Real-Time AI Monitoring of Online Discourse for Crisis Communication

Dmitry Erokhin, PhD

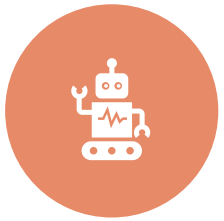
International Institute for Applied Systems Analysis

Laxenburg, Austria

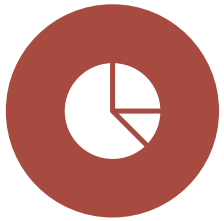
Our Journey Today

THE NEW REALITY: THE CRISIS COMMUNICATION CHALLENGE

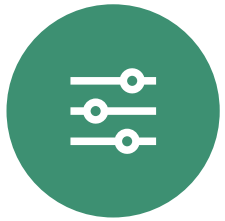
THE PARADIGM SHIFT: FROM REACTIVE TO PROACTIVE WITH AI



HOW IT WORKS: A GLIMPSE INSIDE THE AI ENGINE



THE DASHBOARD: VISUALIZ ING PUBLIC PERCEPTION



PRACTICAL APPLICATIONS

THE PATH TO IMPLEMENTATION: A STRATEGIC FRAMEWORK

NAVIGATING THE CHALLENGES: ETHICS & LIMITATIONS

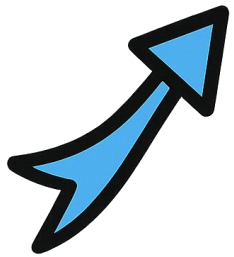
CONCLUSION & THE FUTURE

The Velocity and Volume of Modern Crises

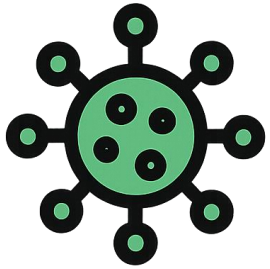
- Crises unfold at digital speed on social media, news sites, and video platforms.
- Public sentiment can solidify before an official statement is even drafted.
- **The central problem:** Information overload + misinformation proliferation = Strategic paralysis.
- Traditional methods of media monitoring are too slow and lack scale.

The Consequences of Lagging Behind the Narrative

Trust



Resources



Misinformation

- **Erosion of trust:** The public loses confidence in institutions that appear silent or out-of-touch.
- **Ineffective resource allocation:** Help is sent to the wrong places. Messages miss their mark.
- **Amplification of harm:** Misinformation fills the vacuum, causing panic and dangerous behavior.

From Listening to Understanding, in Real-Time

- **The goal:** Move from a *reactive* posture to an *adaptive* and *proactive* one.
- Artificial Intelligence acts as a force multiplier for communication teams.
- It continuously processes millions of data points, translating online chatter into structured, intelligible insights.
- This enables organizations to *anticipate* public anxiety and *tailor* their response dynamically.

The Data-to-Insight Pipeline

Data ingestion: Continuous streaming from X, YouTube, Reddit, News Feeds, etc.

NLP processing: The “AI Brain” applies multiple techniques simultaneously.

Insight generation: Actionable metrics are extracted.

Visualization & alerting: Insights are delivered to decision-makers.

The Technical Engine: Natural Language Processing (NLP)

Topic modeling & trend detection

What: Identifies emerging themes and maps their prevalence over time.

Output: Surge in mentions of power outage.

Sentiment & emotion analysis

What: Gauges public mood (positive, negative, neutral) and specific emotions (fear, anger, urgency).

Output: Sharp rise in fearful comments regarding aftershocks.

Misinformation & narrative detection

What: Flags unverified claims and tracks the spread of dominant, often harmful, stories.

Output: Rumor detected: The government is hiding the true death toll.

Visualizing the Pulse Public Discourse

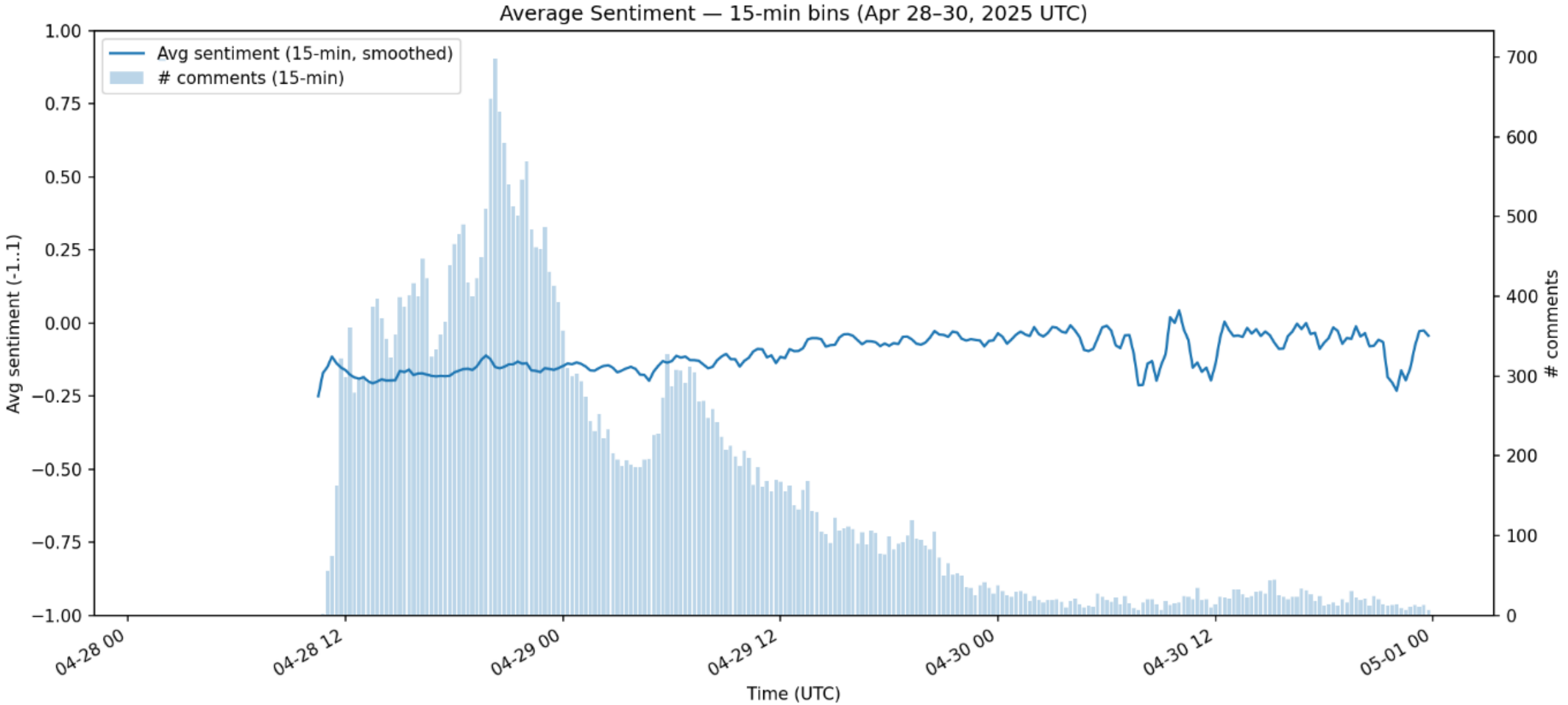
- A centralized view for crisis teams.
- **Key components:**
  - Real-time alert feed for critical events.
  - Sentiment tracker over time.
  - Geomap of discussion hotspots.
  - Top trending topics and hashtags.
  - Misinformation spike alerts.

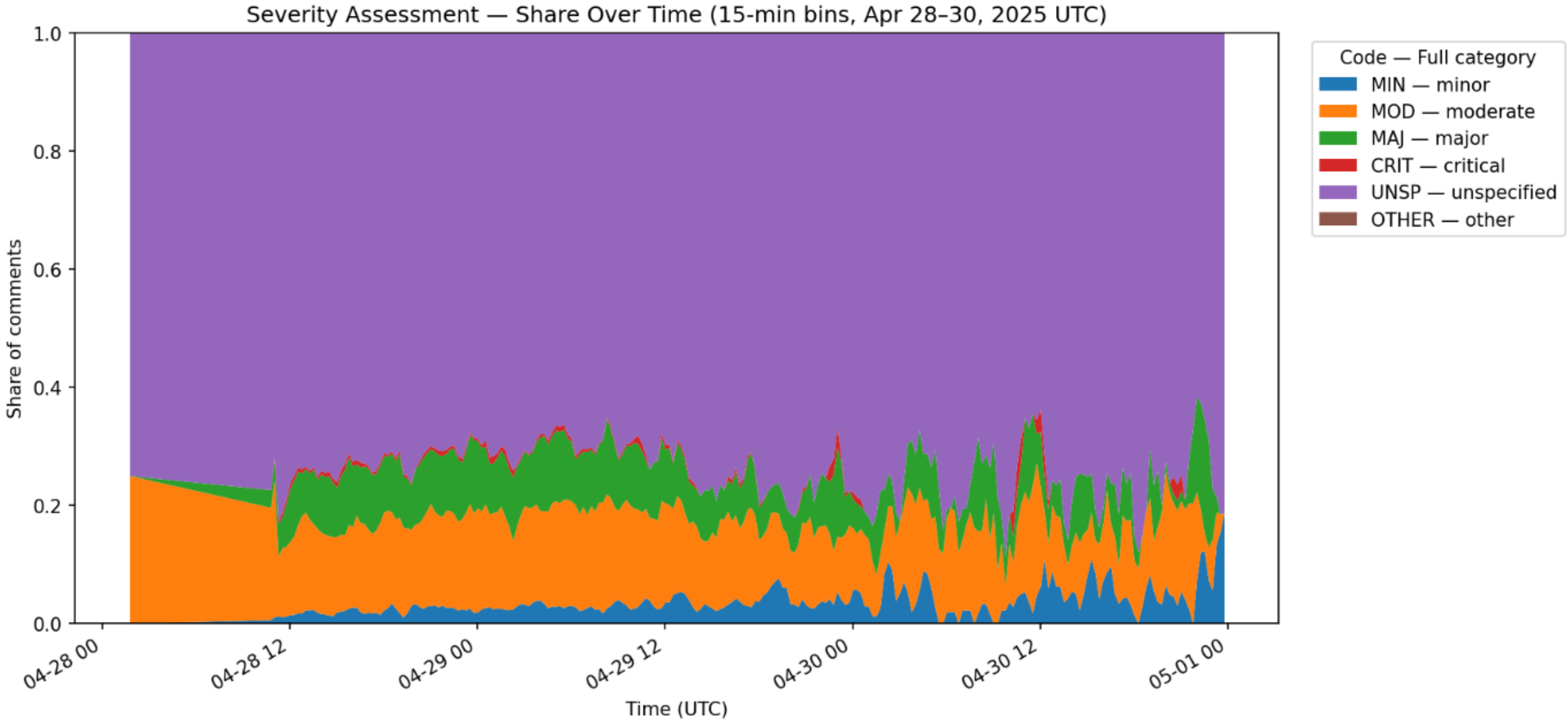
Illuminating Crises with Data: Case Study Preview

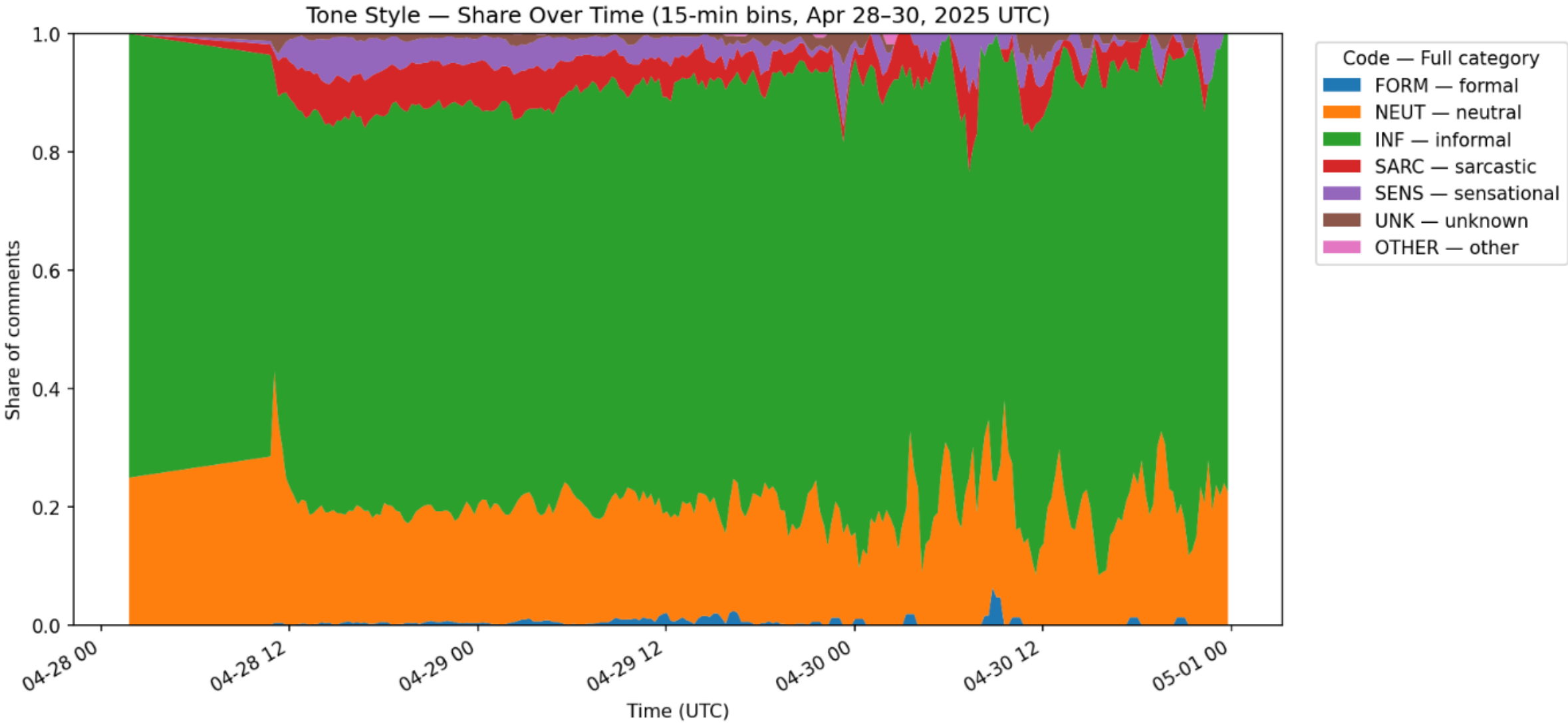
- This section transitions from theory to practice.
- We will examine real-world data from:
  - **The 2023 Turkey-Syria earthquake** (X/Twitter analysis)
  - **The 2025 Iberian blackout** (YouTube comment monitoring)
  - **Nuclear discourse** (Narrative & misinformation tracking)
  - **Peace discourse** (Asian peace perspectives)

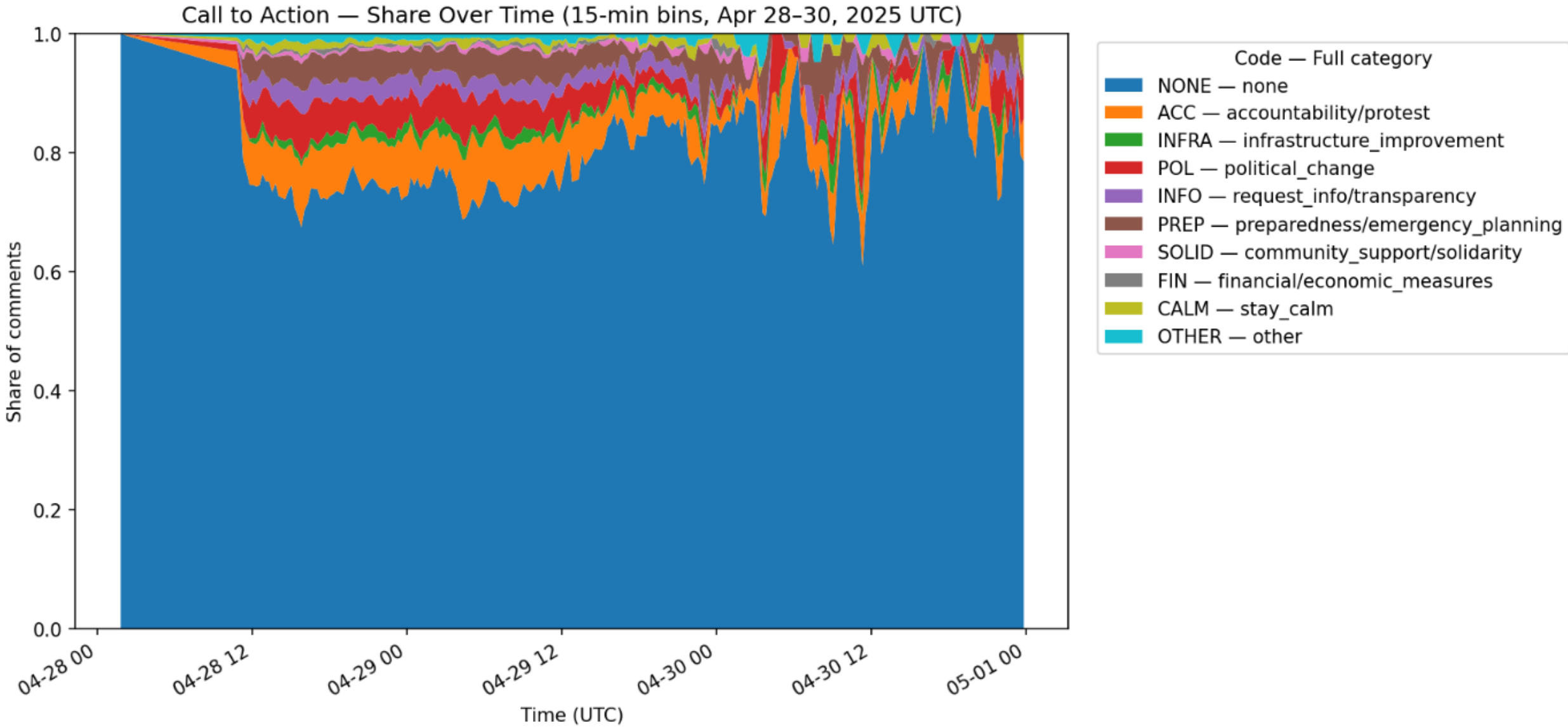
🚨 The April 2025 Iberian Peninsula Blackout

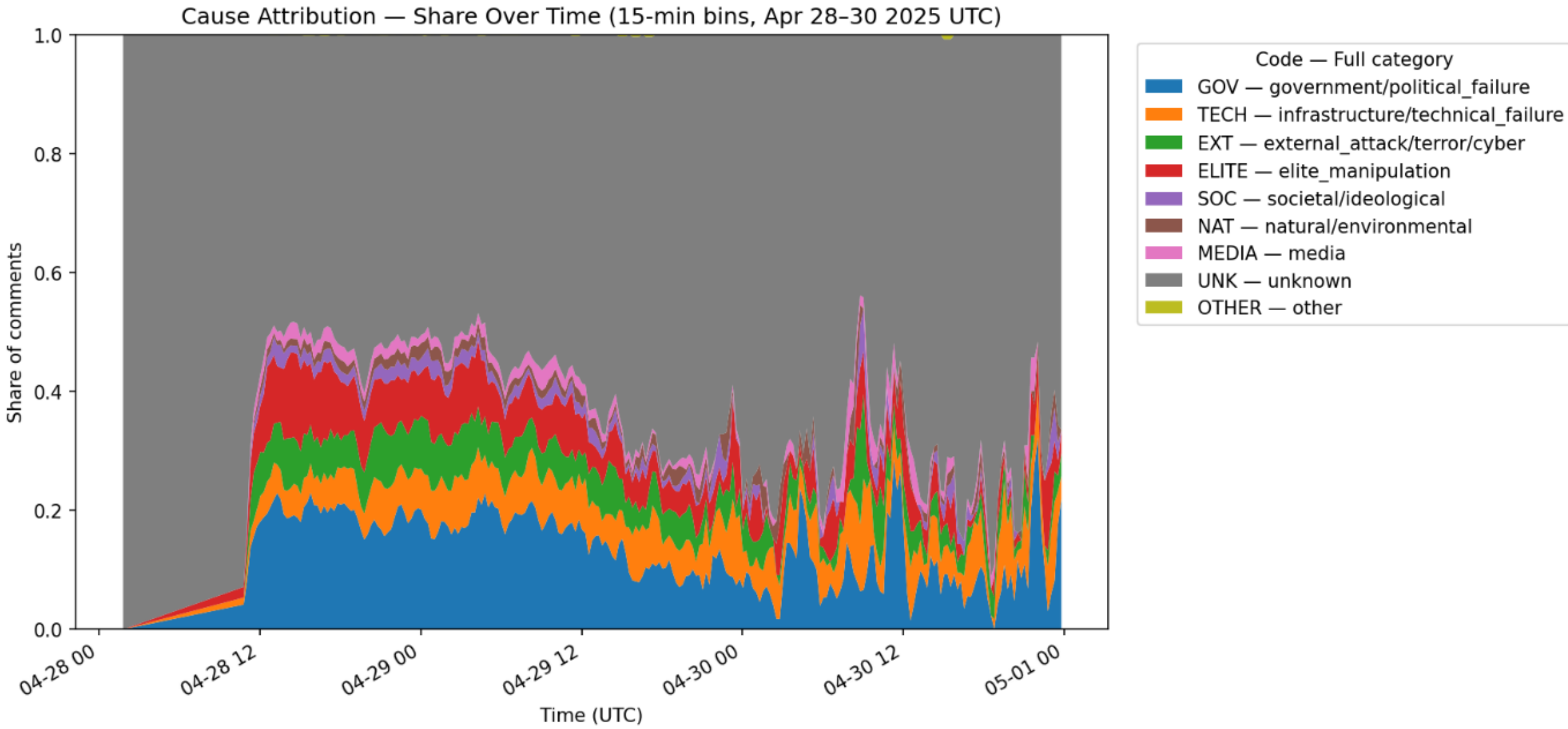
- 🗓️ **Date and time:** 28 April 2025, beginning at 12:33 CEST .
- 📍 **Area affected:** A total blackout across mainland Spain and Portugal, with minor, short-lived disruptions in Andorra and southwestern France .
- 👥 **Scale:** An unprecedented event affecting around 60 million people, described as the most severe power system event in Europe in over 20 years and the first of its kind.
- **Analysis period:** 28 – 30 April 2025
- **Data source:** All relevant YouTube videos covering the “apagón” and their associated comments.

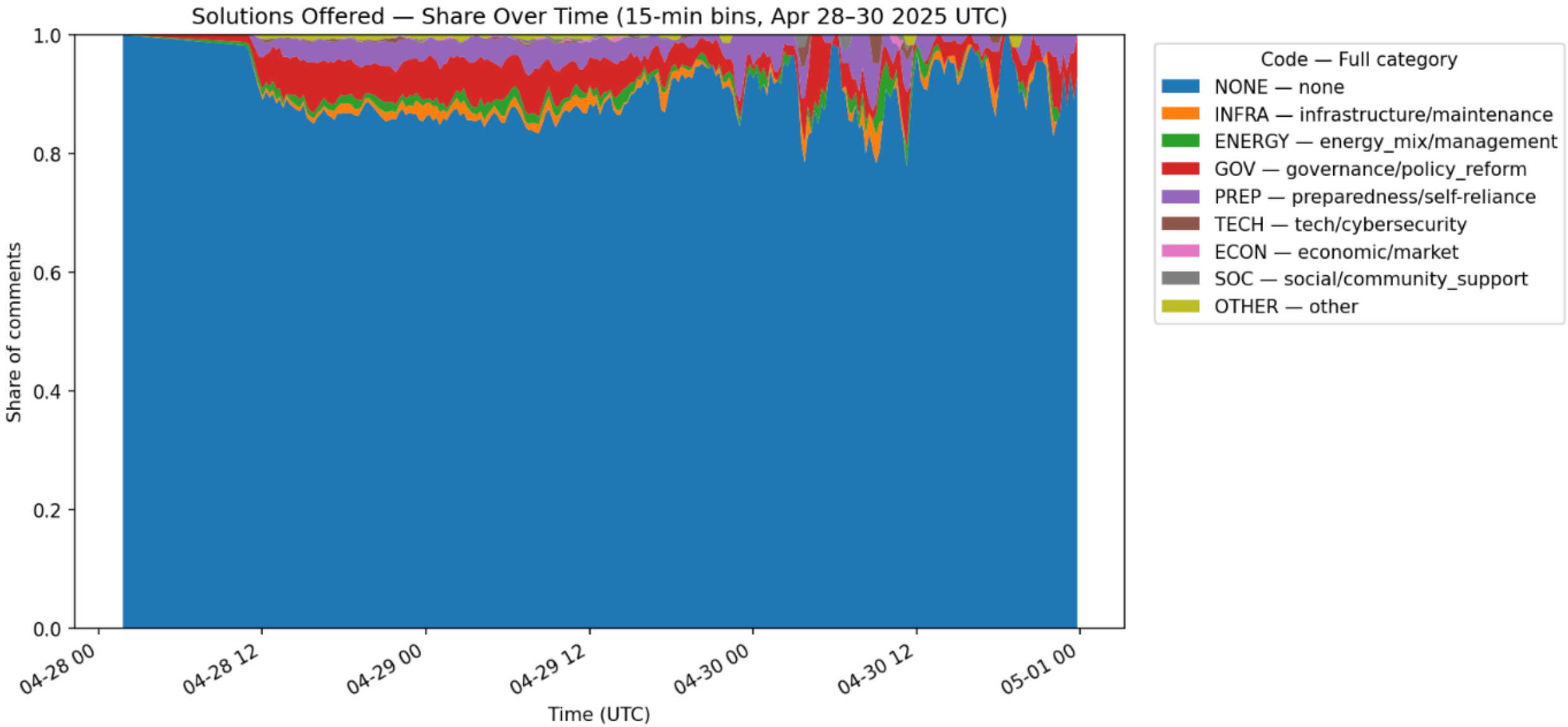

Average Sentiment — 15-min bins (Apr 28–30, 2025 UTC)
Avg sentiment (15-min, smoothed)
comments (15-min)
Avg sentiment (-1..1)
comments
Time (UTC)
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
700
600
500
400
300
200
100
0
04-28 00
04-28 12
04-29 00
04-29 12
04-30 00
04-30 12
05-01 00

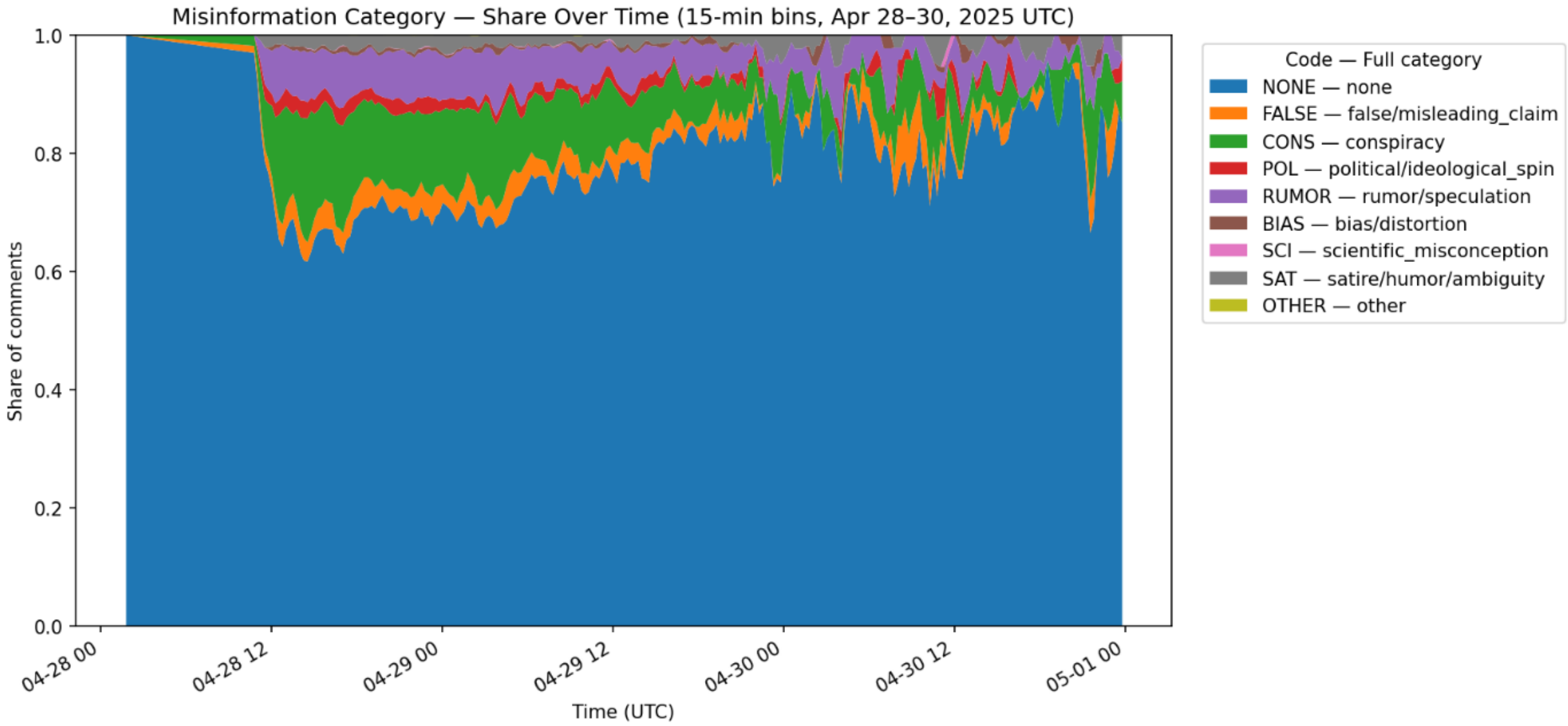

Severity Assessment — Share Over Time (15-min bins, Apr 28–30, 2025 UTC)
Share of comments
1.0
0.8
0.6
0.4
0.2
0.0
04-28 00
04-28 12
04-29 00
04-29 12
04-30 00
04-30 12
05-01 00
Time (UTC)
Code — Full category
MIN — minor
MOD — moderate
MAJ — major
CRIT — critical
UNSP — unspecified
OTHER — other


Tone Style — Share Over Time (15-min bins, Apr 28–30, 2025 UTC)
Share of comments
1.0
0.8
0.6
0.4
0.2
0.0
04-28 00
04-28 12
04-29 00
04-29 12
04-30 00
04-30 12
05-01 00
Time (UTC)
Code — Full category
FORM — formal
NEUT — neutral
INF — informal
SARC — sarcastic
SENS — sensational
UNK — unknown
OTHER — other


Call to Action — Share Over Time (15-min bins, Apr 28–30, 2025 UTC)
Share of comments
1.0
0.8
0.6
0.4
0.2
0.0
04-28 00
04-28 12
04-29 00
04-29 12
04-30 00
04-30 12
05-01 00
Time (UTC)
Code — Full category
NONE — none
ACC — accountability/protest
INFRA — infrastructure_improvement
POL — political_change
INFO — request_info/transparency
PREP — preparedness/emergency_planning
SOLID — community_support/solidarity
FIN — financial/economic_measures
CALM — stay_calm
OTHER — other


Cause Attribution — Share Over Time (15-min bins, Apr 28–30 2025 UTC)
Share of comments
1.0
0.8
0.6
0.4
0.2
0.0
04-28 00
04-28 12
04-29 00
04-29 12
04-30 00
04-30 12
05-01 00
Time (UTC)
Code — Full category
GOV — government/political_failure
TECH — infrastructure/technical_failure
EXT — external_attack/terror/cyber
ELITE — elite_manipulation
SOC — societal/ideological
NAT — natural/environmental
MEDIA — media
UNK — unknown
OTHER — other


Solutions Offered — Share Over Time (15-min bins, Apr 28–30 2025 UTC)
Share of comments
1.0
0.8
0.6
0.4
0.2
0.0
04-28 00
04-28 12
04-29 00
04-29 12
04-30 00
04-30 12
05-01 00
Time (UTC)
Code — Full category
NONE — none
INFRA — infrastructure/maintenance
ENERGY — energy_mix/management
GOV — governance/policy_reform
PREP — preparedness/self-reliance
TECH — tech/cybersecurity
ECON — economic/market
SOC — social/community_support
OTHER — other


Misinformation Category — Share Over Time (15-min bins, Apr 28–30, 2025 UTC)
Share of comments
1.0
0.8
0.6
0.4
0.2
0.0
04-28 00
04-28 12
04-29 00
04-29 12
04-30 00
04-30 12
05-01 00
Time (UTC)
Code — Full category
NONE — none
FALSE — false/misleading_claim
CONS — conspiracy
POL — political/ideological_spin
RUMOR — rumor/speculation
BIAS — bias/distortion
SCI — scientific_misconception
SAT — satire/humor/ambiguity
OTHER — other

Conspiracies and Debunking

🧩 **CONSPIRACY 1: “The blackout was a social experiment or planned by governments / elites (EU, Davos, UN, Agenda 2030, New World Order) to control the population”**

FACT-CHECK / DEBUNKING: There is **no evidence** that any global or European plan exists to deliberately cause blackouts as tests of social control. Power grids are complex, interconnected systems that can fail due to overloads, extreme weather, or technical errors — not coordinated political agendas. The **UN’s Agenda 2030** is a list of *voluntary sustainable development goals* (like reducing poverty or improving education). It does **not** include anything about blackouts or population control.

🧩 **CONSPIRACY 2: “It was a false flag or self-inflicted attack by the EU / Spain / Brussels to justify military spending or distract from corruption”**

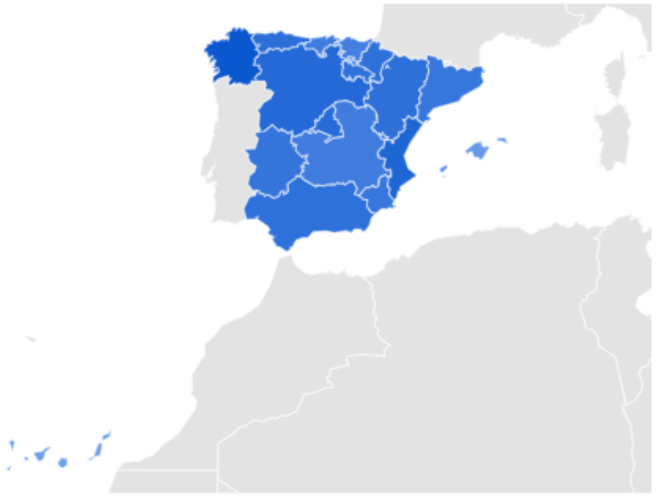
FACT-CHECK / DEBUNKING: False flag operations are *extremely rare* and nearly impossible to organize secretly across multiple governments and agencies. Blackout investigations are handled by **independent technical operators** (e.g., Red Eléctrica, ENTSO-E) whose data are auditable and transparent. No credible link exists between power outages and national budget decisions, military contracts, or corruption scandals.

🧩 **CONSPIRACY 3: “It was a cyberattack by Israel / Morocco / Russia / the US / the CIA / Mossad / etc.”**

FACT-CHECK / DEBUNKING: Attributing a cyberattack requires **digital forensic evidence** (logs, signatures, IP traces, malware analysis). No official agency has confirmed any foreign cyberattack related to this event. Most regional blackouts are caused by **grid synchronization failures or automatic safety shutdowns**, not by hacking or external interference.

Looking at Google Trends

Interest by sub-region



Sub-region

1	Galicia	100
2	Valencian Community	87
3	Castile and León	84
4	Community of Madrid	82
5	Asturias	79

Showing 1–5 of 19 subregions

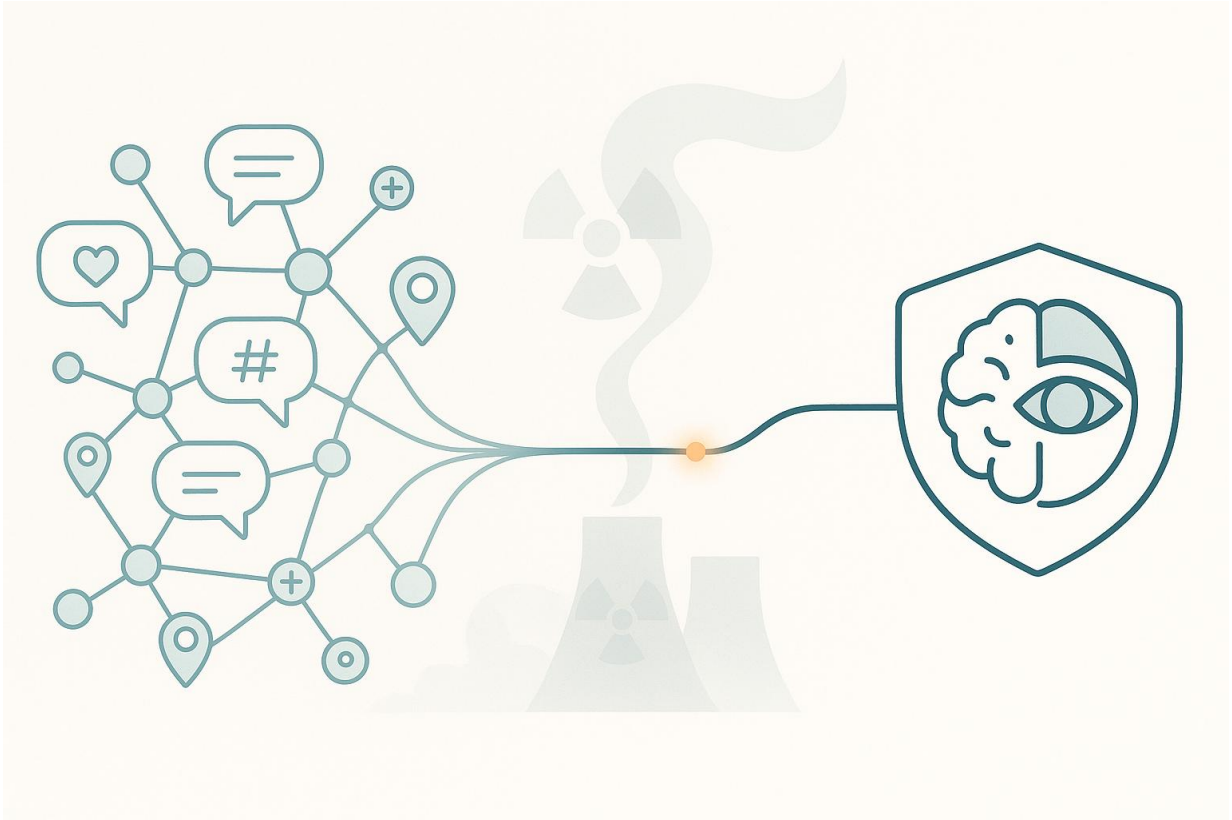
search term	type	trend/score
españa apagon	Rising	Breakout
apagon luz	Rising	Breakout
apagon hoy	Rising	Breakout
apagon de luz	Rising	Breakout
apagon en españa	Rising	Breakout
apagón españa	Rising	Breakout
apagón hoy	Rising	Breakout
apagón luz	Rising	Breakout
apagon madrid	Rising	Breakout
apagon luz españa	Rising	Breakout
apagon europa	Rising	Breakout
apagón de luz	Rising	Breakout
apagón luz hoy	Rising	Breakout
apagon luz hoy	Rising	Breakout
apagon ultima hora	Rising	Breakout
power outage	Rising	Breakout
apagon general	Rising	Breakout
corte de luz españa	Rising	Breakout
apagon electrico	Rising	Breakout
apagon de luz españa	Rising	Breakout
apagón eléctrico	Rising	Breakout
apagón de luz hoy	Rising	Breakout
apagón en españa	Rising	Breakout
apagón españa hoy	Rising	Breakout
barcelona apagon	Rising	Breakout

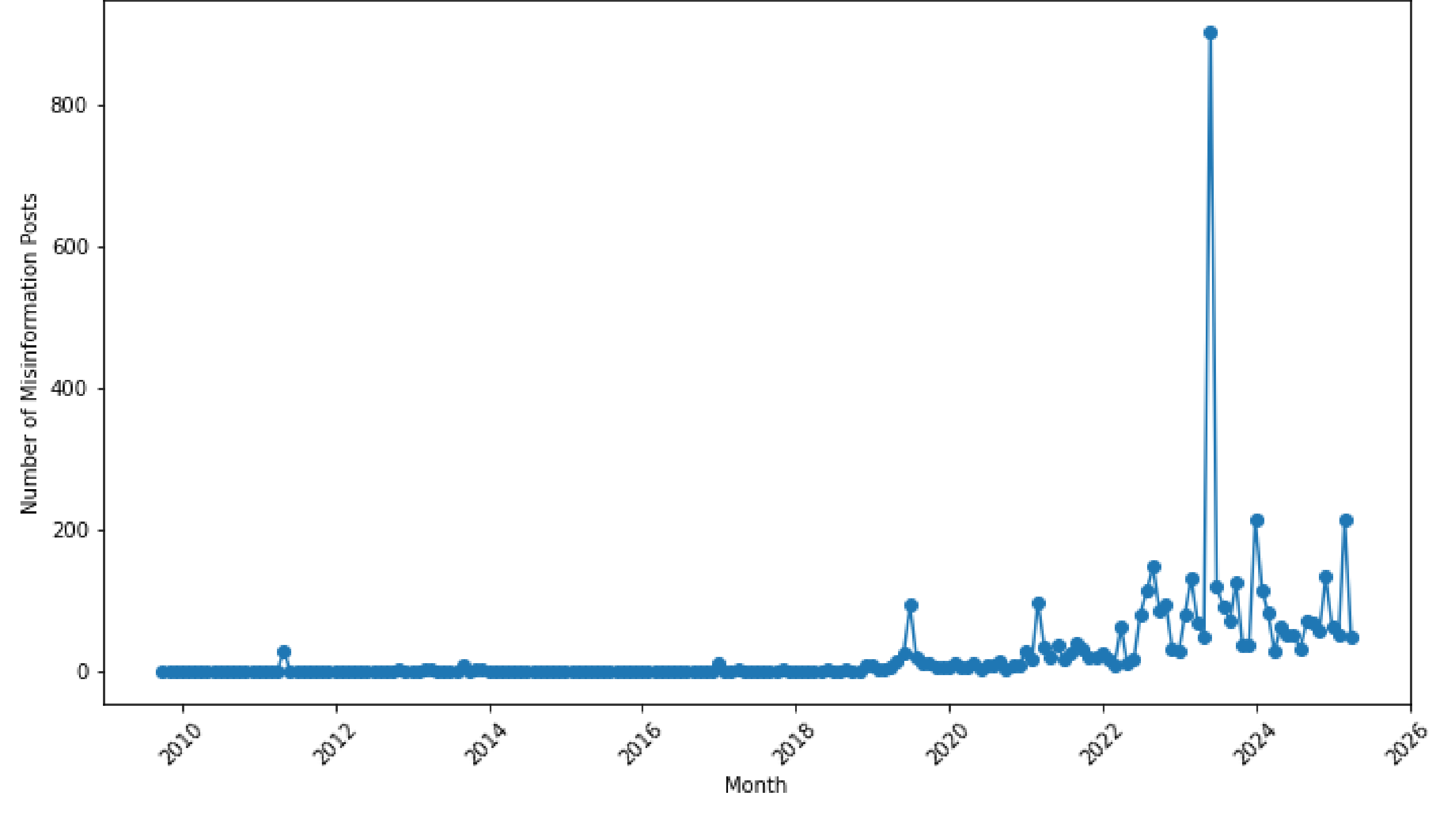
Looking at Google Trends

search term	type	trend/score
españa apagon	Top	100
apagon luz	Top	81
corte de luz	Top	54
apagon hoy	Top	33
apagon de luz	Top	32
apagon en españa	Top	24
apagón españa	Top	23
apagón hoy	Top	23
apagón luz	Top	22
apagon madrid	Top	21
apagon luz españa	Top	20
apagon noticias	Top	19
apagon europa	Top	17
apagón de luz	Top	15
apagón luz hoy	Top	11
apagon luz hoy	Top	10
apagon ultima hora	Top	10
power outage	Top	10
apagon hoy españa	Top	10
corte de luz hoy	Top	9
apagon general	Top	9
corte de luz españa	Top	9
apagon electrico	Top	9
apagon de luz españa	Top	9
apagón eléctrico	Top	9

Nuclear Accident Discourse

- **Data source:** 448 most relevant YouTube videos retrieved by search algorithm for the term **nuclear accident**.
- **Comment volume: 239158** user comments analyzed.
- **Analysis date:** 4 March 2025.




Number of Misinformation Posts
0
200
400
600
800
2010
2012
2014
2016
2018
2020
2022
2024
2026
Month

YouTube AT

Search

Did a Nuclear Accident Just Go Viral?

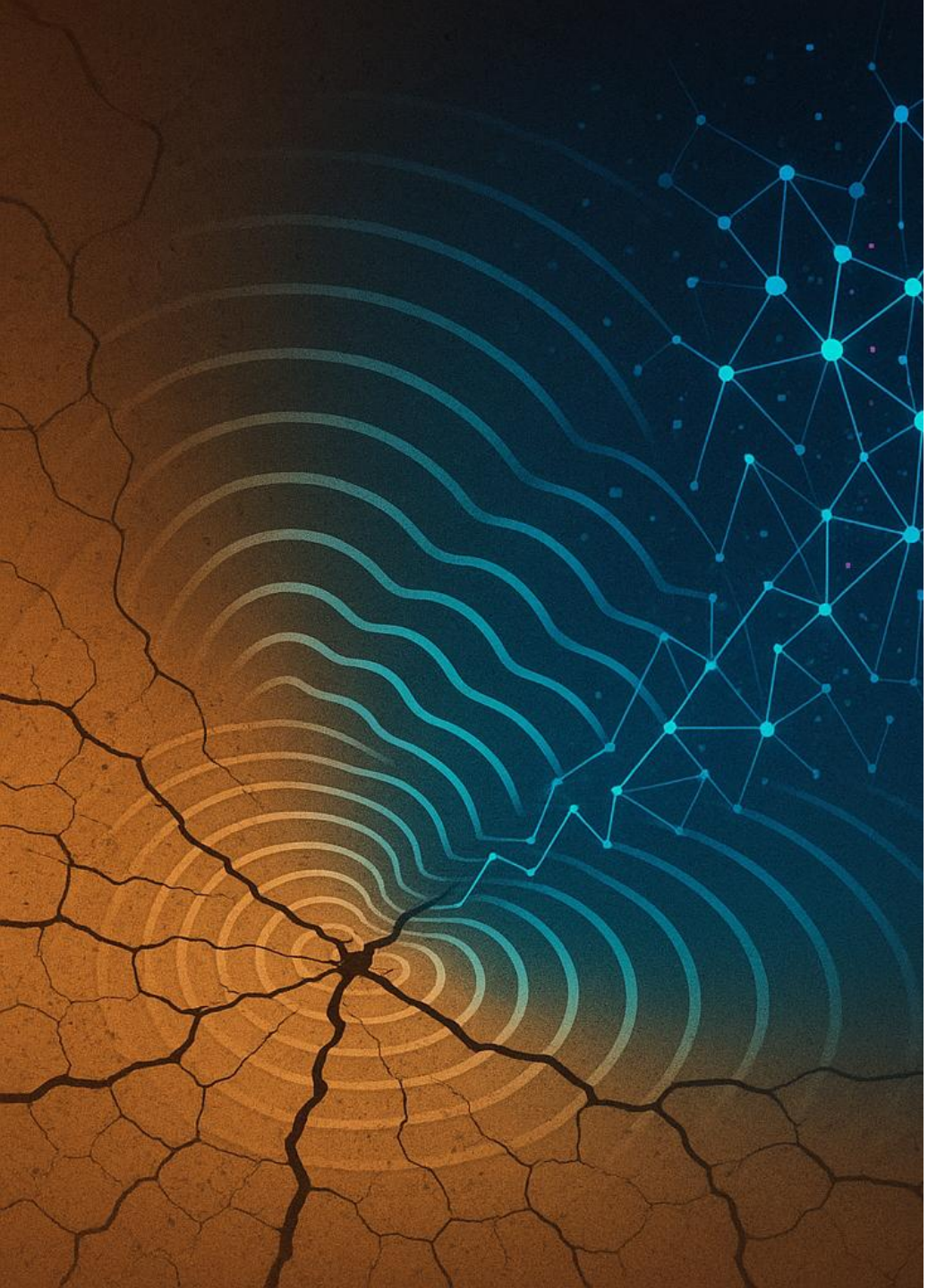
Kyle Hill

2.69M subscribers

Join | Subscribe | 208K | Share | •••

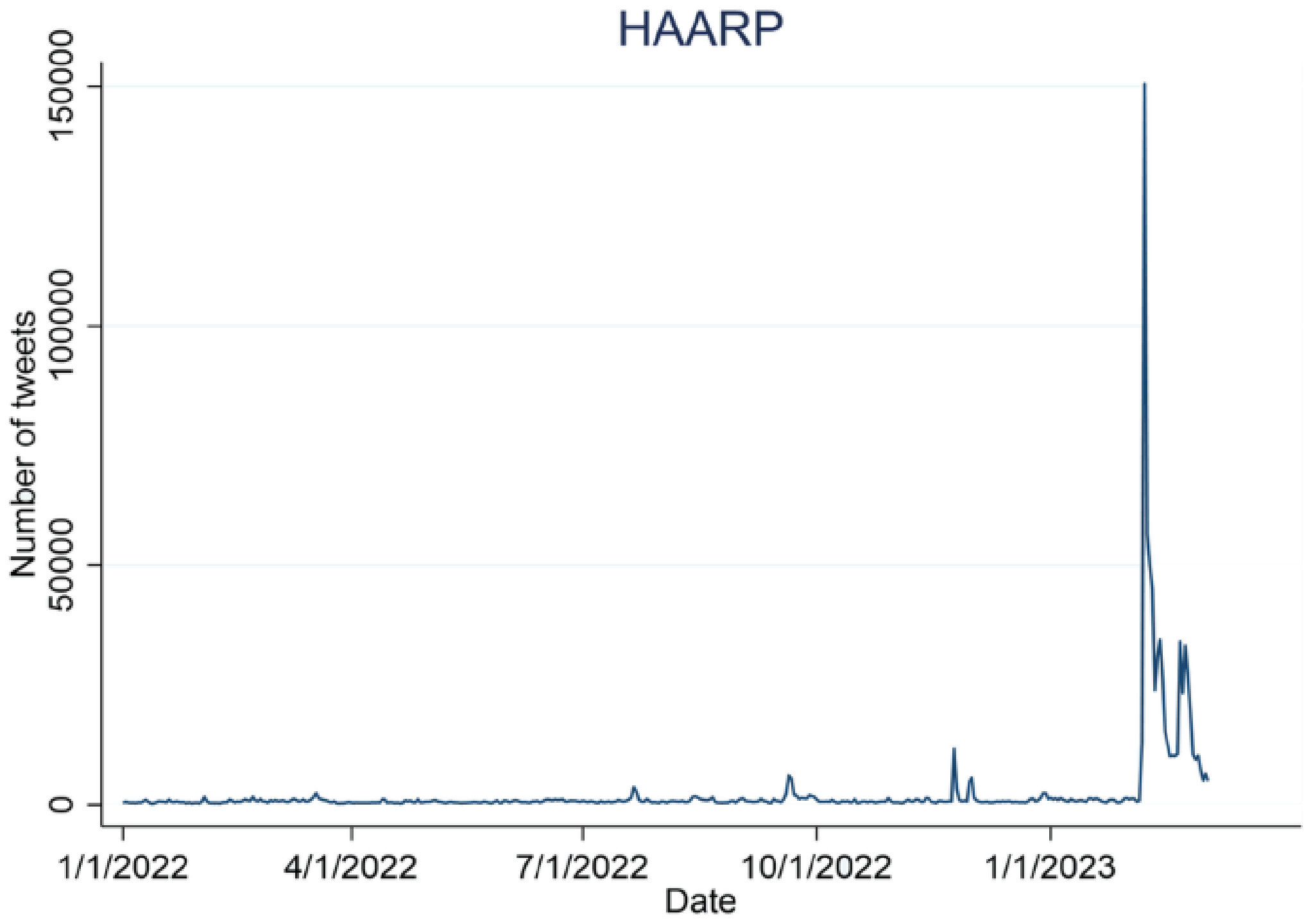
7,722,610 views May 16, 2023

In late 2022, a video of an apparent nuclear accident spread rapidly on Twitter and Tumblr. Was it modern history's first "viral" nuclear accident? Or was it faked for the lolz? This [HALF-LIFE HISTORY] attempts an investigation.



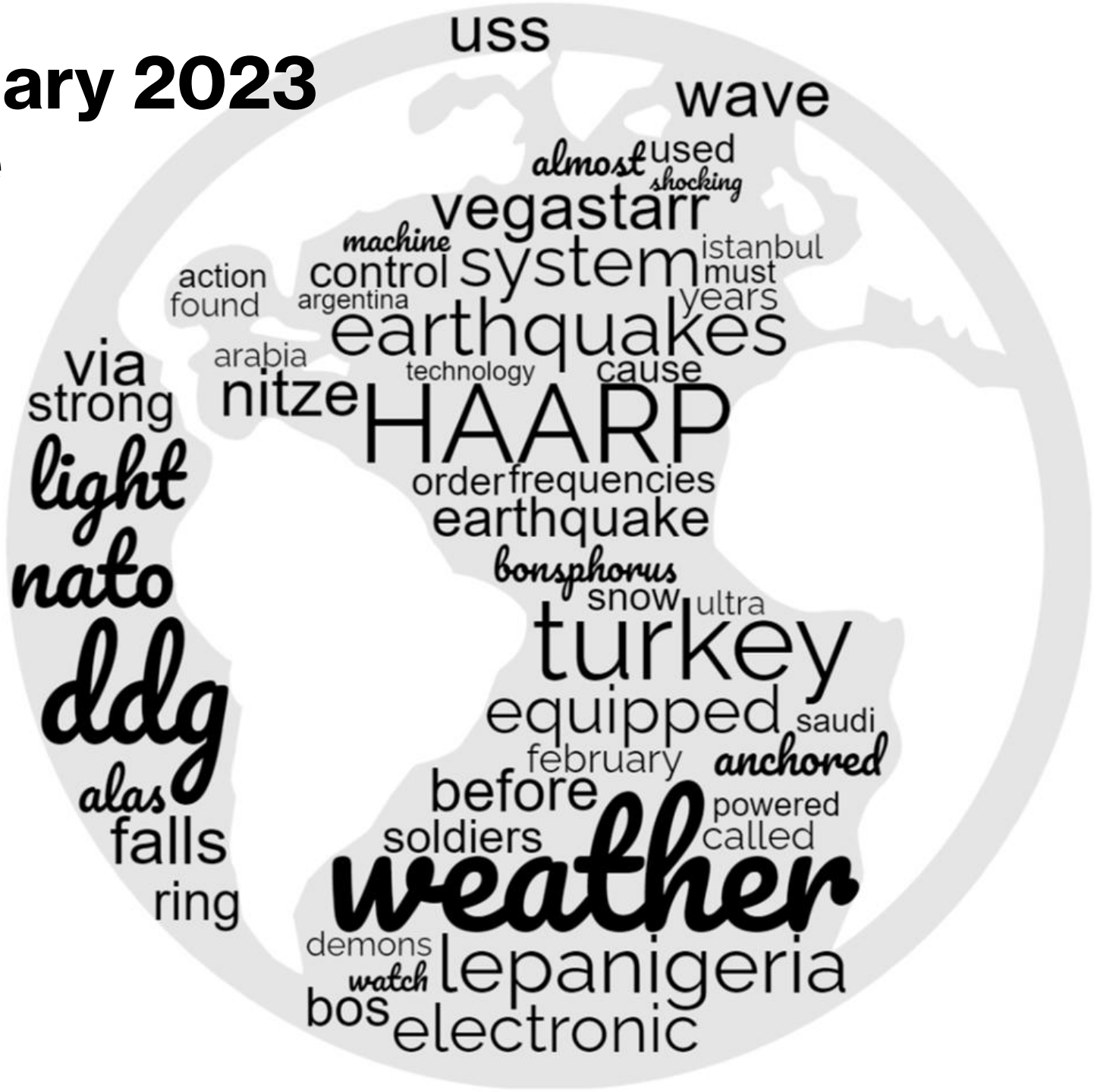
Earthquake Conspiracy Discourse

- **Data source:** Twitter API, collecting tweets containing “HAARP”
  Comment volume: 1041633 tweets analyzed
  Analysis date: Data collected from 1 January 2022 – 4 March 2023


HAARP
Number of tweets
0
50000
100000
150000
1/1/2022
4/1/2022
7/1/2022
10/1/2022
1/1/2023
Date

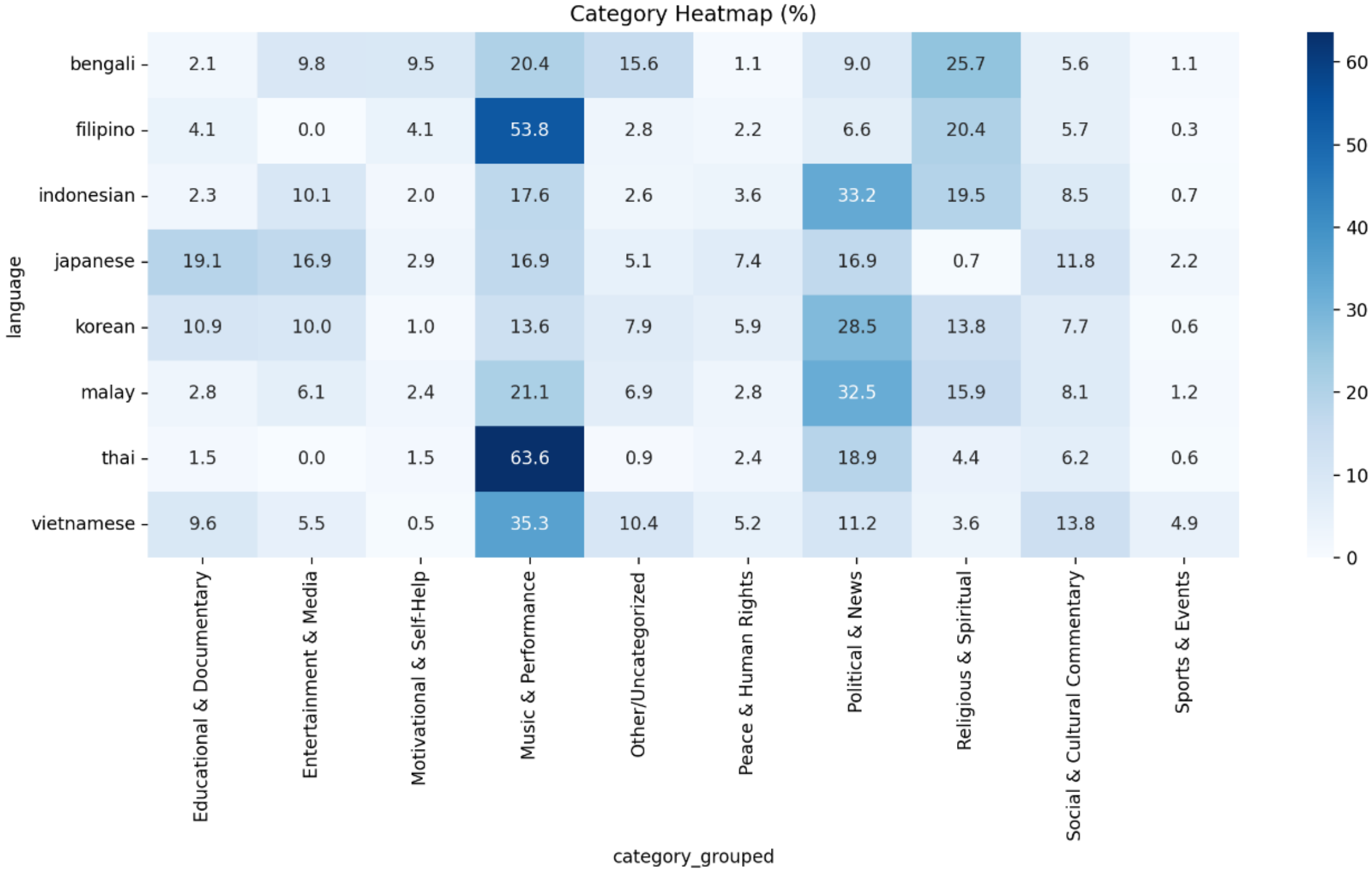
Before February 2023 Earthquake

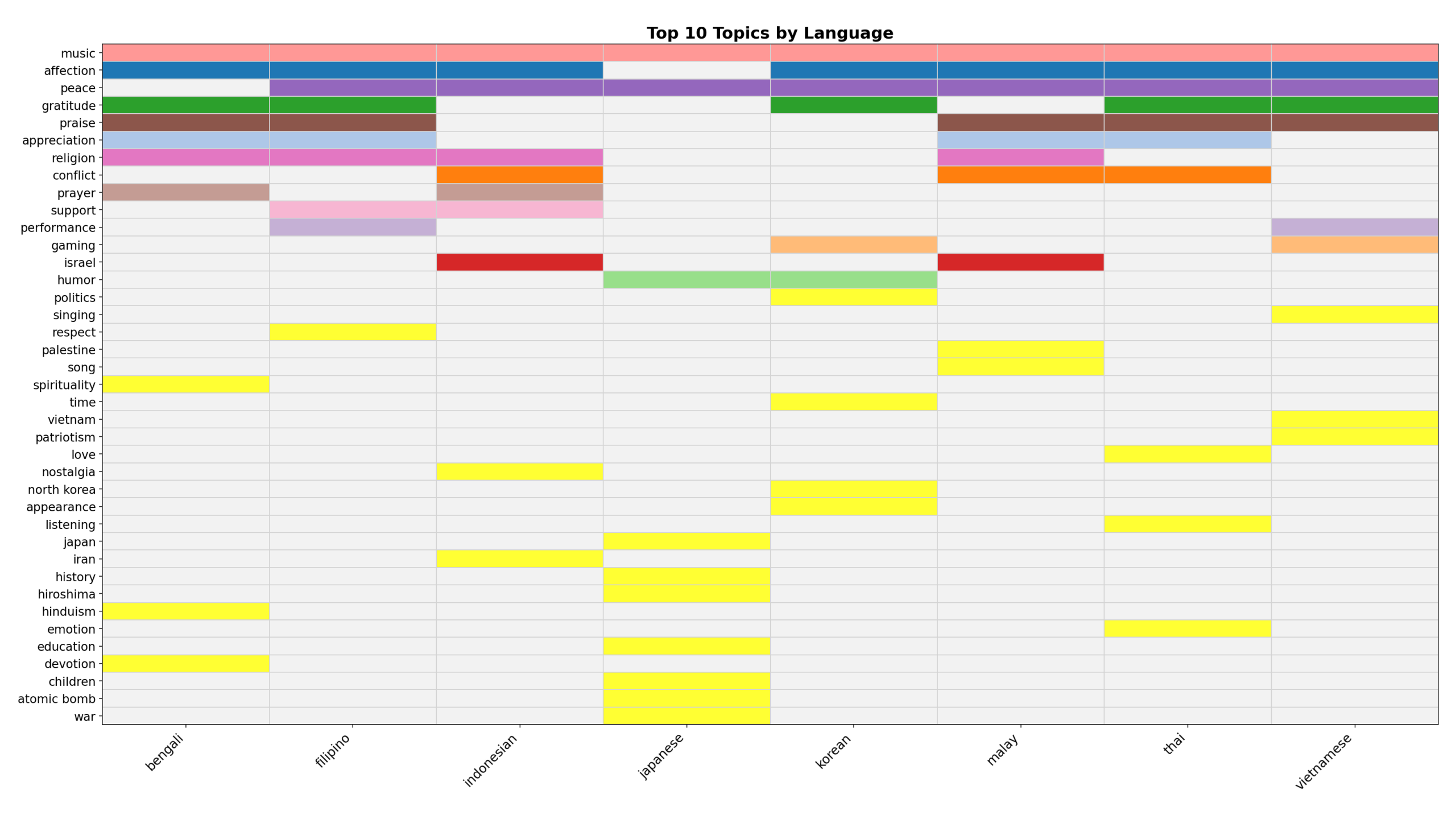
After February 2023 Earthquake



Peace Discourse Analysis

- **Data source:** Most relevant YouTube videos retrieved by search algorithm for the term **peace** in eight Asian languages.
- **Analysis date:** 7-8 July 2025.


Category Heatmap (%)
language
bengali
filipino
indonesian
japanese
korean
malay
thai
vietnamese
2.1 9.8 9.5 20.4 15.6 1.1 9.0 25.7 5.6 1.1
4.1 0.0 4.1 53.8 2.8 2.2 6.6 20.4 5.7 0.3
2.3 10.1 2.0 17.6 2.6 3.6 33.2 19.5 8.5 0.7
19.1 16.9 2.9 16.9 5.1 7.4 16.9 0.7 11.8 2.2
10.9 10.0 1.0 13.6 7.9 5.9 28.5 13.8 7.7 0.6
2.8 6.1 2.4 21.1 6.9 2.8 32.5 15.9 8.1 1.2
1.5 0.0 1.5 63.6 0.9 2.4 18.9 4.4 6.2 0.6
9.6 5.5 0.5 35.3 10.4 5.2 11.2 3.6 13.8 4.9
Educational & Documentary
Entertainment & Media
Motivational & Self-Help
Music & Performance
Other/Uncategorized
Peace & Human Rights
Political & News
Religious & Spiritual
Social & Cultural Commentary
Sports & Events
category_grouped
60
50
40
30
20
10
0


Top 10 Topics by Language
music
affection
peace
gratitude
praise
appreciation
religion
conflict
prayer
support
performance
gaming
israel
humor
politics
singing
respect
palestine
song
spirituality
time
vietnam
patriotism
love
nostalgia
north korea
appearance
listening
japan
iran
history
hiroshima
hinduism
emotion
education
devotion
children
atomic bomb
war
bengali
filipino
indonesian
japanese
korean
malay
thai
vietnamese

Embedding AI Monitoring into Crisis Protocols

Assess and scope

- Define your crisis scenarios and key data sources (e.g., “for a health crisis, we monitor X, Reddit, and news comments”).

Build or partner

- Develop in-house capability or partner with a specialized AI vendor.

Integrate and train

- Embed the dashboard into the emergency operations center. Train communicators on how to interpret and act on the data.

Iterate and refine

- Continuously update models based on post-crisis reviews.

A Responsible and Clear-Eyed Approach

Ethical considerations

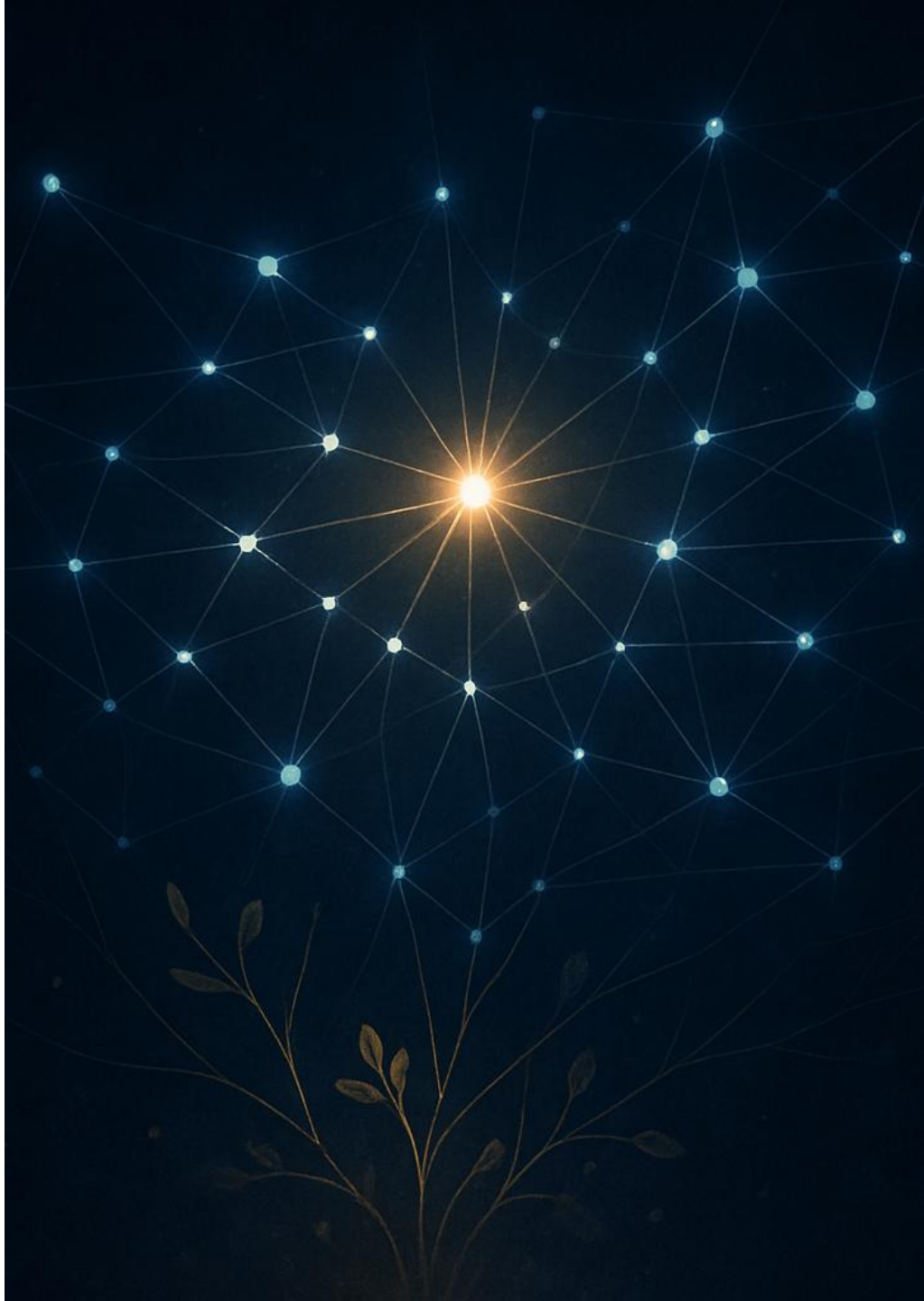
- **Privacy and anonymity:** Using public data, avoiding personal identification.
- **Bias in AI:** Models can reflect societal biases, require constant auditing.
- **Algorithmic transparency:** We must understand why the AI makes a certain classification.

Technical and practical limitations

- **Sarcasm and nuance:** NLP can struggle with complex human language.
- **Data silos:** Private groups and encrypted apps are blind spots.
- **The human-in-the-loop:** AI informs, but humans decide. It is an aid, not a replacement for judgment.

Key Takeaways

- The digital public square moves too fast for manual monitoring.
- AI provides the **scale and speed** to understand public discourse in real-time.
- This leads to more **trustworthy, effective, and agile** crisis response.
- The goal is **strategic resilience**: the ability to anticipate, withstand, and recover from a crisis through superior communication



What's on the Horizon?

- **Multimodal AI:** Analyzing images and videos for crisis evidence (e.g., damage assessment).
- **Predictive analytics:** Moving from real-time monitoring to forecasting crisis spikes.
- **Generative AI integration:** Using LLMs to draft initial public statements based on the detected narrative.

Thank You

Dmitry Erokhin, PhD

Research Scholar

International Institute for Applied Systems Analysis

Laxenburg, Austria

erokhin@iiasa.ac.at