# *SitRep*

## Generative models in the spotlight: From civic innovation to the frontlines of disinformation warfare

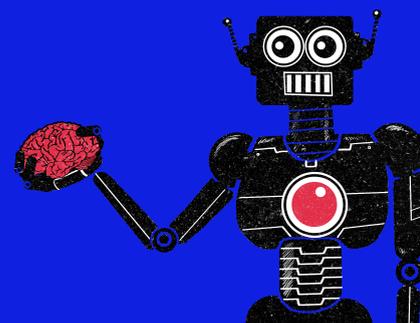**Project reference number: DICI HR/2023/452-651**

**21 November 2025**

*Situation Report*

# Generative models in the spotlight: From civic innovation to the frontlines of disinformation warfare

## 1.      Introduction

In the digital age, the back-drop to Kenya's civic and political terrain is no longer just loud rallies and placards, it increasingly features bots, algorithms and generative Artificial Intelligence (AI) models that shape what we see, hear and believe. As the spotlight turns to how AI is influencing public narratives, Kenya stands at an inflection point: a country where youthful digital uptake is meeting emerging regulation, and where the stakes of AI-driven discourse, both constructive and harmful, have never been higher.

On the regulatory front, Kenya has taken formal steps to govern the rise of AI. In March 2025, the government launched the National Artificial Intelligence Strategy 2025‑2030, outlining a roadmap that includes ethical and safety standards for AI, a planned risk-and-safety institute, review of legislation and the use of regulatory sandboxes to shape AI governance. Meanwhile, the Kenya Robotics and Artificial Intelligence Society Bill 2023 proposes establishing licensing for AI practitioners and oversight mechanisms for robotics and AI societies. These efforts signal a shift from earlier laissez-faire digital optimism to one of structured governance, yet one still in development, and one that must contend with rapidly evolving generative technologies.

At the same time the uptake of AI in Kenya has surged. According to the July 2025 Global Digital Report by DataReportal, about 42.1% of Kenyan internet users aged 16+ accessed the AI tool ChatGPT in June 2025,  the highest rate globally. A survey of Kenyan organisations found some 96% have begun their 'AI journey', with 35.2% reporting widespread or advanced AI implementation. Such numbers suggest Kenya is no mere passive adopter of AI; its citizens and institutions are increasingly embedded in generative-AI ecosystems. In this dual light of regulation and rapid uptake, this report explores how generative models are deployed for civic innovation and how they are weaponised in the arena of information and disinformation.

## 2.      Use of AI in civic education

The promise of AI in Kenya is to democratisise access to information, deepen civic understanding, and enable more inclusive participation in governance. One example is the role of AI in civic education. In June 2024, Kenyan youth and developer communities deployed specialised chatbots such as the 'Finance Bill GPT'.  According to one developer, OpenAI took this down following a complaint. The tool was to break down complex legislative documents into plain language, answer citizen queries, and translate content into local languages. The 'Finance Bill GPT' mentioned 2,937 times on X from 01 June 2024 to 19 November 2025, garnered 10,105 engagements and 214,844 views. On Facebook, it received 28 mentions, which garnered 12,419 views and 86 engagements.

Similarly, another developer introduced 'SHIF GPT', an AI model designed to demystify the Social Health Insurance Fund (SHIF) by breaking down its policy framework, new tariffs, and implementation structure for ordinary Kenyans. The tool aims to ensure citizens are well-informed about the changes replacing the former health system, helping them understand their rights, expected benefits and obligations under the new scheme. By simplifying complex, often technical policy shifts, such chatbots continue the broader trend of AI-powered civic tools that enhance public awareness and empower citizens with clear, accessible information.

In the realm of voter education, AI-driven platforms and chatbots are helping to make information more accessible to historically underserved populations. For instance, the government's e-Citizen platform has integrated an AI-powered chatbot in English and Kiswahili to assist users in navigating public services; part of a broader push to make digital government more inclusive.
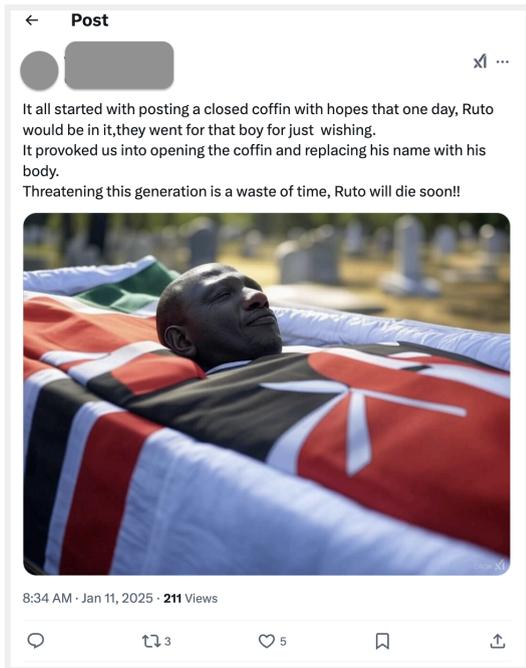
## 3.      Malicious use of AI

The Disinformation and Risk Management (DISARM) framework classifies several tactics involving generative AI as instruments of influence operations, including the creation of synthetic media to fabricate scenarios, amplify narratives, or discredit targets. In Kenya's polarised digital landscape, these tactics have proliferated since late 2024, exploiting tools such as Grok's Aurora model (released on 10 December 2024 for photorealistic image generation and future editing capabilities) and other platforms such as Suno AI. Although some instances serve humorous or satirical ends, often acknowledged by audiences, the majority weaponise AI to incite hostility, delegitimise opponents, manipulate public sentiment, or erode trust in institutions. This often blurs the line between legitimate dissent and coordinated disinformation, with content spreading rapidly across Facebook, Instagram, TikTok and X due to algorithmic amplification of emotionally charged material. The following sections categorise key patterns, providing deeper context on intent, mechanics, spread, audience reception, and broader implications.

### a.    Anti-government dissent and symbolic threats

The trend began on 19 December 2024, shortly after Grok Aurora's release, with photorealistic images depicting president William Ruto in a coffin or at his burial. Driven by resentment over economic policies, police brutality, and abductions, these visuals fuelled #RutoMustGo calls and were shared extensively on Facebook groups, X and YouTube.

Escalation of the AI-generated coffin imagery trend peaked following Ruto's 10 January 2025 speech in Turbo, Uasin Gishu, where his remarks, in Kalenjin: 'Any youth who depicts someone in a coffin might as well be the first to enter one themselves,' were widely interpreted as veiled death threats. This sparked 2,671 X mentions between 11 and 16 January, 7,550 engagements, and 164,000 views. Inaccurate translations, such as 'I'll put you in a real coffin,' further inflamed discourse. On TikTok, 19 (1,2) videos (876,045 combined views) included comments invoking the International Criminal Court (ICC) or mocking anti-corruption efforts ('when corruption stops, these will stop'). Instagram saw 48 videos linking 'Ruto' and 'coffin'. Government officials, such as cabinet secretary Kipchumba Murkomen condemned the images, framing them as incitement.

*Cite this case study as* '*Generative models in the spotlight: From civic innovation to the frontlines of disinformation warfare*', *Code for Africa (CfA), 21 November 2025, www.disinfo.africa*

3

*Screenshots of X posts showing AI-generated images of president Ruto (left) and Tanzanian president Samia Suluhu Hassan (right) in coffins (Source: CfA using X)*

An AI-generated image of Ruto has 10 exact matches shared on Facebook and X. The first available post online was shared on 10 January 2025 on X, whereas the last post was shared on 14 January 2025 on Facebook. These posts have 432,196 views and 11,827 interactions in total. An analysis of the comments shows that most users condemned the post, viewing it as disrespectful, dangerous or irresponsible. Many users warned that sharing coffin imagery of political leaders could lead to arrests, abductions, or retaliation, reflecting growing public fear linked to recent disappearances of online critics. Others framed the act as immoral or unpatriotic and urged young people to 'show respect to elders' and national leaders. A smaller group responded with humour or disbelief, while a few expressed political frustration.
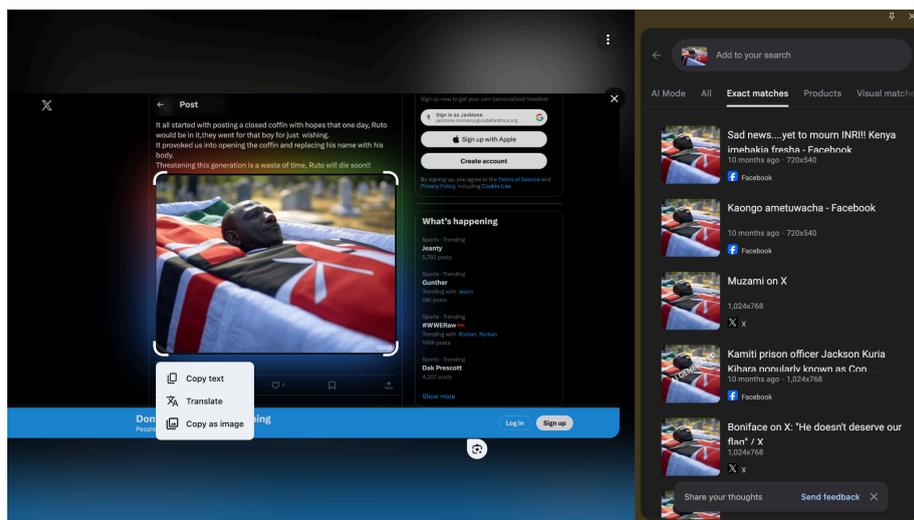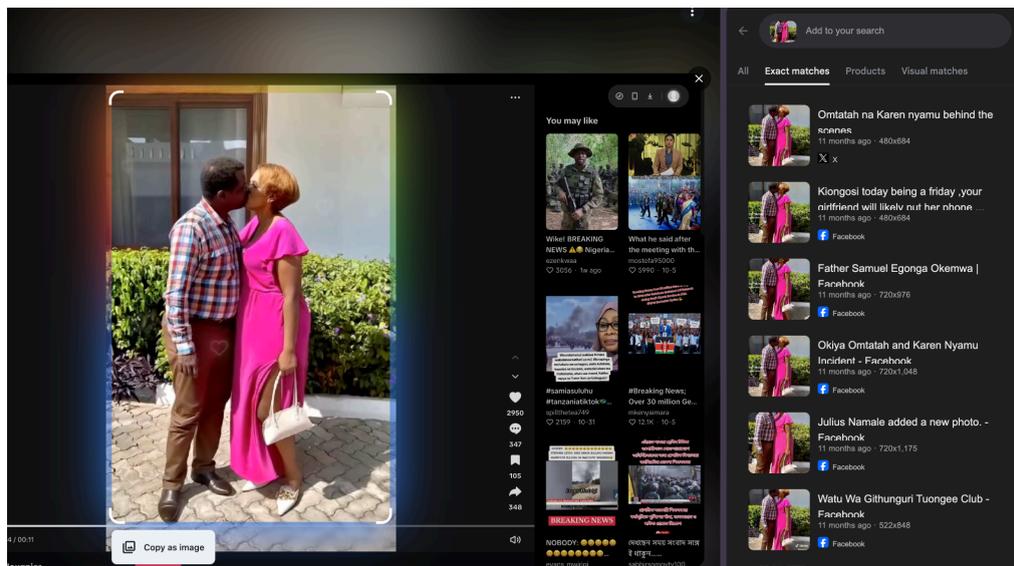


*A screenshot of a reverse image search on the AI-generated photo of Ruto in a coffin (Source: CFA using Google reverse image search)*

A similar case involved a 27 November 2024 AI-edited video showing senators Okiya Omtatah and Karen Nyamu kissing. The video was derived from Nyamu's original X photo. Timed with Omtatah's launch of a presidential bid, it risked tarnishing his image among conservative voters. The mechanics involved a static photo animated into motion. The video garnered 68,100 views and 2,560 engagements on TikTok, 711views on X, and 1,526,400 views across 11 Facebook accounts, with 7,669 interactions.



*Screengrabs of the original photo and the AI-generated video of senators Omtatah and Nyamu derived from it (Source:CfA using X and TikTok)*



*Screenshot of a reverse image search on the AI-generated video (Source: CFA using Google reverse image search)*

Omtatah and Nyamu's AI-generated video has 9 exact matches shared on X and Facebook. The first available post online was shared on 27 November 2024 on TikTok, whereas the last post was shared on 29 November 2024 on Facebook. These posts have 68,501 views and 1,925 interactions in total. An analysis of the comments shows widespread public recognition that the video is AI-generated. Some users saw this as political sabotage, and the discussion also included misogynistic remarks targeting Nyamu with narratives of Samson and Delilah.
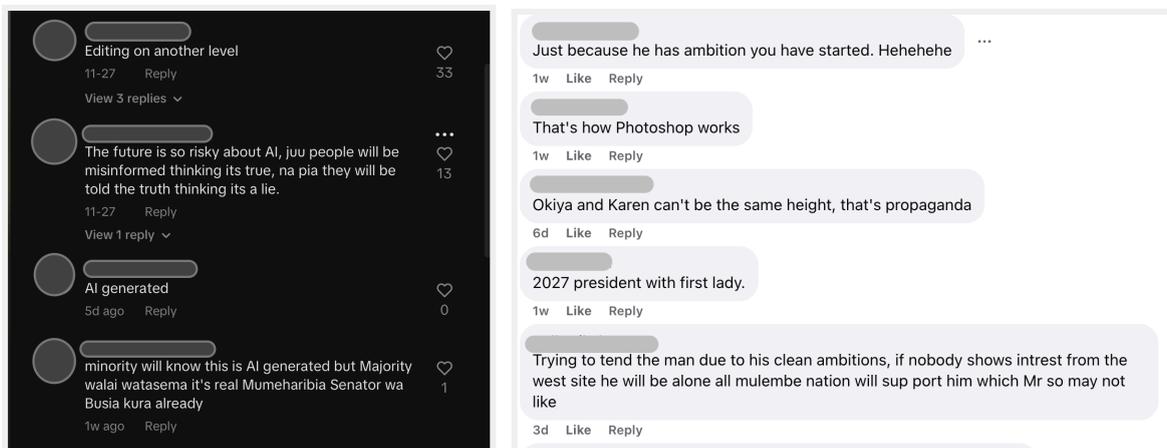


*Screenshots of concerns and comments on Omtatah and Nyamu's AI video (left, right)*
*(Source: CfA using TikTok and Facebook)*

On 08 November 2024, a video mimicked a news broadcast featuring Donald Trump, Barack Obama, Elon Musk, and Andrew Tate, labelling Ruto 'corrupt' and demanding his resignation. With scrolling headlines accusing the government of murders and theft, it amassed 476,937 views on X, with 2,405 mentions, and 592 TikTok views.

The post has two exact matches shared on X. The first available post online was shared on 08 November 2024 on X, while the last post was shared on 28 November 2024 on X. These posts have approximately 677,000 views and 9,804 interactions in total. An analysis of the comments shows that most users recognised the clip as an AI-generated joke, while a smaller group criticised the misinformation or used the post to amplify anti-government sentiment through hashtags like #RutoMustGo.

*Cite this case study as* *'Generative models in the spotlight: From civic innovation to the frontlines of disinformation warfare', Code for Africa (CfA), 21 November 2025, www.disinfo.africa*

6

*Screenshots of the fake news broadcast ([top](#), [bottom](#))*
*(Source: CfA using X and TikTok)*
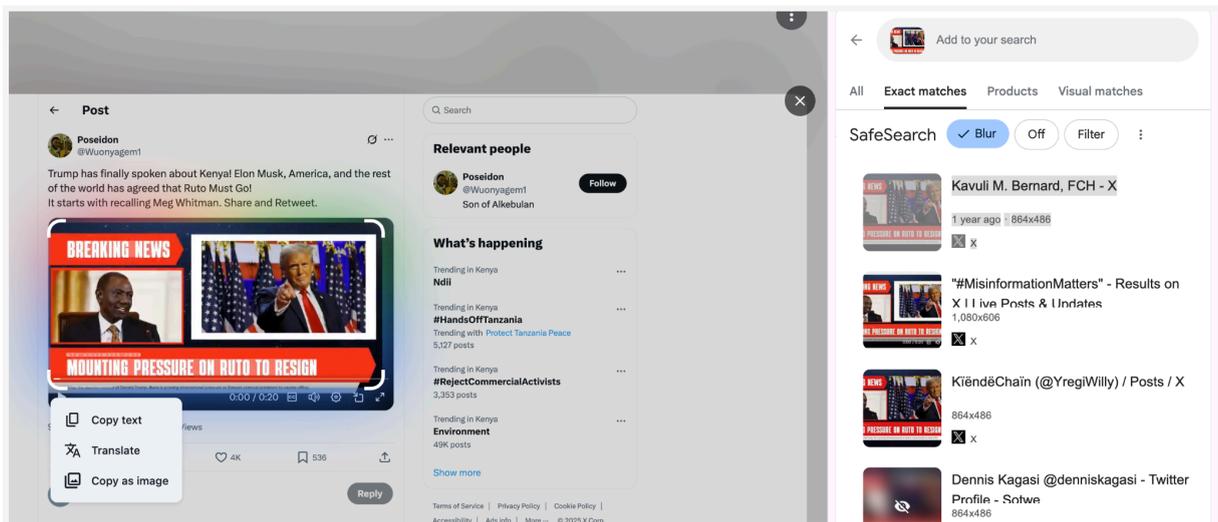


*Screenshot of a reverse image search on the AI-generated 'news broadcast'*
*(Source: CfA using Google reverse image search)*

### b. Pro-government campaigns and delegitimisation

State-aligned actors have systematically countered anti-government narratives by deploying coordinated generative AI campaigns to reframe opponents as failures, traitors, or threats to national stability. Many were previously flagged by Code for Africa (CfA) as sock-puppet or inauthentic accounts.

The #TighterWithWilliam surge, which happened between 20 to 22 February 2025, celebrated the Orange Democratic Movement (ODM) – United Democratic Alliance (UDA) 'broad-based' alliance. The hashtag generated 5,500 X posts in 72 hours, with 99.7% of them on 21 February 2025 alone. The hashtag received 135,000 views, 1,850 engagements. The ethnic phrase appeared in 5,322 posts, garnering about 123,000 views. AI-generated cartoons amplified the words with spud-like figures in traditional attire, their mouths stuffed with potatoes; the crop that symbolises Kikuyu agricultural pride.



*A screenshot of the AI-generated photo portraying Kikuyus as jealous*
*(Source: CfA using X)*

On 18 July 2025, the Facebook page Denis Itumbi Commentary, run by Ruto's former digital strategist, released a six-minute AI video alleging that the 25 June 2025 protests were not organic youth outrage but a foreign-funded coup attempt orchestrated by former deputy president Rigathi Gachagua. The script named high-profile activists - Boniface Mwangi, Wanjira Wanjiru, Okiya Omtatah, Hussein Khalid, Willis Otieno, Jimmy Wanjigi - and accused them of being

'foreign-funded'. The video's narrator, a synthetic black male wearing a Kenyan-flag bandana, delivered lines with visibly mismatched lip-sync over a collage of clips, purportedly showing scenes from the protests. Reverse-image searches of these 'evidence' clips returned no matches to any 25 June 2025 footage. The post garnered 1,400 views on Facebook.
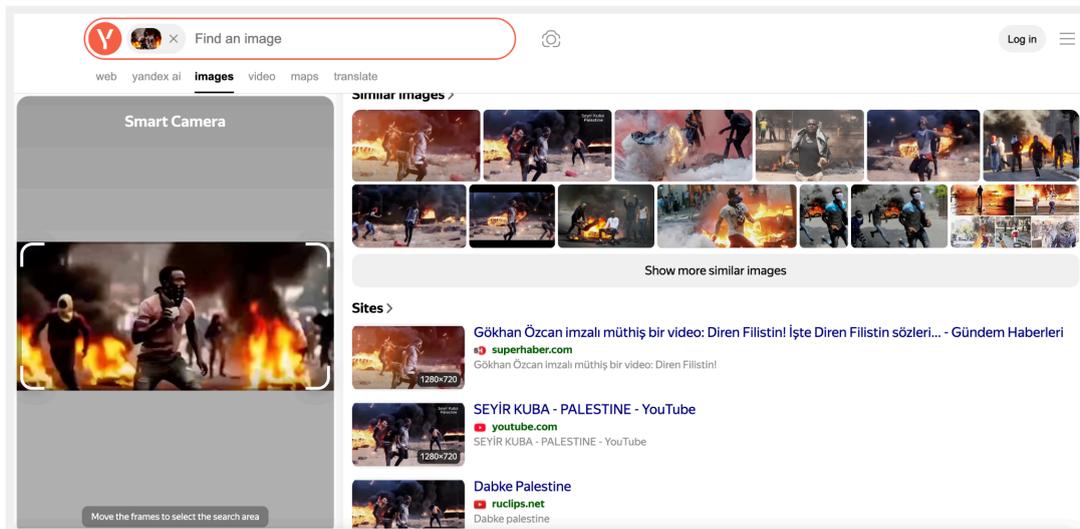


*Screenshot of a reverse image search of the clips allegedly depicting the 25 June 2025 protest.*
*(Source: CfA using Yandex)*

When Kileleshwa member of the county assembly, Robert Alai, reposted it to X two days later, it jumped to 24,900 views and 541 engagements.

One week later, the #RejectAnarchists campaign launched on 24 July 2025. It deployed two longer-form AI videos that stitched protest b-roll with ominous voice-overs warning of 'impending anarchy' and 'foreign-backed chaos'.

The same network simultaneously rolled out #DorcasBoyfriend, an AI deepfake showing pastor Dorcas Rigathi, wife of Gachagua, in a romantic embrace with an unidentified man. The video had irregular hand movements and placement, and it also contained a Hailuo AI video generator watermark at the bottom, proving that it was AI-generated. We identified at least nine X posts containing this video, all published on 24 July 2025. The posts had a combined 8,661 views but received little to no interactions with the three (1,2,3) that had replies from the post authors. We were unable to identify exact matches for the video using reverse image searches.
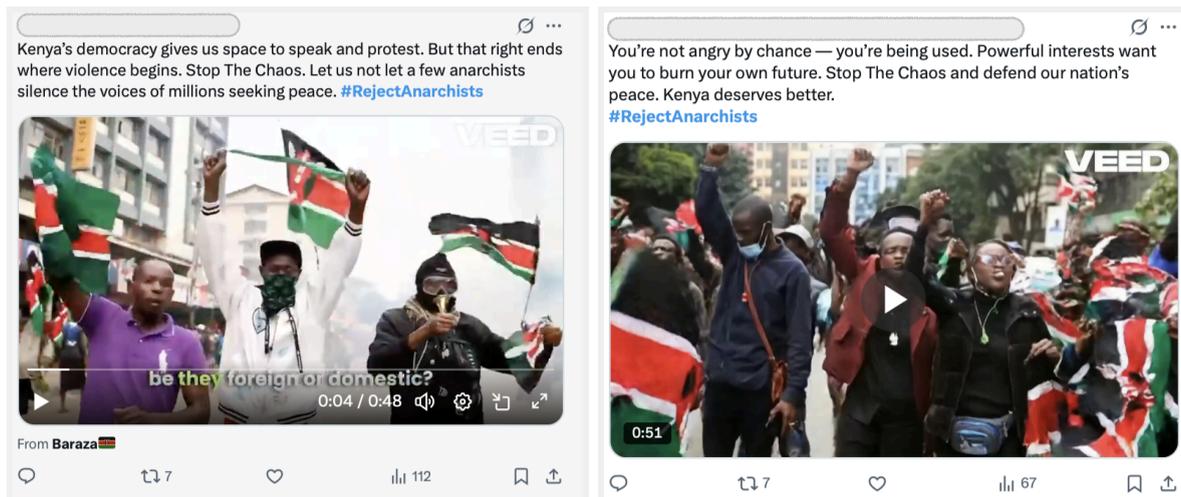
*Screenshots of videos used during #RejectAnarchists ([left](#), [right](#))*
*(Source: CfA using X)*

The twin campaigns, #TighterWithWilliam and #RejectAnarchists, achieved narrative convergence, framing youth protesters as anarchists and the Gachagua household as morally compromised. The most sophisticated operation unfolded on 21 August 2025, the day Gachagua returned early from a scheduled two-month US diaspora tour. Within minutes of his plane touching down, the hashtag [#ExpresswayClosed](#) was combined to push the claims that the Nairobi Expressway had been shut for hours to accommodate a chaotic, Mungiki-orchestrated welcome. [Mungiki](#) is a banned political-religious movement in Kenya, founded in the 1980s by some militant members of the Kikuyu community. Posts were flooded with phrases 'Mungiki at JKIA' and 'Makende Moja' (slang for 'one testicle', implying emasculation). Three accounts: [@tonymkenya](#), [@muneneodg](#), and [@acoco_ke](#), accounted for 616 posts and reposts in under four hours.

The payload included:
- A stitched 'riot' [video](#) combining unrelated Central Nairobi protest clips with traditional Kikuyu songs was used to simulate Mungiki violence at the airport. A [reverse image search](#) of a screenshot from the video showed that the protest was in Kenyatta Avenue, not at the Jomo Kenyatta International Airport (JKIA) as the video claimed. The location was also verified using [Google Street](#) View.
- An AI-edited Citizen TV [news segment](#) featuring anchor Jeff Koinange announcing that Gachagua had held a 'discreet meeting' with former US ambassador Michael Ranneberger, who allegedly refused to back his 2027 ambitions. The post, which was published on the morning of 21 August 2025, alluded to the previous night's Citizen TV 9 pm bulletin. One [reply](#) correctly identified that the 20 August 2025 9 pm [news](#) was not presented by Koinanga but rather by his colleague, Olive Barrows.

The hashtag generated 904,000 views, 7,570 mentions and 8,740 engagements in 24 hours. MOJA Expressway, the private operator, issued a real-time statement confirming the road remained fully open, with no closures or disruptions. The official statement debunking the claims barely registered against the deluge. Network analysis revealed that 17 of the top amplifiers in #ExpresswayClosed had also driven the hashtag #RejectAnarchists the previous month, forming a persistent core of coordinated inauthentic behaviour.

### c. Homophobic smears and image-based abuse

People are increasingly using AI to exploit Kenya's cultural sensitivities around sexuality and gender, making it easy to damage someone's reputation. These attacks aren't meant to spark real policy discussions.

The pattern began with Ruto. On 16 July 2025, after Kenya voted to renew the United Nations' LGBTQ+ rights mandate, a deepfake image surfaced showing Ruto in a full rainbow sequin outfit, wearing pink-orange lipstick and earrings, appearing to be at a Pride parade. The image, generated with Grok Aurora's photorealistic engine, was shared with captions like: 'Our President has finally come out' and 'Rainbow William leading the parade'. Though widely dismissed as satire on urban X timelines, the visual was engineered for rural WhatsApp forwards, where media literacy is lower and religious conservatism runs high. The timing was surgical: it reframed a diplomatic vote at the UN as a personal betrayal of 'African Christian values'; a narrative that resonates strongly in Central and Rift Valley constituencies. Between 08 and 16 July 2025, this image was posted at least 15 times on Facebook, garnering about 1,231 interactions. On X, it attracted about 415,000 views and 12,900 interactions.



*Screenshot of a deepfake depicting Ruto as a supporter of LGBTQ+ rights*
*(Source: CfA using X)*

On 02 January 2025, a 12-second AI clip depicted Ruto's long-time personal aide, Farouk Kibet, leaning in to peck the president on the forehead during what was framed as a private State House moment. The video, posted to a low-follower Instagram account, garnered 90 views and no comments.
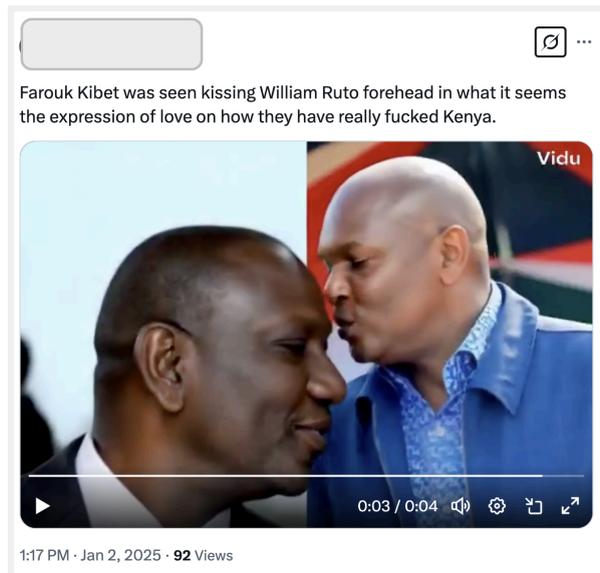
*A screenshot of the AI-edited video of Kibet pecking Ruto*
*(Source: CfA using X and Instagram)*

Three weeks later, on 29 January 2025, a more polished deepfake showed Ruto in a full lip-lock with the majority leader of parliament, Kimani Ichung'wah, complete with synchronised head tilts and background cheers. Shared by accounts on X with the caption: 'Gen Z upgraded their fight', it exploded to 7,000 views and 120 interactions within 48 hours. On Facebook, the video garnered 24,000 views and 305 interactions. Ichung'wah issued a rare public rebuttal, calling it a malicious attempt to undermine the strong working relationship between the president and the Legislature. One viral reply captured the intended ripple effect: If this reaches the villages, voter regret will hit hard.' The comment was liked 1,800 times.



*Screenshot of the AI-edited video of Ruto and Ichung'wah supposedly kissing*
*(Source: CfA using X)*

Activist Morara Kebaso faced similar treatment in late 2024, as his rapid rise as a Gen Z anti-corruption voice threatened establishment narratives. Between 04 and 09 September 2024, 320 X posts flooded timelines with hashtags like #MoraraSponsor, #WhereIsMyMoraa, and #VampireDiaries. They received 47,600 views, 520 interactions.

The core payload:
- An AI-edited [photo](#), from an account that has now been suspended, showing Kebaso holding hands with prominent lawyer Donald Kipkorir. The [image](#) was originally of Kipkorir and Kenya's ambassador to South Korea, Mwende Mwinzi, holding hands during the celebration of her 50th birthday.
- A [forged](#) letter on a GALCK+ (Gay and Lesbian Coalition of Kenya) letterhead, stamped and signed, 'confirming' Kebaso's paid membership since 2022.
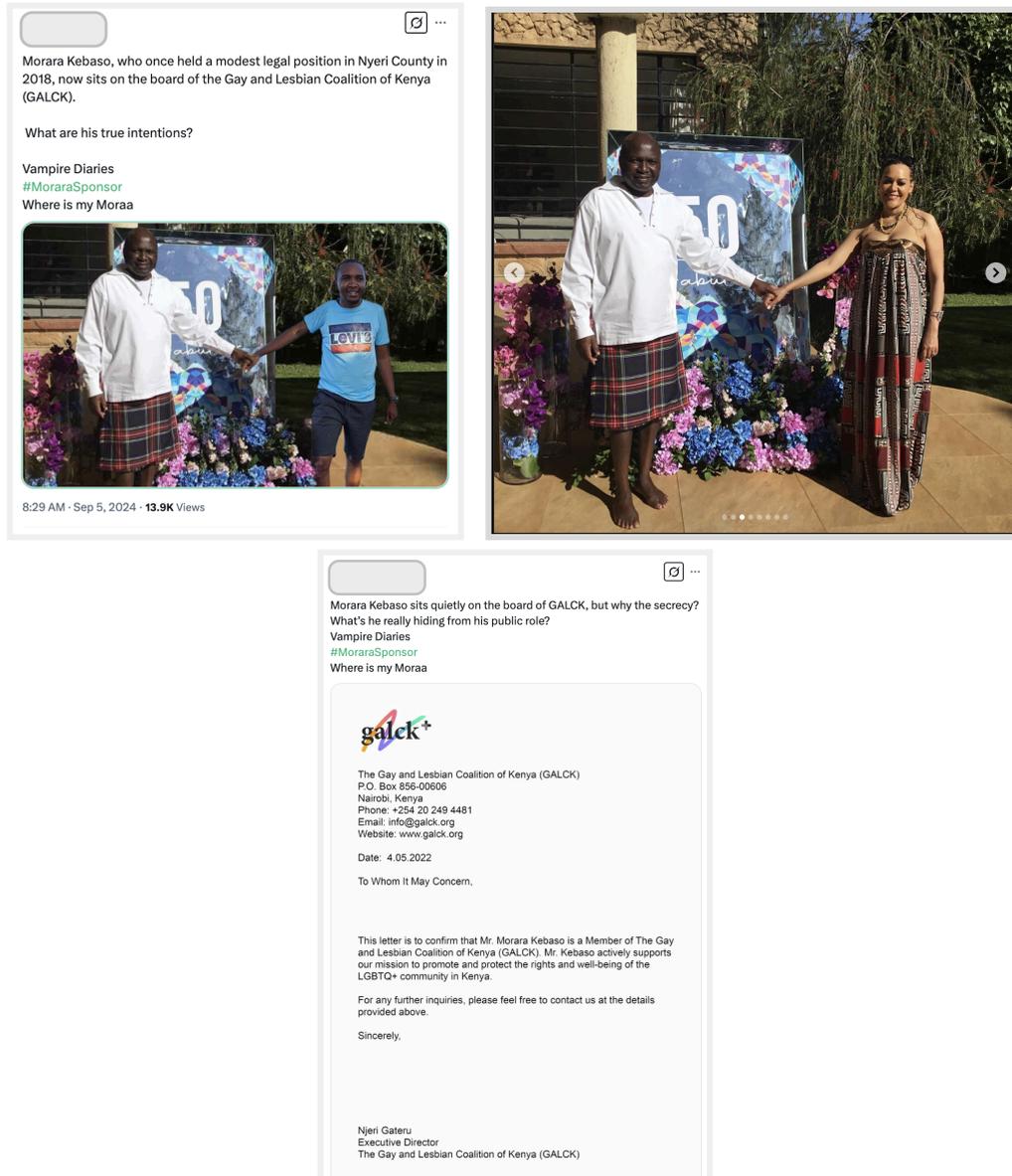


*Screenshots of the AI-edited photo of Kebaso and Kipkorir, the original image and the fabricated letter of GALCK+ ([left](#), [right](#), [bottom](#)) (Source: CfA using X and Instagram)*

The campaign questioned not just his sexuality but his funding, his 'flashy car,' which the campaign alleged was bought by 'sponsors', and his ties to Gachagua. PesaCheck [debunked](#) the letter within 36 hours, GALCK+ confirmed no such member existed and the stationery font was incorrect.

Women in public life were attacked the most using AI-generated content. For instance in early 2025 there was an AI-generated image of the first lady Rachel Ruto in black lace lingerie and red stilettos, leaning against a neon-lit wall with the caption: 'From Sabina Joy to State House: The real hustle.' Sabina Joy, a notorious Nairobi brothel, became the hook for a narrative that reframed her evangelical persona as hypocritical. The image was shared in Kikuyu-language Facebook groups where it fuelled sermons about 'immoral leadership'.

Journalist and activist Hanifa Farsafi Adan faced a sustained, multi-phase assault starting in April 2025.

A single X thread by an anonymous account prompted Grok to:
● Remove her hijab and replace it with a red bikini top;
● Add a durag and oversized sunglasses;
● Expose her hair entirely which is a violation of her Islamic practice; and
● Digitally undress her below the waist.

The thread went viral with 18,000 views before deletion. Adan responded publicly: 'This use of AI to womanise incompetence by dressing corrupt male politicians in bras and panties is a disgusting behavior. This isn't how we win this war. This isn't how we fight for good governance.'
Her statement was shared 2,300 times.

Following multiple reports, Grok halted all prompts involving digital undressing and the removal of religious symbols on 07 May 2025, issuing an on-platform notice. The fix was reactive, not preventative, and did nothing to scrub the existing copies. The pattern is the same every time. A usder types a short prompt into an easy-to-use AI tool like Grok, Midjourney, or a free deepfake app and within a minute they have a polished fake. They post it from a disposable account, bot networks pick it up and it spreads into closed groups where fact-checks rarely reach. The damage is completely one-sided and the target may spend days or weeks trying to protect their name, while the person who created the fake has already moved on. And in a country where 70% of voters say 'moral character" shapes their vote, these AI-driven smears aren't just online noise — they're highly targeted attacks with real consequences ahead of the 2027 election.

### d. AI: the new frontier of foreign influence

Tanzania has also been a site for AI-generated content. For example, on 26 June 2025, an AI-generated video appeared on TikTok claiming that the 'baby of Samia and Ruto was handsome', referring to Tanzanian president Samia Suluhu Hassan. The video, which showed a baby being handed to Hassan, gathered 78,800 views and 174 interactions. The account behind this post, @samiahruto, focuses exclusively on defamatory AI-generated content targeting Hassa. The videos have the PixVerse.ai watermark. The account has 31,200 followers and follows 703 other accounts, with its earliest videos dating back to 28 May 2025. One of the videos, posted on 29 May 2025, falsely alleged that Hassan and the president of Namibia, Netumbo Nandi-Ndaitwah, were in a lesbian relationship; it received 9,523 views and 79 interactions.
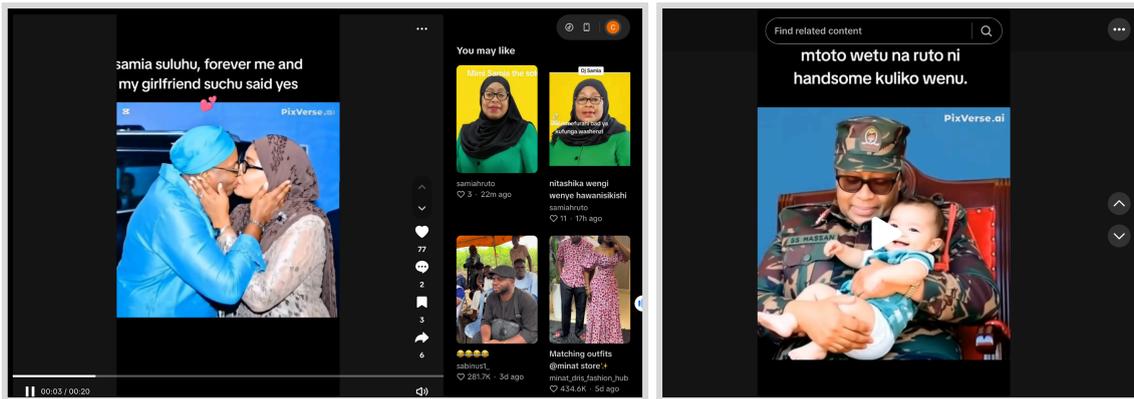
*Sample screenshots of some of the AI videos made by [@samiahruto](#) about Hassan ([left](#), [right](#)).*
*(Source: CfA using TikTok)*

On 28 October 2025, another AI-generated [video](#) surfaced on TikTok, showing Hassan running away after Tanzanian citizens turned their backs on her. The video was shared by[@shot_clipsvideos](#) and accumulated 2.4 million views and 33,654 interactions. The account [@shot_clipsvideos](#) focuses on AI-generated content mainly regarding the three East African presidents: Yoweri Kaguta Museveni of Uganda, Ruto of Kenya, and Hassan of Tanzania. The latest video, for example, portrays Museveni and some of his security with a caption that reads: 'Museveni hijacking the Indian ocean'. This video has 1,456 views and 33 interactions.



*Sample screenshots of some of the AI videos made by @shot_clipsvideos ([left](#), [right](#))*
*(Source: CfA using TikTok)*

Notably, a grammar, vocabulary and video sound review of the accounts @samiahruto and @shot_clipsvideos showed they are likely operated by Kenyan users rather than Tanzanians. The Swahili captions contain forms typical of a Kenyan dialect. For instance, a [post](#) from 09 September 2025 uses the phrase 'monoma' or 'munoma' instead of the Tanzanian 'mimi ni noma' ('I'm tough or unbeatable'), a phrase popularised by Kenya's deputy president, Kithure Kindiki. Similarly, a [post](#) from 10 August 2025 reads 'wakenya hunitamani sana juu ya my beautiful legs' ('Kenyans admire me a lot because of my beautiful legs'), where a Tanzanian speaker would more naturally say 'kwa sababu ya' rather than 'juu ya'. The account, @samiahruto, also primarily engages with content by other Kenyans.

## 4.     Conclusion

Kenya's encounter with generative AI in 2025 shows that the same technology that empowers citizens has been repurposed to create division. From coffin imagery that is used to criticise leaders to deepfake kisses that exploit homophobia and from potato-stuffed caricatures of Kikuyu voters to digitally undressing a Muslim activist, the malicious playbook is used to destroy trust and democratic legitimacy. In the run-up to 2027, the question is no longer whether generative AI will help shape Kenya's public space; it already has.

***Do you know of incidents of suspected information manipulation, coordinated influence operations or other forms of disinformation or weaponised toxic content that you think we should investigate? Share your leads with us, securely, on** [ilab@codeforafrica.org](mailto:ilab@codeforafrica.org).*