

**moz://a**



# **From Dance App to Political Mercenary**

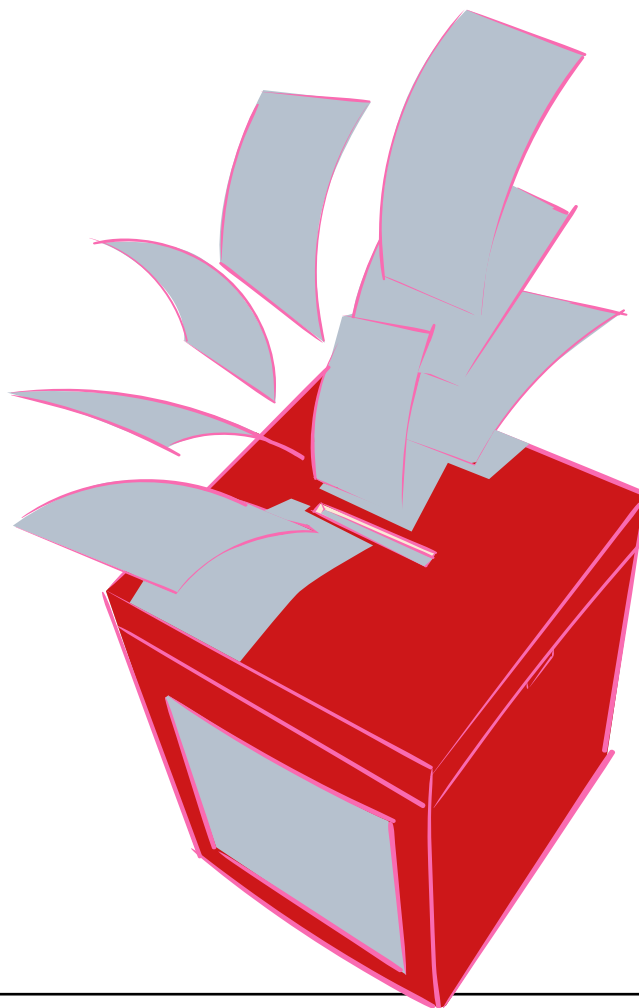
*How disinformation on TikTok gaslights  
political tensions in Kenya*

# Contents

Abstract	3
Executive Summary	4
1. Introduction	5
2. Extensive Disinformation and TikTok Policy Violations	6
2.1 Data sample and size	6
2.2 Incitement against communities	6
2.3 Use of synthetic and manipulated content	8
2.4 Context bias	9
3. Looking Back: Familiar Terrain for Kenyans	10
4. Recommendations/ Areas of Concern	12
5. Conclusion: TikTok is Failing its First Real Test in Africa	14

## Abstract

Kenya's 2022 general election is slated for August 9, and in the run-up to voting, social media platforms in the country are teeming with political disinformation. While more mature platforms like Facebook and Twitter receive the most scrutiny in this regard, TikTok has largely gone under-scrutinized — despite hosting some of the most dramatic disinformation campaigns.



---

Indeed, research by Mozilla Fellow Odanga Madung reveals that the disinformation being spread on TikTok violates the platform's very own policies. This disinformation is similar in tone and quality to the Cambridge Analytica and Harris Media content that spread on Kenyan Facebook in 2017.

## Executive Summary

Kenya's general election is scheduled for August 9, 2022. The date is an important moment for its democracy, but also one that carries with it a fraught history. In recent elections, Kenya's democratic process has sometimes led to violence.

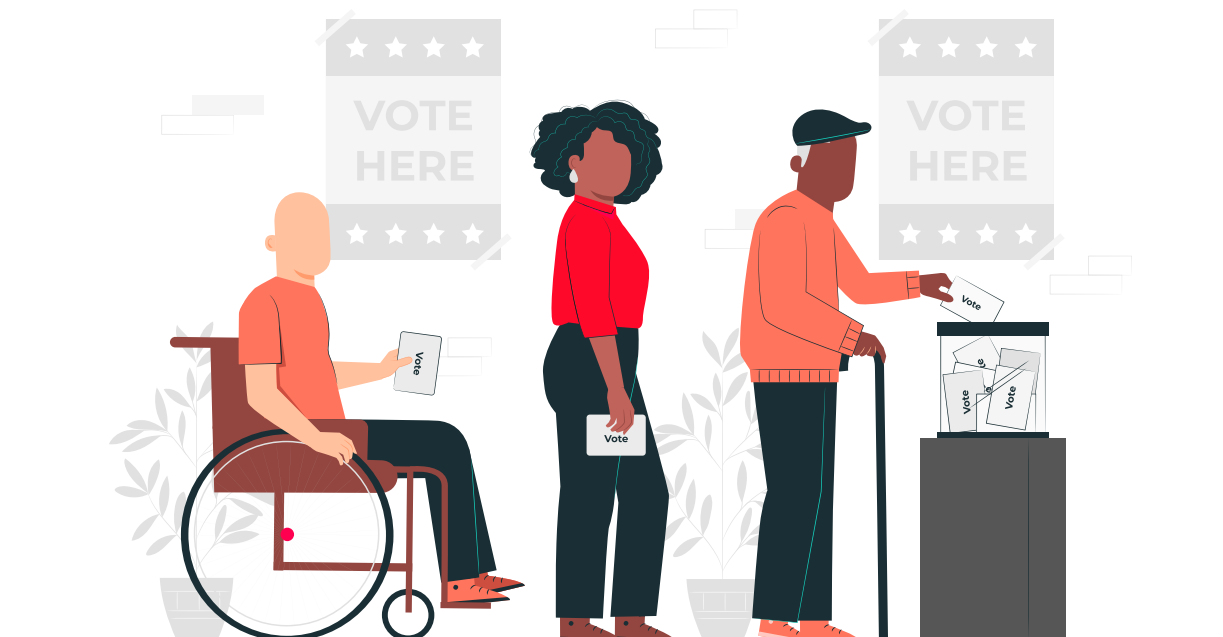
Now, as August approaches, millions of Kenyans are learning about and discussing the election online. Platforms like Facebook, Twitter, and increasingly TikTok are forums for political speech. However, our research shows that TikTok is acting as more than a forum for political speech — it is also a forum for fast and far-spreading political disinformation.

In this report, we examine a sample of problematic political content on the platform: over 130 videos from 33 accounts which have been viewed collectively over 4 million times. Our analysis reveals that hate speech, incitement against communities, and synthetic and manipulated content — despite being in violation of TikTok's very own policies — is both present and spreading on the platform.

Our research also entails interviews with past content moderators at TikTok, which help illuminate why this content may be present on the platform, despite its violations. The interviews reveal a moderation ecosystem that lacks both the context and resources to adequately engage with election disinformation in Kenya.

***The upshot is that TikTok is failing its first real test in Africa. Rather than learn from the mistakes of more established platforms like Facebook and Twitter, TikTok is following in their footsteps, hosting and spreading political disinformation ahead of a delicate African election.***

The final portion of this report offers TikTok suggestions for improvement. While moderating political content will always be difficult and complex, the platform can still take steps forward — like partnerships with fact checkers and civil society, more clear policies and guidelines, and algorithmic transparency.



# 1. Introduction

In August, the world will turn its attention to Kenya's general election, one of the most significant political events in Africa in 2022. The country's recent history features hotly contested and sometimes violent elections. Competing [camps on either side of the political divide](#) have [been noted](#) to use reckless, incendiary rhetoric amid escalating tribal tensions among Kenyan citizens.

Kenya's election is TikTok's first real test in an African democratic process. In past elections, most attention has focused on platforms [like Facebook and Twitter](#) and their outsized role in Kenyan politics. Meanwhile, TikTok's thriving user base and influence has shaken up Kenya's social media landscape. It's the most downloaded in the country ([according to AppFigures](#)) and has even launched several [influencers into stardom](#).

As a relatively new platform, TikTok has largely escaped scrutiny. Despite the fact that the platform bans political ads, politics and political content still dominate user feeds. Past Mozilla research on the [2021 German election](#) and [U.S. TikTok influencers](#) has shown that TikTok struggles to enforce its policies around elections and political influence.

Indeed, it's not just a dancing and lip-syncing app: The Chinese-owned platform has emerged as one of the most popular social media apps for sharing political content. For example, the hashtags #siasa and #siasazakenya (which translate to "politics" and "Kenyan politics," respectively) have over 20 million views on the platform. The top videos on those hashtags have close to a million views, with many in the hundreds of thousands. In contrast, the same hashtag on Instagram has fewer than 100 posts and the most popular videos were viewed only hundreds of times. In spite of TikTok's claims that it is not a place where political conversations take place, the platform is rapidly becoming a hotbed for politics.

Our previous research has shown how platforms like Twitter have failed to recognize the intricacies of Kenya's democracy, leaving the country vulnerable to [influence peddling from outside groups](#), attempts to [consolidate power by the elite](#) and [neutralization of public outcry](#).

**Here, we argue that TikTok is no different.** This research provides an in-depth analysis of political disinformation on TikTok based on a sample of 132 videos. Our research suggests that Kenyan TikTok has become a breeding ground for propaganda, hate speech, and disinformation about Kenya's election. A highly sophisticated disinformation campaign is underway on the platform, which includes slickly produced video content and attack ads spewing false claims about candidates, while also threatening various ethnic communities. Many of the videos are getting outsized viewership in comparison to their followership — and [according to researchers](#), this suggests that the content may be gaining amplification from TikTok's For You Page algorithm.

## 2. Extensive Disinformation and TikTok Policy Violations

### 2.1 Data sample and size

Note: On June 7, 2022, after reviewing this research, TikTok removed several of the posts in question.

We reviewed over 130 videos from 33 accounts which have been viewed collectively over 4 million times on the platform.

*Our research sample:*

---

**>130**  *from* **33**  *accounts* *with* **>4,000,000**  **views**

---

We parsed this content into two broad categories of disinformation:

(1) content that could fall into the category of **hate speech and incitement against communities**; and (2) **synthetic and manipulated content**.

The data show that TikTok is enabling the rapid spread of disinformation and incendiary rhetoric about the Kenyan election. We obtained the content using TikTok's search function with a keyword list comprising phrases and names of political candidates, key locations, political parties, ethnic communities, and other terms related to the election.

TikTok's [terms of service](#), [community guidelines](#), and [policies](#) state that its platform is designed with safety in mind. And yet, many of the videos we gathered in the course of this study potentially violate, while others appear to be in clear violation of TikTok's policies — and were still allowed to flourish.

In this section, we outline how the categories of content mentioned above come into conflict with TikTok's rules. While we do not have data about which videos appeared on people's FYP feed, many of the videos we reviewed had a large number of views beyond the following of the accounts that published them, suggesting that there may have been some algorithmic amplification involved.

### 2.2 Incitement against communities

Across many social media platforms, we have observed efforts to manipulate engagement mechanisms to evoke the ghosts of Kenya's violent electoral past for political gain. This is especially true with regards to the use of images surrounding [the events that followed the 2007 election](#), when widespread dissatisfaction over the result erupted in extreme violence in several parts of the country and brought life to a standstill. Fear of post-electoral violence is something [that is still very real](#) to Kenyans. Our research

suggests that actors may be using TikTok to take advantage of this and polarize voters on the platform.

In their policies, [TikTok states](#) that they do not allow posts praising, promoting, or supporting any hateful ideology: “TikTok is a diverse and inclusive community with no tolerance for discrimination. We do not permit content that contains hate speech or involves hateful behavior, and we remove it from our platform.” Additionally, the platform prohibits users from posting content that incites hate, prejudice, or fear.

However, we found content on the platform which, in the context of Kenya’s electoral history, is problematic and could fall into the category of incitement and hate speech along ethnic lines. Many of the videos we reviewed contained **explicit threats of ethnic violence** specifically targeting members of ethnic communities that are based within the Rift Valley region. Similar narratives stoked the post-election violence of 2007/2008, where over 1000 Kenyans died and thousands more were displaced.

In one instance, a video clip showing William Ruto (the current Deputy President and presidential candidate) giving a speech at a rally had the caption “Ruto hates Kikuyus and wants to take revenge come 2022.” The video was widely distributed on TikTok and it got over 445,000 views on its platform.

Another video took on the form of a detergent infomercial and had the narrator say that “UDA can be used to remove madoadoo such as Kikuyus, Luhyas, Luos, and even Kambas (all these are tribes in Kenya).” Alongside the mentions of the communities were graphic images obtained from previous post electoral skirmishes in Kenya.

### Hate speech/ incitement against communities



Looking at TikTok’s policies, it is not clear whether the rhetoric featured in some of this content fits the platform’s definition of hateful ideology or incitement of hate, but it is problematic either way when we consider Kenya’s context. The content targets specific communities with threats and uses past violence as a tool of fear. All this means that TikTok’s current approach to content moderation in Kenya may not take into account the full cultural context to review and police such kinds of content that are moving through their platform. As a result the videos we identified that fell into this category garnered a total of close to 1.2 million views on the platform.



[TikTok also explicitly states](#) that videos depicting things that may be shocking to a general audience may not be eligible for recommendation. However, of the videos we reviewed, the content with the most gruesome images often got a higher number of views in comparison to those that didn't. For example, we [identified a video](#) containing a manipulated image of one of the political candidates as its thumbnail. It showed him in a shirt covered in blood holding a knife to his own neck with a caption alleging that he is a murderer. This video garnered over 505,000 views on the platform.

### 2.3 Use of synthetic and manipulated content

During an election season, synthetic content can spread with disastrous effects. It often erodes the public's sense of trust in the news – or even the very idea of truth – upon which an informed electorate depends. Since 2017, the [appropriation of the identities](#) of credible media outlets and manipulation of their media content to misinform citizens has been a growing trend in Kenya. Social media platforms have often been used to distribute this fake content.

TikTok's [policies](#) “prohibit synthetic or manipulated content that misleads users by distorting the truth of events in a way that could cause harm.” However, our investigation shows that this type of false content is thriving on TikTok, and in some cases, it is receiving higher engagement as compared to other platforms.

One example we identified was a rip off of the Netflix documentary [“How to Become a Tyrant.”](#) It mashes up clips from the film along with Kenyan mainstream media news clips and is accompanied by slick narration. Whereas [multiple versions](#) of the [videos](#) distributed across other platforms received only hundreds of views, the TikTok versions we identified received over 8,000 views.

This method has been highly effective at allowing misleading information to be distributed to Kenyans on the platform. We identified several manipulated pieces of content on the platform that were widely viewed: a fake Kenya Television Network (KTN) news bulletin with a fake opinion poll and dubbed narration; a video showing a fake Joe Biden tweet; and various false newspaper covers. These videos garnered over 342,000 views on TikTok.

#### Synthetic Content





## 2.4 Context bias

Why is such content allowed to fester and thrive on the platform? We spoke to former TikTok content moderators including [Gadear Ayed](#). She went public about her time while there. According to her, it was common practice for moderators to be put on workstreams with unfamiliar contexts and languages. Says Ayed: “Sometimes the people moderating the platform don’t know who the entities in the videos are and therefore the videos can be left to spread due to lack of knowledge of context. It’s common to find moderators being asked to moderate videos that were in languages and contexts that were different from what they understood. For example I at one time had to moderate videos that were in Hebrew despite me not knowing the language or the context. All I could rely on was the visual image of what I could see but anything written I couldn’t moderate.”

This could explain why content deemed as problematic in Kenyan contexts is being allowed to thrive on TikTok.

Gadear also explained that the need for speed in the content moderation process likely gets in the way of moderators evaluating a video properly, something that was also echoed by other moderators we spoke to. Says Ayed: “When I worked at TikTok I was reviewing 1,000 videos a day. There was no time limit for the videos that we would moderate. Instead, what we had were targets of videos to moderate per day. So you wouldn’t want to watch a video too much because that will get in the way of you achieving your target. Sometimes we would watch a video at two- to three-times the speed to get around this problem.”

This need for quick processing therefore meant that fact checking of viral content could be trivial, and identification of manipulated content would also be difficult. Says Ayed:

---

***“We didn’t have any way to identify whether a video was real or fake. The moderation process is very fast and TikTok didn’t want us spending too much time checking if the content is real or not. If the content is false or fake it can still spread across the platform unless it breaches another aspect of the policies we were told to watch out for.”***

---

### 3. Looking Back: Familiar Terrain for Kenyans

On TikTok, it appears that bad actors are using well-trodden tactics to spread specific political narratives. Political rhetoric by Kenyan politicians during election periods tends to veer toward character assassination. Opposing sides will [focus on creating the specter of “the opponent”](#) as an existential threat to stability and security. This is the case with platformized propaganda as well, and Kenyans have walked down this path in the past.

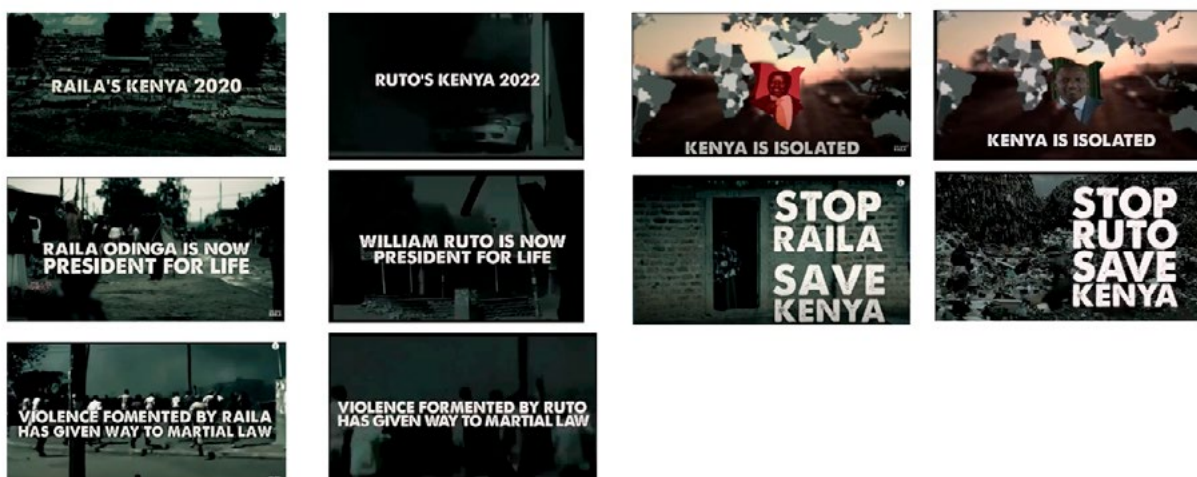
In 2017, it is alleged that the Jubilee party ran [the “Real Raila” campaign](#) in the lead-up to that year’s election and sought to manipulate Google Search, YouTube, and Facebook. The campaign was executed by [Harris Media LLC](#), a Texas-based media firm that was also used by former U.S. President Donald Trump in the 2016 presidential election, as well as by several other far-right European parties.

Through several dark videos and websites, Harris Media [attempted to paint](#) Raila Odinga as a monster who would destroy Kenya if he became president. Many of those videos received millions of views on YouTube and Facebook because [the platforms were paid to distribute](#) them. Some of that content is [still online to this day](#).

As a result, Harris Media injected a new campaign tactic into Kenya’s electoral landscape — one that its political scene has had trouble shaking off ever since. We witnessed similar smear campaign tactics on display when [we investigated the disinformation apparatus](#) around the BBI campaigns in 2021.

In what could perhaps be an attempt at replicating the model of 2017’s Real Raila campaign, our investigation uncovered a video on TikTok that mimicked one of the most popular [pieces of content](#) from it. However, this time the names of the individuals in the video have been switched up. This version shows Kenya as a post-apocalyptic hellscape after William Ruto has become president. Prior versions that were run in 2017 had Raila’s name and depiction instead.

#### The Return of the Real Raila campaign



Members of civil society in Kenya see the entrenchment of such campaign practices in Kenya as a worrying trend. TikTok is highly popular among younger audiences in Kenya who are still forming their political identities and value bases. Irungu Houghton, the Executive Director at Amnesty international, said that “TikTok’s demographic is much younger and it worries me because they don’t have the levels of political maturity or a clear value base that may allow them to sift through such information.” He went on to point out that some of the effects of this kind of content on the platform will persist far beyond August’s election: “TikTok need to recognize that the demographic they are dealing with is a formative generation and therefore the impacts of such campaigns are not things that we’re likely to see immediately — but we may see its effects in decades to come.”

## 4. Recommendations/ Areas of Concern

Our findings parallel our prior research into how [Twitter](#) has been consistently manipulated in Kenya for the purposes of spreading disinformation. Both platforms have enabled bad actors to publish malicious content with an intent to incite audiences by using references to past electoral violence, and have gained wide distribution.

As it was with Twitter and now with TikTok, we see [context bias](#) rear its ugly head. It is clear that TikTok's current moderation guidelines are failing to help moderators identify misleading and harmful content. The presence of explicit and targeted ethnic references in a threatening manner makes us ask the question of whether the people working for the platform know about the political sensitivities of Kenya's election.

Such problems can only be solved by being transparent about their approach to moderation and doubling down on local partnerships with CSOs, in order to get a thorough sense of what the stakes are for Kenyans in this electoral process. It is only then that they will be able to identify the most common vectors for disinformation, and then develop active monitoring systems and early warning triggers around candidates, vulnerable geographies, minority groups and ethnicities.

Additionally, given the rampant use of historical references of past electoral skirmishes to threaten communities, an area of concern for us is that TikTok's policies on violence and incitement don't address this behavior or outline steps on how to curtail it. An example worth looking at is what [Facebook have](#) done in their violence and incitement guidelines.

It is also concerning to us that many problematic videos got widespread viewership on the platform way beyond the count of their followership. Algorithmic transparency from TikTok is necessary here in terms of enabling researchers to understand how problematic content is seeded and distributed on the platform, despite breaking its own community guidelines in some cases.

Otherwise, if the platform cannot get a handle of the dangerous content its algorithms are spreading, TikTok may need to start looking into ways of switching off the function altogether as a circuit breaker during sensitive times. For example, [as journalists have pointed out](#), the war in Ukraine served as a showcase of what can happen when TikTok fails to get a handle on its features. Both Twitter and Facebook have made consequential decisions about algorithmic elements on their platform at critical junctures; Twitter has at times switched off the trending section and Facebook has switched off group recommendations. There's no reason why TikTok shouldn't consider such interventions as well.

Finally, our research suggests that TikTok as a platform is vulnerable to the spread of falsehoods and synthetic content. As per our interviews with former moderators, there appears to be a prioritization of moderation speed, with little consideration for truth. This is similar to problems [moderators of other platforms](#) such as Facebook have pointed out.

TikTok says it has fact checking partnerships to cover Kenya and even more broadly, for the Sub-Saharan African region. But it is not clear how many moderators TikTok has dedicated towards their operations in Kenya. In our research, Mozilla did not encounter

any labeling policies for falsehoods or synthetic content that applies to Kenya at a critical time in its electoral process. We recommend that TikTok gets this apparatus in motion and perhaps goes even further, using the [template of their COVID response](#) to promote high quality information sources wherever they are showing news and other election related content.

## 5. Conclusion: TikTok is Failing its First Real Test in Africa

TikTok appears to be vulnerable to various forms of manipulation aimed at swaying Kenyans' opinions. Kenya has [a political culture of incendiary speech](#) that is spilling over into platforms, which have the potential to amplify it and make it worse. All this is happening in an environment where the Permanent Secretary for Interior decreed that politicians [will not be prosecuted](#) for hate speech until after the election — a move he said would keep the government from “wasting resources.”

TikTok needs to acknowledge their responsibility in fostering healthy debate. The type of content we identified is a threat to the integrity of Kenya's electoral process, and TikTok's shortcomings in terms of moderation of the platform only adds fuel to the fire.

Academics have previously pointed out that Kenyans' [reliance on rumor](#) as an information source and the [inflammatory language of politicians](#) were major reasons why the 2007 election turned violent. In light of this, the challenge of moderating political content on the platform will be difficult and complex. The upshot is that TikTok is failing its first real test in Africa. Rather than learn from the mistakes of more established platforms like Facebook and Twitter, TikTok is following in their footsteps, contributing to the pollution of an information environment ahead of a delicate African election.





---

*The Mozilla Foundation is a nonpartisan charitable organization that fights against misinformation and lack of transparency in online political and election-related messaging. It does not favor or oppose particular candidates or parties in elections, and any comments it makes about social media platforms' handling of particular election-related messages or speakers should not be taken as support for or opposition to those messages or speakers themselves. Voters should consider a wide variety of factors outside those addressed by the Foundation's work in deciding how to vote.*

---



**moz://a**