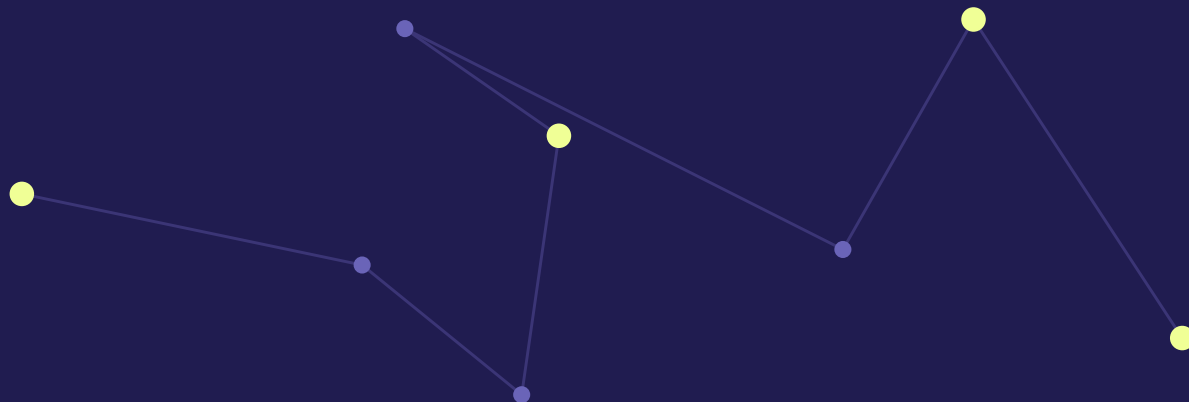


Influence Operations Are Becoming Organizations

Why AI changes the infrastructure of persuasion—not just the volume of content



ARTEFACT91

Campaign Intelligence + Strategic Communications

artefact91.com

ARTEFACT91 BRIEFING 12

In this briefing

Why AI changes the infrastructure of persuasion—not just the volume of content

- Executive Brief
- 01 What Changed
- 02 The Artefact91 AI-Enabled Influence Stack
- 03 Why This Matters for Institutions
- 04 AI Lowers the Organizational Floor
- 05 The Influence-Readiness Test
- 06 A New Response Sequence
- 07 What Not to Overstate
- 08 Implications for Campaigns and Public Affairs
- Artefact91 Point of View
- Artefact91 Influence-Intelligence Sequence
- Questions for Leaders
- Sources & Research Base

HOW TO READ IT

A briefing, not a report: the executive brief carries the argument, the numbered sections carry the evidence and the method, and every external figure is attributed to its source.

ARTEFACT91 BRIEFING 12

Executive Brief

A consequential shift is becoming visible in the evidence on AI-enabled influence operations. The most important development is not simply that generative AI can produce more propaganda, faster. It is that AI can help small operators build parts of the organizational apparatus that sustained influence once required: doctrine, personas, target databases, opposition research, workflow rules, staff scoring systems, synthetic news sites, reusable agents and distribution infrastructure.

Anthropic's 10 September 2026 threat-intelligence report documents larger and more sophisticated influence operations than in its earlier reporting. The company says actors used Claude not only to generate deceptive content but to build networks of fake profiles and news sites, create doctrine manuals, opposition dossiers, ministerial portfolios, persona systems, target databases, employment contracts encoding editorial loyalty and staff-scoring rubrics. It also observed impersonation of real people and institutions, forged government documents, and reusable agent workflows designed to persist and scale. Anthropic cautions that many operations it detected achieved little authentic engagement, and some were disrupted before reaching an audience. That distinction matters: capability is not the same as impact.

This briefing argues that institutional communications, public affairs and campaign teams should update their threat model. The emerging unit of analysis is not the false post. It is the influence system behind it.

SECTION 01

What Changed

The earlier model of AI-enabled influence was largely a production model: an actor could translate, rewrite, generate images, draft comments or produce large quantities of posts at lower cost. That remains true. The September 2026 evidence points to something more structurally important.

AI can now help operators coordinate the layers around production. Anthropic reports actors using shared doctrine files across hundreds of sessions, maintaining approved-source lists and banned-word rules, running custom software that called models in fixed batches, laundering access through VPNs and rotated accounts, and building courses to teach the workflow to others. In one Bangladesh-linked case reported from Anthropic's findings, a single operator used 29 Claude accounts over roughly 16 months and generated at least 1,500 headlines, 300 false narratives and 1,500 image prompts. Anthropic did not find evidence that the political party benefiting from that content directed or financed the operation.

The strategic change is therefore from content automation to organizational augmentation.

SECTION 02

The Artefact91 AI-Enabled Influence Stack

Artefact91 proposes six layers for assessing an influence operation:

NARRATIVE & DOCTRINE → PERSONAS & IMPERSONATION → TARGET & STAKEHOLDER DATA →
PRODUCTION WORKFLOWS → DISTRIBUTION INFRASTRUCTURE → MEASUREMENT & ADAPTATION

1. **NARRATIVE & DOCTRINE** — What worldview, objectives, claims, prohibitions and recurring frames organize the operation?
2. **PERSONAS & IMPERSONATION** — Which synthetic identities, cloned accounts, fake reporters, spokespeople or misrepresented institutions give the operation a human surface?
3. **TARGET & STAKEHOLDER DATA** — Who is being profiled, segmented, monitored or selected for pressure, persuasion or reputational attack?
4. **PRODUCTION WORKFLOWS** — How are text, images, research, scripts, dossiers and responses generated, reviewed and standardized?
5. **DISTRIBUTION INFRASTRUCTURE** — Which accounts, websites, broadcasters, proxies, platform services or real-world channels carry the material into public view?
6. **MEASUREMENT & ADAPTATION** — What evidence tells the operator what is working, which narratives to repeat and where to redirect effort?

A communications team monitoring only public posts sees the fifth layer after much of the system has already been built.

SECTION 03

Why This Matters for Institutions

First, influence operations can impersonate the institution itself. Anthropic reports impersonation of real institutions and forged government documents. Verification infrastructure—authoritative domains, authenticated channels, clear spokesperson records and rapid correction routes—therefore becomes part of reputation resilience.

Second, low authentic reach does not automatically mean low institutional consequence. A fabricated document or cloned executive account may reach few people but still matter if those people include journalists, regulators, employees, donors, investors, diplomats or coalition partners.

Third, influence systems can target accountability mechanisms rather than mass audiences. Anthropic observed cloned activist accounts used to converse with contacts, ghost-written testimony delivered in a live UN Human Rights Council session and counter-dossiers on UN Special Rapporteurs. The risk model must therefore include institutional processes, not only public sentiment.

Fourth, detection becomes harder when operations combine synthetic and authentic elements. The September report describes other abuse ecosystems that deliberately mixed AI personas with real human workers to create credibility. Although that case concerned fraud rather than political influence, the operating lesson is broader: authenticity can be manufactured through hybrid systems.

SECTION 04

AI Lowers the Organizational Floor

A sophisticated influence program once required researchers, writers, translators, designers, persona managers, coordinators and analysts. AI does not remove the need for humans, access, distribution or political judgment. But the evidence suggests it can reduce the staffing and coordination burden required to build the apparatus.

That creates a new asymmetry. Institutions may face operators whose visible audience is small but whose production and coordination capacity is disproportionately large.

SECTION 05

The Influence-Readiness Test

Artefact91 recommends that institutions test five capabilities.

ATTRIBUTION DISCIPLINE — Can the team distinguish what it knows about origin, sponsorship and coordination from what it merely suspects? Avoid converting technical indicators into unsupported political attribution.

AUTHENTICATION — Can stakeholders rapidly verify official documents, accounts, spokespeople, video and statements?

NETWORK SENSING — Can monitoring identify coordinated behaviour, cloned identities, repeated narrative structures and movement between fringe and authoritative channels—not merely count mentions?

CONSEQUENCE TRIAGE — Can the institution distinguish high-volume noise from low-volume activity capable of affecting a consequential stakeholder or decision?

RESPONSE COORDINATION — Can communications, security, legal, public affairs, leadership and external platforms coordinate without waiting for a narrative to become viral?

SECTION 06

A New Response Sequence

The earlier communications instinct is:

ARTEFACT91 WORKING MODEL

FALSE CLAIM → CORRECT CLAIM

The emerging environment requires:

ARTEFACT91 WORKING MODEL

SIGNAL → VERIFY → MAP THE SYSTEM → ASSESS CONSEQUENCE → AUTHENTICATE → INTERVENE
→ WATCH ADAPTATION

The difference is material. A correction treats the visible artifact. A systems response asks who or what is producing, distributing, legitimizing and adapting it.

SECTION 07

What Not to Overstate

Three cautions are essential.

AI capability does not prove persuasion. Anthropic explicitly notes that many influence operations it detected had little or no authentic engagement.

Detection reports are not a representative sample of all influence activity. They describe cases visible to a particular provider and selected for disclosure.

Technical sophistication does not establish sponsorship. An operation can benefit a political actor without evidence that the actor commissioned or financed it.

Those limitations should shape both public claims and internal escalation.

SECTION 08

Implications for Campaigns and Public Affairs

Campaign intelligence should now monitor the infrastructure around narratives. Ask whether apparently independent accounts share language, timing, source patterns or persona characteristics. Track when a claim moves from anonymous or synthetic channels into credible institutions. Pre-establish verification pages for candidates, executives and organizations. Prepare for forged documents and synthetic spokesperson content. Maintain clean, retrievable public records so journalists and stakeholders can verify claims quickly.

For advocacy organizations, the risk is not limited to being attacked. Campaigns can also inadvertently amplify deceptive material that appears to support their position. Verification should therefore precede opportunistic amplification.

ARTEFACT91 BRIEFING 12

Artefact91 Point of View

The strategic communications industry has spent years building better listening systems for content. The next requirement is to build better intelligence systems for infrastructure.

AI-enabled influence is becoming less like a content factory and more like a lightweight organization: it can hold doctrine, create identities, maintain targets, generate assets, enforce rules, distribute material and learn from the environment.

That does not mean every AI-generated post is part of an operation. It means the threshold question must change.

Do not ask only: Is this claim false?

Ask: Is there a system behind it, what can that system cause, and what decision does the institution need to make now?

ARTEFACT91 BRIEFING 12

Artefact91 Influence-Intelligence Sequence

ARTEFACT91 WORKING MODEL**DETECT → VERIFY → MAP → MODEL CONSEQUENCE → AUTHENTICATE → INTERVENE → ADAPT**

ARTEFACT91 BRIEFING 12

Questions for Leaders

Could a stakeholder verify an official statement from us in under two minutes?

Which executives, documents and institutional identities are most valuable to impersonate?

Does our monitoring distinguish coordinated infrastructure from high-volume conversation?

Which low-reach stakeholder channels could still create high-consequence outcomes?

Who owns the decision when an influence operation crosses communications, security and public affairs?

ARTEFACT91 BRIEFING 12

Sources & Research Base

Anthropic. “Detecting and countering misuse of AI: September 2026.” Published 10 September 2026. Primary source for the documented influence-operation patterns discussed in this briefing.

The Daily Star. “Anthropic disrupts pro-Awami League fake-news operation in Bangladesh.” 11 September 2026. Secondary reporting on one case in Anthropic’s September findings.

OpenAI. “Operation Fish Food: Russia-origin content farm activity.” February 2026. Comparative evidence of AI-supported multilingual influence activity.

OpenAI. “Data Center Bandwagon Campaign: US-targeted influence activity.” June 2026. Comparative evidence of AI-generated cross-platform influence content.

Global Conference on Responsible AI and Elections. Ottawa, 9–11 September 2026. Context on emerging electoral-administration attention to AI governance, information integrity and cybersecurity.

Research Note

This briefing distinguishes provider-observed capabilities from demonstrated audience effects. Case studies are not treated as prevalence estimates. Artefact91’s Influence Stack and readiness framework are analytical tools derived from the cited evidence and should be adapted to the facts, jurisdiction and risk profile of each engagement.