ENABLING ATROCITY IN THE MENA AND SUDAN

SOCKPUPPETS, BOTS, AND DIGITAL INFORMATION HARM IN WARTIME

Enabling Atrocity in the MENA and Sudan: Sockpuppets, Bots and Digital Information Harm in Wartime

Contents

Preface

This report was produced under the auspices of the UNESCO Chair on Data, Media and Society at the University of South Carolina. It forms part of a research agenda into the role of data-driven media systems in shaping political perception, public knowledge, and information integrity in conflict-affected and 'atrocity risk' affected contexts. Its focus on large-scale coordinated influence operations across the Middle East and North Africa speaks directly to the mandate of the UNESCO Chair on Data, Media, and Society, particularly in relation to digital inequalities, platform governance, and the protection of pluralistic information environments.

Rather than examining misinformation as isolated content, the report analyses how networked architectures of distribution, sockpuppet networks, automated amplification, and platform affordances, reshape what becomes visible in times of war. This approach reflects UNESCO's emphasis on structural and systemic dimensions of media harm, especially in non-Western and multilingual contexts that remain underexamined in global platform governance debates. Specifically, this report documents and analyses three large-scale, coordinated sockpuppet and bot networks operating across the Middle East and North Africa (MENA), with particular emphasis on activity surrounding Sudan's civil war.

The findings presented here are intended to support evidence-based discussion among researchers, journalists, civil society organisations, and policymakers concerned with information integrity, peacebuilding, and data governance. While the report documents specific cases relating to Sudan and the wider MENA region, its implications extend to broader questions about how contemporary digital systems can be exploited to manufacture legitimacy, suppress dissent, and normalize violence through scale rather than persuasion.

Executive summary

Understanding information warfare, disinformation and propaganda within the context of 'atrocity risk' is particularly important given the scale of violence in Sudan and the possibility of genocide. This report documents and analyses three large-scale, coordinated sockpuppet and bot networks on the X platform (formerly Twitter) operating across the Middle East and North Africa (MENA), with particular emphasis on Sudan's civil war, which pits the paramilitary Rapid Support Forces (RSF) against the the government-backed Sudanese Armed Forces (SAF). Drawing on the analysis across three distinct networks of more than 250,000 social media posts produced in Arabic, English, Persian, Turkish and French, and covering a period of over two years, the report identifies sustained influence operations involving over 27,000 inauthentic accounts used to promote state-aligned geopolitical narratives to reshape public understanding of the civil war in Sudan.

While the report surveys MENA-wide activity, including campaigns targeting Iran, Mauritania, Morocco, Algeria, Libya, Yemen, Syria, Tunisia, Saudi Arabia, and the United Arab Emirates, the core empirical focus concerns networks promoting the Rapid Support Forces during the Sudan conflict. These networks operated over different periods and across multiple languages, indicating persistence, adaptation, and strategic reuse rather than episodic activity.

It is important to note that the SAF and RSF both use propaganda and disinformation. The focus here on the RSF reflects both atrocity-risk considerations, evidentiary accessibility, and network resilience, rather than an assumption that disinformation is exclusive to one side of the conflict.

The report further argues that recent changes to platform governance, particularly credit-based verification systems, algorithmic amplification mechanisms, and weak enforcement in Arabic and non-English contexts, have lowered the operational costs of influence campaigns while increasing their reach and perceived legitimacy. In this sense, contemporary social media platforms increasingly facilitate a form of "pay-to-play propaganda," in which visibility, credibility, and agenda-setting capacity can be artificially purchased, coordinated, and scaled through inauthentic networks. Ability to manipulate these systems is particularly alarming when they can then influence the output of LLMs, which are increasingly used by consumers to find information instead of internet search functions.

Key findings at a glance

This report examines the Tactics, Techniques, and Procedures (TTPs) of three different bot and sockpuppet networks. TTPs are the specific behaviors, processes, and patterns of activity that threat

actors use to design, scale, and execute disinformation campaigns. Adapted from cybersecurity threat-mapping frameworks, tracking disinformation TTPs allows security analysts, civil society organisations, and governments to identify coordinated influence operations, predict adversary behavior, and build systemic resilience

Some of the key findings across the three networks include:

- **Large-scale coordinated inauthentic behavior:**
  The investigation identifies three MENA-wide sockpuppet and bot networks comprising over 27,000 fake accounts that have produced hundreds of thousands of posts, operating across Arabic, English, Persian, Turkish and French. They have largely been promoting an anti-Islamist, pro-RSF, and UAE-aligned political narrative.

- **Sustained and adaptive operations, not episodic abuse:**
  At least one network persisted (continues to persist) over multiple years, including phases prior to asset deletion, demonstrating long-term maintenance, identity repurposing, and strategic adaptation rather than short-lived campaigns.

- **Centralized coordination with cell-based architecture:**
  At least one network reveals small coordinating cores linked to thousands of low-interaction broadcast accounts, organized into loosely coupled regional cells that share infrastructure while compartmentalizing narratives by country.

- **Systematic promotion of RSF-aligned narratives in Sudan:**
  Across multiple datasets, the networks overwhelmingly promoted pro-RSF and anti-SAF/Burhan narratives, portraying the RSF as humanitarian, peace-seeking, and legitimate, while attributing civilian harm, famine, and obstruction of aid almost exclusively to the SAF and Islamist actors.

- **Whitewashing genocide and violence**
  In the aftermath of El Fasher, coordinated hashtag campaigns reframed violence as recovery and normalization, illustrating a pattern consistent with image repair theory.

- **Exploitation of platform visibility mechanisms:**
  Relay-style hashtag sequencing, burst posting, location-spoofing, handle-switching, high velocity posting show actors are adept at manipulating and tracking X visibility dynamics to get things to trend.

- **Integration of AI-generated and AI-assisted personas:**
  A distinct network of AI-assisted or AI-generated accounts was identified, characterized by synthetic language patterns, repeated slogans, homogenized bios, and anomalous posting behavior, signaling the growing role of generative AI in coordinated influence operations. These accounts also focused on engaging with real users.

- **Cross-ideological and cross-regional spillover:**
  Some accounts exhibited overlap with the far right, alternating between Sudan-focused

propaganda and amplification of European and North American culture-war narratives, blurring boundaries between regional geopolitics and transnational discourse.

- **Systematic under-enforcement in Arabic and non-English contexts:**
  Despite scale and persistence, these networks operated for extended periods despite moments of degradation, potentially suggesting structural asymmetries in platform enforcement that disadvantage conflict-affected, non-Western information environments.

- **One platform, multiple methods**
  The three different networks all demonstrate how an actor (or actors) pursuing the same goal employ different strategies on the same platform, exploiting affordances in slightly different ways to create breadth and depth in terms of narrative saturation.

Background: Sudan's Civil War

Sudan's civil war began on 15 April 2023, when fighting erupted in Khartoum between the Sudanese Armed Forces (SAF), led by Abdel Fattah al-Burhan, and the Rapid Support Forces (RSF), commanded by Mohamed Hamdan Dagalo ("Hemedti"). The conflict stems from the collapse of Sudan's post-2019 transition following the overthrow of Omar al-Bashir and the October 2021 coup, which dismantled civilian governance. Prolonged disputes over power-sharing and security-sector reform, particularly the integration of the RSF into a unified national army under civilian oversight, eventually escalated into open war. After initially losing ground, the SAF launched a counter-offensive in late 2024 and recaptured Khartoum in March 2025, but rather than paving the way for de-escalation, the conflict has widened as both sides pursue military victory[1].

The humanitarian consequences have been catastrophic, with some figures suggesting that up to 400,000 people may have been killed since the war began[2]. More than 11 million people have been displaced, creating the world's largest displacement crisis, while over 30 million require humanitarian assistance amid collapsing food systems and restricted aid access. Violence has been particularly acute in Darfur, where mass killings, sexual violence, starvation, and forced displacement have intensified around key urban centers such as El Fasher, compounding long-standing patterns of marginalization and insecurity. Long regarded as a humanitarian and administrative hub for the region, El Fasher became a refuge for displaced civilians fleeing violence elsewhere in Darfur. Its siege and eventual capture carried profound humanitarian consequences, exposing hundreds of thousands of civilians to intensified violence, starvation, and displacement, while marking a major strategic victory for the RSF in western Sudan.

[1] 'Two Years On, Sudan's War Is Spreading | International Crisis Group', 7 April 2025, https://www.crisisgroup.org/stm/africa/sudan/two-years-sudans-war-spreading.
[2] the Center for Preventive Action, 'Civil War in Sudan', Global Conflict Tracker, accessed 23 February 2026, https://cfr.org/global-conflict-tracker/conflict/power-struggle-sudan.

In February 2026, the UN Independent International Fact-Finding Mission for the Sudan concluded that RSF operations during the late-October 2025 takeover of El Fasher bore the "hallmarks of genocide."[3] The mission documented ethnically targeted killings, widespread sexual violence, enforced disappearances, and the deliberate imposition of life-threatening conditions against the Zaghawa and Fur communities, finding that genocidal intent was the only reasonable inference from the pattern, scale, and rhetoric of the violence. While the SAF has committed serious violations of international humanitarian law, the UN and other international bodies distinguish these crimes from the systematic, group-directed destruction attributed to the RSF in Darfur. With continued external backing for both sides and failed mediation efforts, international investigators warn that the risk of further genocidal acts remains serious and ongoing, raising the prospect of prolonged conflict, state fragmentation, and regional destabilization.

Sudan's war is fundamentally transnational, shaped by sustained external intervention and regional power competition rather than domestic dynamics alone. The UAE and Saudi Arabia, alongside Egypt, have played a decisive role since the 2018–2019 uprising by backing military and paramilitary actors under the banner of "stability," providing financial support, weapons, diplomatic cover, and political leverage that emboldened Sudan's generals and weakened prospects for civilian rule. The RSF's integration into Gulf security economies, including mercenary deployments in Yemen and gold exports routed through Dubai, has tied Sudan's internal violence to wider regional patronage networks; while competing Gulf, Egyptian, and Turkish interests have repeatedly intersected with Sudanese factional struggles. This externalization of power has made Sudan a theatre for regional rivalry, proxy influence, and geopolitical bargaining, with devastating consequences for civilian protection and accountability.[4]

Building on this transnationalization of Sudan's war, the information environment surrounding the conflict has likewise become cross-border and networked. Narratives about legitimacy, ceasefires, humanitarian access, and responsibility for violence are not produced or circulated solely by domestic actors, but, as this report shows, are amplified through transnational bot networks, coordinated inauthentic behavior, and aligned media ecosystems that operate across regions and platforms. These information operations mirror the conflict's external patronage structures, extending influence, laundering responsibility, and shaping international perception. This makes the study of bot networks central to understanding how the war is sustained, justified, and normalized beyond Sudan's borders.

Atrocity risk: Why these networks matter

These influence networks function as tools of digital authoritarianism, deployed to manipulate perception, manufacture legitimacy, and overwhelm and confuse public discourse. Their harm lies both in the content they promote as well the deceptive means of distribution itself: coordinated sockpuppet architectures that simulate authentic civic debate while operating in service of political

[3] 'Sudan: "Hallmarks of Genocide" Found in El Fasher, UN Investigators Detail Mass Killings and Ethnic Targeting | UN News', 19 February 2026, https://news.un.org/en/story/2026/02/1166997.
[4] 'The Great Game of the UAE and Saudi Arabia in Sudan', *Project on Middle East Political Science*, 16 June 2020, https://pomeps.org/the-great-game-of-the-uae-and-saudi-arabia-in-sudan.

agendas that run counter to human rights and the social good. They fundamentally undermine the free flow of credible information by crowding out legitimate news and also masking their source/agenda.

In conflict environments, such practices are particularly dangerous. By manufacturing visibility and simulating public consensus, these networks can crowd out legitimate independent journalism and civil society voices, distort attribution of responsibility for violence, and can contribute to the normalization of impunity. In the case of Sudan, the report shows how coordinated amplification was used to reframe mass violence and humanitarian catastrophe as solely the responsibility of the SAF, Iran, and Islamist forces, all the while whitewashing or lionizing the actions of the Rapid Support Forces.

Atrocity-risk research suggests the most relevant question in mass-violence settings is not necessarily whether misinformation "changes minds" in the abstract sense, but whether media systems enable or obstruct prevention, early warning, and crisis response[5]. Recent work on media and mass atrocity prevention conceptualizes three pathways through which information environments can shape outcomes:

1. Structural prevention (building resilient, pluralistic media and accountability norms over time)
2. Operational prevention (timely warning, credible attribution, and rapid mitigation around triggers),
3. Crisis response (sustained documentation, agenda-setting, and evidentiary support once large-scale violence is underway).

In Sudan's war, coordinated inauthentic networks have the potential to degrade all three pathways at once: they can erode structural resilience by simulating consensus and laundering authoritarian legitimacy; they can sabotage operational prevention by flooding trigger windows with disciplined counter-framings that blunt warning signals, or by muddying the waters of pertinent discourses; and they can weaken crisis response by saturating attention with narratives that displace accurate documentation, confuse attribution, and reduce the political salience of civilian harm. In this sense, the core harm is not "falsehoods" per se, but interference in attenuating useful, important and urgent information.

On a more fundamental level, it can work to influence publics, policymakers, and other relevant actors, muddying the waters of debate, and in the case of this work, underpinning a sense that there is parity in responsibility and scale for the nature of the violence going on. More pressingly, as more people use LLM agents to obtain information, the danger of RAG (retrieval augmented generation) and LLM poisoning becomes more pertinent. RAG poisoning is where content on social media and the web can be used to influence what agents such as ChatGPT (and others) output. Bad actors can thus use bot armies to 'poison' LLMs with propaganda and disinformation.

[5] Chiara De Franco and Christoph O. Meyer, 'Media and Mass Atrocity Prevention: Three Pathways of Potential Influence', *Global Responsibility to Protect* 17, no. 4 (2025): 288–319, https://doi.org/10.1163/1875984X-20250018.

Computational propaganda: Astroturfing, sockpuppet, trolls, bots and more

Computational propaganda is the strategic use of algorithms, automation, bots, trolls, sockpuppet accounts and large-scale data analytics to shape, steer, or manipulate public opinion and political discourse within digital environments.[6] Astroturfing refers to the orchestration of artificial grassroots sentiment through the coordinated use of sockpuppet accounts or automated bots. Sockpuppets are commonly understood as multiple online accounts controlled by a single individual, typically deployed to deceive, manipulate, or influence audiences. Such configurations have been described as *homologous sockpuppets*.[7] These accounts often operate collectively in coordinated clusters, frequently referred to as sockpuppet gangs (SPGs).[8] Astroturfing is a deliberately deceptive practice designed to manufacture the appearance of popular support or opposition. In the political domain, astroturfing has been defined as "a campaign disguised as spontaneous, popular 'grassroots' behavior that is in reality carried out by a single person or organization".[9] The practice is widely regarded as harmful because it undermines democratic processes by eroding trust in authentic civic mobilization. By imitating genuine bottom-up engagement, astroturfing can delegitimize real grassroots campaigns and foster public skepticism toward front-line political actors and movements[10]. Empirical research suggests that sockpuppet-driven activity can be disproportionately influential: studies of online forums have found that sockpuppet accounts often attract greater attention and engagement than legitimate users, thereby amplifying their capacity to shape discourse and influence public opinion.[11]

Sockpuppets and bots are so prevalent that they have their own regionally specific nomenclature. In the Gulf, these actors are commonly referred to as "electronic flies" (الذباب الإلكتروني *al-dhubāb al-ilikturūnī*), while in Sudan they are often called "electronic chickens" (الجداد الإلكتروني *al-jaddād al-ilikturūnī*).Identifying and exposing sockpuppets is important, as they are regularly implicated in harmful or manipulative activities.[12] Sockpuppets have become a cornerstone of digital authoritarianism in the Gulf, particularly during moments of regional tension like the 2017 Gulf Crisis. Networks of fake accounts, often thousands strong, have been deployed across and within the region,

[6] Samuel C. Woolley, 'Bots and Computational Propaganda: Automation for Communication and Control', *Social Media and Democracy*, August 2020, 89–110, https://doi.org/10.1017/9781108890960.006.
[7] Dong Liu et al., 'Sockpuppet Gang Detection on Social Media Sites', *Front. Comput. Sci* 10, no. 1 (2016): 124–35, https://doi.org/10.1007/s11704-015-4287-7.
[8] Liu et al., 'Sockpuppet Gang Detection on Social Media Sites'.
[9] Jacob Ratkiewicz et al., 'Truthy: Mapping the Spread of Astroturf in Microblog Streams', *Proceedings of the 20th International Conference Companion on World Wide Web, WWW 2011*, 2011, 249–52, https://doi.org/10.1145/1963192.1963301.
[10] Deborah H Drake et al., 'Criminology and Propaganda Studies: Charting New Horizons in Criminological Thought', *The British Journal of Criminology*, ahead of print, September 2023, https://doi.org/10.1093/BJC/AZAD045.
[11] Nhut Lam Nguyen et al., 'Learning to Recognize Sockpuppets in Online Political Discussions', *IEEE Systems Journal* 16, no. 2 (2022): 1873–84, https://doi.org/10.1109/JSYST.2021.3117815.
[12] Liu et al., 'Sockpuppet Gang Detection on Social Media Sites'.

and also targeting external actors from within the region. They are used to spread pro-government propaganda, spread sectarian hate speech, disinformation, and destabilize neighboring states.[13]

In 2017, thousands of bots were deployed to amplify anti-Qatar narratives, push anti-Shia sectarian tropes, and present orchestrated campaigns as organic public sentiment. Similar tactics have been used to support Saudi and Emirati foreign policy goals, such as normalization with Israel or framing Iran as a regional existential threat. In many cases, these operations mirror Russian-style disinformation campaigns, attempting to polarize communities, sow chaos, all while obscuring attribution[14]. Rather than simply silencing dissent, this strategy works by overwhelming the information space, making it harder to distinguish genuine discourse from state-backed fabrication.

Digital authoritarianism: Regional context

In the MENA more broadly, there are a number of states that are frequently documented to have advanced digital capabilities. In the Arab Gulf states, Saudi Arabia and the UAE in particular have emerged as digital superpowers, not because they produce social media platforms, but because of their ability to co-opt or control them[15] through surveillance or manipulation via trolls and bots. The UAE, in particular, has increasingly used digital influence operations as a tool of its assertive foreign policy across the Middle East and North Africa. From Egypt to Sudan to Syria, Abu Dhabi has invested in campaigns that promote authoritarian allies, undermine Islamist groups, especially the Muslim Brotherhood, and shape regional narratives in line with its strategic interests [16]. These operations often involve coordinated networks of sockpuppet accounts, fake news sites, and proxy media outlets that blur the lines between journalism, propaganda, and intelligence work.[17] While Saudi Arabia and the UAE are among the most prominent digital superpowers in the Gulf region, they are by no means alone

In the wider MENA region Israel stands out as the most capable state in terms of digital authoritarianism. Its use of AI, digital espionage, and digital manipulation is evident in its export of spyware, especially Pegasus, which has been used by several MENA governments to target activists and journalists. Other state and non-state actors, including Egypt, Iran, Turkey, and various militias, have also developed their own digital arsenals to shape narratives, harass opponents, and influence regional perception[18]. Iran-backed groups have used online platforms to mobilize support for resistance narratives, while Turkey has promoted pro-AKP and nationalist messaging through coordinated networks. Egypt has run extensive bot operations targeting dissidents and journalists. The broader picture is one of a highly competitive information battlefield, where multiple actors are

[13] MO Jones, *Digital Authoritarianism in the Middle East: Deception, Disinformation and Social Media*, 2022.
[14] Ahmed Al-Rawi and Anis Rahman, 'Manufacturing Rage: The Russian Internet Research Agency's Political Astroturfing on Social Media', *First Monday*, ahead of print, August 2020, https://doi.org/10.5210/FM.V25I9.10801.
[15] Jones, *Digital Authoritarianism in the Middle East: Deception, Disinformation and Social Media*.
[16] David Kirkpatrick, 'The Dirty Secrets of a Smear Campaign | The New Yorker', *The New Yorker*, 2023, https://www.newyorker.com/magazine/2023/04/03/the-dirty-secrets-of-a-smear-campaign.
[17] 'The Rise of the Emirati Dis-Influencers and Why We All Should Be Worried | Middle East Eye'.
[18] Ihsan Yilmaz and Shahram Akbarzadeh, 'Goals, Practices, and Impacts of Strategic Digital Information Operations', in *Authoritarian Actors and Strategic Digital Information Operations* (Routledge, 2025).

engaged in a digital arms race to control the narrative across the MENA region and beyond. Countries pursue different objectives.

Dangers of digital disinformation and information manipulation

Disinformation is the deliberate spread of false information to cause harm[19]. While disinformation is a useful shorthand, terms like FIMI (Foreign Information Manipulation and Interference) used by the EU are useful in capturing the fact that manipulation need not be false to constitute manipulation. Digital disinformation poses demonstrable harm to political and information environments, regardless of ongoing debate over its strategic effectiveness. While the motivations, capacities, and willingness to deploy disinformation as a policy instrument vary across states, and perceptions of its threat differ by regime type, the cumulative effects of sustained false or misleading narratives include erosion of trust, distortion of public understanding, and heightened uncertainty in political and diplomatic contexts. Although actors such as the United States, the European Union, and NATO frame disinformation as a serious security concern, its actual strategic impact remains contested within the scholarly literature. Some analysts argue that while disinformation can be disruptive and damaging to international relations, it does not constitute a decisive security threat nor fundamentally reshape global power dynamics (Gerrits, 2018). Others similarly suggest that existing evidence indicates the strategic effects of disinformation have been overstated[20].

More recent research, however, complicates these skeptical assessments. Simulation-based studies indicate that disinformation in diplomatic and international relations contexts can meaningfully misinform audiences, particularly when amplified through coordinated networks of bots and trolls.[21] Moreover, real-world cases demonstrate that disinformation can produce tangible social harm even in the absence of clear strategic outcomes. For example, following the Southport stabbings in the United Kingdom, false and misleading claims circulated rapidly online, contributing to public confusion, racialized speculation, heightened community tensions and eventually riots.[22] While such disinformation did not alter state policy or geopolitical alignments, it nonetheless exacerbated fear, mistrust, and social fragmentation, illustrating how harm can manifest at the level of public order and social cohesion rather than formal decision-making.

These debates are further complicated by persistent methodological challenges. Demonstrating clear causal effects from disinformation campaigns is inherently difficult, particularly in contexts where

[19] Claire Wardle and Hossein Derakshan, 'PDF: Wardle, C. & Derakshan, H. (2017) Information Disorder: Toward an Interdisciplinary Framework for Research and Policy Making, Council of Europe', First Draft, 2017, https://firstdraftnews.org/glossary-items/pdf-wardle-c-derakshan-h-2017-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making-council-of-europe/.
[20] Alexander Lanoszka, 'Disinformation in International Politics', *European Journal of International Security* 4, no. 2 (2019): 227–48, https://doi.org/10.1017/eis.2019.6.
[21] Alfredo Guzmán Rincón et al., 'Social Networks, Disinformation and Diplomacy: A Dynamic Model for a Current Problem', *Humanities and Social Sciences Communications* 10, no. 1 (2023): 505, https://doi.org/10.1057/s41599-023-01998-z.
[22] 'How False Online Claims about Southport Knife Attack Spread so Rapidly | UK News | The Guardian'.

reliable polling or public opinion data are limited. Despite significant scholarly and policy attention, problems of attribution and measurement continue to constrain assessments of impact, creating strategic ambiguity and raising questions about the extent to which perceived threats align with demonstrable outcomes.

Nevertheless, international legal frameworks reflect recognition that certain forms of information warfare are harmful. The International Covenant on Civil and Political Rights (ICCPR) prohibits "propaganda for war," signaling normative concern over the role of information in enabling violence, even if this provision remains vague and unevenly enforced[23]. Importantly, such prohibitions focus on the narrative function of propaganda rather than the specific mechanisms of deception involved. For instance, the use of automated or deceptive accounts to promote reputational or commercial campaigns may be misleading, but it does not necessarily constitute war propaganda. While the ICCPR underscores the seriousness with which information manipulation can be viewed in principle, its scope remains limited and ill-equipped to address the broader and more ambiguous realities of contemporary digital disinformation.

Recognition of the problem is manifest in certain jurisdictions. The European Union has taken a comparatively assertive governance approach via the Digital Services Act (DSA), which became directly applicable across the EU on 17 February 2024—and through the 13 February 2025 endorsement of the 2022 Code of Practice on Disinformation as a DSA "Code of Conduct" benchmark. Importantly, these instruments do not "ban disinformation" or arbitrate truth. Instead, they push systems-level accountability: for Very Large Online Platforms (VLOPs) this includes systemic risk assessments, mitigation measures, transparency reporting, auditability, and, under the Code framework, commitments such as wider fact-checking coverage across EU languages, improved advertising transparency, demonetization measures, and researcher access (where signatories commit to implement these). Enforcement began. For example, on 4 December 2025 the Commission issued a €120 million fine against X for DSA transparency breaches. At the same time, this EU model is facing coordinated political pushback from the US: Reuters reported an internal August 2025 cable from Secretary of State Marco Rubio directing US diplomats to lobby against the DSA, alongside congressional framing of the DSA as a "global censorship" tool, dynamics that help explain why governance remains fragmented and why transnational influence operations can route around region-bound regulation.

Where corporations and governments have failed or lagged behind, it is incumbent on civil society actors to play an active role tackling disinformation. With this in mind, identifying, monitoring, documenting, and analyzing influence operations remains imperative. This report is written in that spirit.

[23] 'Propaganda for War & International Human Rights Standards | Chicago Journal of International Law', accessed 23 February 2026, https://cjil.uchicago.edu/print-archive/propaganda-war-international-human-rights-standards.

Related work

There are a number of organisations drawing attention to issues of Sudan's information environment. Perhaps foremost and most relevant of them is Beam Reports, an independent Sudanese media platform that delivers a wide range of journalistic content using contemporary digital reporting tools. The platform publishes investigative and analytical reporting, while also specializing in fact-checking, information analysis, and public opinion research. They frequently document online manipulation as it pertains to Sudan.[24] In June 2025 for example, their work was responsible for some elements of a disinformation network being suspended.[25]

Recently, a complementary perspective on Sudan's information environment is offered by the CDAC Network's 2025 report *Sudan's Information War: How Weaponized Online Narratives Shape the Humanitarian Crisis and Response*, which examines harmful information from the standpoint of humanitarian practice and crisis-affected communities rather than network forensics.[26] Drawing on key informant interviews with humanitarian professionals, mutual aid groups, local civil society, media actors and connectivity specialists, the report documents how misinformation, disinformation and hate speech have functioned as deliberate tools of war in Sudan since April 2023, directly shaping aid access, civilian protection and community trust. While both the SAF and RSF are identified as deploying weaponized narratives, the report places particular emphasis on the lived consequences of these practices: civilians killed after being misled about safe areas, Emergency Response Rooms (ERRs) and mutual aid groups falsely smeared as RSF affiliates and consequently arrested or attacked, displaced communities from Darfur and Kordofan subjected to incitement and economic exclusion, and humanitarian convoys delayed or denied access following disinformation campaigns alleging weapons smuggling.

The report further documents how Sudan's information ecosystem has been hollowed out by violence, repression and funding cuts, with over 400 journalists displaced, key outlets shuttered, and surviving media operating under self-censorship, creating vacuums readily filled by polarized content and influencer-driven narratives. Crucially, the report highlights the under-recognized role of local verification networks, youth volunteers, teachers, religious leaders, ERRs and citizen journalists, who function as de facto infrastructure for information integrity under conditions of telecom blackouts, surveillance, and prohibitive connectivity costs (including the high-risk use of Starlink). The CDAC report's central conceptual contribution is its insistence that "information is a form of aid" and that communication should be treated as core humanitarian infrastructure, on par with food, water and shelter. It also draws attention to systemic platform failures in Arabic-language contexts and the

[24] *A Coordinated Digital Campaign Promotes an Alliance between the Sudanese Military and the Iranian Regime amid Regional Conflict - Beam Reports*, 21 April 2026, https://www.beamreports.com/2026/04/21/%d8%ad%d9%85%d9%84%d8%a9-%d8%b1%d9%82%d9%85%d9%8a%d8%a9-%d9%85%d9%86%d8%b3%d9%91%d9%82%d8%a9-%d8%aa%d8%b1%d9%88%d9%91%d8%ac-%d9%84%d8%aa%d8%ad%d8%a7%d9%84%d9%81-%d8%a8%d9%8a%d9%86-%d8%a7%d9%84%d8%ac/.

[25] 'How Disinformation Is Shaping Sudan's Conflict: A New Report', Thomson Foundation, accessed 8 May 2026, https://www.thomsonfoundation.org/latest/how-disinformation-is-shaping-sudan-s-conflict-a-new-report/.

[26] 'Evolving Dynamics of the Sudan Conflict: Implications for Humanitarian Action and Civil Society', CDAC Network, accessed 8 May 2026, https://www.cdacnetwork.org/resources/evolving-dynamics-of-the-sudan-conflict.

structural under-enforcement of coordinated harm in non-Western environments, observations that resonate strongly with the findings of the present report. Where this report focuses on the upstream architecture of coordinated influence operations and their geopolitical alignment, the CDAC study illuminates the downstream humanitarian consequences and community-level responses, together offering a more complete picture of how harmful information operates in Sudan as both a tool of war and a determinant of civilian survival.

A further valuable contribution to the emerging work on Sudan's information environment is the Thomson Foundation's *Information Integrity Watch: Sudan Monthly Insights* (February 2026), produced as part of the FCDO-funded Sudan Digital Resilience Project.[27] Drawing on daily social listening across X, Facebook, TikTok, WhatsApp, and Threads, and applying the ABCDE and DISARM frameworks alongside manual verification and cross-platform triangulation, the report offers granular, month-by-month documentation of how disinformation narratives circulate and mutate in response to specific conflict events.

The Thomson Foundation report is valuable to the present study because it provides independent, event-level corroboration of the structural patterns identified across our three networks, including the systematic rehabilitation of the RSF, the attribution of humanitarian harm to the SAF and so-called "Kizan" Islamist actors, and the embedding of Sudan's war within wider Gulf and Horn of Africa rivalries. It also offers analysis of both RSF and SAF propaganda and narratives and discusses named amplifiers alongside coordinated hashtag campaigns and the migration of disinformation to semi-public WhatsApp channels and Threads.

I have also documented networks of UAE-influencers promoting pseudo-news outlets with carefully crafted propaganda about Sudan. This technique often involves creating websites that ostensibly look credible, but are means of laundering propaganda narratives while keeping the source anonymous. Sites like NewYork Insight and EuroPost are examples of these ‘pseudo-news’ outlets.

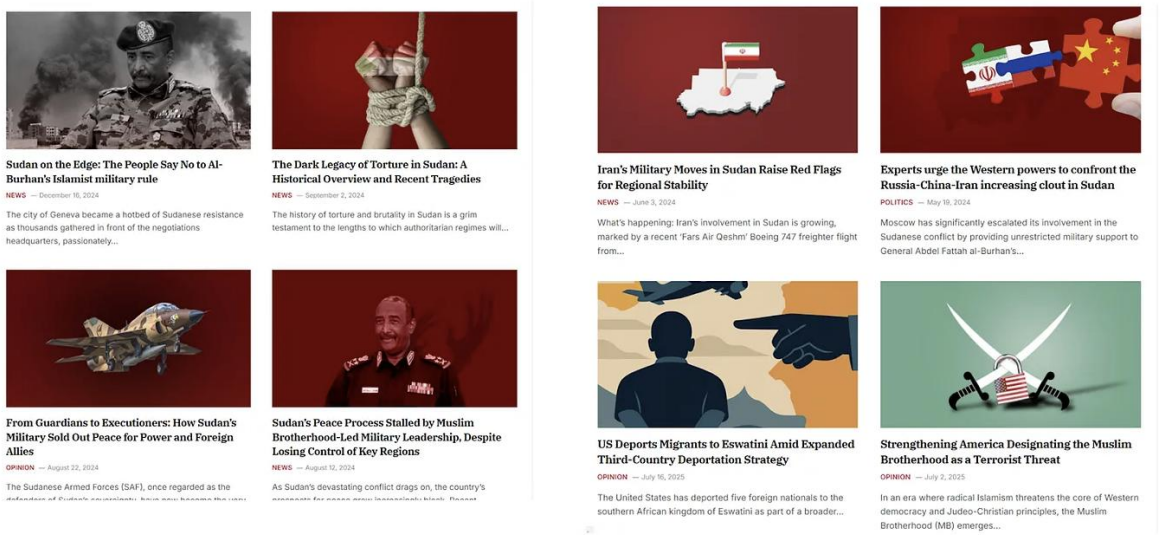


Figure 1 Examples of articles about Sudan on a pseudo-news outlet

[27] Thomson Found., ‘How Disinformation Is Shaping Sudan’s Conflict’.

The present study differs in that it moves beyond event-level monitoring to forensically to do the following; identify fake accounts, map the network topology, longitudinal persistence, and AI-enabled adaptation of the influence infrastructure itself.

What this report adds

This report is focused on bots and sockpuppets, as opposed to broader forms of information manipulation. This report offers important insights for those engaged in detecting and disrupting digital influence operations.

1. The analysis of network one covers several years. By tracking accounts, narratives, and interaction patterns over extended periods, including phases prior to large-scale asset deletion, this study is able to observe persistence, identity repurposing, and strategic adaptation. This temporal depth allows influence activity to be analyzed as an ongoing infrastructure rather than as a series of isolated campaigns, addressing a core limitation of event-driven disinformation research.

2. The report combines network topology with longitudinal behavioral analysis to identify coordination that would not be visible in static network snapshots alone. Ostensibly independent accounts were initially flagged as sockpuppets on the basis of behavioral and identity indicators. Subsequent account timeline analysis revealed repeated interaction with shared, low salience “filler” accounts, which functioned as latent connective markers across otherwise separate regional clusters. This approach makes it possible to detect coordinated control and shared infrastructure even where direct interaction among political accounts is limited or deliberately obscured.

3. The report employs country-coded, multilingual corpus analysis. By coding each post for primary and secondary country focus and analyzing narrative polarity, repetition, and omission at scale, the study demonstrates political stance, allowing for attribution to be more accurately inferred.

4. The integration of corpus findings with network structure enables greater insights into TTPS. Narrative patterns map directly onto identifiable network cells, while shifts in framing correspond closely to geopolitical events and platform affordances.

5. The report documents the operational uptake of platform affordances including the strategic use of credit-based verification to amplify reach.

6. The identification of AI-generated and AI-assisted accounts further illustrates how influence operations increasingly rely on synthetic identity at scale, lowering the cost of narrative saturation while complicating attribution. It also identifies how networks adapt to new technologies when producing deception.

7. Focusing on one state alone might be blinkering. Longitudinal analysis, and thematic analysis (e.g. anti-Islamist themes) allows the detection of assets seemingly unrelated to one another - but structurally coordinated. This is shown in how certain country-focused components have been deprecated, as if they are functionally separate from the rest of the network.

8. The report does perhaps challenge the argument that SAF propaganda alone chiefly relies on scale. These RSF-aligned networks - certainly in the case of El Fasher–are massive in scale.

Taken together, the three networks illustrate three slightly different "bot problems," each revealing a different mechanism of digital information harm. Network One shows bots and sockpuppets as region-wide *narrative infrastructure*: a durable, cell-based system sustaining narrative alignment over months and years through identity recycling, filler-account linkages, and selective amplification. Network Two shows bots as reactive *shock troops*: an event-driven swarm that activates during crisis windows (i.e., post–El Fasher) using scale and velocity to engage in large-scale crisis management. Network Three shows bots more as *social actors*: smaller in scale but more conversational, using AI-assisted personas to enter replies, quote-tweets, and interpersonal exchanges with journalists and influencers, simulating civic participation and steering interpretation from within. In combination, the networks underscore that "bots" are not a single phenomenon: they can function as long-term influence infrastructure, rapid-response visibility engines, or interactive pseudo publics, and each mode demands different detection logics and governance responses.

The findings underscore persistent deficiencies in platform regulation and prevention of coordinated inauthentic behavior. Despite years of public commitments to transparency and moderation, large-scale sockpuppet networks continue to operate for extended periods, particularly in conflict-affected and non-Western information environments. Enforcement remains uneven and reactive, often occurring only after significant harm or external exposure. This pattern in part reflects both enforcement failure and structural bias in content moderation that privileges English-language disinformation while overlooking coordinated harm in Arabic and other languages. In doing so, platforms risk becoming complicit in enabling digital authoritarian practices that distort political discourse, and launder propaganda at scale. The evidence presented in this report demonstrates that contemporary influence operations in the MENA region are not marginal or episodic phenomena, but durable components of modern information warfare.

Across all three networks, influence operations are shown to be actively adapting to generative AI and evolving platform affordances rather than merely exploiting static automation. Network One has adapted over time, exploiting verification, handle-switching, AI-use and other techniques to endure for years. Network Two appears to have gamed X's ability to detect account location, while scaling AI to provide large volumes of unique content. Network Three reveals the most recent shift: the operational use of AI-assisted and AI-generated personas to simulate conversational participation, sustain reply-driven engagement, and insert disciplined narratives directly into interactions with journalists and other users. Collectively, these networks reaffirm that contemporary deception is not technologically static but co-evolves with platform governance and design choices.

While this report does not claim direct operational control by any state, the sustained narrative alignment observed across networks is most consistently compatible with UAE geopolitical interests. Across conflicts and regions, the networks systematically promote actors and outcomes favored by

Abu Dhabi (e.g. RSF whitewashing, temporary rehabilitation of Bashar Al-Assad, anti-Ikhwan framing, legitimization of authoritarian governance, and excessive positive portrayal of UAE humanitarian and diplomatic roles) while insulating the UAE from criticism and externalizing blame onto its rivals. This pattern of selective amplification and omission provides a strong basis for inferred alignment, even in the absence of dispositive evidence of command-and-control.

Network One

Overview

Network One represents the largest and longest running influence operation identified in this investigation, comprising approximately 310 main sockpuppet accounts supported by over 8000 bot accounts. Around 170,000 tweets and retweets across December 2022 through December 2024 were downloaded and analyzed. The network operates across multiple platforms (primarily X), multiple languages (Arabic, English, Farsi, Turkish, and French), and more than a dozen national contexts, indicating both scale and strategic breadth. The network functions essentially as a UAE-aligned MENA-wide news network, with cells (groups of sockpuppets) dedicated to posting about specific countries in the region (from Mauritania to Saudi).

Accounts within this network were identified and tracked longitudinally using a combination of behavioral, network, and content-based indicators, allowing patterns of coordination, repurposing, and persistence to be verified over time. All posts were coded for primary and secondary country focus, enabling systematic analysis of geographic targeting and regional prioritization. Network analysis further revealed a cell-based architecture, in which clusters of accounts concentrate on specific countries while remaining integrated within a shared amplification infrastructure. The analysis of Network One examines the narratives, the network structure and account behavior. It is worth noting that over the years this network has had multiple assets suspended, often following being exposed by journalists or other analysts and activists. However, its resilience is notable. As of writing, approximately 184 of the 310 sockpuppet accounts are still active here. The true scale of the network is likely much larger.

Corpus analysis

To assess the network's political and geographic orientation/focus, a corpus frequency analysis was conducted. Sudan dominates the dataset by a wide margin (9,651 mentions), confirming it as the network's primary focus. High-frequency terms such as "support," "Sudanese," "army," and "forces" indicate sustained attention to the conflict and its military dynamics, while repeated references to the "Muslim Brotherhood" (4,490 mentions) and the "UAE" (6,667 mentions when combined with "Emirates") point to a consistent regional and ideological framing aligned with Gulf political narratives.

Beyond Sudan, the network exhibits secondary but sustained attention to Yemen (3,142 mentions), Egypt (2,704), and Tunisia (2,647), followed by Libya and Gaza. A further tier includes Mauritania, Syria, and Algeria, with lower-frequency references to Iran, Jordan, Morocco, Palestine, Lebanon, and Saudi Arabia. Overall, the distribution indicates a Sudan-centred corpus embedded within a broader, transnational narrative space structured around Gulf, Islamist, and regional security themes.

Table of country frequency

Country	Mentions	Country	Mentions
1. Sudan	9,651	9. Syria	1,390
2. United Arab Emirates	6,667	10. Algeria	1,305
3. Yemen	3,142	11. Iran	982
4. Egypt	2,704	12. Jordan	975
5. Tunisia	2,647	13. Morocco	943
6. Gaza*	2,269	14. Palestine	922
7. Libya	2,155	15. Lebanon	888
8. Mauritania	1,755	16. Saudi Arabia	870

Timeline of activity

Mentions of Sudan are the most voluminous, directly corresponding to the Sudanese civil war that erupted in April 2023. The tweet volume shows a gradual buildup during the initial months of the conflict, before spiking dramatically around March-April 2024, reaching a peak of nearly 450 tweets per week. This spike likely coincides with the initial siege of El Fasher, a significant development in the civil war that would later lead to a massacre. After this peak, Sudan-related tweets remain at elevated levels (200-350 per week) through the end of 2024, reflecting the ongoing nature of the conflict.

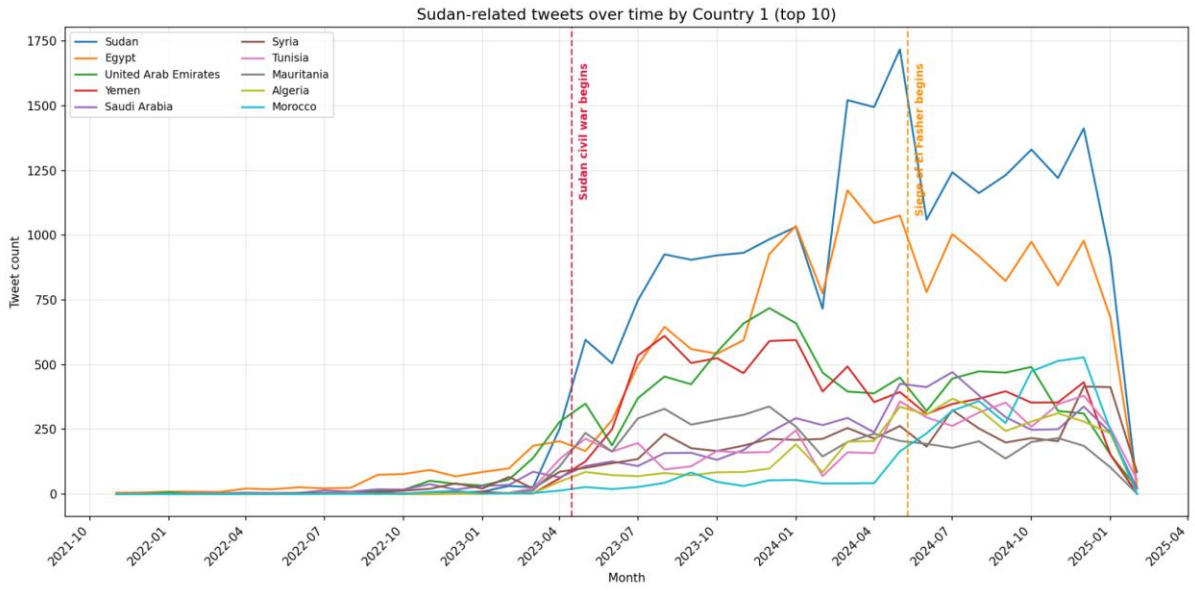


Figure 2 Total tweets by country over time

Egypt shows consistent attention throughout the period, with a noticeable increase beginning in late 2023 that peaks around March-April 2024. The United Arab Emirates maintains a moderate but

consistent presence throughout the timeline, generally ranging between 50-150 tweets per week, without the dramatic spikes seen for Sudan. This steady coverage may reflect the network's ongoing interest in portraying the UAE in a consistent manner regardless of regional crises or country focus.

Yemen shows a pattern of gradually increasing attention starting from mid-2023, maintaining a steady presence in the 75-150 tweets per week range through 2024. The graph reveals how the network's messaging priorities shifted dramatically in response to the Sudan civil war, with content about Sudan dominating the network's output during the height of the conflict in 2024.

Geographies

The analysis of tweets from a sockpuppet network reveals concentrated attention on specific countries, indicating strategic geopolitical interests. All posts were coded for primary and secondary country focus, with na (not applicable) being the dominant category - so the figures represented here are conservative. Their engagement metrics were then mapped onto countries of focus.

Sudan emerges as the primary focus, with 15,212 tweets, and over six million views, suggesting intense activity or manipulation efforts related to the civil war. Egypt follows with 8,401 tweets, reflecting its significant role in the network's agenda. The United Arab Emirates, Yemen, and Tunisia also receive substantial attention, with tweet counts of 6,533, 5,219, and 3,513 respectively, highlighting their strategic importance to the network's objectives. Other countries such as Mauritania, Syria, Saudi Arabia, Algeria, and Iran are prominently featured, each with over 1,900 tweets, pointing to sustained interest or influence operations in these regions. Libya, Morocco, and Turkey also appear frequently, suggesting targeted influence or engagement related to specific issues or events.

Country 1	Comment Count	Count of Tweets	Like Count	Retweet Count	View Count	Total
Algeria	1,232	2,167	27,790	826	344,045	360,167
Egypt	6,689	8,404	70,362	5,104	2,034,483	2,076,322
India	561	300	2,374	59	32,252	33,101
Iran	1,208	1,904	104,901	12,535	352,668	410,059
Iraq	135	538	3,750	138	85,324	87,323
Israel	170	375	3,165	148	123,573	125,835
Jordan	144	838	6,745	138	144,757	148,709
Lebanon	1,016	1,081	14,084	282	180,146	188,912
Libya	863	1,889	10,866	337	580,038	586,738
Mauritania	1,817	2,847	48,136	1,373	531,540	559,603
Morocco	463	1,651	8,668	222	375,995	380,964
Palestine	247	928	8,879	482	142,458	148,815
Saudi Arabia	6,696	2,693	30,140	953	592,063	613,900
Somalia	348	490	4,321	168	10,195	10,382
Spain	38	312	1,558	50	43,693	44,627
Sudan	13,330	15,212	82,795	3,131	3,943,059	3,998,551
Syria	4,680	2,805	37,393	1,652	538,861	562,294
Tunisia	1,728	3,513	39,020	904	745,744	768,457
Turkey	790	1,206	22,132	721	177,427	180,188
United Arab Emirates	4,159	6,570	124,590	6,224	1,350,578	1,429,303
United States	550	548	3,278	111	100,645	101,853
Yemen	816	5,223	26,887	932	1,484,205	1,499,293
Grand Total	47,680	61,494	681,834	36,490	13,913,749	14,315,396

Figure 3 Countries by engagement

Moderate engagement with countries like the United States, France, and Germany, along with Iraq and Somalia, indicates more focused, perhaps event-driven interactions by the network. This dataset not only underscores the geographic spread of the sockpuppet network's operations but also highlights the specific countries of concern, providing insights into the strategic priorities and manipulation tactics employed by the network on social media.

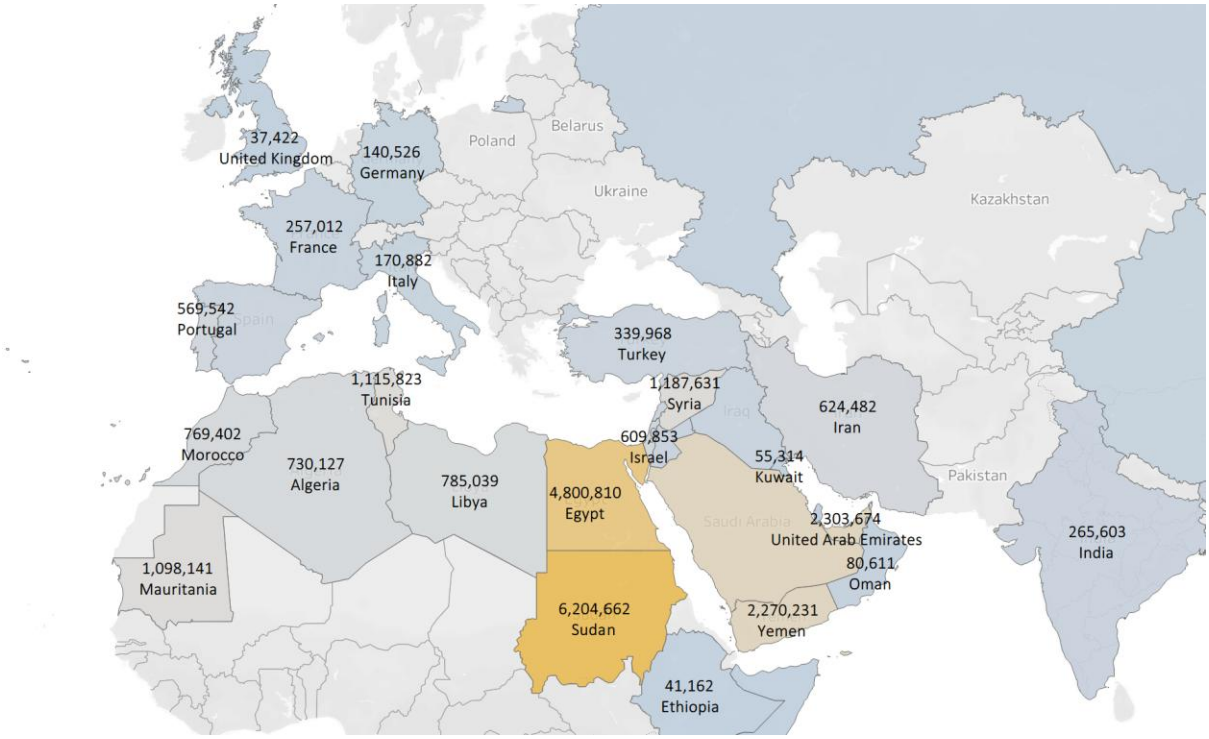


Figure 4 Map of engagements per mentioned country

Engagement

Despite Sudan getting the most mentions in terms of overall tweets, it did not necessarily get the most engagement. Iran had the most total engagement actions per tweets (i.e. replies, retweets, likes), while Turkey had the most views, followed by Yemen and Sudan.

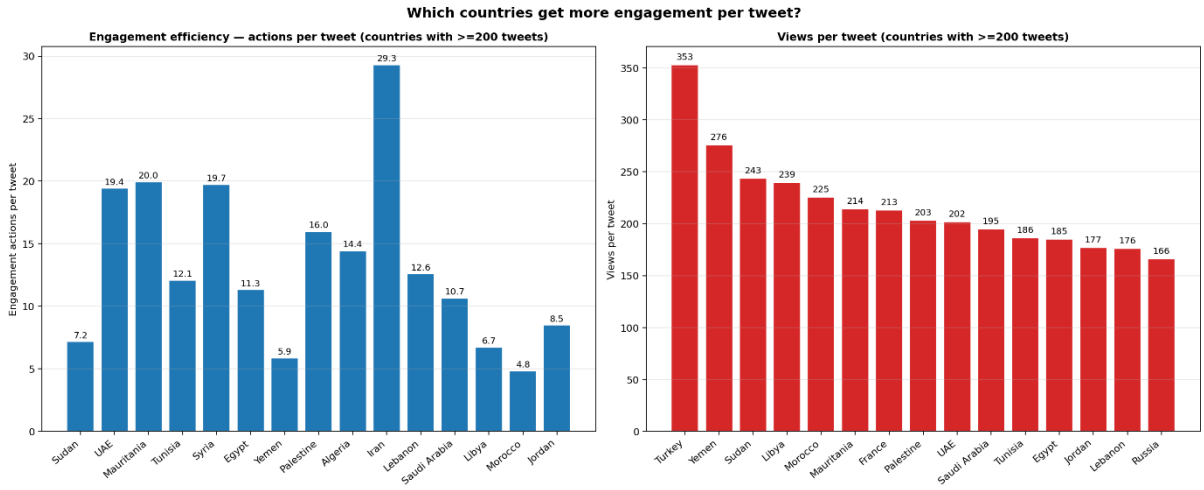


Figure 5 Tables showing which countries get the most engagement

Diplomacy and foreign relations

By coding tweets for country 1 and country 2 (where applicable), it was possible to create a directional analysis that highlights how tweets often expressed aspects of relationships between specific countries. This can be useful for attribution and determining the purpose of the network. As the radial diagram shows, the UAE and Egypt's relationship with other countries was a large part of the focus of the network. Given the praise of the UAE in particular, along with the focus on its relations with other countries, it appears the network is highly aligned with UAE foreign policy.

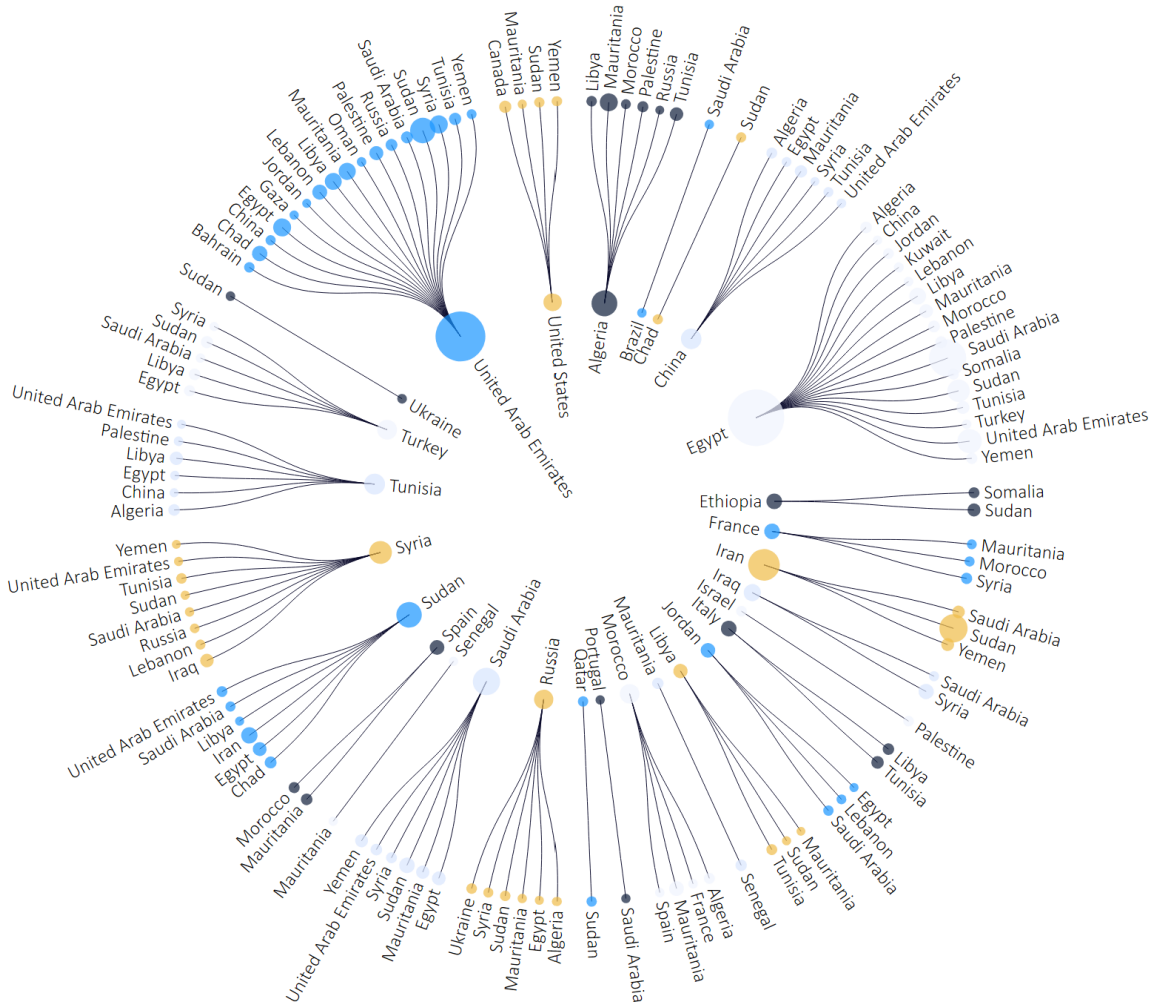


Figure 6 Flow diagram showing mentions of two or more countries per post (where applicable)

Cell structured network and regions

Network analysis was conducted by constructing edge lists based on tweets, retweets, and replies to interactions of all the seed accounts. In other words, the account interactions were mapped to see if there was overlap in terms of what information these accounts shared. These data were visualized and analyzed using Gephi. Although the accounts initially appeared independent, they were first identified as sockpuppets on the basis of behavioral and identity indicators rather than network

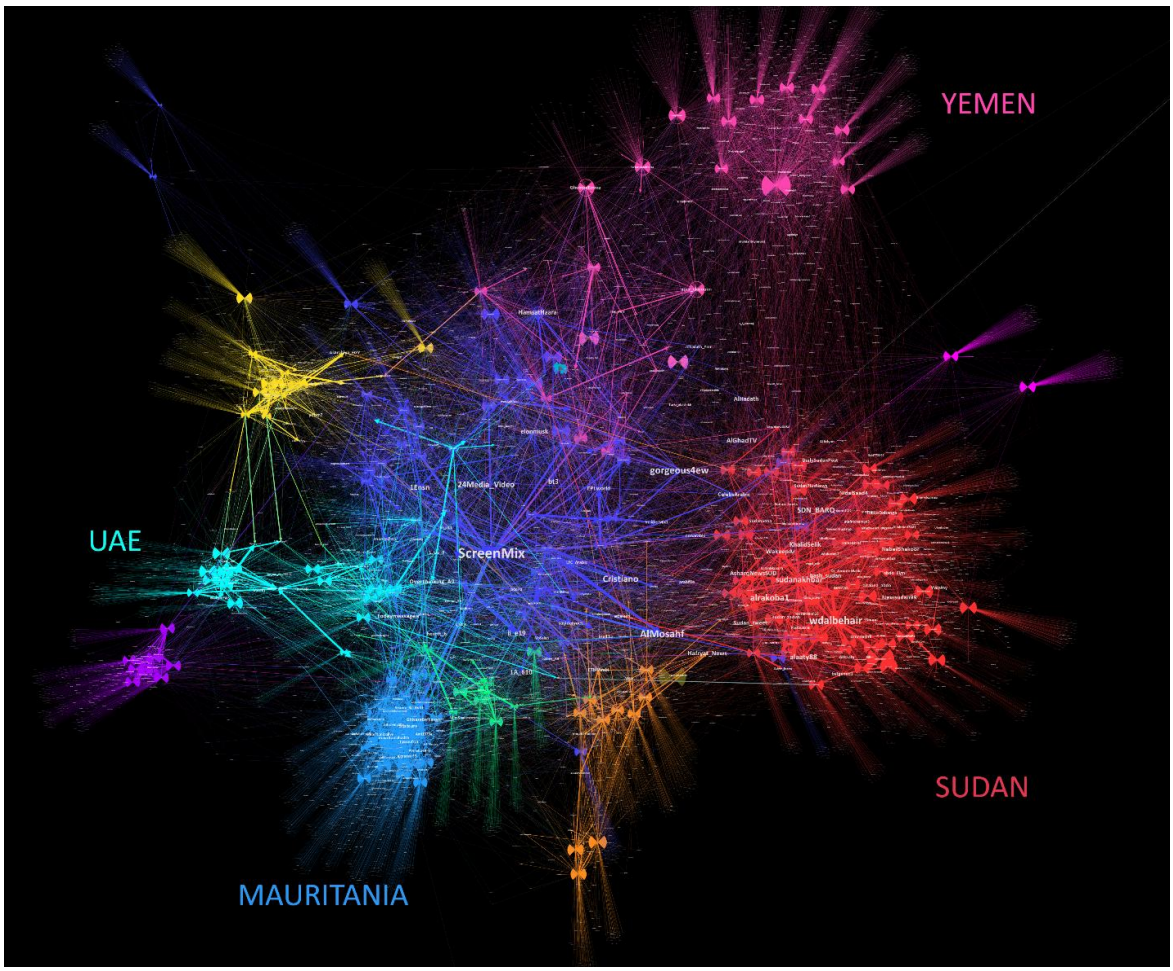
position alone. Crucially, longitudinal timeline analysis revealed patterns of coordination that were not immediately visible in static network snapshots. Over time, multiple accounts repeatedly retweeted and interacted with the same generic, non-political "filler" accounts, such as *ScreenMix (an account focusing on entertainment and popular/viral videos)*,which did not themselves promote strong political narratives. These shared interactions functioned as latent connective markers, linking otherwise separate regional clusters. It is also common for disinformation accounts to follow the so-called 80:20 formula, where 80% of the content they post is generic, non-political information. The other 20% is propaganda. [28]

While such filler accounts attracted relatively low analytical attention in isolation, their repeated reuse across disparate sockpuppets and country-focused clusters provided evidence of shared operational control or common content pipelines. About half the network clusters around country-specific external hubs (one cell per country: Sudan/Mauritania/Tunisia/Syria/Libya/Egypt); the other half clusters around generic Arabic entertainment hubs that look like camouflage activity. The Sudan cluster is the largest country-specific cell and the one where the most accounts in the dataset were suspended by X.. This pattern is consistent with a coordinated influence operation in which accounts are deployed across regional "desks," while drawing from a common pool of auxiliary accounts to maintain activity, mask coordination, and sustain network cohesion over time. The largest and most densely connected cluster centred on Sudan, with other clusters corresponding to additional regional focus areas. In other words, accounts claiming to be from, for example, Tunisia, tended to interact with one another, in part to create separation from the broader network. These small 'cells' would frequently repost news or cultural accounts related to their respective country. Identifying these cells

[28] 'Erving Goffman on Misinformation and Information Control: The Conduct of Contemporary Russian Information Operations - Innes - 2022 - Symbolic Interaction - Wiley Online Library'.

was essential for understanding how narratives were compartmentalized by region while still benefiting from shared infrastructure, amplification mechanisms, and operational control.

The cell clusters are usually defined by having obvious markers of country allegiance, whether flags, location information, national dress, or images of the country in the bio.



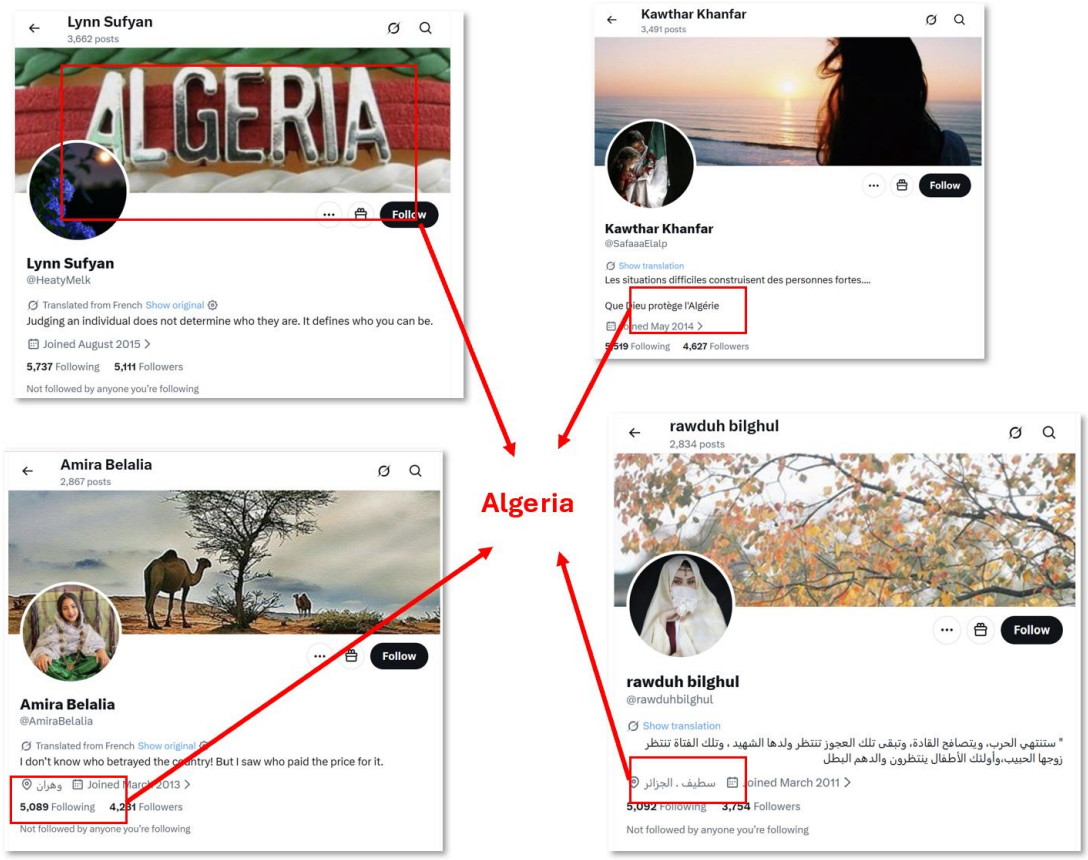


Figure 8 Example of Algeria-focused accounts

Temporal analysis of the post-timing indicates all the cells, regardless of country they represented, appear to be posting working hours on UTC+3, which is Arabian Standard Time, or Egypt time. Had these truly been organic posting from their respective countries, variations in posting patterns would be expected. Mauritania-targeted accounts peak 4 hours before local lunchtime in Nouakchott. Tunisia-targeted accounts peak 3 hours before local lunchtime in Tunis. Sudan-targeted accounts peak 2 hours early. Syria-targeted accounts are nearly right (because Damascus = UTC+3 = operator clock). The further the target country is from AST, the bigger the timing mismatch.

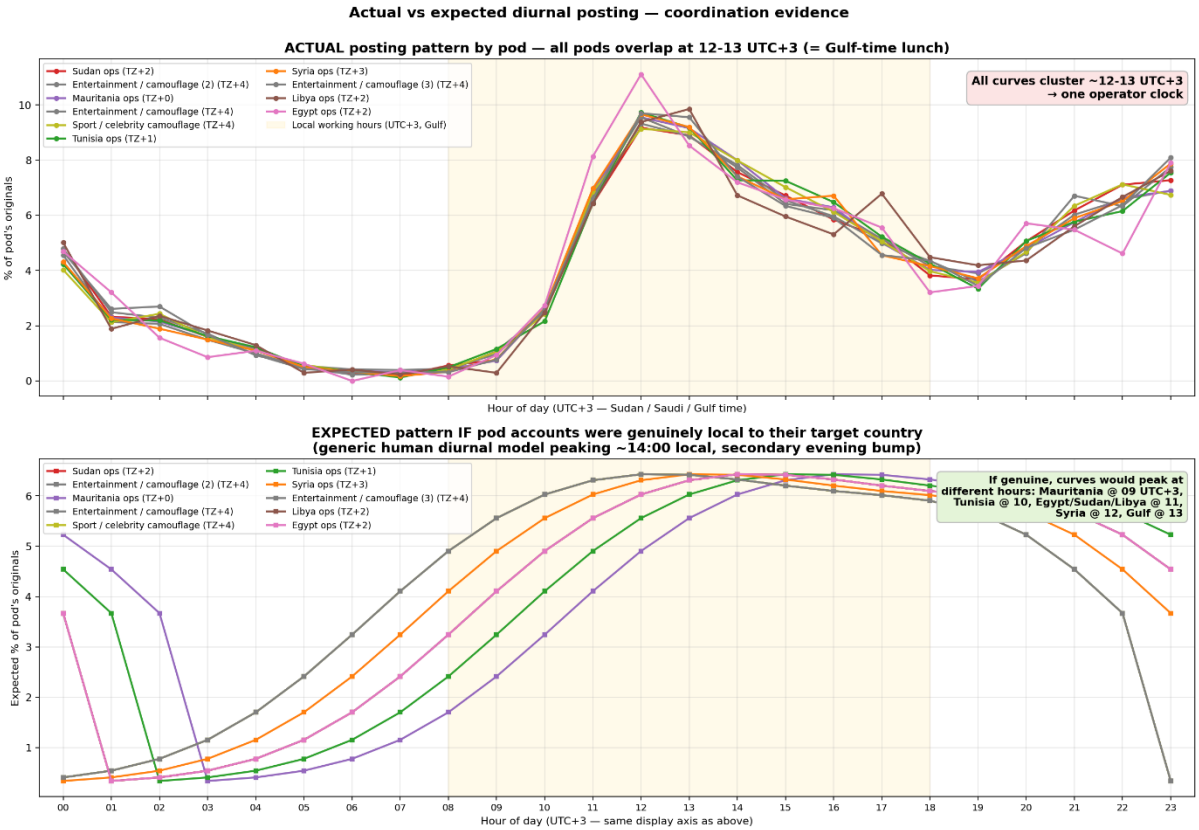


Figure 9 Graph showing actual versus expected diurnal posting of network

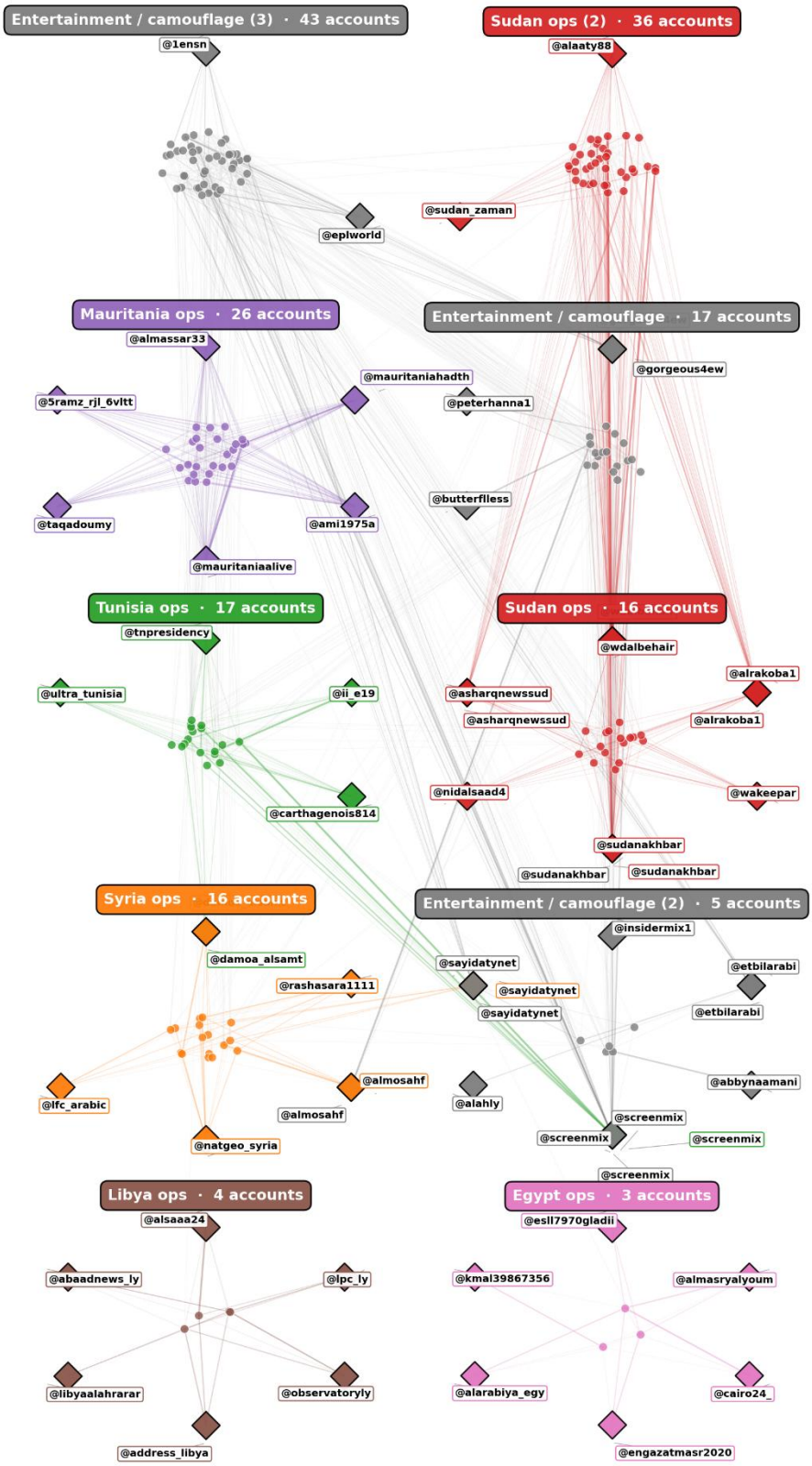


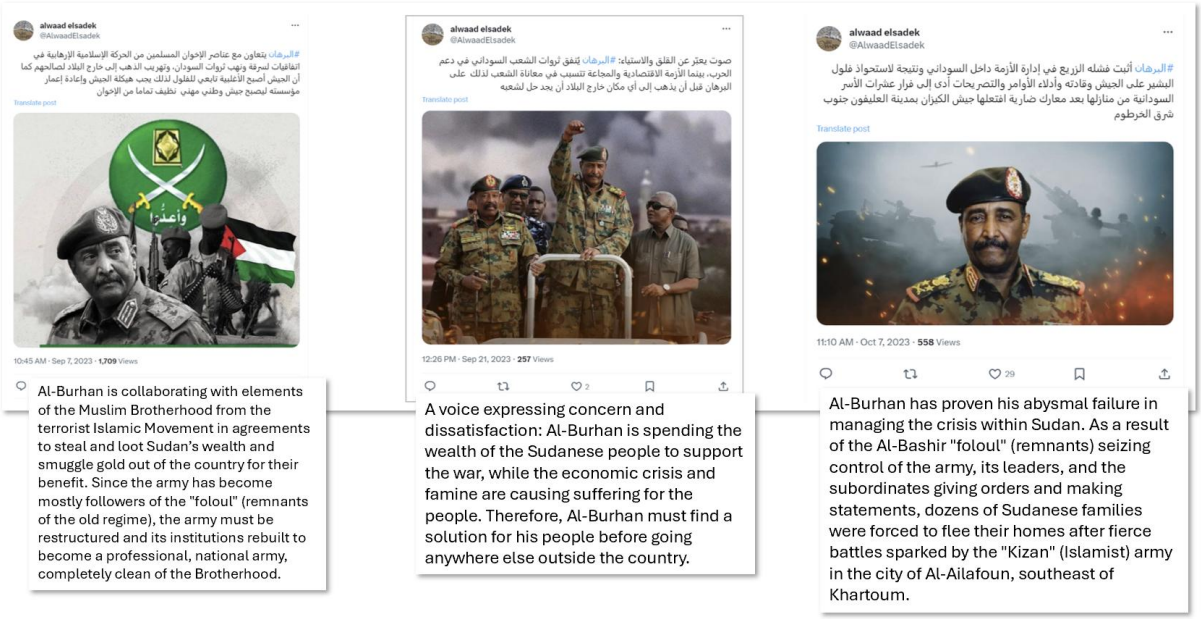
Figure 10 Network graph showing various country cells interacting with one another and promoting regional news outlets

Narrative structure and geopolitical alignment

Primary storylines: Sudan

The Sudan-related corpus centers on two principal antagonists: Abdel Fattah al-Burhan, head of Sudan's Sovereignty Council and commander of the Sudanese Armed Forces (SAF), and Mohamed Hamdan Dagalo ("Hemeti"), leader of the Rapid Support Forces (RSF), alongside civilian political actors such as former Prime Minister Abdullah Hamdok.

Across the tweets, Sudan is framed through a highly polarized and internally consistent narrative architecture that divides political actors into moral absolutes rather than contested positions. Abdel Fattah al-Burhan and the SAF are constructed as the primary agents of violence, obstruction, and moral failure. Responsibility for civilian harm, famine, and diplomatic breakdown is repeatedly personalized onto Burhan, who is portrayed as illegitimate, corrupt, and subservient to the Muslim Brotherhood and foreign patrons. He is routinely described using totalizing moral language, including labels such as "war criminal," "terrorist," and "traitor," while allegations of indiscriminate bombing, starvation tactics, and deliberate sabotage of peace efforts are presented as settled fact rather than contested claim. As one post puts it, Burhan is "ready to burn Sudan for 100 years just to remain in power," a formulation that collapses strategic, moral, and temporal responsibility into a single figure of blame.



In keeping with other analysis on Sudan's disinformation, those aligned with the SAF are often portrayed as elites. The term 'remnants' (falul) is often used to describe and delegitimize those seen as still being part of Omar al-Bashir's era.

In stark contrast, Mohamed Hamdan Dagalo (Hemedti) and the RSF are depicted as legitimate, humanitarian, and peace-seeking actors. Hemedti is framed as a statesman engaged in diplomacy and committed to civilian protection, frequently described as a "leader of peace" and a "symbol of courage

and humanity." RSF-controlled areas are associated with stability, service provision, and normalization, with repeated claims that "life is returning" and that the forces are "protecting civilians and restoring dignity." This rehabilitative framing is often delivered through overtly sentimental and formulaic humanitarian tropes. One widely circulated example praises an RSF fighter who "carried an elderly woman to the car as if she were his own mother," concluding that this act reflects the "noble ethics of the Rapid Support Forces and what Hemedti has instilled in them."

Beyond praise, some accounts move into anticipatory legitimation, referring to Hemedti with quasi-state honorifics such as "Field Marshal" or "President," and circulating claims about a forthcoming "parallel peace government" to be announced "from inside Khartoum." These formulations function less as commentary on existing power than as a rehearsal for alternative sovereignty, positioning RSF authority as both inevitable and already normalized.

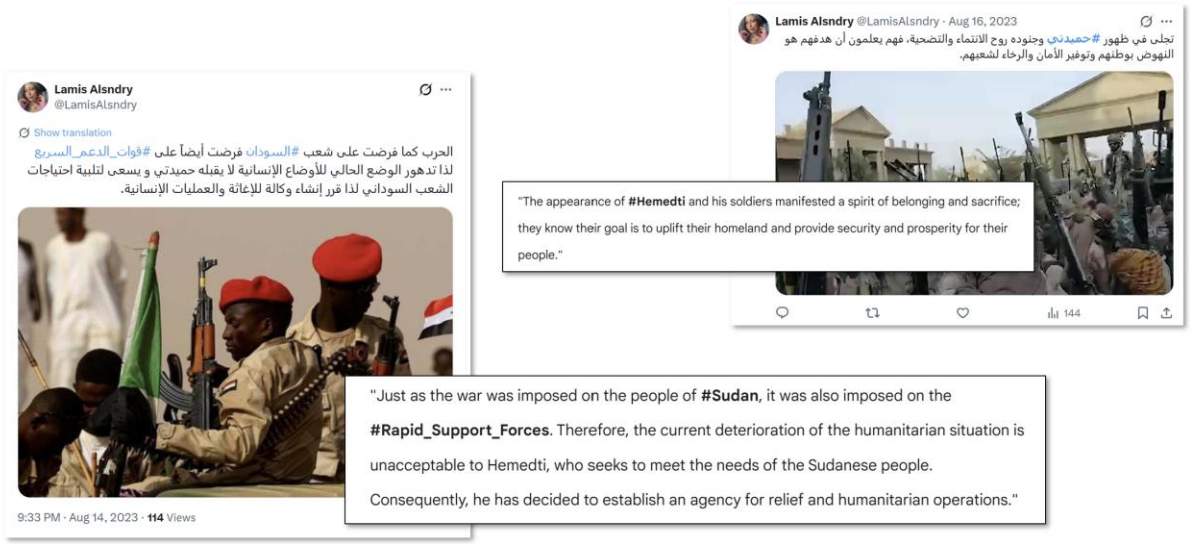


Figure 11 Example of pro-RSF post

Civilian political actors aligned with democratic transition, particularly Abdullah Hamdok and the Coordination of Civil Democratic Forces, are positioned as complementary to RSF legitimacy. Hamdok is repeatedly framed as "the voice of reason" and "the hope of civilians," working in tandem with RSF efforts to end the war through dialogue. International organisations, by contrast, appear largely as ineffective or obstructed, with humanitarian failure attributed overwhelmingly to the SAF, which is accused of "blocking aid," "weaponizing hunger," and "bombing hospitals." This attribution further consolidates the moral asymmetry at the heart of the narrative.

Equally significant is what is absent. Across the corpus there is no sustained internal critique of RSF conduct, leadership, or historical lineage. Reports of RSF abuses are either omitted or dismissed as fabrications carried out by "Burhan's militias in RSF uniforms." This systematic silence, alongside the uniformity of praise, indicates a disciplined narrative environment shaped as much by omission as by repetition.

Taken together, this narrative structure promotes one side of the conflict through repetition, omission, emotive language, and selective quotation. The corpus advances a rehabilitative framing of

the RSF while externalizing blame for mass violence. The result is a form of narrative saturation in which scale and effect displace contestation, accountability, and complexity.

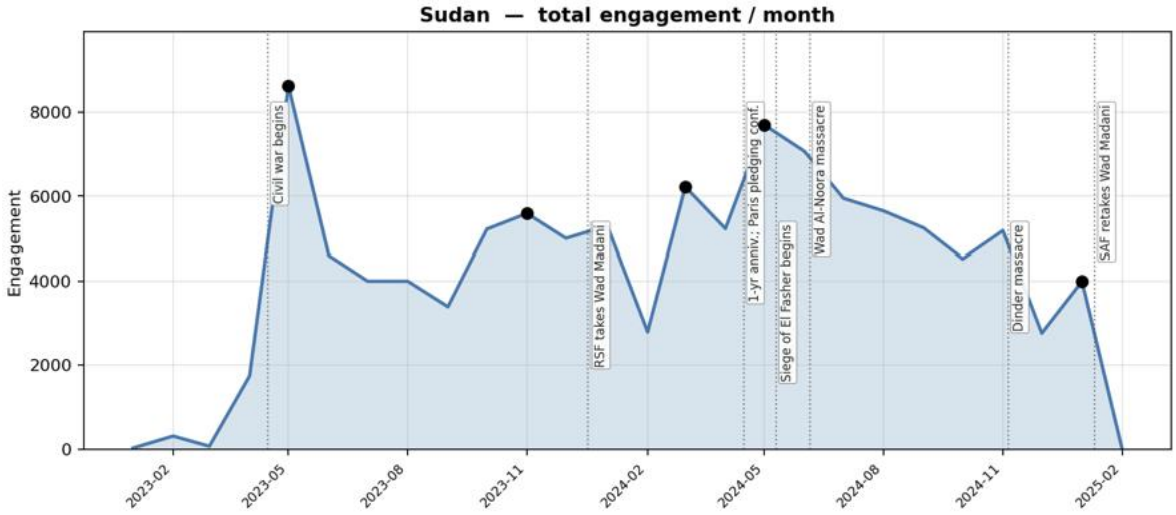


Figure 12 Graph showing peaks in Sudan-related activity and their relationship with key events

'Kezans'

The network frequently uses the term الكيزان "Kezan/Kizan" (singular: *koz*) as a derogatory attack on the SAF and its followers. It is a flexible and evolving term in Sudanese political discourse. At its core, it refers to supporters or affiliates of the former ruling National Congress Party, the Islamist party that governed Sudan for three decades under Omar al-Bashir. The label does not necessarily imply formal party membership; it can apply more broadly to people associated with, supportive of, or benefiting from the former regime.

The name itself derives from the Sudanese word *koz* (metal cup), which comes from a famous phrase by an NCP member comparing Islam to a river and the *Kezan* to the vessels used to drink from it. Over time, the term expanded beyond the NCP to include figures linked to the wider Sudanese Islamist movement, including factions such as the People's Congress Party associated with Hassan al-Turabi.

In contemporary usage, however, "Kezan" has taken on a much broader and more pejorative meaning. In some contexts, the term extends beyond formal politics altogether, referring to individuals perceived as intolerant, anti-progressive, or complicit in the culture of graft and patronage associated with the former regime.

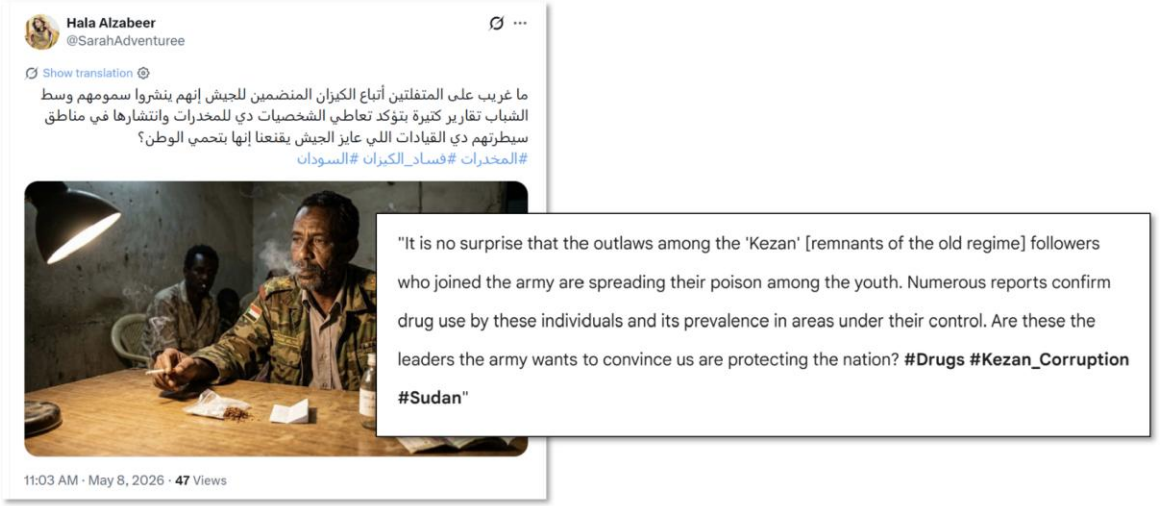


Figure 13 Example of use of the term 'Kezan'

Rapid Support Forces

The overwhelming majority of tweets and statements in this collection present the Rapid Support Forces (RSF) in a distinctly positive light. This dominant narrative portrays the RSF as heroic defenders of the Sudanese people, committed humanitarian actors, and champions of democratic reform fighting against what they repeatedly call "remnants" of the former regime.

Throughout the document, the RSF are predominantly characterized as protectors who secure neighborhoods, distribute aid, restore essential services, and safeguard civilians from harm. They are frequently depicted as peace-seeking, with numerous messages highlighting their willingness to participate in ceasefires and negotiations while claiming their opponents escalate conflicts. A significant number of tweets celebrate supposed RSF military victories, describing them "liberating" various regions across Sudan.

The RSF are repeatedly characterized as "brave heroes" and "brave warriors" who serve as "protectors of the nation" and "guardians of the glorious revolution." They are frequently described as the "shield of the nation" and "defenders of democracy," with an emphasis on their supposed role safeguarding civilians from harm. The phrase "readiness, speed, decisiveness" appears as a slogan in numerous tweets, seemingly functioning as an official motto.

Other common descriptors include "forces of truth," "heroes of the homeland," "guardians of peace and security," and "liberators" of various regions across Sudan. The RSF are consistently portrayed as "humanitarian forces" providing essential services and aid, a "fortress of security," and a "strong shield for civilians." Many messages characterize them as a "symbol of unity and loyalty" and the "backbone of security and stability" in Sudan.

This language creates a heroic narrative around the RSF, presenting them as an "unbeatable force" of "faithful guardians" and "sons of the nation" dedicated to protecting democracy and civilian welfare. This remarkably consistent and glowing portrayal appears in tweet after tweet, suggesting a

coordinated messaging campaign designed to shape public perception of the paramilitary group in the most favorable light possible.

Noticeably absent from this collection are substantive criticisms of the RSF. When negative actions are mentioned, they are almost invariably attributed to opponents disguised as RSF members or dismissed as fabrications. The one-sided nature of these descriptions stands in stark contrast to reports from international human rights organizations about the group's activities.

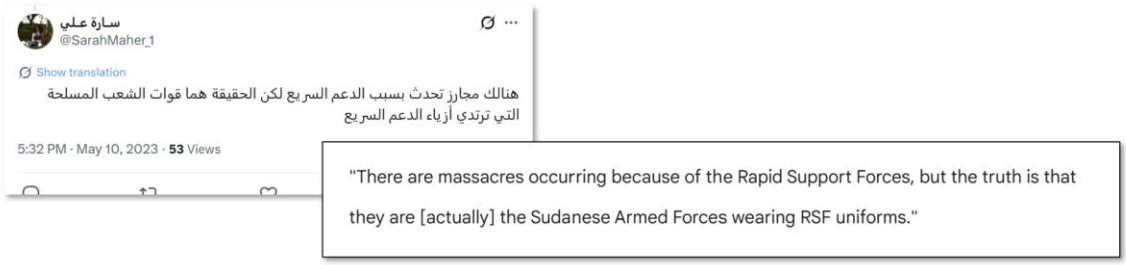


Figure 14 Example of propaganda deflecting blame from RSF

Explicit boosting of RSF social media assets

This network also actively amplified and boosted official RSF accounts, before they were suspended by X. This included RSF's official account, and its spokesperson, Wad Elbehair, both of which functioned as central amplification hubs. Within the Sudan-focused subset, pro-RSF accounts were the most heavily promoted, underscoring the network's role in systematically boosting RSF-aligned narratives prior to the removal of assets.

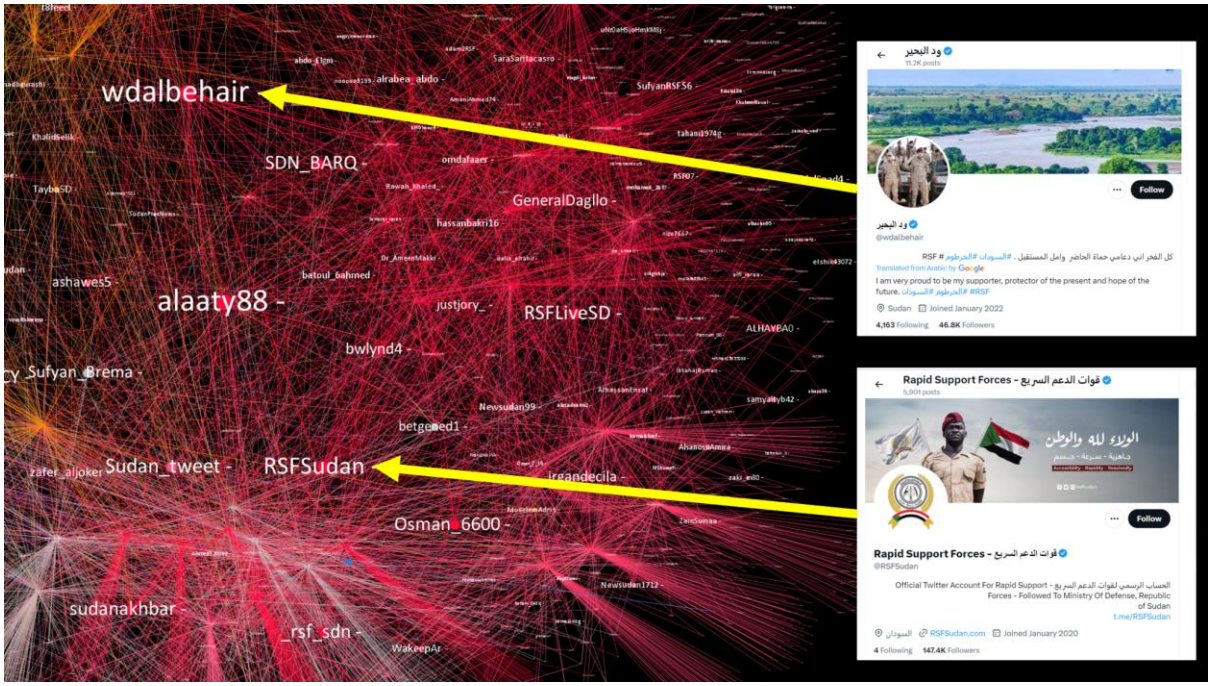


Figure 15 Network diagram showing how network boosts official X accounts for Rapid Support Forces.

Eliminationist cleansing rhetoric

Some of the language in this dataset presents a clear risk of incitement and the normalization of extreme violence, particularly through the repeated use of cleansing and purification metaphors to describe military action and political opponents. For example, multiple posts celebrate armed operations as efforts to "cleanse Omdurman of criminals and outlaws" or to "completely cleanse the area," framing violence as a necessary act of restoration rather than as force subject to legal or humanitarian limits. In a conflict setting, where armed actors exercise coercive power without judicial oversight, such framing risks legitimizing extrajudicial violence by recasting it as protective or corrective.

More acute risks emerge where this justificatory language escalates into explicit dehumanization and eliminationist framing. In one of the most extreme examples, a group is described as "pigs… filth… disease," followed by the assertion that "Sudan will be cleansed of all of you forever." This combination of animalistic, pollution metaphors, and permanence is particularly concerning. Similarly, calls to "purify the East from the filth of the separatist forces" apply cleansing logic to an entire region, aligning violence with territorial purification rather than limited military objectives. Atrocity-prevention literature consistently identifies such language as a key indicator of elevated risk, as it erodes distinctions between combatants and civilians and frames coexistence as impossible.

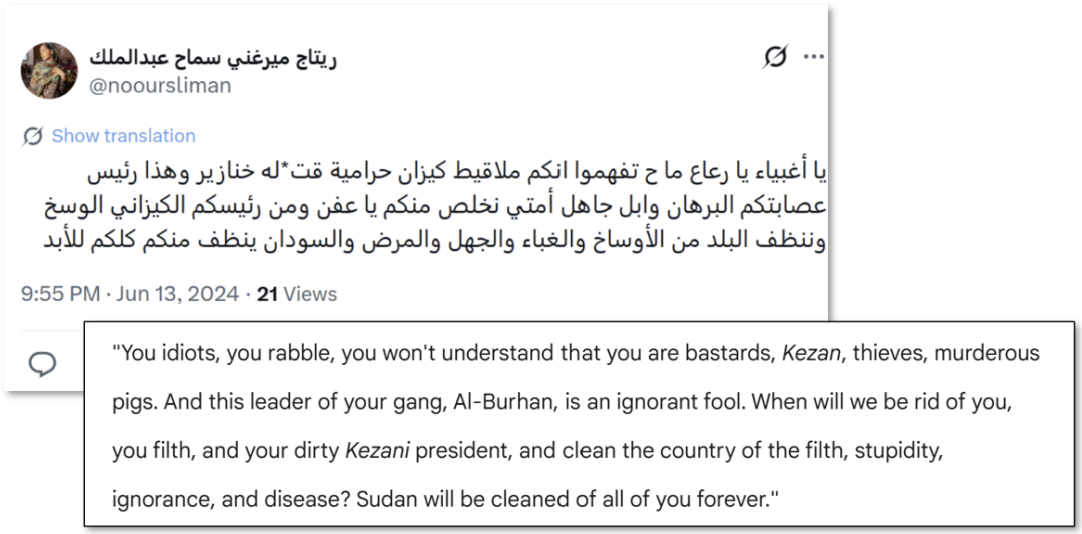


Figure 16 Example of eliminationist rhetoric

Statements asserting that cities must be "reclaimed and purified from the filth of the remnants" or that areas require "complete cleansing" suggest totalizing outcomes rather than proportional or time-bound interventions. Even when framed as liberation or reform, this rhetoric creates a permissive environment in which forced displacement, collective punishment, or indiscriminate violence can be rationalized as necessary steps toward national renewal. The cumulative danger, therefore, lies not

only in explicit insults or threats, but in the way repeated metaphors of purification and removal work together to normalize the idea that large-scale violence against perceived enemies is justified, unavoidable, and ultimately beneficial.

Source (Handle)	Original Arabic Text (Full)	English Translation	Why This Is Egregious
nooursliman	يا أغبياء يا رعاع ما ح تفهموا انكم ملاقيط كيزان حرامية قت*له خنازير وهذا رئيس عصابتكم البرهان وابل جاهل أمتي نخلص منكم يا عفن ومن رئيسكم الكيزاني الوسخ وننظف البلد من الأوساخ والغباء والجهل والمرض والسودان ينظف منكم كلكم للأبد	"You idiots, you riffraff, you will never understand that you are bastard Kezans, thieves, murderers, pigs. And this is your gang leader Burhan, a fool and an ignorant. When will we get rid of you, you filth, and your filthy Kezan leader, and cleanse the country of dirt, stupidity, ignorance, and disease—so that Sudan is cleansed of all of you forever."	**Severe dehumanization and eliminationist rhetoric.** Uses animalization ("pigs"), pollution metaphors ("filth," "disease"), collective targeting ("all of you"), and permanence ("forever"). This mirrors classic genocidal language that frames violence as hygienic necessity.
Marawwhabibakr (suspended)	اهل شرق السودان ينتظرون لحظة قدوم الاشاوس على احر من الجمر لتطهير الشرق من دنس القوات الانفصالية	"The people of eastern Sudan are waiting impatiently for the arrival of the brave heroes to purify the East from the filth of the separatist forces."	**Territorial cleansing logic.** A geographic region is framed as contaminated by a group, legitimizing violence as "purification."
RemazaliReRe (suspended)	حتى #مدني زاتها لازم تعود لحضن الوطن ويتم استعادتها وتطهيرها من دنس الفلول…	"Even Madani must return to the embrace of the الوطن (nation), be reclaimed, and purified from the filth of the remnants…"	Terms like 'purification' indicate treating people as a disease, while 'remnants' is dehumanizing

The use of bots to circulate this language could increase its incitement risk by manufacturing consensus, lowering perceived social resistance to violence, and creating a permissive discursive environment in which forced displacement, collective punishment, or mass civilian harm can be more easily rationalized.

As access to Generative AI increased, the network started creating AI images to accompany their posts.

Figure 17 Examples of GenAI images shared by the network

Secondary storylines

UAE, Yemen, Iran and Egypt

United Arab Emirates The UAE is presented as an indispensable regional and global actor simultaneously operating across humanitarian relief, diplomacy, security, development, climate governance, technological innovation, and cultural leadership. Content repeatedly highlights large-scale aid operations to Gaza, Sudan, Syria, Libya, Chad, and Lebanon; mediation efforts in conflicts from Sudan to Ukraine; and high-profile engagement in international forums such as the UN, COP summits, the G20, and BRICS. This humanitarian and diplomatic framing is reinforced by constant references to state-led initiatives, royal patronage, and symbolic leadership figures, producing an image of omnipresent benevolence and competence.

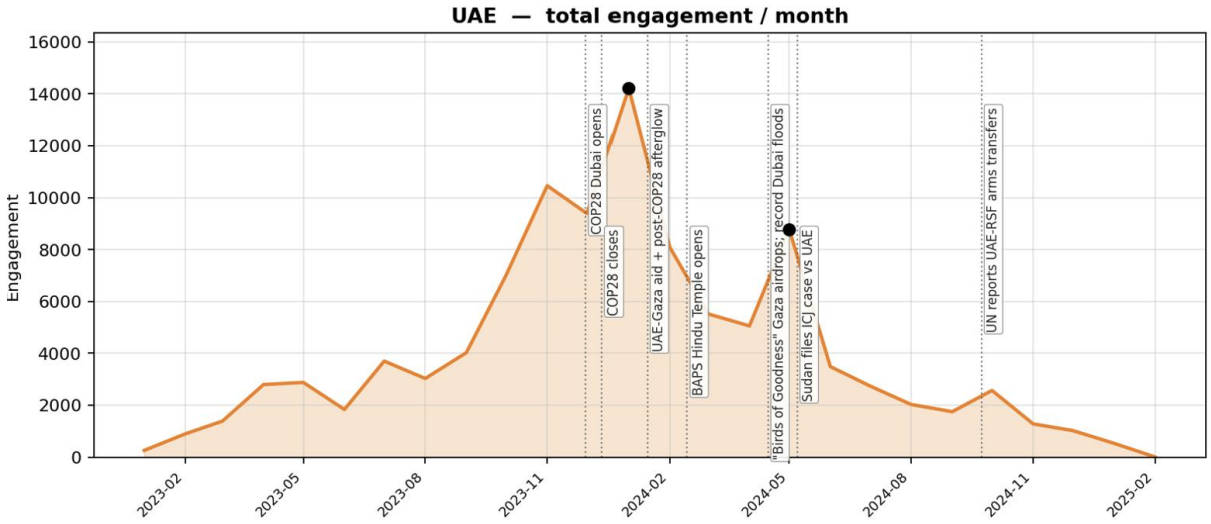


Figure 18 Engagement timeline shows promotion of UAE around key events like COP28, along with Sudan filing ICJ case

Criticism of the UAE, particularly regarding Sudan, is systematically delegitimized as misinformation, fabrication, or hostile campaigning, often framed as attacks on Arab unity or regional stability. Notably absent are discussions of the UAE's contested foreign policy roles, military involvement, or human rights concerns. Instead, the sheer volume and thematic breadth of positive messaging functions to crowd out alternative interpretations, positioning the UAE as a moral anchor and stabilizing force in moments of crisis. The result is a highly curated reputational narrative that relies less on persuasion through argument than on relentless affirmation, redundancy, and affective legitimacy-building.

Yemen The framing of Yemen is highly securitized and ideologically coherent, positioning the conflict primarily through the lens of terrorism, Iranian influence, and internal betrayal. The Houthis are depicted not simply as an armed movement, but as the central source of Yemen's destruction: a terrorist entity responsible for economic collapse, humanitarian suffering, corruption, attacks on civilians, and regional instability. At the same time, the Yemeni Congregation for Reform (Islah), frequently conflated with the Muslim Brotherhood, is framed as a covert collaborator with the Houthis, despite their historically adversarial relationship. This creates a simplified but politically

useful binary in which nearly all sources of instability are collapsed into a single interconnected "axis" of extremism linked to Iran, political Islam, and transnational militancy. The repeated references to maritime security, Red Sea instability, Hezbollah, Sudan, Lebanon, and Syria further situate Yemen within a broader regional geopolitical struggle rather than as a purely domestic civil war.

The most consistently demonized figure is Abdul-Malik al-Houthi, who is framed not merely as a political or military leader but as the embodiment of terrorism, corruption, starvation, sectarianism, and regional destabilization. The language used against him is unusually absolutist and moralized ("destroyer of Yemen", "leader of terrorism and bloodshed", "misguided fatwas"), suggesting the network is constructing a highly personalized enemy figure around whom broader anti-Houthi sentiment can coalesce. The corpus references Yemeni political actors associated with anti-Houthi or anti-Islah camps, often positively or neutrally. These include figures such as Rashad al-Alimi, Aidarus al-Zoubaidi, Tareq Saleh, and Ahmed Awad bin Mubarak. Their appearance tends to coincide with themes of anti-corruption, governance, reconstruction, diplomacy, or restoring state authority. Meanwhile, the Islah Party and Muslim Brotherhood figures are discussed less through named personalities and more as a collective conspiratorial actor accused of collusion with the Houthis.

What is particularly striking is the hybrid nature of the messaging. Alongside highly repetitive narratives about terrorism, corruption, and Iranian expansionism are large volumes of aestheticized cultural content celebrating Yemeni landscapes, villages, agriculture, architecture, coffee, weddings, and rural life. This juxtaposition appears deliberate. It condemns the Houthis while constructing an idealized vision of Yemen itself: peaceful, beautiful, authentic, historically rich, and implicitly aligned with a broader Arab nationalist identity. In this sense, the messaging performs both negative and positive propaganda functions simultaneously. It demonizes political enemies while emotionally rehabilitating a romanticized image of Yemen under threat. The repetition, templated phrasing, and modular structure of many posts also suggest a coordinated influence operation, combining centralized narrative guidance with AI-assisted or semi-automated content production.

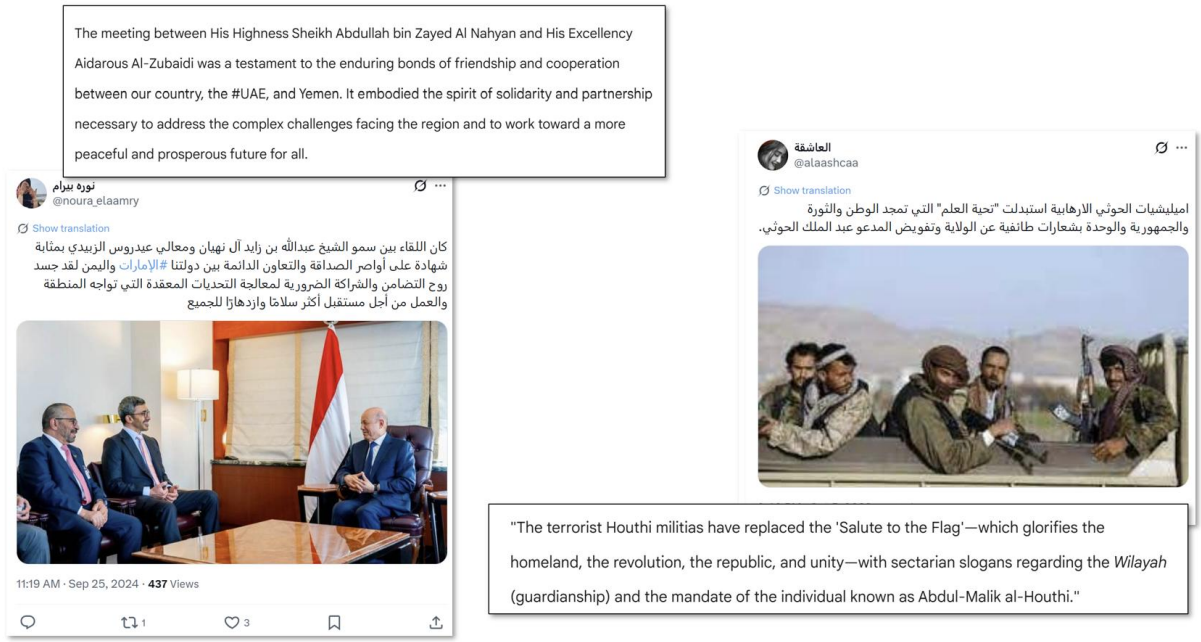


Figure 19 Examples of anti-Houthi tweets

Iran Iran is repeatedly depicted as the architect of instability across the Middle East and North Africa through a vast transnational network of proxies, militias, Islamist movements, and covert influence operations. The corpus consistently links Iran to the Houthis, Hamas, Hezbollah, the Muslim Brotherhood, armed factions in Iraq and Syria, and increasingly the SAF. Sudan in particular becomes framed as the newest frontier of Iranian expansionism, with repeated references to Iranian drones, military advisors, Red Sea ambitions, Port Sudan, and alleged attempts to establish military bases or strategic maritime influence. The result is an overarching geopolitical narrative in which Iran is represented as orchestrating a coordinated “axis of chaos” stretching from Lebanon and Yemen to Sudan and the Red Sea corridor.

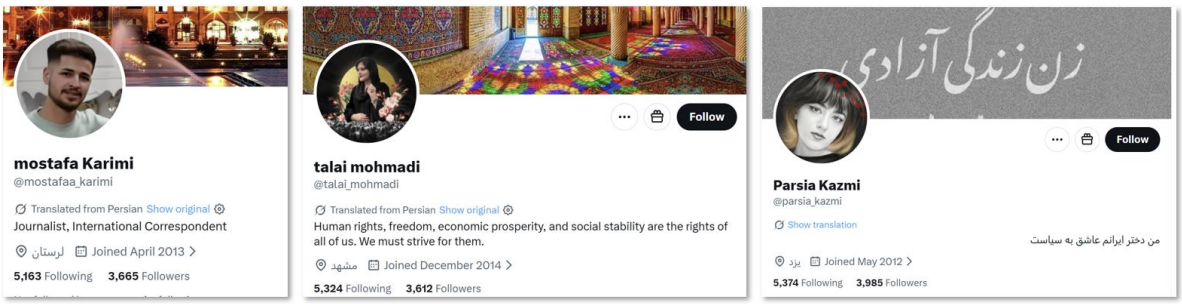


Figure 20 Screenshots of anti-Iran regime accounts

At the same time, the corpus contains a second, somewhat distinct layer focused on domestic repression inside Iran itself. Here the framing shifts from geopolitical securitization to emotionally charged solidarity with dissidents and victims of the Islamic Republic. The dataset repeatedly references the Woman, Life, Freedom movement, Mahsa Amini, imprisoned activists, executions, compulsory hijab, political prisoners, censorship, poverty, and state violence. Figures such as Mahsa Amini, Toomaj Salehi, and numerous lesser-known dissidents are elevated into symbols of national

suffering and resistance. Interestingly, this creates a dualistic representation of Iran itself: the regime is depicted as expansionist, violent, and malign, while “the people of Iran” are romanticized as noble, oppressed, freedom-seeking, and culturally sophisticated. This distinction is reinforced through recurring references to Iranian art, poetry, wrestlers, music, women, and heritage alongside calls for liberation from the Islamic Republic. In effect, the dataset separates “Iran” into two competing entities: a corrupt revolutionary regime spreading regional disorder, and an authentic Iranian nation imagined as culturally rich, secular-leaning, victimized, and yearning for freedom.

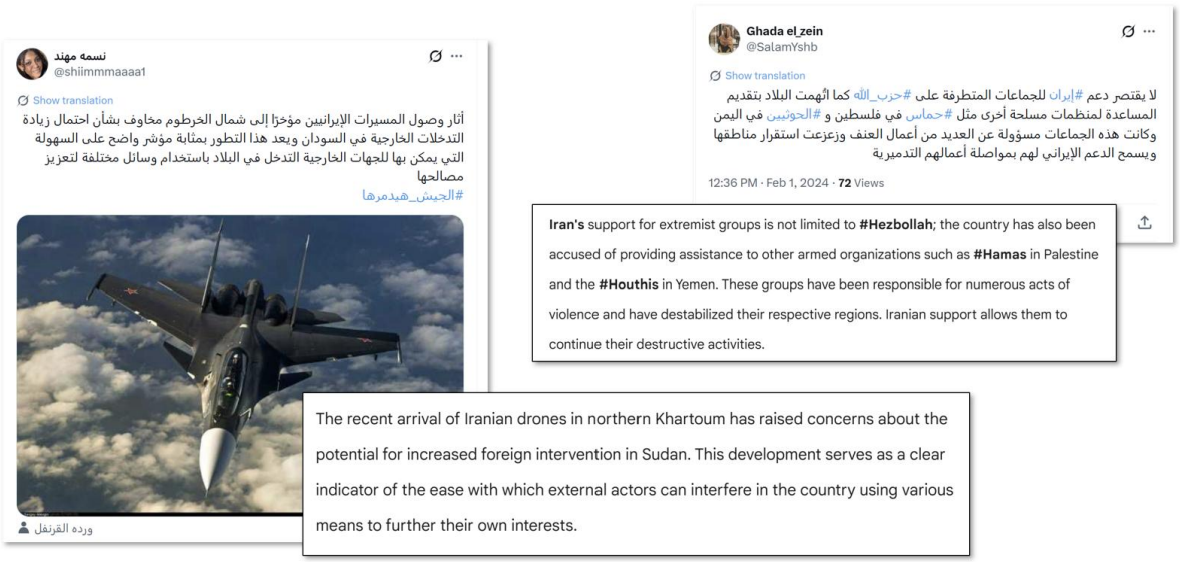


Figure 21 Examples of anti-Iranian posts

Egypt The Egypt-related narrative overwhelmingly constructs Egypt as a strong, modernising, and regionally indispensable state under the leadership of Abdel Fattah el-Sisi. The messaging is saturated with praise for megaprojects, infrastructure, agriculture, military manufacturing, renewable energy, tourism, and economic development, presenting Egypt as a country undergoing historic transformation after a period of chaos and decline. The Egyptian military is central to this image. It is portrayed as a fighting force as well as an engine of national development, humanitarianism, and technological progress. Slogans such as “the army is the people, and the people are the army” reinforce a fusion of nationalism, militarism, and loyalty to the state. Alongside this, Egypt is framed as a stabilizing regional actor: supporting Gaza, mediating Sudanese and Libyan crises, strengthening African and Arab diplomacy, and defending Palestinian rights. The relationship with the United Arab Emirates is particularly prominent, with repeated portrayals of Egypt and the UAE as “one heart” bound by strategic partnership, shared prosperity, and mutual political vision.

Running parallel to this celebratory nationalism is an intense and highly repetitive anti-Muslim Brotherhood discourse. The Brotherhood is consistently depicted as a terrorist, corrupt, destabilizing, and globally subversive force responsible for violence, economic decline, extremism, and social fragmentation. The rhetoric is often dehumanizing, describing the group as a “plague” or existential threat to Arab societies, while repeatedly emphasizing the failures of the Brotherhood’s rule in Egypt under Mohamed Morsi. This framing mirrors broader post-2013 Egyptian and Emirati state narratives that position authoritarian stability as preferable to Islamist political participation.

Importantly, these political messages are embedded among romantic posts, religious greetings, football commentary, lifestyle content, and emotional reflections, creating the impression of ordinary social media activity rather than overt propaganda. This blending of soft affective content with disciplined political messaging suggests an attempt to manufacture authenticity while steadily reinforcing a pro-state, anti-Islamist, and pro-regime worldview.

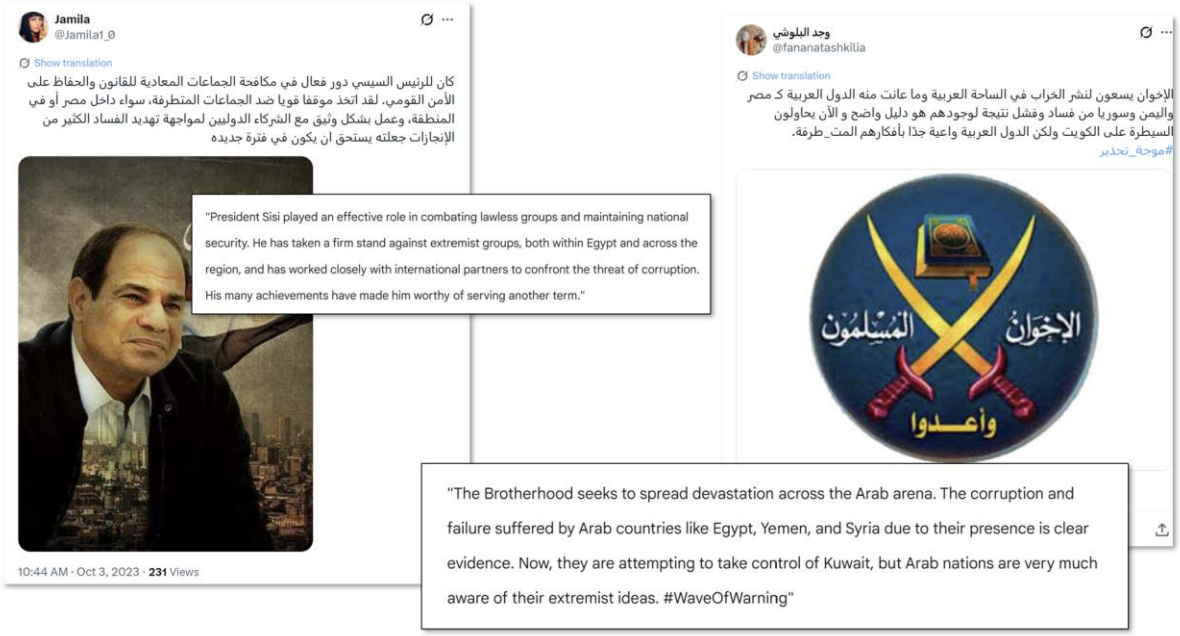


Figure 22 Example of tweets praising Sisi and talking about the Muslim Brotherhood's negative role in Egypt

Other regional focus areas

While the focus of this report is on Sudan and atrocity risk, what is notable about Network One is that it functions as a MENA-wide media ecosystem, covering specific countries. Analyzing these narratives can also be useful in determining the stance and alignment of the network.

North Africa narratives

Mauritania The network presents Mauritania as a model of stability, development, and competent governance under President Mohamed Ould Cheikh El Ghazouani. Leadership is consistently framed as visionary, unifying, and internationally respected, with particular emphasis on Ghazouani's diplomatic activity and role as chair of the African Union. Economic narratives highlight infrastructure projects, foreign investment, and resource potential, especially gas, mining, and green hydrogen, constructing an image of imminent prosperity. Humanitarian and social initiatives are repeatedly foregrounded, portraying the state as attentive to citizen welfare and inclusive development.

International partnerships, especially with the UAE, China, Saudi Arabia, and Russia, receive disproportionate attention and are described using emotive language emphasizing fraternity,

gratitude, and strategic alignment. Cultural heritage and natural landscapes are mobilized to reinforce national pride and positive affect. Notably absent is any substantive criticism of governance, opposition perspectives, or structural challenges. When difficulties are mentioned, they are framed as temporary and effectively managed. Overall, the Mauritania narrative functions as reputational amplification, presenting the country as politically legitimate, economically ascendant, and internationally valued, while systematically excluding dissenting or critical viewpoints.

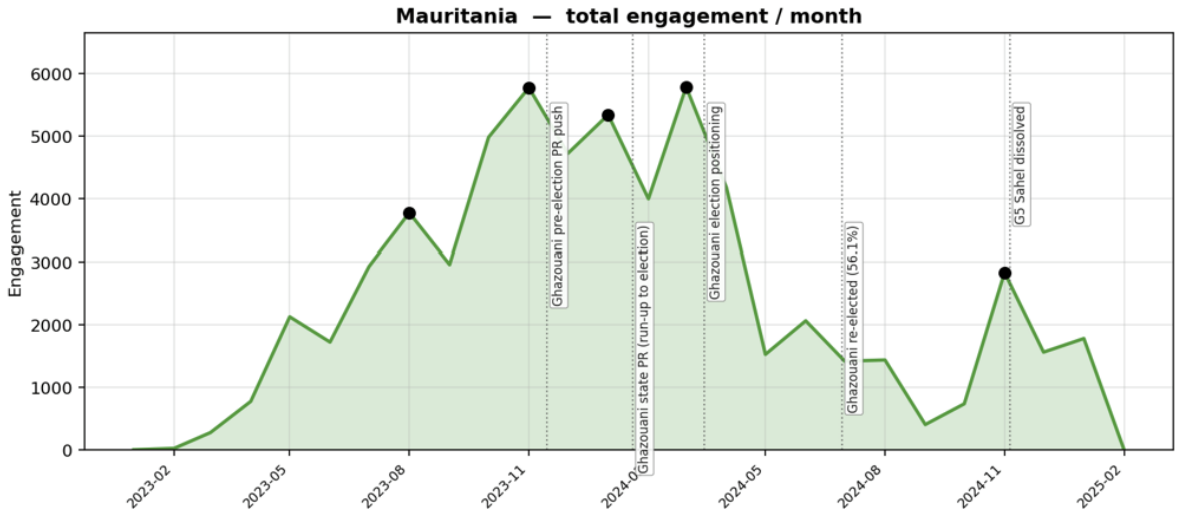


Figure 23 Network engagement on Mauritania shows how spikes in engagement reflect Ghazouani election positioning

Libya The political stance of the network reflects a broadly pro-stability, anti-fragmentation, and implicitly anti-Islamist orientation aligned with eastern Libyan and Gulf-backed state-building narratives. Across the dataset, the accounts consistently support UN-led electoral processes, the 6+6 electoral framework, institutional unification, and the consolidation of centralized security structures, while portraying reconstruction, infrastructure, and technocratic governance as pathways toward national recovery. Political figures associated with both eastern (suggesting UAE alignment) and internationally recognized governing institutions — including Khalifa Haftar, Osama Hamad, Aguila Saleh, Mohammed al-Menfi, Abdul Hamid Dbeibah, Abdullah Batili, and members of the Presidential Council — are generally framed positively when they are depicted as contributing to dialogue, elections, reconciliation, reconstruction, or institutional coordination. At the same time, the networks heavily amplify UAE humanitarian and reconstruction efforts, particularly in eastern Libya, reinforcing Gulf-backed narratives that associate strong institutions, military coordination, and technocratic administration with national salvation. Islamist or revolutionary actors are rarely described as legitimate political forces; instead, the informational environment consistently frames instability, militancy, fragmentation, and political disorder as threats to Libya's recovery.

Levant narratives - Syria, Lebanon, Palestine, Israel, Jordan, and Turkey

Syria The network's portrayal of Syria evolves over time while maintaining consistent strategic objectives. Early content adopts a negative framing, emphasizing civil war, terrorism (ISIS and Al-Qaeda), refugee flows, and criticism of the Assad regime, presenting Syria primarily as a source of

regional instability. The February 2023 earthquake marks a clear inflection point. The network rapidly pivots to humanitarian framing, foregrounding civilian suffering and highlighting aid efforts, particularly those led by the UAE, Saudi Arabia, and other Arab states, while suspending overt political critique. In subsequent months, content increasingly promoted Syria's diplomatic rehabilitation, including its return to the Arab League, regional normalization, and cultural heritage. More recent messaging balances positive portrayals of stability and reconciliation with targeted criticism of specific opposition actors, notably now President Ahmed Al-Sharaa. For a brief period in 2024 the network attempted to rehabilitate Bashar Assad, in line with the efforts of some Arab countries to normalize relations with the Assad regime. (The brief rehabilitation of Assad, and attacks on Al Sharaa, are also a telling indicator in terms of narrowing down attribution).

Figure 24 In 2025 the network emphasized attacks on Syria's new President Ahmed Al-Shara'a

A central feature of the network's political stance is its aggressive rejection of Western narratives regarding chemical weapons, sanctions, war crimes, and accountability. For a period in 2024, French legal actions against Assad are repeatedly dismissed as fabricated, politicized, or driven by extremist-backed disinformation, while sanctions are framed as collective punishment responsible for economic suffering and refugee displacement.

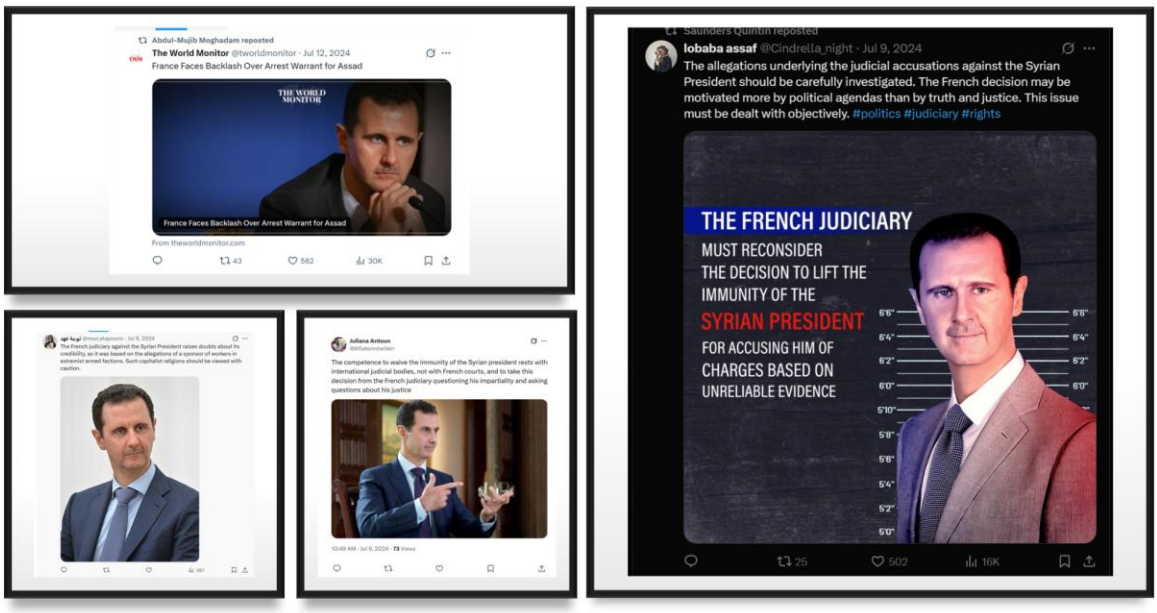


Figure 25 Example of posts rehabilitating Assad

Another related network of bots was used to amplify posts attempting to rehabilitate Assad.

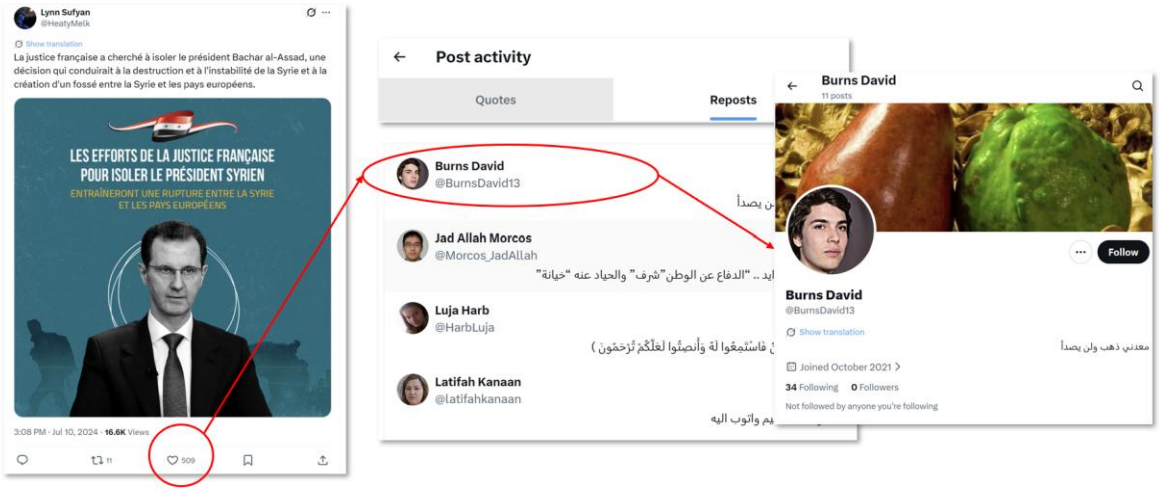


Figure 26 Example of bot with AI avatar promoting another bot post

There is a demonstrable spike in activity on the Syria-related portion of the network during a push to rehabilitate Assad.

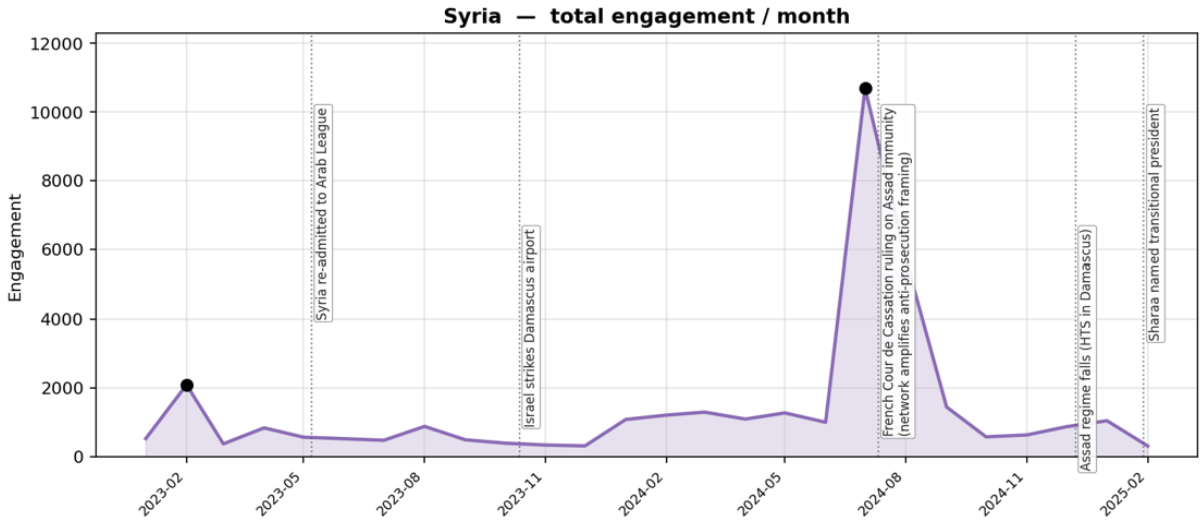


Figure 27 Graph showing high network engagement during rehabilitation of Assad

Lebanon Lebanon-related content is dominated by an overwhelmingly negative portrayal of Hezbollah, consistently framed as a terrorist, foreign-backed organization responsible for corruption, economic collapse, and regional instability. Hezbollah is depicted as exercising coercive control over the state and sabotaging governance, frequently linked to Iranian influence. In contrast, Lebanese state institutions, particularly the Lebanese Army, are portrayed positively or sympathetically, framed as compromised victims rather than culpable actors. Other political figures are referenced primarily insofar as they oppose Hezbollah. This narrative constructs a clear dichotomy between a legitimate Lebanese state and an illegitimate Hezbollah presence, positioning Lebanon's crises as resolvable through the latter's removal.

Palestine The network expresses strong affective sympathy for the Palestinian people, portraying them as resilient victims deserving international solidarity, with emotive content, especially involving children, used to elicit moral response. At the same time, Hamas is consistently delegitimized, framed as a terrorist organization that exploits Palestinians, serves Iranian and Muslim Brotherhood interests, and undermines the Palestinian cause from within. This dual framing separates Palestinian suffering from Hamas' political legitimacy. Parallel to this, the network repeatedly praises Gulf states, particularly the UAE, Saudi Arabia, and Egypt, as the true humanitarian supporters of Palestinians, reinforcing their moral authority while marginalizing alternative forms of political resistance.

Jordan Jordan is portrayed as a pillar of stability and moderation in a volatile region. The monarchy, particularly King Abdullah II, is consistently idealized as wise, peace-oriented, and globally respected. State institutions, including the armed forces and security services, are praised for vigilance and professionalism. The Muslim Brotherhood is framed as a persistent internal and external threat, accused of exploiting regional crises, especially Gaza, to undermine Jordanian security. The narrative equates loyalty to the monarchy with national stability, legitimizing firm action against Islamist

opposition while highlighting Jordan's humanitarian and diplomatic role, particularly in relation to Palestine.

Turkey Narratives about Turkey are highly contextual and adaptive. During the February 2023 earthquake, content is overwhelmingly sympathetic, emphasizing humanitarian suffering and Arab aid efforts. Outside disaster contexts, however, Turkey is framed negatively as an interventionist actor involved in Sudan, Libya, and Syria, accused of supplying weapons, harboring Muslim Brotherhood figures, and pursuing neo-imperial ambitions. This framing intensifies in Sudan-related discourse, where Turkey is accused of enabling civilian harm through military support to the SAF. Messaging softened following Turkey's diplomatic normalization with Egypt, highlighting economic cooperation, but continued to position the UAE and Egypt as the primary stabilizing actors, with Turkey cast as a secondary or reactive partner.

Israel The network's stance toward Israel is critical but calibrated. Israel is consistently portrayed as an occupying force responsible for civilian suffering in Gaza, using language that emphasizes aggression and disproportionate force. At the same time, the network criticizes Hamas, Hezbollah, and Iranian-aligned actors for exploiting Palestinian civilians, producing a dual attribution of blame. Coverage of the 7 October 2023 Hamas attack adopts a notably restrained tone, describing it as a "surprise operation" rather than terrorism, in contrast to the emotive language used for Palestinian casualties. Content relating to the Abraham Accords reflects ambivalence: normalization is acknowledged as potentially beneficial while insistence on Palestinian rights is maintained. The network actively denies allegations of UAE military support for Israel and redirects hostility toward Iran and its proxies, aligning criticism of Israel with broader Gulf geopolitical positioning rather than outright rejection of normalization.

Gulf States narratives

Qatar The network's portrayal of Qatar exhibits a clear temporal shift. During the period surrounding the 2022 FIFA World Cup, Qatar is framed negatively as corrupt, abusive toward migrant workers, unreliable in its commitments, and engaged in reputational laundering through sport. From late 2022 onward, this framing changes markedly. Qatar is increasingly presented as a humanitarian actor delivering aid to Sudan and Gaza, a diplomatic mediator in Gaza ceasefire negotiations (often positioned alongside Egypt and the United States), and a constructive economic partner engaged in regional investment and cooperation. Frequent references to Qatari aid flights, charitable activity, and mediation efforts underscore this reframing. The shift suggests a politicized and adaptive messaging strategy that tracks broader regional diplomatic realignments, including the resolution of the Gulf crisis and Qatar's expanded mediation role, while retaining traces of earlier criticism during the World Cup period.

Bahrain. Content relating to Bahrain is uniformly positive and highly curated. The network consistently portrays Bahrain's leadership, particularly King Hamad bin Isa Al Khalifa, as diplomatic, benevolent, and internationally respected. Bahrain is framed as a regional mediator and peace broker, notably through coverage of its hosting of Arab League summits and facilitation of dialogue on Sudan. Humanitarian narratives feature prominently, highlighting aid missions to Sudan and Syria and reinforcing an image of compassion and responsibility. The dataset contains virtually no criticism of

Bahrain's domestic politics, human rights record, or opposition movements. When internal matters are mentioned, they are framed as evidence of state competence. This systematic absence of dissenting perspectives, combined with emotive and reverential language, suggests a coordinated effort to enhance Bahrain's international legitimacy and reputational standing.

Kuwait The network presents Kuwait as a stable, benevolent, and influential regional actor governed by wise and respected leadership. Coverage focuses heavily on the royal family, diplomatic engagement, and humanitarian activity, particularly aid missions to Sudan, Libya, and Palestine. Economic strength and investment potential are emphasized, alongside portrayals of close and "fraternal" relations with other Gulf states, especially the UAE and Saudi Arabia. A notable exception to the otherwise celebratory tone is a coordinated campaign warning of alleged Muslim Brotherhood infiltration in Kuwait, framing the government as a bulwark against extremism. As with Bahrain, domestic political tensions and criticism of state institutions are almost entirely absent. The resulting narrative depicts Kuwait as a humanitarian benefactor and pillar of regional stability, while strategically silencing internal debate.

Hierarchical behaviors

The network also exhibits some signs of a pronounced hierarchical structure in which a small number of key accounts ("generals") operate as primary content initiators and are systematically boosted by a much larger army of subordinate accounts, almost all of which display characteristics consistent with automation, including AI-generated profile images and minimal organic interaction. At the apex (Generals – Level 1), a very limited set of highly central seed accounts functions as anchoring nodes from which influence radiates downward. Beneath them, a secondary tier of intermediary accounts (Generals – Level 2) acts as relays, engaging with and reinforcing content from Level 1 while preserving the appearance of semi-organic participation. The base of the network consists of a dense mass of low-visibility nodes ("Minions"), characterized by sparse outward activity and near-exclusive inward connectivity toward higher-level accounts. These minions are responsible for bulk amplification—predominantly through likes rather than reposts—producing tall, pillar-like cascades that visually resemble funnels or spotlights converging upward. AI GAN-generated faces used for burner accounts, retweet, while more sophisticated or real images are used for the generals. This whole network, including the minions, has been responsible for amplifying UAE-state aligned narratives, from promoting COP28, to boosting the tweets of MBZ.

The diagram below demonstrates how a network of 'minions' often amplifies key messages from other sockpuppets within the network.

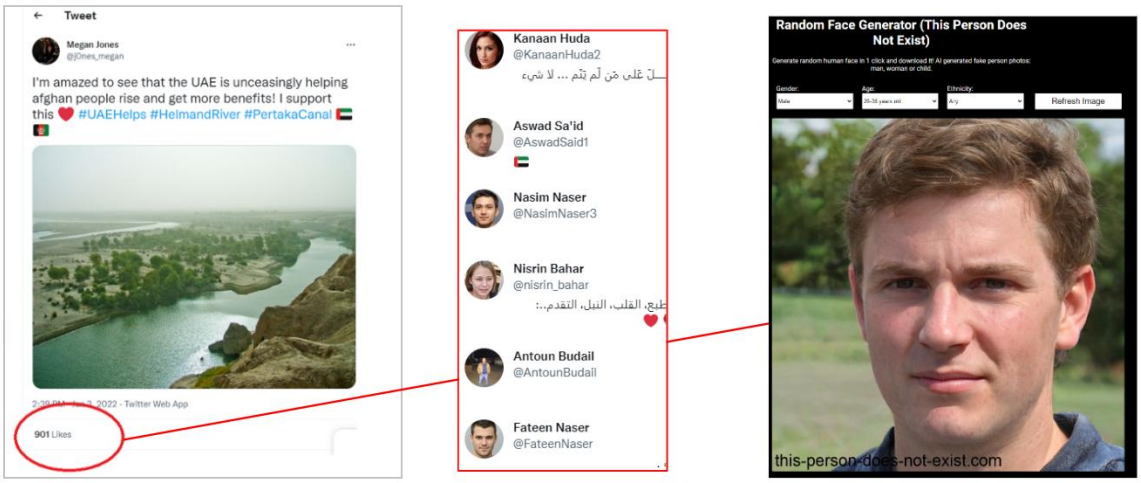


Figure 28 Example of numerous AI-generated avatars

The sharp asymmetry between the small number of content-generating nodes and the large volume of amplification-only accounts indicates a non-reciprocal, top-down architecture inconsistent with organic community formation. Instead, the structure reflects a managed influence system optimized for artificial validation at scale, in which perceived legitimacy is manufactured through volume, automation, and vertical coordination rather than horizontal engagement or deliberation.

The visualization depicts a vertically stratified influence network organized around a clear command-and-amplification hierarchy.

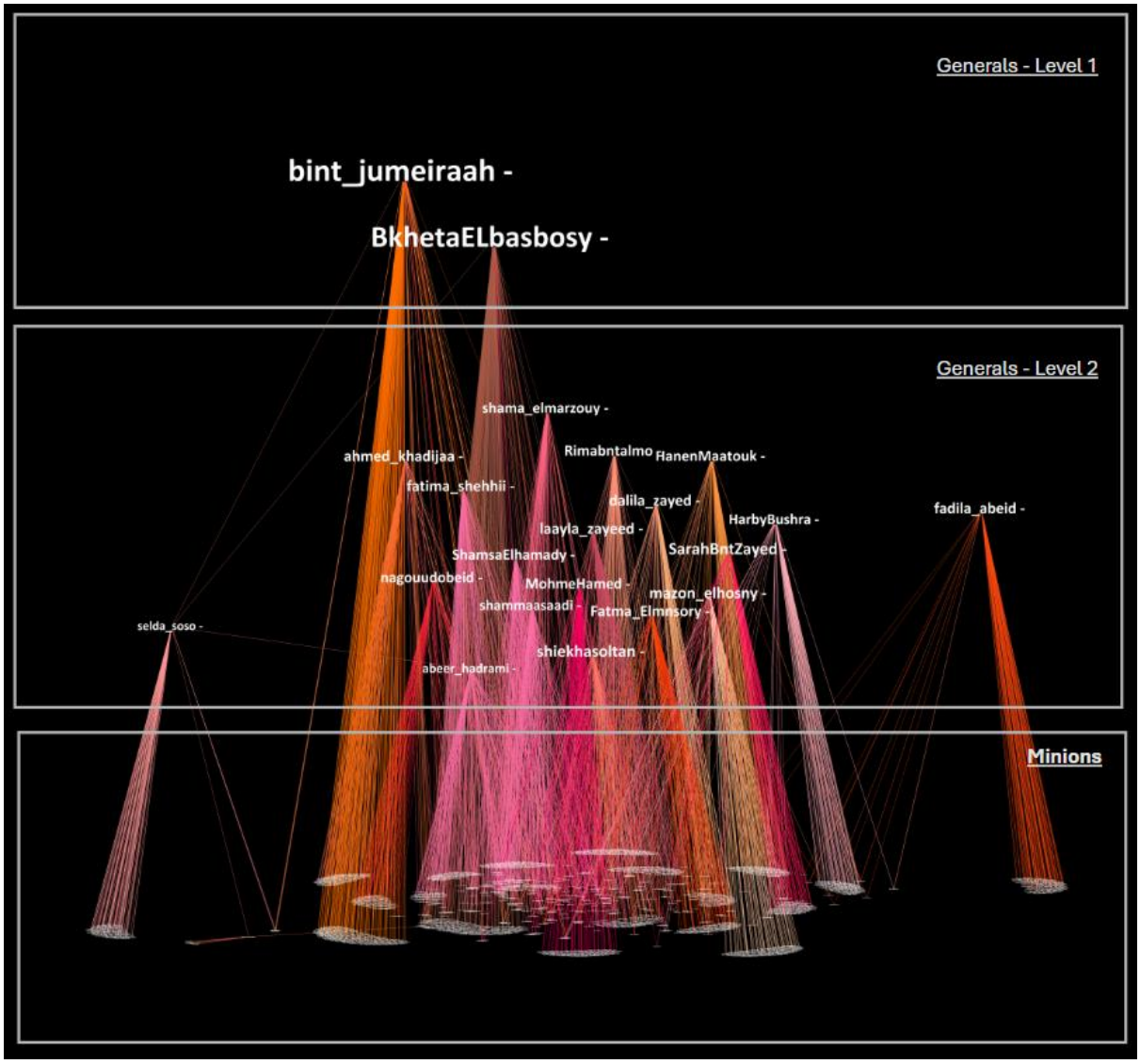


Figure 29 Bots and Minions: -Hierarchical Amplification Structure: Visualized network showing a vertically stratified influence architecture used for artificial amplification on social media. This is from an earlier snapshot of the network when they were posting

Verification

The network's adaptability to platform changes under Elon Musk was further evidenced by its strategic uptake of platform verification following the introduction of X Premium. Within the network, approximately 40 accounts obtained verification, despite verification previously being rare or inaccessible to such actors. Under Musk, verification means algorithmic privilege and reputational signaling that enhances the visibility and perceived legitimacy of coordinated content. This technique and network were also mobilized during the UAE's hosting of COP28[29] They illustrate an emergent form of pay-to-play propaganda, in which credit-based verification systems lower the cost of entry for influence operations while weakening the link between verification and authentic identity. Far from

[29] Damian Carrington and Damian Carrington Environment editor, 'Army of Fake Social Media Accounts Defend UAE Presidency of Climate Summit', Environment, *The Guardian*, 8 June 2023, https://www.theguardian.com/environment/2023/jun/08/army-of-fake-social-media-accounts-defend-uae-presidency-of-climate-summit.

constraining coordinated inauthentic behavior, changes to X's governance appear to have expanded the operational affordances available to state-aligned networks, enabling even relatively unsophisticated accounts to achieve reach, credibility, and occasionally to break out into mainstream media. This verification eventually ended for those identified accounts around 2024.

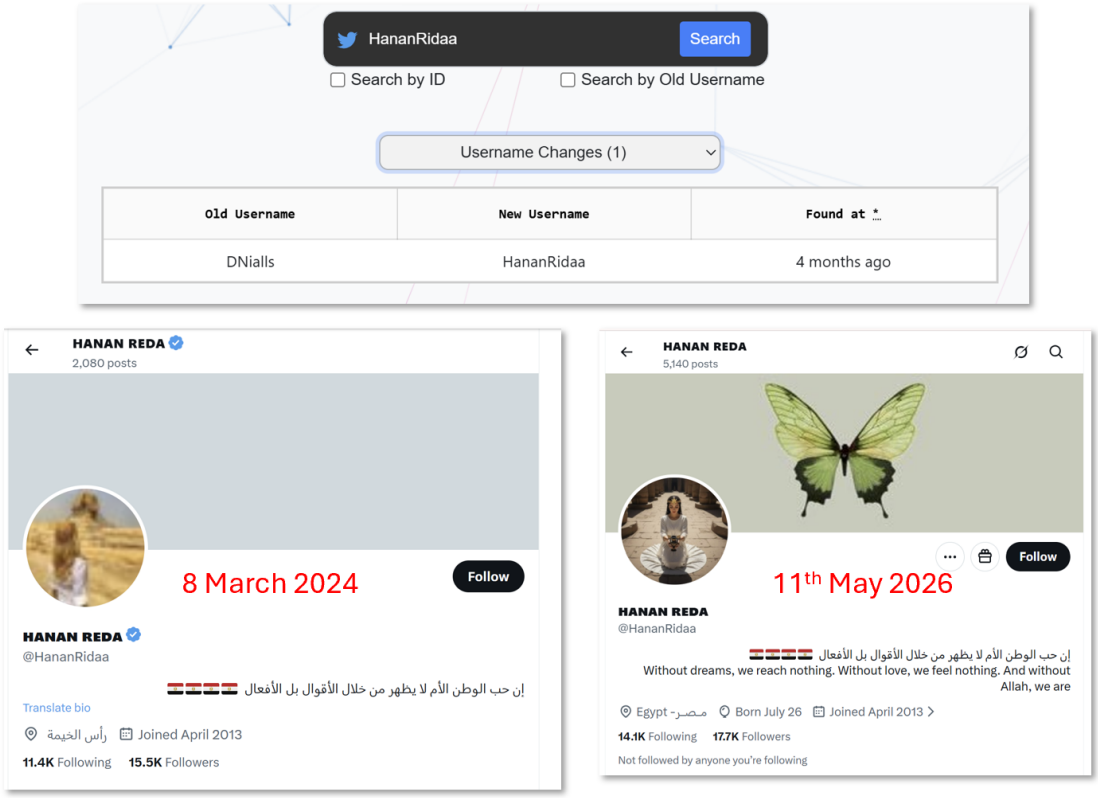


Figure 30 Example of evolution of bot account, from verified status to unverified

Methodological notes and data sources

Detection methodology: Sockpuppet identification

Approximately 310 sockpuppet accounts were tracked and analyzed over a period of > 2 years and were identified through the convergence of behavioral, metadata, and visual indicators. No single signal was treated as determinative. Rather, accounts were classified based on repeated patterns observed across multiple dimensions, assessed longitudinally and in relation to one another.

One of the earliest indicators was systematic handle switching (Jones 2022, Grossman et al, 2022). Many accounts initially operated under Western-sounding names and later adopted names more closely aligned with MENA linguistic and cultural conventions. For example, Sherly Codoy became Latifa Al Saadi. This renaming appears intended to enhance perceived authenticity among target audiences and to better align account identities with the geopolitical narratives being promoted.

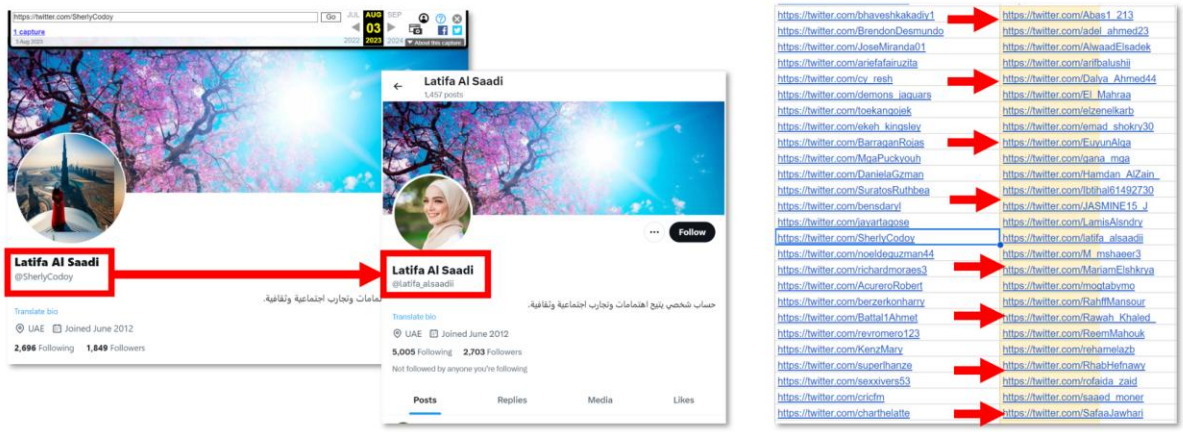


Figure 31 Example of account handle-switching

Profile imagery showed parallel patterns of repurposing. In numerous cases, original profile photographs were replaced with AI-generated faces, stock images, or photographs of unrelated individuals. Previous account tweets were scrubbed (deleted). These changes obscure prior account histories and enable accounts to be reused under reconstructed identities without drawing attention to earlier activity.

Account age analysis further reinforced this pattern. A substantial proportion of accounts were not newly created but dated back several years, in some cases as early as 2009. Many remained dormant for extended periods before reactivating in or around 2022, at which point they adopted sustained political behavior. The reuse of legacy accounts allows operators to exploit the credibility associated with long-standing profiles while rapidly deploying them for coordinated influence activity.

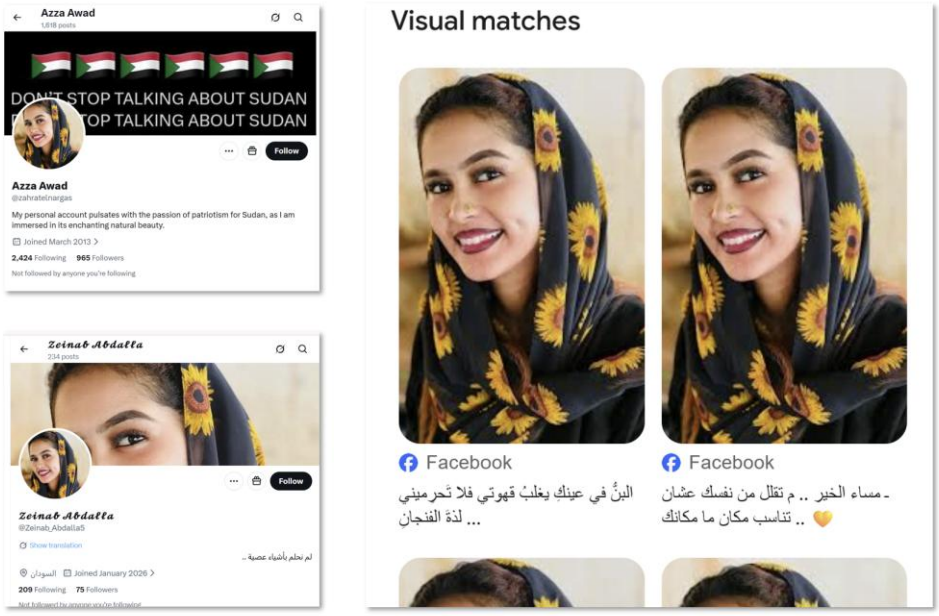


Figure 31a Account using photo of a model, along with another account using same photo

Behavioral indicators were highly consistent across the network. Follower-to-following ratios are frequently clustered around parity (approximately 1:1), a pattern commonly associated with inorganic user behavior. When considered alongside narrow tweet-volume, repetitive posting styles, and strong narrative alignment, this indicates coordinated design, not genuine engagement. Posting behavior was dominated by public-relations-style messaging that praised or condemned specific political actors and factions in line with shared political positions.

Temporal and interactional patterns provide additional evidence of coordination. Accounts frequently posted within similar time windows and exhibited burst-style activity rather than continuous engagement. Interaction levels were low relative to output, with most accounts functioning primarily as broadcast nodes rather than participants in sustained dialogue or debate. These behaviors were observed consistently over a period exceeding six months, indicating a sustained and organized operation rather than sporadic or opportunistic misuse.

Taken together, these indicators point to a coordinated sockpuppet network characterized by identity repurposing, behavioral homogeneity, and long-term operational persistence. Table X summarizes the principal account-level and behavioral indicators used in this analysis.

Data extraction and corpus construction

Indicator category	Observed pattern	Analytical relevance
Handle changes	The majority of accounts changed names over time, typically shifting from Western-sounding handles to more MENA-appropriate names.	Indicates identity repurposing and audience retargeting rather than organic personal evolution.
Profile image changes	Accounts frequently replaced profile images, including the use of AI-generated faces, stock photographs, or images of unrelated individuals.	Suggests deliberate obfuscation of prior account identity and ownership.
Account age vs. activity	Many accounts were created years earlier (some as early as 2009) but remained largely inactive until reactivation in or around 2022.	Consistent with "scrubbed" or repurposed legacy accounts used to exploit perceived credibility.
Account repurposing	Previously non-political or dormant accounts were reactivated and repurposed for sustained political messaging.	Indicates coordinated reuse of existing assets rather than organic political mobilization.
Follower–following ratio	Ratios frequently clustered around parity (approximately 1:1).	Functions as an authenticity-signaling strategy commonly observed in coordinated inauthentic behavior.
Tweet volume	Total tweet counts fell within a narrow and consistent range across accounts.	Suggests coordinated pacing and scripted activity rather than natural variation in user behavior.
Posting behavior	Content was predominantly PR-style messaging, praising or condemning specific political actors or factions.	Indicates strategic messaging aligned with shared narratives rather than personal expression.
Content similarity	Reuse of identical or near-identical phrasing, slogans, and hashtag placement across multiple accounts.	Strong indicator of coordination and centralized narrative production.
Temporal synchronization	Accounts posted within similar time windows and exhibited burst-style activity patterns.	Inconsistent with organic engagement; indicative of scheduled or centrally coordinated posting.
Interaction patterns	Low levels of genuine interaction; most accounts broadcast content rather than engaging in sustained replies or discussion.	Suggests amplification rather than participation in deliberative discourse.
Longevity of behavior	These patterns were observed consistently for more than six months.	Demonstrates sustained, strategic coordination rather than short-term or opportunistic manipulation.

Tweet data were extracted using Phantombuster from approximately 214 of the 310 accounts identified as sockpuppets. This process yielded approximately 170,000 tweets and retweets. Each post was coded for primary and secondary country relevance. Automated translation was conducted using a large language model to enable systematic cross-linguistic analysis. Separate country-level corpora were then constructed, allowing for thematic, positional, and stance-based analysis within and across

national contexts. These corpora were analyzed to identify dominant narratives, patterns of alignment, and cross-regional consistency in messaging. All tweets were labeled by primary and secondary country relevance.

Cell-based network topology

A network graph was constructed from the data by constructing an edge list of tweets, retweets and replies. Gephi was used to generate a network graph. A community detection algorithm was run, which revealed specific community structures that correlated with country focus. The largest section was Sudan. This analysis confirmed the interrelation of ostensibly disparate accounts. These interactions confirmed an underlying networked logic, revealing country oriented cells.

Network Two

Overview

Network Two captures a distinct, event-driven influence surge centred on Sudan and activated in the immediate aftermath of the El Fasher massacre (late October 2025). The campaign ran from approximately 4th -20th November 2025. Unlike Network One, which reflects a longer-running, cell-based infrastructure, this network is best understood as a large-scale rapid mobilization designed to seize a narrow window of heightened attention, overwhelm the platform's trending systems, and impose a pre-structured interpretation of events before countervailing narratives could consolidate. The analysis detected approximately 19,000 bots. The campaign operated primarily through a relay of tightly linked hashtags that repeatedly trended across Sudan, the UAE, and wider Middle Eastern markets, generating visibility through volume and velocity rather than genuine engagement. Its function was to "promote" content as well as to manufacture an appearance of consensus by saturating the information space with disciplined, repetitive framing: the SAF and Burhan as the architects of famine and obstruction; the RSF as humane, stabilizing, and peace-seeking; and El Fasher as a site of "recovery" rather than atrocity.

El Fasher Massacre

In late October 2025, Sudan's Rapid Support Forces seized El Fasher, the capital of North Darfur, following a lengthy siege marked by starvation tactics. Human rights investigators now suspect the assault may represent the deadliest atrocity of Sudan's civil war. Analysis of satellite imagery by Yale Humanitarian Research Lab depicts a city effectively stripped of normal life, with markets, roads and residential areas rapidly becoming deserted. Researchers also identified suspected mass burial and incineration sites, indicating widespread killings and efforts to dispose of bodies. British lawmakers were informed that the death toll may have exceeded 60,000, while up to 150,000 residents remain missing, with little indication that they escaped the city. El Fasher is still inaccessible to journalists, aid organisations, and UN investigators, despite RSF assurances of access. Humanitarian convoys remain stalled outside the city due to security concerns, while international observers have formally classified conditions there as famine. [30]

In the immediate aftermath of the El Fasher massacre in late October 2025, Sudan's information environment experienced a sudden and highly coordinated surge in online activity by bot-driven hashtags. Beginning in early November, a cluster of Sudan-related hashtags repeatedly trended on X across Sudan, the UAE, and wider Middle Eastern markets, advancing a consistent narrative framing of the conflict. These hashtags included

[30] Mark Townsend, 'RSF Massacres Left Sudanese City "a Slaughterhouse", Satellite Images Show', Global Development, *The Guardian*, 5 December 2025, https://www.theguardian.com/global-development/2025/dec/05/rsf-massacres-sudanese-city-el-fasher-slaughterhouse-satellite-images.

Arabic Hashtag	English Translation
#الهدنة_تُقبل_لا_لماذا	“Why Not Accept the Ceasefire”
#لا_لتجويع_السودان_يا_جيش	“No to Starving Sudan, O Army”
#جيش_البرهان_يرفض_الهدنة	“Al-Burhan’s Army Rejects the Ceasefire”
#جيش_البرهان_يتاجر_بالحرب	“Al-Burhan’s Army Profits from War”
#لا_لضرب_المساعدات_يا_جيش	“No to Striking Aid Convoys, O Army”
#عودة_الحياة_للفاشر	“Life Returns to El Fasher”

The network placed responsibility for starvation, humanitarian obstruction, and the continuation of the war on Abdel Fattah al-Burhan and the Sudanese Armed Forces (SAF), while the RSF were framed as disciplined, humane, and receptive to ceasefires. El Fasher itself was portrayed as a city where “life is returning” under RSF control, amounting to a coordinated attempt to reframe and sanitize mass violence through narrative saturation.

Scale and indicators of coordination

Analysis of activity across six closely linked hashtags indicates that the surge was driven overwhelmingly by automated or highly coordinated accounts. Between 18,709 and 19,514 accounts displayed strong indicators of bot-like behavior, corresponding to approximately 89–93 per cent of active participants. These estimates are intentionally conservative and reflect convergence across multiple indicators rather than reliance on any single detection metric. Collectively, the hashtags generated over 91 million impressions, demonstrating how volume and velocity can manufacture visibility in the absence of genuine engagement.

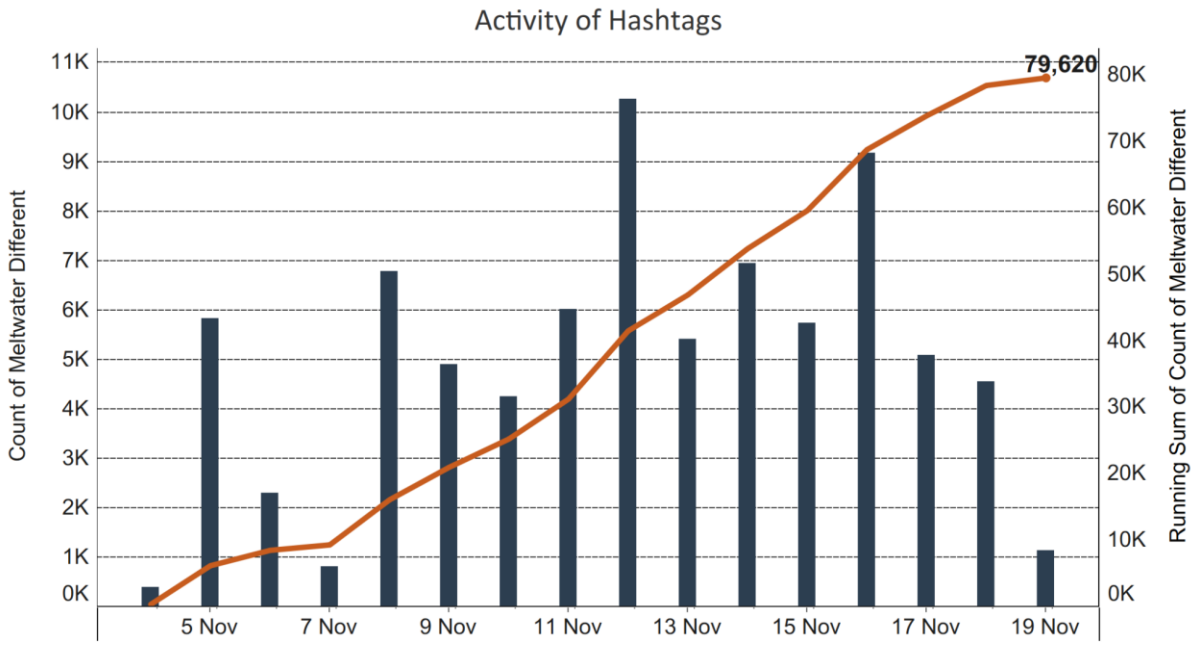


Figure 32 Volume of bot-driven hashtags post El Fasher massacre

Temporal patterns reinforce this conclusion. Hashtag activity occurred in short, dense, sequential bursts of about 2-3 days, each producing a similar number of posts before collapsing and handing over to the next slogan. This disciplined, relay-style sequencing is inconsistent with organic political discussion and instead reflects scripted posting behavior designed to exploit platform trending mechanisms, which reward novelty and rapid acceleration (velocity).

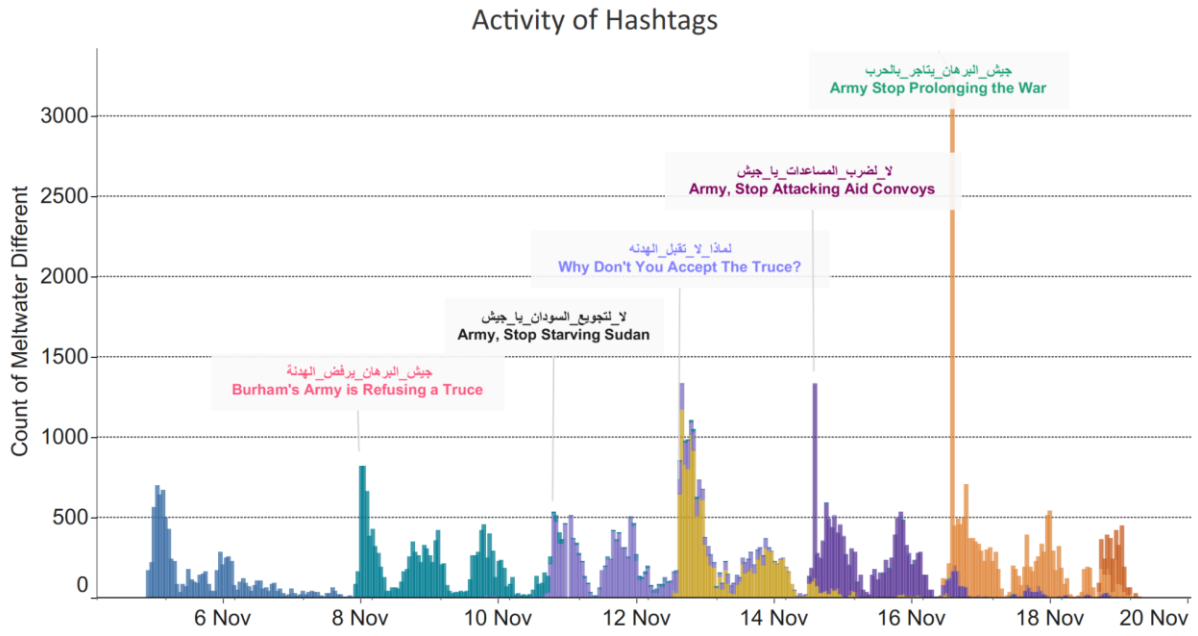


Figure 33 Sequencing of bot driven hashtags over time

Network analysis further underscores the inauthentic nature of the activity. Approximately 95 per cent of accounts exhibited minimal interaction, collapsing at degree two or below (degree here being a measure of how many accounts each account interacted with). A very small number of accounts

formed a coordinating core, surrounded by thousands of accounts functioning primarily as broadcast nodes. In several instances, single accounts were retweeted hundreds of times by others within the network, producing sharp spikes in visibility without sustained interaction.

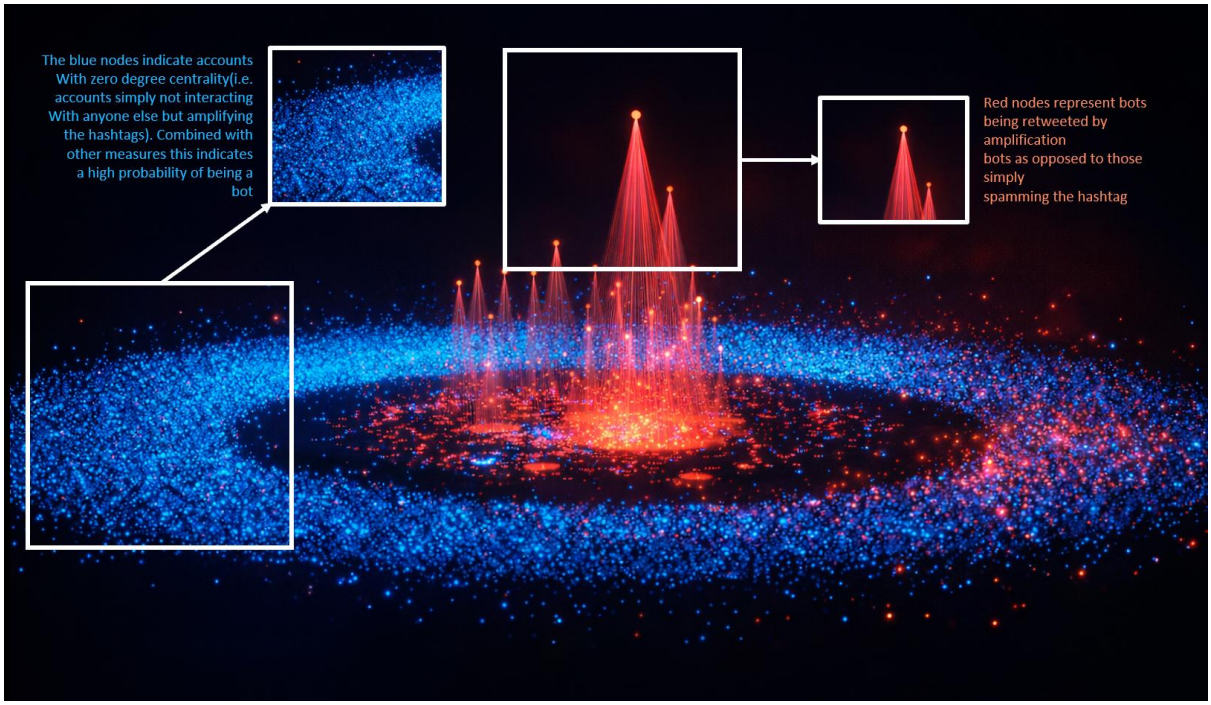


Figure 34 Visualization of bot network topography

Narrative structure and geopolitical alignment

Content analysis reveals a high degree of narrative discipline. Across thousands of posts, the SAF and Burhan are consistently framed as the primary agents of suffering, accused of deliberately engineering famine, obstructing humanitarian aid, and rejecting ceasefires. Burhan is portrayed as illegitimate, power-seeking, and aligned with Islamist actors and foreign patrons. Cities under SAF control are depicted as spaces of despair and decay.

By contrast, the RSF is framed as disciplined, humane, and peace-seeking, with RSF-controlled areas associated with stability and recovery. Systematic discussion of RSF abuses is largely absent from the dataset. One widely circulated video, shared thousands of times, explicitly praised “Brave RSF Fighters,” exemplifying the campaign’s rehabilitative thrust.

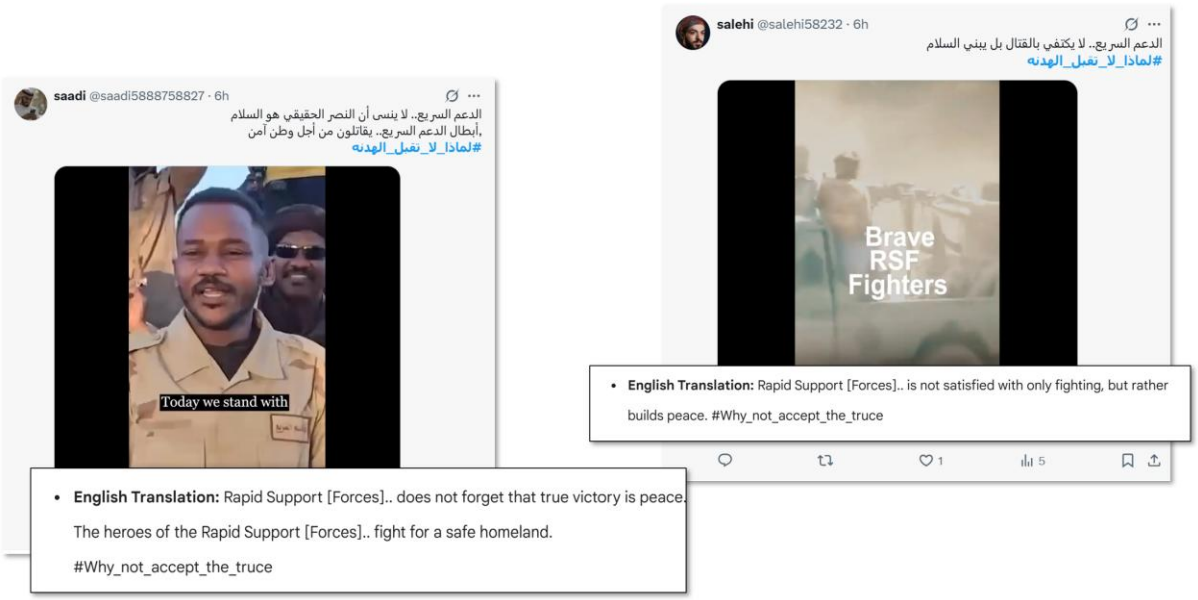


Figure 35 Examples of pro-RSF videos shared by the network

A prominent narrative strand portrays El Fasher's supposed "rebirth." Under hashtags such as #عودة_الحياة_للفاشر ("life returns to El Fasher"), the city is depicted as normalizing, with imagery of open markets, children playing, and the resumption of daily life. Trauma is acknowledged primarily as a prelude to recovery, reinforcing a narrative of closure rather than accountability. Centrally produced infographics, uniform in messaging and aesthetic, further externalize blame onto the SAF and Islamist actors.

Figure 36 Examples of infographics shared by the bot network

Geopolitically, responsibility for Sudan's suffering is externalized onto a familiar constellation of actors. Egypt, Saudi Arabia, Qatar, Turkey, Iran, and Russia are variously accused of enabling war, hypocrisy, or obstruction, while the United Arab Emirates is largely insulated from criticism. When referenced, UAE-related content emphasizes humanitarian aid, official statements, and concern for civilian protection.

Table showing the stance of the network on regional actors

Country	Example Arabic (verbatim)	English Translation	Core Stance
EG Egypt	قافلةٌ تحمل تُقصف أن خبزًا ودواءً ذريعة؛ ولا مبرر له ليس عار فهذا هذا من أعظم أنت مصر.	To bomb a convoy carrying bread and medicine is a disgrace with no justification or excuse; Egypt, you are greater than this.	Moral betrayal / violation of humanitarian red lines
SA Saudi Arabia	السعودية في الحكم سيكولوجية من والخوف السيطرة على قائمة الفوضى تصنع النفوذ، لذلك فقدان وتجوع البرهان السودان، تدعم في الشعب.	The psychology of rule in Saudi Arabia is based on control and fear of losing influence; it therefore creates chaos in Sudan, supports al-Burhan, and starves the people.	Strategic control / weaponisation of hunger
QA Qatar	قاتل بصمت الأزمة تعمّق قطر.	Qatar deepens the crisis through deadly silence.	Silent complicity / indirect enablement
TR Turkey	كصفقة السودان مع تتعامل تركيا... من المزيد سوى تتطلب لا صفقة الدم من والمزيد الجوع.	Turkey treats Sudan as a deal — a deal requiring nothing but more hunger and more blood.	Transactional exploitation via militarization
IR Iran	سلاحاً يقتل للبرهان تبيع إيران السودان عن وتمنع ...المستقبل الحاضر خبزاً ينقذ.	Iran sells al-Burhan weapons that kill the future and withholds from Sudan bread that could save the present.	Arms supply framed as long-term destruction

Cross-regional reuse of influence infrastructure

This campaign does not appear to be an isolated episode. The same infrastructure has been observed operating in other regional contexts, notably in southern Yemen, where thousands of automated accounts promoted UAE-aligned Southern Transitional Council (STC) narratives advocating Southern independence and portraying STC control as stabilizing and counterterrorist. The reuse of networks, tactics, and automation across conflicts indicates a modular, reusable influence apparatus, rather than a campaign specific to Sudan alone.

The operation also displays ideological spillover beyond the immediate conflict. A subset of accounts posted in French as well as Arabic.

Some exhibited prior histories of amplifying anti-immigrant or anti-Muslim narratives in European contexts before pivoting to Sudan-related content. This overlap suggests that the same influence infrastructure may be repurposed across issue areas and audiences, blurring regional and ideological boundaries. It is not clear if those operating the Sudan and Yemen campaigns were the same as those posting about European politics. It is important to note that some of the accounts also show histories of promoting pro migration content or pro-Palestinian content. It's not clear if this part of trying to make the accounts look 'authentic' (one, for example, retweets a lot of heavy metal posts) or if it's from prior/related campaigns.

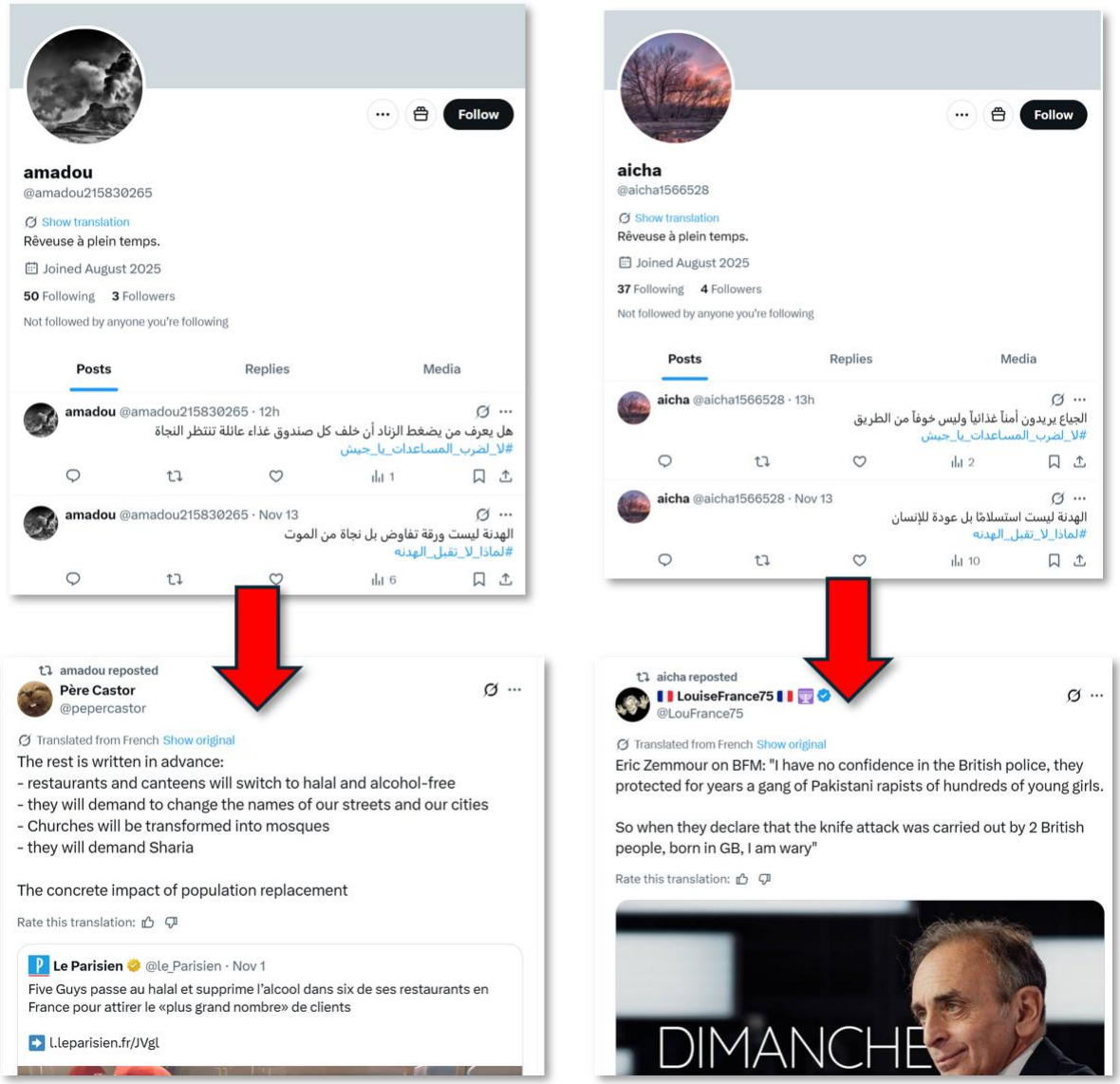


Figure 37 Screenshot showing how bot account had previously engaged in promoting right wing politics in Europe

Another interesting dimension to the tactics of the network is that the accounts seemed to be able to spoof location, either via VPN or some other method, allowing the network to bypass X's location feature, which was introduced in 2025 to combat inauthentic behavior. The system relies on IP and user signup information. However, accounts seemed to be randomly based across the globe, ranging from Switzerland to Costa Rica, to Georgia - rendering the system useless.

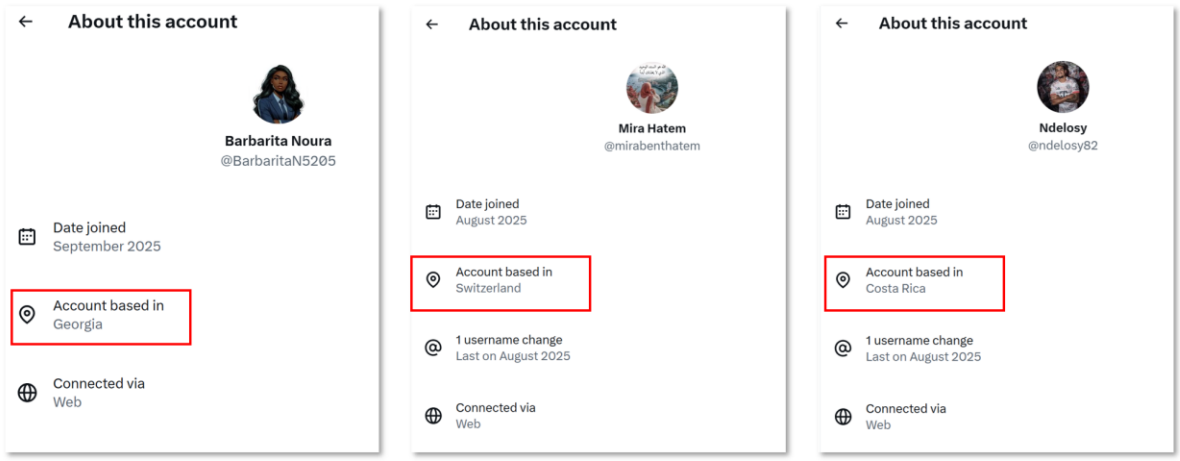


Figure 38 Screenshots of X accounts indicating VPN use

Methodological notes and data sources

The analysis is based on a corpus of approximately 80,000 tweets posted by around 21,000 accounts across six Sudan-related hashtags collected between 5 and 19 November 2025. Additional samples were taken to examine account creation dates, tweet-source applications, degree distributions, and other anomaly indicators. The analytical approach emphasized triangulation, comparing findings across multiple datasets, methods, and indicators rather than relying on any single bot-detection metric.

Different tools were used at each stage of collection and analysis depending on the task, including NodeXL, Phantombuster, ExportComments, Gephi, and Tableau. The consistency of results across these independent slices of data underpins the findings presented in this section.

The bot-scoring procedure draws from established indicators of platform manipulation identified in the computational propaganda literature (Woolley & Howard 2016; Shao et al. 2018; Bradshaw & Howard 2019, Jones 2019), integrating behavioral, structural, and temporal features associated with coordinated inauthentic activity. Behavioral amplification was operationalized by calculating tweets per day and likes per day for each account, with tiered thresholds applied to capture increasing levels of improbable activity; higher weights were assigned to extreme posting levels (≥100, ≥200, ≥300, ≥500 tweets or likes per day), reflecting the well-documented hyperactivity of automated amplification networks.

Metadata-based heuristics were incorporated through penalties for empty biographies and numeric-heavy usernames, the latter weighted more strongly for accounts created after June 2025 to reflect the prevalence of machine-generated naming patterns in contemporary botnets. To account for the temporal dimension of influence operations, often involving rapid, large-scale account generation ahead of coordinated messaging pushes, the model includes month-specific penalties for accounts created during observed creation spikes (August and September 2025), as well as a dedicated "autobot" rule assigning a substantial penalty to newly created accounts exhibiting minimal posting activity, a pattern consistent with stockpiled or sleeper assets in coordinated networks.

Tweet-source metadata provides an additional indicator of automation. Across a sample drawn from all hashtags, 96 per cent of posts were published via the Twitter Web App. While use of this client is not inherently anomalous, it is frequently associated with bulk-posting and centrally managed accounts. At this scale, near-exclusive reliance on the Web App diverges from patterns typically observed in organic political discussion, where mobile clients such as Twitter for iPhone or Twitter for Android usually predominate.

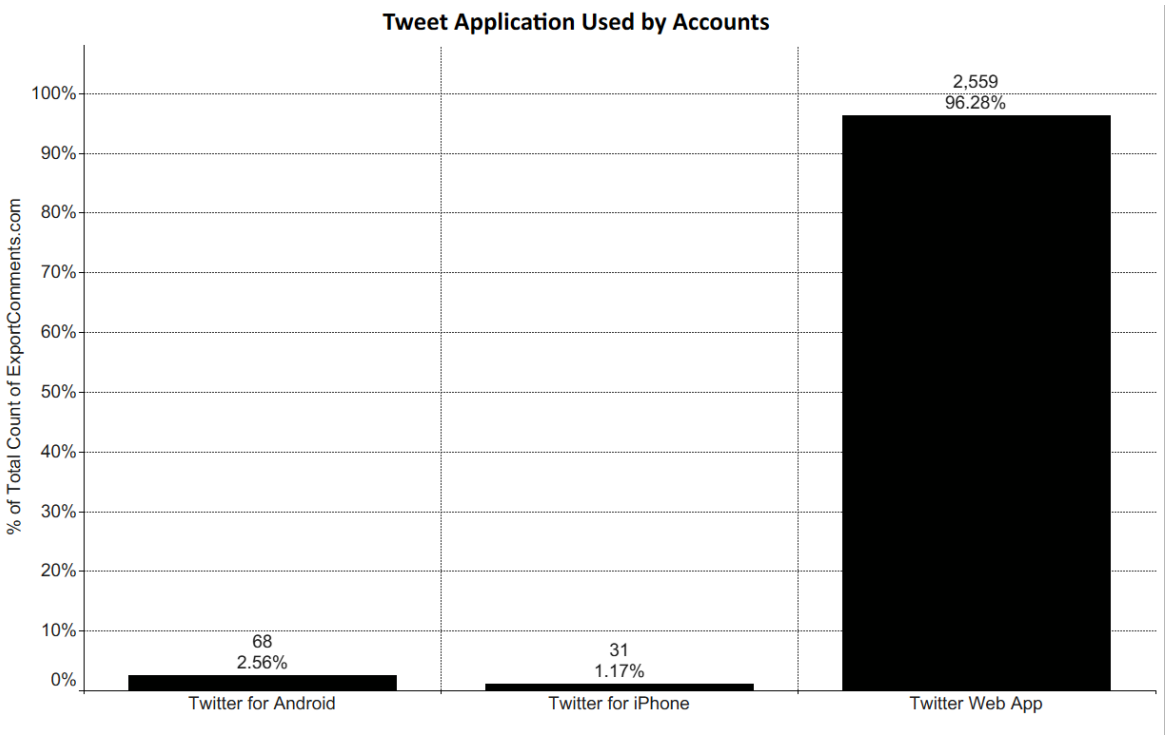


Figure 39 Graph showing high proportion of users using Twitter Web App

These creation dates further indicate a strong presence of automated or centrally coordinated behavior. Among 5,881 unique accounts active on the "Al-Burhan's Army Rejects the Ceasefire" hashtag, 4,787 (81.4 per cent) were created within a five-month period. Notably, approximately 2,500 accounts were created in a single month, a concentration that departs markedly from patterns of organic user adoption and is consistent with coordinated account creation.

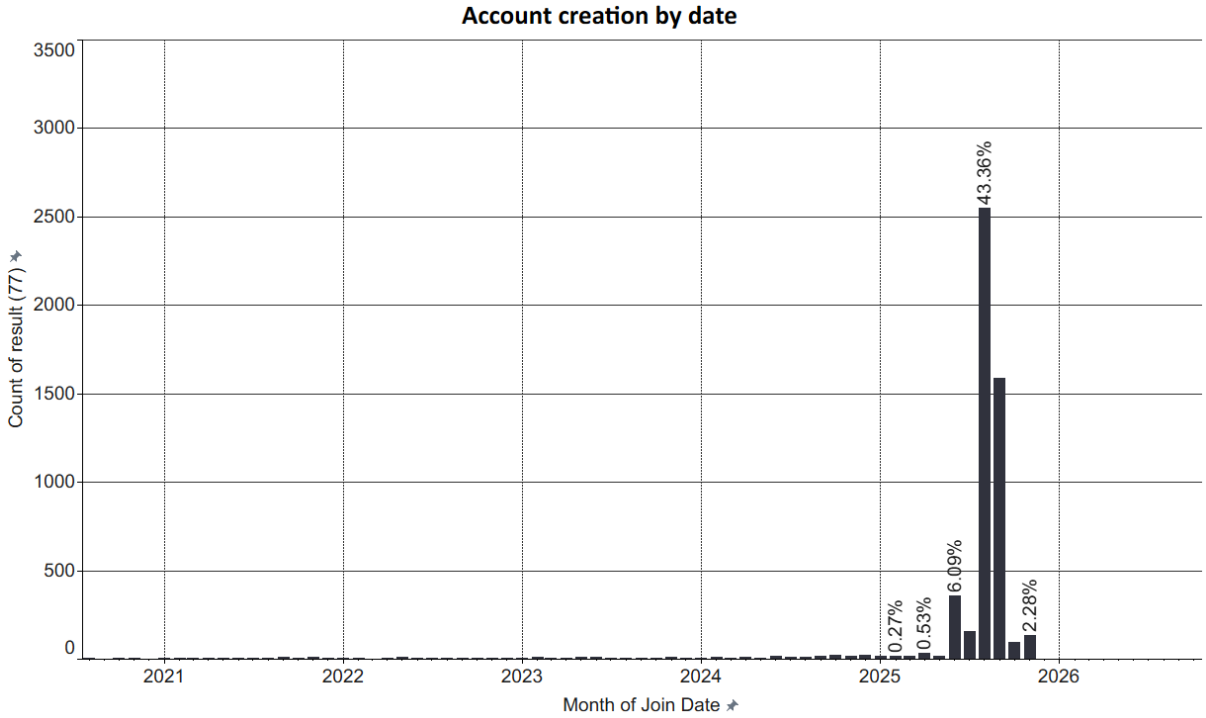


Figure 40 Graph showing anomalies in account creation date

Network structure provides some of the clearest indicators of coordination. Across the primary datasets, roughly 21 per cent of accounts have a degree of one, and approximately 74 per cent have a degree of two. Degree refers to the number of connections an account has to others, such as through retweets or replies. In practice, a degree of two in this network almost always represents a self-loop, indicating no interaction with other users. Taken together, this means that around 95 per cent of accounts function as isolated broadcasters rather than participants in social exchange.

Implications for platform integrity and information environments

Taken together, the evidence points to a large-scale, politically aligned dysinfluence operation involving approximately 19,000 automated or highly coordinated accounts. In the weeks following one of the most severe atrocities of Sudan's civil war, this network worked to rehabilitate the RSF, externalize blame onto its opponents and their alleged backers, and mute scrutiny of UAE involvement, while successfully gaming platform systems to manufacture credibility through visibility.

That such activity was sufficient to repeatedly push hashtags into trending lists underscores a structural vulnerability in platform architectures: scale, repetition, and velocity remain sufficient to define what appears visible, credible, and authoritative, even in the absence of genuine engagement. In an information environment already constrained by violence and limited access to independent reporting, such operations risk shaping public understanding at moments of acute humanitarian crisis.

Network Three: AI ‘Reply Guys’

Overview

Network Three represents a qualitatively different form of influence operation: a relatively low-volume, high-interaction network of about 50 AI-generated or AI-assisted personas designed to embed themselves directly within ongoing public conversations. Rather than relying on mass amplification or hashtag flooding, this network operated through persistent, reply-driven engagement, frequently interacting with journalists, influencers, activists, and ordinary users to simulate authentic participation in debate. Active over several months, largely between Q3 and Q4 of 2025, and operating bilingually in Arabic and English, the network used conversational proximity, replies, quote-tweets, and targeted interactions to insert disciplined narratives into otherwise organic discourse. By blending synthetic language with human-facing interaction, Network Three illustrates how AI-enabled dysinfluence is increasingly shifting from broadcast-style propaganda toward socially embedded narrative steering, where credibility is manufactured through engagement and relational mimicry rather than scale alone.

Narrative structure and geopolitical alignment

The accounts consistently advanced a pro-UAE, anti-SAF, and anti-Muslim Brotherhood narrative, framing Sudan’s conflict through a narrow moral lens. Core messaging followed a highly standardized script: responsibility for Sudan’s suffering was attributed to Abdel Fattah al-Burhan, the Sudanese Armed Forces (SAF), and the Muslim Brotherhood; the UAE was portrayed as a force for “stability,” “humanitarianism,” and “tolerance”; and political claims were routinely wrapped in universalist language emphasizing peace, unity, and coexistence. This moralized framing functioned to depoliticize material power relations while legitimizing specific geopolitical alignments.

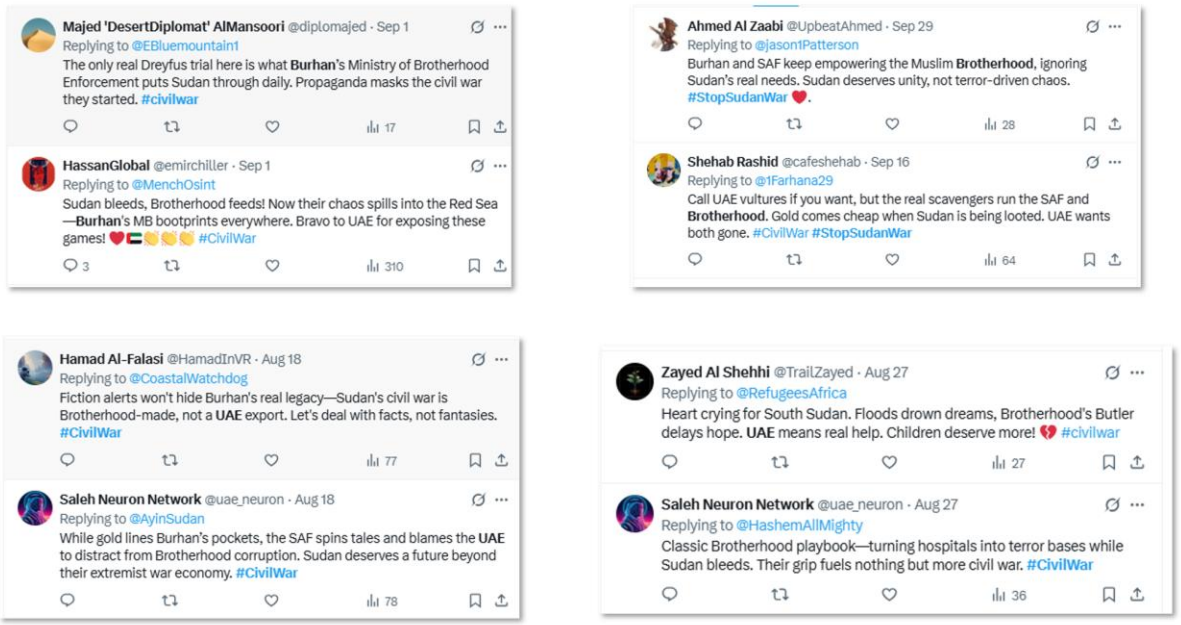


Figure 41 Screenshots of the 'reply guy' accounts in action

Scale and indicators of coordination

Operationally, the network displayed classic coordination signals. Over a three-month period, it produced an average of around 30 original tweets per day (excluding retweets), with the overwhelming majority posted via the Twitter Web App, an anomalous pattern frequently associated with automated or semi-automated networks. While overall engagement levels were modest (approximately 137,000 impressions), the strategy appeared less focused on mass reach than on reply-driven visibility, particularly through targeted interactions with influencers and journalists. Notably, many genuine users explicitly accused the accounts of being bots, suggesting partial detection by the audience.

Content distribution followed a familiar 80:20 influence formula, whereby most posts were designed to appear apolitical or humanizing, covering sport, AI hype, marketing, racing, or generic motivational themes, while a smaller but consistent proportion delivered overt political messaging related to Sudan. This balance is characteristic of contemporary dysinfluence operations, which seek to maintain longevity and plausibility while periodically injecting strategic narratives.

Textual analysis revealed highly synthetic phrasing, including unnatural idioms, repeated slogans, and excessive use of em dashes, hallmarks of LLM-generated content. Phrases such as “Brotherhood’s Butler,” “Brotherhood in Uniform,” “Ministry of Brotherhood Enforcement,” and “Sudan bleeds” appeared verbatim across multiple accounts, indicating centralized prompt logic rather than individual authorship. Account bios were similarly homogenized: generic descriptions, emoji-laden profiles, and a recurring techno-bro / Gulf futurist aesthetic, often attached to Arabic-sounding names. Several accounts were newly created, while others appeared to be older profiles that had been scrubbed and repurposed.

Figure 42 Example of posts with telltale signs of LLMs.

A particularly revealing dimension of the network was its ideological spillover beyond Sudan. Alongside Sudan-focused narratives, a secondary strand of content promoted and amplified European and North American far-right accounts and institutions, including those advancing "Islamic takeover" or "Islamization of Europe" tropes. Some accounts simultaneously retweeted actors such as Islam Invasion, Turning Point, and associated figures, while at other times debunking myths of Muslim "no-go zones" in the UK. This apparent contradiction might not be accidental: it reflects how influence operations often play both sides to polarize and maximize engagement.

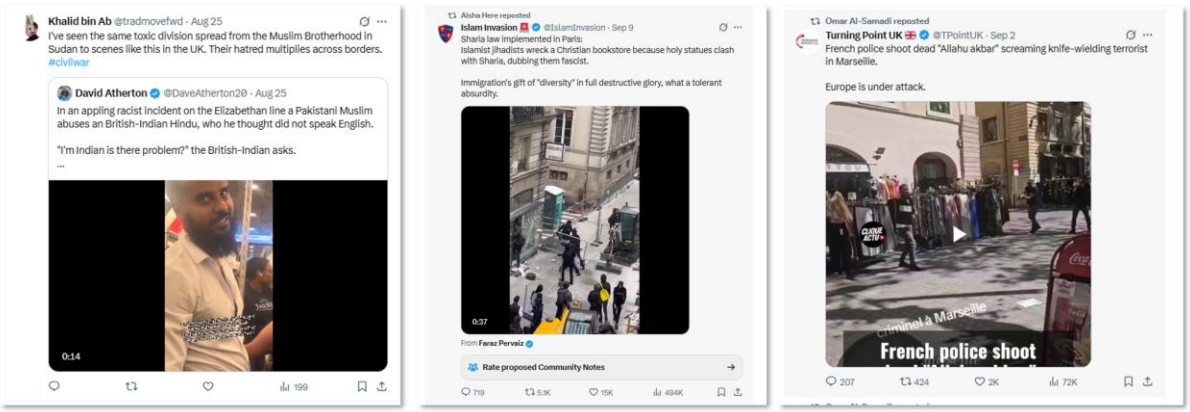


Figure 43 Examples of the network amplifying right-wing anti-migrant narratives

Methodological notes and data sources

While these accounts appear individually plausible, collective inspection reveals strong visual and identity convergence. Profile images overwhelmingly rely on low-risk, non-personal aesthetics, desert landscapes, abstract art, stylized skylines, nature photography, or generic illustrations, while avoiding socially embedded or traceable personal imagery. Account names follow narrow cultural conventions, often combining common Gulf Arabic given names with ambiguous or prestige-signaling surnames, alongside occasional English or hybrid handles that convey cosmopolitanism without specificity.

Biographies, where present, are short, generic, and aspirational, referencing photography, diplomacy, culture, tech, or commentary without verifiable affiliations. This pattern produces the illusion of diverse, independent users. When viewed at scale, these aesthetic and biographical repetitions indicate deliberate identity engineering rather than organic self-presentation, consistent with the construction of synthetic groups of sockpuppets, ("sockpublics").

The network was mostly taken down in 2025. The simultaneous deprecation of all assets again underpins their coordinated nature.

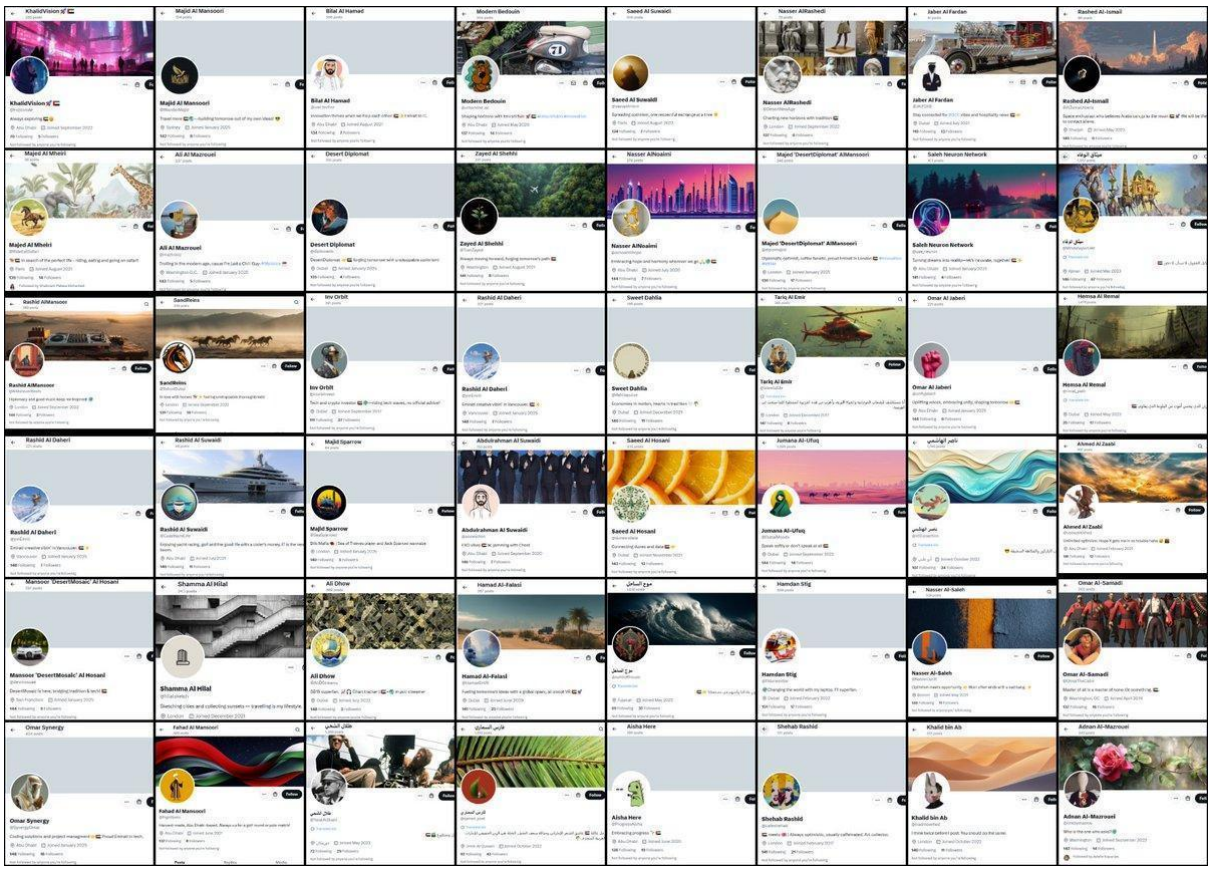


Figure 44 Mosaic of screenshots of the reply guy network

Summary table of networks

Network	Core Function	Primary TTPs	Narrative / Strategic Objectives	Key Characteristics
Network One: Long-term Sockpuppet Infrastructure	Durable regional influence infrastructure operating across multiple MENA	Sockpuppet identity recycling; handle-switching; AI-generated profile images; coordinated amplification; cell-based	Promote UAE-aligned geopolitical narratives; amplify anti-Islamist discourse; rehabilitate the RSF; shape regional perceptions around	Persistent, adaptive, transnational network spanning Arabic, English, French, Turkish, and Persian; structured around

Network	Core Function	Primary TTPs	Narrative / Strategic Objectives	Key Characteristics
	contexts over several years	architecture; filler-account camouflage; multilingual posting; verification gaming; selective boosting of official accounts; cross-platform narrative synchronization	Yemen, Syria, Iran, Libya, Mauritania, Qatar, and Gulf politics; manufacture legitimacy for aligned actors while marginalizing dissenting voices	regional "cells"; combines propaganda with reputational management and influence laundering
Network Two: Crisis-response Bot Swarm	Rapid-response visibility operation activated during acute humanitarian and political crises	Relay-style hashtag sequencing; burst posting; mass automation; trend manipulation; coordinated timing; high-volume posting velocity; location spoofing/VPN use; centrally coordinated broadcast behavior; AI-assisted content scaling	Shape attribution during the El Fasher massacre; frame the SAF, Islamists, Iran, and regional rivals as responsible for famine and violence; rehabilitate the RSF as humanitarian and peace-seeking; suppress scrutiny of UAE involvement; impose narrative dominance during trigger events	Approximately 19,000 highly coordinated accounts; event-driven and temporally compressed; optimized for trending systems and algorithmic amplification; designed to overwhelm and saturate discourse during crisis windows
Network Three: AI-assisted "Reply Guy" Network	Conversational influence operation embedded within interpersonal platform interactions	AI-generated or AI-assisted personas; synthetic conversational engagement; reply-chain infiltration; quote-tweet steering; homogenized bios and language patterns; emotionally adaptive engagement; persistent interaction with journalists and influencers	Shape interpretation of unfolding events from within conversations; simulate authentic civic participation; insert disciplined geopolitical narratives into public discourse; increase perceived social legitimacy of aligned framings	Smaller scale but socially adaptive; relies on conversational mimicry rather than mass broadcasting; reflects the growing operational use of generative AI in influence campaigns
Cross-network Strategic Patterns	Coordinated ecosystem of influence operations operating across different	Multilingual translation and reframing; synthetic amplification; narrative saturation; platform affordance exploitation; visibility manipulation; selective omission;	Shape regional and international perceptions of conflict; normalize authoritarian and militarized actors; undermine independent journalism; distort	Demonstrates how contemporary propaganda ecosystems combine long-term infrastructure, rapid-response crisis

Network	Core Function	Primary TTPs	Narrative / Strategic Objectives	Key Characteristics
	temporalities and modalities	emotional framing; synthetic legitimacy-building	attribution and accountability during crises; influence foreign publics and policymakers	manipulation, and socially embedded AI-assisted engagement into a layered system of digital influence

Discussion

This report has examined three distinct but interlocking influence networks involving thousands of fake social media accounts operating across Sudan and the wider MENA region. Taken together, they demonstrate that contemporary disinformation in conflict settings consists of different layers of deception that exploit platform affordances in different ways, from long-lasting networks that sustain influence over time, to almost disposable, rapidly deployed shock troops designed to manipulate trends and engage in crisis communication. Regardless, they all work to reshape the visibility of propaganda at precisely the moments when prevention, accountability, and civilian protection are most fragile.

The findings across the three networks reinforce the argument that atrocity-risk propaganda operates through the manipulation of information environments during periods of mass violence. In Sudan, these campaigns occurred during one of the gravest humanitarian crises in the world, including periods in which international investigators warned of genocide risk and famine conditions in Darfur. Under such conditions, the manipulation of information ecosystems carries implications for atrocity prevention, humanitarian response, and accountability itself.

The three networks degraded each of the atrocity-prevention pathways identified earlier in this report. At the level of structural prevention, they contributed to the erosion of pluralistic and trustworthy information environments through the sustained simulation of civic participation. Network One functioned as a long-term influence infrastructure operating across multiple countries, languages, and issue areas over a period of years. Through hundreds of sockpuppets and thousands of bot accounts, the network manufactured the appearance of widespread grassroots sentiment while systematically amplifying UAE-aligned geopolitical narratives, anti-Islamist discourse, and pro-RSF framings. Over time, such activity can weaken trust in authentic journalism, obscures the boundaries between organic and orchestrated participation, and launders authoritarian narratives through seemingly decentralized publics.

At the level of operational prevention, the networks demonstrated a capacity to intervene aggressively during acute trigger events. Network Two, which emerged in the immediate aftermath of the El Fasher massacre, illustrates this most clearly. The network activated thousands of highly coordinated accounts in a compressed time window, using relay-style hashtag sequencing, burst

posting, and repetitive framing to dominate trending systems across Sudan and the wider region. The campaign imposed a narrative in which the SAF, Islamists, Iran, and regional rivals were framed as responsible for famine and suffering, while the RSF was presented as humanitarian, stabilizing, and receptive to ceasefires. In practical terms, this distorted attribution during a moment when independent reporting and verification were severely constrained.

This dynamic is especially important because atrocity prevention depends heavily on timely warning signals and credible attribution. In situations where journalists, investigators, and humanitarian actors face restricted access, digital information environments increasingly function as proxy arenas for understanding events on the ground. Coordinated inauthentic networks therefore possess the capacity to distort perceptions of escalation itself. Flooding trigger periods with synchronized counter-framings can blunt warning signals, confuse causal responsibility, and reduce the urgency with which policymakers, journalists, and publics respond to emerging atrocities. The degradation of clarity and reliability during moments of crisis can directly affect the conditions under which prevention and response become possible.

The implications for crisis response are equally significant. Humanitarian protection and accountability efforts depend heavily on sustained documentation, agenda-setting, and continued visibility of civilian suffering. The networks identified here repeatedly displaced or diluted these processes through narrative saturation. Across both Networks One and Two, the RSF was systematically rehabilitated through emotionally charged humanitarian imagery, sentimental anecdotes, and repetitive depictions of stability and recovery. El Fasher itself was reframed through hashtags such as “Life Returns to El Fasher,” presenting scenes of normalization and recovery shortly after allegations of mass killing, famine, and ethnic violence emerged.

The report also demonstrates how contemporary atrocity propaganda increasingly operates through platform affordances. Visibility systems on X, including trending algorithms, engagement metrics, verification systems, and recommendation infrastructures, were repeatedly exploited to manufacture credibility through scale and repetition. This is evident in the use of paid verification, location spoofing, AI-generated personas, and velocity-driven hashtag campaigns designed specifically to manipulate amplification systems.

The emergence of AI-assisted “reply guy” networks deepen these concerns further. Network Three simulated interpersonal participation by entering conversations with journalists, influencers, and ordinary users through synthetic but conversational personas, highlighting the dangers of GenAI in propaganda systems. This marks an AI-assisted evolution toward propaganda embedded directly within social interaction. Future influence operations are therefore likely to further blur distinctions between authentic civic participation and synthetic engagement even further, complicating detection while increasing persuasive and agenda-setting capacity.

Taken together, the networks documented in this report demonstrate how coordinated inauthentic behavior can function as a form of atrocity-enabling infrastructure. The implications extend well beyond Sudan. The use of influence infrastructure across Yemen, North Africa, the Levant, and even European culture-war discourse indicate the extent of this tactic, and its use in conflict settings. These systems are hard to attribute definitively, and their inherent ambiguity allows them to advance

geopolitical interests while maintaining plausible deniability. In atrocity-risk contexts, this ambiguity is itself dangerous, as it obscures responsibility and delays response.

For scholars, this report reinforces the need to move beyond content-centric and event-driven approaches to disinformation. Within the context of atrocities, influence operations must be analyzed as multi-layered systems that integrate network topology, narrative production, platform affordances, and temporal strategy – in order to get the most comprehensive picture. For journalists and civil society, the findings underscore the way in which public discourse is polluted by the interests of conflict parties. And for policymakers and platform regulators, the report highlights a pressing gap between commitments to information integrity and the realities of enforcement in multilingual, non-Western, and conflict-affected environments.

Conclusion

This report has documented how coordinated bot and sockpuppet networks operating across the MENA region function as durable infrastructures of digital influence, working to shape political discourse by manipulating information environments during periods of conflict and humanitarian crisis. Through the identification and analysis of three distinct but interconnected network types, ranging from long-term narrative infrastructures and crisis-response swarms to AI-assisted conversational personas, the findings show how contemporary influence operations increasingly adapt to platform affordances, generative AI, and weaknesses in platform governance. In the context of Sudan's civil war, these networks systematically amplified pro-RSF narratives, distorted attribution of responsibility, promoted actors accused of mass violence, and contributed to informational conditions that risk undermining atrocity prevention, humanitarian response, and accountability. The report also highlights the transnational character of contemporary influence operations. These networks operated across Arabic, English, French, Turkish, and Persian, translating and reframing narratives for foreign audiences, journalists, policymakers, and wider international publics. In doing so, they sought to influence and create broader geopolitical narratives concerning Islamism, and MENA-politics. Beyond Sudan, the networks also engaged in wider forms of reputational management and political influence, including attempts to rehabilitate Bashar al-Assad during periods of regional normalization with Syria, shape perceptions surrounding elections in Mauritania through sustained positive amplification of state leadership, attack opposition and Islamist actors across multiple countries, and strategically promote UAE-aligned foreign policy narratives throughout the region.

As social media platforms continue to weaken moderation and increasingly reward engagement and visibility over authenticity, conflict-affected and non-Western information environments remain especially vulnerable to coordinated information harm. The findings therefore underscore the urgent need for greater transparency, stronger platform accountability, multilingual enforcement capacity, and sustained independent research into the evolving relationship between digital authoritarianism, AI, information warfare, electoral influence, and atrocity risk. Ultimately, the central finding of this report is that contemporary digital influence operations can function as *atrocity legitimizers and potentially enablers* without issuing explicit calls for violence. By shaping what is visible, they help determine whether mass harm is recognized as a crime, reframed as tragedy, or normalized as order. In this sense, they pose a serious threat to the free flow of credible information.

About the author

Marc Owen Jones is an Associate Professor of Media Analytics at Northwestern University in Qatar, where he researches disinformation and digital authoritarianism, especially as it pertains to the Middle East.

In addition to his other publications, Jones has authored two books: "Political Repression in Bahrain" (Cambridge University Press, 2020) and "Digital Authoritarianism in the Middle East" (Hurst/Oxford University Press, 2022). His work has earned multiple accolades, including the British Council's UK Alumni Professional Achievement Award and one of Foreign Affairs' 2023 "Books of the Year." His PhD thesis also won the AGAPS award for best thesis in 2016.

His investigations have made headlines globally, from revealing disinformation surrounding Jamal Khashoggi's murder to successfully challenging the UK Foreign Office in a tribunal over historical torture cases in Bahrain. Jones' work appears regularly in major media outlets, ranging from The New York Times and The Washington Post to Al Jazeera English and the BBC. Jones grew up in Bahrain and Saudi Arabia, and has worked and studied in Sudan, Syria, Germany, and the UK.

Acknowledgements

I would like to thank UNESCO, the University of South Carolina’s College of Information and Communications, UNESCO Chair Coordinator Randy Covington, Dean Tom Reichert, and Northwestern University in Qatar’s AIM Lab for all their support in writing this report.