



How Twitter's Negligence is Harming Kenya's Democracy

By Odanga Madung



On the 24th of June 2021 at around 6 a.m., an insidious hashtag, #KatibaMbichi, appeared on Kenyan Twitter timelines. Its trend seemed to be driven by a number of faceless bots, and retweeted by a series of catfishes that sent it to the number one spot on the Kenyan Twitter trends.

Our investigations have uncovered how such malicious, coordinated, inauthentic attacks that seek to silence members of civil society, muddy their reputations and stifle the reach of their messaging, is a growing problem in Kenya. Twitter, especially, has been central to these operations due to the influence it has on the country's news cycle.

The proliferation of digital media platforms in Kenya carries the promise of a renewed definition of freedom of speech. Moreover, Twitter has been a vital tool of expression for many Kenyan citizens, many of whom use it to hold their leaders to account and to call out their failures. But civil society members and journalists have increasingly come under attack thanks to disinformation campaigns in the country.

Through a series of interviews with anonymous influencers involved in these campaigns, we accessed their inner workings and gained crucial insights into how they are organized.

An examination of the campaigns has provided our team with a window into the shadowy world of

Twitter influencers for political hire in Kenya. Many of the accounts and individuals involved promote brands, causes and political ideologies without disclosing that they are part of paid campaigns.

Twitter features such as the trending algorithm are exploited to achieve the goals of these campaigns by amplifying them. Certain verified accounts on the platform are complicit in leading these attacks. The goal of these campaigns is to exhaust critical thinking and poison the information environment by muddying the truth.

Our investigations examined two months' data between 1 May 2021 and 30 June 2021, with a particular focus on the Constitutional Amendment Bill—famously known as the Building Bridges Initiative—that was being promoted in Kenya at the time.

With the aid of Twint, Sprinklr and Trendinalia, we trailed the attacks by mapping and analysing specific hashtags that the influencers used on Twitter. This involved mapping certain accounts that posted malicious content targeting Kenya's activists and judicial officers. The flagged hashtags often displayed synchronized publishing timestamps within the metadata, with a lack of content on most days, followed by one very sharp burst of activity and then fizzling out.

In total, using Sprinklr, which has access to Twitter's full historical archive, we flagged 23,606 tweets and retweets released by 3,742 accounts under the 11 hashtags. We also obtained 15,350 of these tweets using the Twint package on Github to carry out further analysis of the content.

How disinformation is spread

The Twitter campaigns we looked at were those that were pro-BBI and directly attacked citizens and prominent civil society activists that were vocally opposed to the proposed reforms, and also sought to discredit civil society organisations and activists by portraying them as villains who were being funded by Kenya's Deputy President William Ruto—he opposes the BBI process.

The well-coordinated attacks are launched through WhatsApp groups to avoid detection. The WhatsApp group admins give direction about what to post, the hashtags to use, which tweets to engage with, and whom to target. They also synchronize the posting to enable the tweets to trend on twitter.

There is money to be made in attacking civil society. Our sources confirmed that they get paid between US\$10 and US\$15 to participate in three campaigns per day. Those higher up the ranks are on a monthly retainer that can go as high as US\$500. Those who are on a retainer supervise the hashtags and ensure that they trend on the days they are posted.

Who the disinformation targets

From our analysis, the top three most frequent victims were Kenyan journalists, judges, and known activists. Prominent anti-BBI activists under the *Linda Katiba* movement who petitioned the courts against the BBI were the targets of some of the most vicious attacks.

The attacks peaked in early May with the specific goal of trying to discredit the anti-BBI campaign. Jerotich Seii, a key member of the *Linda Katiba* campaign who was targeted, said in interview that she had to spend a lot of time trying to prove that her activism efforts were genuine and that she was not a front for someone else. "The disinformation attacks against me focussed on painting me as someone with ulterior motives who isn't interested in the welfare of Kenyans. I had to spend a good chunk of my time defending my position as someone who is actually a patriot who does what they do out of love for their country," said Seii.

From our analysis, the top three most frequent victims were Kenyan journalists, judges, and known activists.

All this is leading to self-censorship by some of the activists on the platform as they feel that it is pointless to use a platform that cannot deliver any meaningful engagement. One activist we spoke to said that she had significantly scaled down her Twitter activity because of all the trolling she had experienced.

The Kenyan High court struck down the BBI on 14 May on the grounds that the initiative was unconstitutional and the Court of Appeal followed suit on August 20th. The ruling not only strained the already bad relationship between Kenya's Judiciary and the Executive, it also led to wave after wave of disinformation attacks seeking to question the judges' judicial independence and the accuracy of their decision.

A notable change in these attacks was how the visual aesthetics of the content within the campaigns evolved; newspaper editorial cartoon-style caricatures and memes were employed, a likely indication of a change of leadership or strategy at the top that sought to make the content more palatable and shareable.

What is the impact of the slander?

The data that we gathered from Trendinalia (which collects data on Twitter trends in Kenya) shows that sufficient amplification was achieved for 8 of the 11 hashtags we identified that became trending topics. This amplification was achieved partly through the use of verified accounts. One anonymous influencer we spoke to said that owners of certain verified accounts involved in these campaigns would often rent them out to improve the campaign's chances of trending. "The owner of the account usually receives a cut of the campaign loot from the person that rented it from them once it's over," the influencer said.

The demand for this service by the political class in Kenya is markedly strong. During the months of May and June alone, we counted at least 31 artificial political hashtags, including the ones linked to the BBI process. This translates to at least one manipulated disinformation campaign that Kenyans have to deal with every two days.

Curiously, there is little evidence that these operations actually sway people's opinions. However, they do have an effect on how Twitter users interact with their information environment. The goal of such operations is to overwhelm, to create an environment where nobody knows what is true or false anymore. The objective is to exhaust critical thinking and muddy the truth.

During the months of May and June alone, we counted at least 31 artificial political hashtags, including the ones linked to the BBI process.

Typically, a post by any of the prominent activists or judicial officers is bombarded with so much aggression, insults, and dismissive comments that the space for a good conversation is lost. The point is always to ensure that sober-minded people are disincentivized from amplifying the topic after encountering so much aggression in the replies and the quote tweets.

The role of Twitter Inc.

To many Kenyans, Twitter matters. The platform has become a very critical avenue of expression,

networking, running ads, and a means of obtaining information. It is also an important avenue for active citizenship as #KOT (Kenyans on Twitter) is one of Africa's loudest and most lively internet communities.

On the darker side however, some of the features on Twitter are being exploited for nefarious purposes. The platform is failing Kenyans—and Africans more broadly. Political actors are using it to try to control political narratives by poisoning the platform and harassing dissenting voices.

Specifically, Twitter's trending algorithm, which selects and highlights content without examining its potential for harm, often serves as an on-ramp for users who are trying to find information on the platform. Our sources said that Twitter trends is the primary key performance indicator by which most of their campaigns are judged. They admitted that without it their jobs would not exist. "The main goal is to go trending on Twitter. I'm not sure what our jobs would look like without that target," said one source.

The evidence available points to the fact that, for the executives at Twitter, this is not a new phenomenon. The trending algorithm in particular, which is a big part of how Twitter works, has been abandoned to disinformation campaigns and attacks.

Twitter's Moderation Team should pay close attention, keenly monitor and regulate its trending section. Activists, such as Sleeping Giants, have repeatedly called for Twitter to "untrend" itself. This could be done by either removing the feature completely or by disabling it during critical times such as during election periods.

The evidence available points to the fact that, for the executives at Twitter, this is not a new phenomenon.

Arguably, Twitter does not have an incentive to fix this. It sells ads for "promoted trends" and "promoted tweets" within the feeds of hashtags on its trending topics section to business clients. This puts Twitter squarely in the middle of the mess as it profits from this harmful activity.

Ad Dynamo, an agency that sells Twitter Ads in Kenya, currently offers promoted trends for US\$3,500 per day within the country. The overall message this sends is that it is ok to sow hate on the platform so long as Ad Dynamo owners can place ads next to the trending content and make a profit from it.

As Kenya heads towards elections in 2022, the demand for these services will increase and many political parties will seek out malicious coordinated trending models and create the risk of a repeat of the 2007 political violence.

Published by the good folks at [The Elephant](#).

The Elephant is a platform for engaging citizens to reflect, re-member and re-envision their society by interrogating the past, the present, to fashion a future.

Follow us on [Twitter](#).