lindell, M.K.; Perry, R.W. The protective action decision model: Theoretical modifications and additional evidence. *Risk Anal. Int. J.* **2012**, *32*, 616–632. [CrossRef]

37. World Health Organization. *Heatwaves and Health: Guidance on Warning-System Development*; Technical Document; World Health Organization: Geneva, Switzerland, 2016. Available online: https://www.who.int/publications/m/item/heatwaves-and-health-guidance-on-warning-system-development (accessed on 20 May 2026).
38. World Health Organization. *Heat and Health*; World Health Organization: Geneva, Switzerland, 2024. Available online: https://www.who.int/news-room/fact-sheets/detail/climate-change-heat-and-health (accessed on 20 May 2026).
39. Imran, M.; Castillo, C.; Diaz, F.; Vieweg, S. Processing social media messages in mass emergency: A survey. *ACM Comput. Surv.* **2015**, *47*, 67. [CrossRef] [PubMed]
40. Abatzoglou, J.T.; Kolden, C.A.; Cullen, A.C.; Sadegh, M.; Williams, E.L.; Turco, M.; Jones, M.W. Climate change has increased the odds of extreme regional forest fire years globally. *Nat. Commun.* **2025**, *16*, 6390. [CrossRef] [PubMed]
41. Ruffault, J.; Curt, T.; Moron, V.; Trigo, R.M.; Mouillot, F.; Koutsias, N.; Pimont, F.; Martin-StPaul, N.; Barbero, R.; Dupuy, J.-L.; et al. Increased likelihood of heat-induced large wildfires in the Mediterranean Basin. *Sci. Rep.* **2020**, *10*, 13790. [CrossRef]
42. Intergovernmental Panel on Climate Change. *Summary for Policymakers*; Intergovernmental Panel on Climate Change: Geneva, Switzerland, 2022. Available online: https://www.ipcc.ch/report/ar6/wg2/chapter/summary-for-policymakers/ (accessed on 20 May 2026).
43. Cebrián, E.; Domenech, J. Addressing Google Trends inconsistencies. *Technol. Forecast. Soc. Change* **2024**, *202*, 123318. [CrossRef]
44. Gummer, T.; Oehrlein, A.S. Using Google Trends data to study high-frequency search terms: Evidence for a reliability-frequency continuum. *Soc. Sci. Comput. Rev.* **2025**, *43*, 814–826. [CrossRef]
45. Hölzl, J.; Keusch, F.; Sajons, C. The (mis) use of Google Trends data in the social sciences-A systematic review, critique, and recommendations. *Soc. Sci. Res.* **2025**, *126*, 103099. [CrossRef]
46. Liu, K. The measurement errors of Google Trends data. *Discov. Data* **2024**, *2*, 7. [CrossRef]
47. Efstratiou, A. On YouTube search API use in research. In Proceedings of the 2025 ACM Internet Measurement Conference, Madison, WI, USA, 28–31 October 2025; pp. 919–927. [CrossRef]
48. Rieder, B.; Padilla, A.; Coromina, Ò. Forgetful by design? A critical audit of YouTube's search API for academic research. *Inf. Commun. Soc.* **2025**, 1–20. [CrossRef]
49. Cakmak, M.C.; Agarwal, N.; Oni, R. The bias beneath: Analyzing drift in YouTube's algorithmic recommendations. *Soc. Netw. Anal. Min.* **2024**, *14*, 171. [CrossRef]
50. Dargin, J.S.; Fan, C.; Mostafavi, A. Vulnerable populations and social media use in disasters: Uncovering the digital divide in three major US hurricanes. *Int. J. Disaster Risk Reduct.* **2021**, *54*, 102043. [CrossRef]
51. Fan, C.; Esparza, M.; Dargin, J.; Wu, F.; Oztekin, B.; Mostafavi, A. Spatial biases in crowdsourced data: Social media content attention concentrates on populous areas in disasters. *Comput. Environ. Urban Syst.* **2020**, *83*, 101514. [CrossRef]
52. Wiegmann, M.; Kersten, J.; Senaratne, H.; Potthast, M.; Klan, F.; Stein, B. Opportunities and risks of disaster data from social media: A systematic review of incident information. *Nat. Hazards Earth Syst. Sci. Discuss.* **2020**, *21*, 1431–1444. [CrossRef]
53. Ouyang, T.; MaungMaung, A.; Konishi, K.; Seo, Y.; Echizen, I. Stability analysis of chatgpt-based sentiment analysis in ai quality assurance. *Electronics* **2024**, *13*, 5043. [CrossRef]
54. Paullada, A.; Raji, I.D.; Bender, E.M.; Denton, E.; Hanna, A. Data and its (dis) contents: A survey of dataset development and use in machine learning research. *Patterns* **2021**, *2*, 100336. [CrossRef]
55. Törnberg, P. Best practices for text annotation with large language models. *arXiv* **2024**, arXiv:2402.05129. [CrossRef]
56. Dcruz, J.G.; Zolotas, A.; Greenwood, N.R.; Arana-Catania, M. Structured AI decision-making in disaster management. *Sci. Rep.* **2025**, *15*, 32093. [CrossRef]
57. Sculley, D.; Holt, G.; Golovin, D.; Davydov, E.; Phillips, T.; Ebner, D.; Chaudhary, V.; Young, M.; Crespo, J.-F.; Dennison, D. Hidden technical debt in machine learning systems. In *Advances in Neural Information Processing Systems 28*; Curran Associates, Inc.: Red Hook, NY, USA, 2015. Available online: https://proceedings.neurips.cc/paper_files/paper/2015/hash/86df7dcfd896fcaf2674f757a2463eba-Abstract.html (accessed on 20 May 2026).
58. World Meteorological Organization. *Common Alerting Protocol*; World Meteorological Organization: Geneva, Switzerland, 2026. Available online: https://wmo.int/common-alerting-protocol (accessed on 20 May 2026).
59. Beatson, A. *Social Media Monitoring—An Assessment for Emergency Management: 2015*; Opus International Consultants Ltd.: Wellington, New Zealand, 2015. Available online: https://www.civildefence.govt.nz/assets/Uploads/documents/resilience-fund/2014-15/06/2014-06-Public-Alerting-and-social-Media-monitoring.pdf (accessed on 20 May 2026).
60. United Nations Office for Disaster Risk Reduction. *Global Status of Multi-Hazard Early Warning Systems 2024*; United Nations Office for Disaster Risk Reduction: Geneva, Switzerland, 2024. Available online: https://www.undrr.org/reports/global-status-MHEWS-2024 (accessed on 20 May 2026).
61. Mitchell, M.; Wu, S.; Zaldivar, A.; Barnes, P.; Vasserman, L.; Hutchinson, B.; Spitzer, E.; Raji, I.D.; Gebru, T. Model cards for model reporting. In Proceedings of the Conference on Fairness, Accountability, and Transparency, Atlanta, GA, USA, 29–31 January 2019; pp. 220–229. [CrossRef]

62. National Institute of Standards and Technology. *NIST AI RMF Playbook*; National Institute of Standards and Technology: Gaithersburg, MD, USA, 2025. Available online: https://www.nist.gov/itl/ai-risk-management-framework/nist-ai-rmf-playbook (accessed on 20 May 2026).
63. Covington, P.; Adams, J.; Sargin, E. Deep neural networks for youtube recommendations. In Proceedings of the 10th ACM Conference on Recommender Systems, Boston, MA, USA, 15–19 September 2016; pp. 191–198. [CrossRef]