



moz://a

F
A
K
E
N
E
W
S

F
A
K
E
N
E
W
S

F
A
K
E
N
E
W
S

F
A
K
E
N
E
W
S

F
A
K
E
N
E
W
S



**EXPORTING DISINFORMATION:
HOW FOREIGN
GROUPS PEDDLE
INFLUENCE IN
KENYA THROUGH
TWITTER**

CONTENTS

Executive Summary	5
Abstract	4
Introduction	7
Who Are CitizenGO?	10
Tapping into Kenya's Disinformation Industry	11
Where does this leave Twitter?	17
Conclusion	20



Abstract

In recent years, people around the globe have watched social media platforms fall short of their idyllic promises and become potent tools for bad actors.

This is especially pronounced in the non-western world, where platforms lack the cultural context, the staff, and the will to confront issues of misinformation, disinformation and hate speech. This phenomenon was explored by the Mozilla Foundation in [September](#) and [November](#) of 2021.

This new research demonstrates the extent to which platforms like Twitter can be weaponized by bad actors from afar – sometimes continents away – to manipulate a country’s public discourse.



Executive Summary

Between 2020 and 2021, Kenyans discussed and debated two pieces of legislation online: The Reproductive Healthcare Bill 2019 and the Assisted Reproductive Technology Bill 2019 (aka the Surrogacy Bill). The bills aimed to outlaw forced sterilizations; make prenatal, delivery and postnatal services free to every woman in the country; and develop standards, regulations and guidelines on assisted reproduction. (The first bill was withdrawn; the second remains up for debate.)

These were important and nuanced conversations that frequently unfolded on Twitter. But evidence collected by researcher and Mozilla Fellow Odanga Madung suggests a political organization thousands of miles away – the right-wing, Spain-based organization CitizenGO – played an outsized role in steering these Twitter conversations by inserting misinformation and inflammatory rhetoric. And it appears that Twitter’s own trending topics feature acted as an unwitting accomplice, amplifying at least 10 related hashtags.

These Twitter campaigns attacked Kenyan politicians and activists who advocated for the bills and reproductive rights more broadly. The campaigns also made false and misleading claims around abortion, surrogacy, and other areas of reproductive health. For example, they claimed children of surrogacy are more likely to display behavioral and emotional problems – a claim the African Population Health Research Center (APHRC) has cited as “lacking any scientific backing.”

Madung’s research shows that these Twitter campaigns had several key markings of inauthentic, coordinated behavior. Using Twitter’s Firehose API, Madung revealed that thousands of the problematic tweets repeated identical hashtags, phrases, and memes; came from accounts that tweeted nothing but the hashtags; and were carefully synchronized for certain times of day.

Further, participants in Kenya’s “disinformation industry” corroborated CitizenGO’s involvement. In his previous research, Madung plumbed Kenya’s disinformation industry, a network of Twitter users in the country who receive cash and instructions to make certain topics and opinions trend. Most recently, these sources said they received money, content, and instructions from CitizenGO over WhatsApp.

Indeed, upon presenting this research to Twitter in February 2022, the platform permanently suspended over 240 accounts associated with the campaigns, deeming it a violation of their Terms of Service.

While these particular Twitter campaigns have previously gone unreported on, the broader phenomenon is not new. Kenya is not the only place Twitter’s trending algorithm has allowed for manipulation by external entities. In the case of South Africa, the UK-based public relations company [Bell Pottinger ran a covert campaign](#) aimed at sewing seeds of racial tension.

This research ultimately reveals how Twitter has not sufficiently invested the resources – from understanding the cultural context, to adequate staff, to internal prioritization – to meaningfully address the ways in which the platform can be weaponized by bad actors. And, how this grows more and more precarious as Kenya’s 2022 federal election approaches.

Introduction

In our [previous research](#), we covered how the features of social media platforms have been exploited by malicious actors (particularly governments) for nefarious ends. And we've shown how the failure by these platforms to pay attention and respond to the potential harm of their features has had very real consequences. The platforms have been used to quash dissent and consolidate power across Kenya. This time, we wish to show how this negligence also leaves the Kenyan ecosystem of platforms open to manipulation from outside groups. Particularly, European right-wing movements that may be seeking to manipulate conversations about reproductive rights and health.

In 2020, [the Reproductive Healthcare Bill](#) was tabled in Parliament. It asked Kenyan politicians to vote to "provide a framework for the right to reproductive healthcare and the decisions regarding them and to set the standards of reproductive health." The bill sought to outlaw forced sterilizations and make prenatal, delivery and postnatal services free to every woman in the country. It made clear that everyone is entitled to reproductive healthcare services, including adolescents. Almost immediately, anti-abortion campaigners launched objections, calling it "[The Abortion Bill](#)." Additionally, they claimed the bill opened the way for legalized abortion and use of contraceptives among adolescents.

In Kenya, abortion is illegal unless a woman's life or health is in danger and emergency treatment is necessary. According to a 2012 study, [40% of pregnancies](#) in Kenya are not planned. Since not all women have access to contraception, unregulated and unsafe abortions are common. In the same year, researchers found out that nearly [half a million abortions](#) were carried out in the country.

In 2021, the Surrogacy Bill, was put to Kenyan politicians. It proposed the formation of the Assisted Reproductive Technology authority to develop standards, regulations and guidelines on assisted reproduction; and establish and maintain a confidential national database of persons receiving or providing services, sperm, eggs and embryos, among other functions. Again, staunch objections to the bill were launched. This time, anti-abortion campaigners claimed that the bill would transform Kenya into "a baby manufacturing hub" and that "the practice of surrogacy is the equivalent of buying and selling of children." However, unlike the Reproductive Healthcare Bill, this one was passed by parliament and will likely be debated at Kenya's Senate in February 2022.

Kenya is not the only place Twitter's trending algorithm has allowed for manipulation by external entities. In South Africa, now [disgraced and defunct](#) London-based public relations company Bell Pottinger took on a notorious client in 2016 that was the catalyst for its demise. In January of that year, the firm met with the owners of Oakbay Investments – Ajay, Atul and Tony Gupta, a family accused of [corrupt dealings](#) under the leadership of then president Jacob Zuma. Also at the meeting was Duduzane Zuma, the former president's son. According to leaked documents, Duduzane and the Guptas briefed Bell Pottinger to run a campaign targeted at "grassroots"

South Africans. The idea was to direct attention away from alleged [state capture by the Gupta](#) family by distracting the largely marginalized black population and stirring up conversations about “economic apartheid” in the country since the advent of democracy in 1994.

It was found that Bell Pottinger’s campaign infiltrated a mixed bag of communication streams including Twitter. “White Monopoly Capital” became the most popular phrase synonymous with the campaign. It [was reported](#) that over 100 fake accounts shared over 220,000 tweets with some of the most popular hashtags being #WhiteMonopolyCapital and #RespectGuptas. The coordinated attack was intended to unite South Africans against the common enemy which these bots said was white-owned businesses and not the Guptas – whose reputation the campaign sought to sanitize. While the campaign blew up in the faces of the Guptas, Bell Pottinger and Jacob Zuma, it cannot be denied that it was effective in sowing seeds of racial tension in the country.

Cases like this show that the trending algorithm is far too easy to manipulate by bad actors seeking to drive conversation.

In this research, we explore how a far-right group in Europe may have co-opted conversations in Kenya about reproductive rights and health, inserting disinformation and inflammatory rhetoric into an important and nuanced regional conversation. Not only did these conversations lead to dissent on Twitter, they may have influenced parliamentary legislative decision-making. This research illuminates how CitizenGO appears to have conducted its disinformation campaigns, the consequences in Kenya, and Twitter’s failure to address the problem.

These two scenarios have something in common. On each occasion, opposition to legislation was accompanied by online campaigns launched on Twitter. Additionally, in both cases there were attacks on the women politicians who sponsored the legislation. Finally, in both cases these campaigns were amplified by Twitter’s trending algorithm to millions of Kenyans.

Who, however, was behind these online campaigns? And were the accounts tweeting these messages real and authentic? According to our research, the Spain-based group CitizenGO appears to be the instigator. We established this through an analysis of tweets from CitizenGO’s [Twitter account](#) and [staff](#), along with interviews with influencers who claimed to have been paid to participate in the campaigns. Our research suggests that these campaigns should be considered [coordinated inauthentic activity](#), and that Twitter, through its trending algorithms, ended up amplifying the messages of these campaigns.

Indeed, when we presented this evidence to Twitter in 2022, the platform responded by permanently suspending over 240 accounts associated with this activity, which they deemed to have violated their [Terms of Service](#). Specifically, the accounts

violated Twitter's rules on [artificial engagement](#) and [coordination to amplify or disrupt conversations](#).

Who Are CitizenGO?

CitizenGO is a Spanish organization that presents itself as a community of active citizens working together to defend anti-choice movements. Its founder is Ignacio Arsuaga. The organization has close ties with Spain's far-right through its [association](#) with the Vox Party. [It has been labeled an anti-immigrant, anti-gender, climate denialist party](#) by the BBC, and was found to have been running a [Super PAC during the 2019](#) Spanish national elections. It has also been [linked to](#) far-right movements in Italy, Hungary and Germany. Additionally, one of CitizenGO's board members was [sentenced to four years in prison](#) in January 2021 on corruption charges, and the organization [has also been accused](#) of presenting misleading information regarding how it funds its operations.

In Kenya, Nigeria, Tanzania and other African countries, there is evidence that CitizenGO [is exporting its model of petitions, campaigns and offline lobbying, which seeks to deny women their human rights](#). This includes their rights to health and family life. Kenya has been a key target for CitizenGO, given that it is a conservative society that [identifies as 85% Christian](#). Its tactics in Kenya echo its actions in Spain and the rest of the EU. In this way, Kenya may serve as a petri dish for CitizenGO to test out messaging, activities, and tactics throughout the African continent.

According to British human rights campaigner Peter Tatchell, right-wing movements see Africa as one of their bastions of hope. "The religious and far right have lost the fight against gender rights in the West. They are turning their focus to poorer countries where they use their millions to buy influence, manipulate the political agenda and expand their base of adherents." By their own admission, right-wing groups look long term at Kenya "[because it is such an influential country throughout Africa.](#)"

The face of right wing and conservative activism in Kenya has also started to shift. Rights campaigners have said that the people behind these movements are no longer the white saviors of ages past, but they essentially now "[look like us and talk like us,](#)" Such is the case with CitizenGO's Africa director, Ann Kioko. She is routinely at the forefront of CitizenGO's activities in Kenya and the rest of Africa.

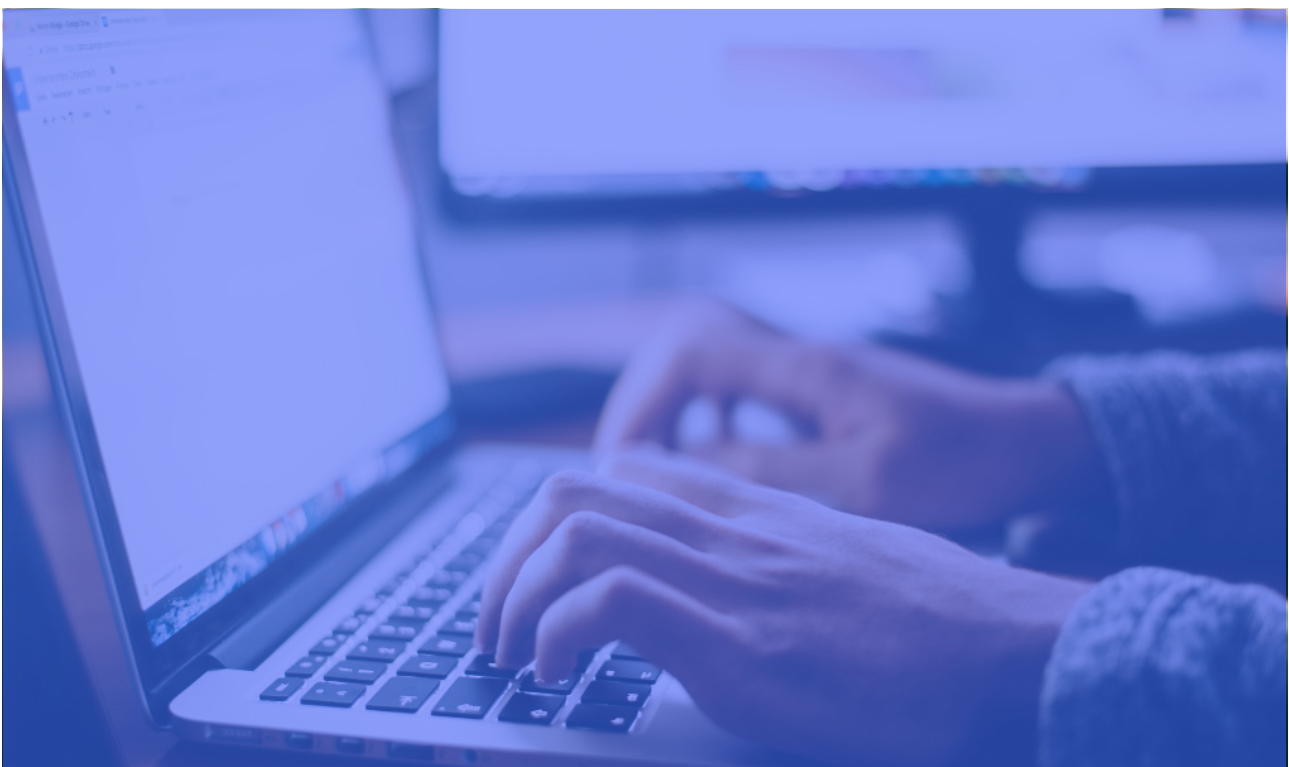
Tapping into Kenya's Disinformation Industry

In order to seed its narratives into the mainstream, one of the tools CitizenGO used was Twitter. In particular, it appears they tapped into Kenya's [flourishing disinformation industry](#). We managed to extensively interview three Kenyan "disinformation influencers" via phone interviews and WhatsApp, who claim to have been paid between \$10 -\$15 per campaign by individuals on behalf of CitizenGO. According to them, about 15 individuals who hold multiple sock puppet Twitter accounts participated in the campaigns.

The influencers shared with us screenshots from WhatsApp groups that bore CitizenGO's name and branding, and in which they would coordinate. From our review of the content within them, these groups appeared to serve two functions: (1) they were the command center of the campaign where instructions and media packs were shared with influencers; and (2) they were where influencers would report the progress of the campaign as to whether it achieved its intended goal or not.

We were not able to confirm whether members of CitizenGO itself were in these groups, but CitizenGO's Africa account participated in seven of the 11 campaigns we identified. What's more, many of the media assets in these campaigns had CitizenGO's branding.

Overall, these manipulation campaigns may have successfully used a large number of accounts to tweet what appear to be predetermined hashtags, ensuring that its problematic messages were amplified by Twitter's trending algorithm. After examining the activities of CitizenGO and its staff's accounts between 2019-2021, we identified 11 such campaigns and reviewed a total of 20,811 tweets gathered from Twitter's firehose. Twitter promoted 10 of these campaigns on its trending section, and according to the influencers we spoke to, these were the campaigns they were paid to participate in.

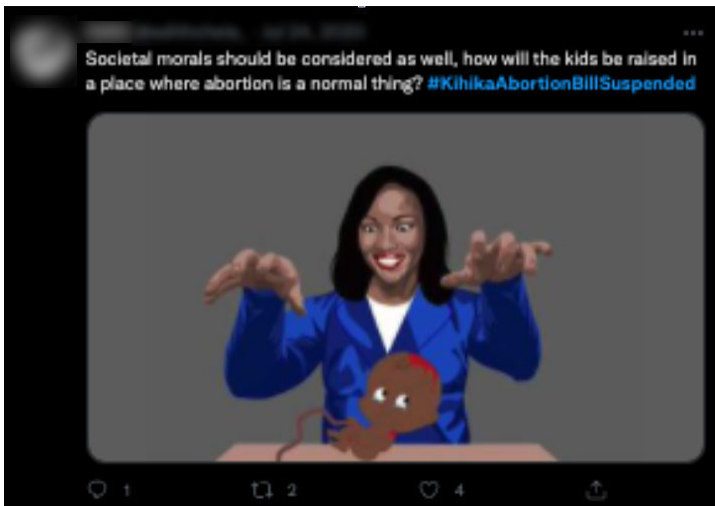


Hashtag	Link	Dates	Trending
#celebratingAnnkioko	https://twitter.com/search?q=%23celebratingannkioko&src=typeahead_click	04/06/2021	https://archive.twitter-trending.com/kenya/12-09-2021
#Istandwithcitizengo	https://twitter.com/hashtag/IstandWithCitizenGO?src=hashtag_click	09/11/2020	http://www.trendinalia.com/twitter-trending-topics/kenya/kenya-200911.html
#Annkiokoat33	https://twitter.com/hashtag/AnnKio-koAt33?src=hashtag_click	05/02/2021	https://archive.twitter-trending.com/kenya/02-05-2021
#stopkihikaabortionbill	https://twitter.com/hashtag/stopkihikasabortionbill?src=hashtag_click	06/25/2020	https://archive.twitter-trending.com/kenya/25-06-2020
#Notoabortionin-covid19	https://twitter.com/hashtag/NoToAbortionInCovid19?src=hashtag_click	05/09/2020	https://archive.twitter-trending.com/kenya/08-05-2020
#passarisabortionkills	https://twitter.com/hashtag/PassarisAbortionKills?src=hashtag_click	03/12/2020	https://archive.twitter-trending.com/kenya/12-03-2020
#stopsurrogacybillKE	https://twitter.com/hashtag/stopsurrogacybillke?src=hashtag_click	11/19/2020	https://www.google.com/url?sa=i&url=https%3A%2F%2Fmobile.twitter.com%2Fhashtag%2F3stopsurrogacybillke&psig=AOvVaw08y-UsGdwodQVzP22qLEWn-j&ust=1642607256730000&source=images&cd=vfe&ved=0CAwQjhxqF-woTCJj66oLTu_UCFQA-AAAAAdAAAAABAD
#notoabortionKE	https://twitter.com/hashtag/NoToAbortionKE?src=hashtag_click	09/12/2021	https://archive.twitter-trending.com/kenya/12-09-2021
#citizengoatthe-senate	https://twitter.com/hashtag/CitizenGOAtTheSenate?src=hashtag_click	09/07/2020	https://archive.twitter-trending.com/kenya/07-09-2020
#STOPabortionUNCSW	https://twitter.com/search?q=%23stopabortionUNCSW&src=typed_query&f=live	03/23/2021	https://drive.google.com/file/d/1IXbEMdbaSkGmg8rb-jInbiqXVOKBjZek/view?usp=sharing

Hashtag	Link	Dates	Trending
#KihikaAbortionBill-Suspended	https://twitter.com/hashtag/KihikaAbortionBillSuspended?s-rc=hashtag_click	06/24/2020	https://archive.twitter-trending.com/kenya/24-07-2020

Our research suggests that the goal of the campaigns in particular may have been to frustrate any attempts to have factual conversations on the bills that were put forward and to create moral panic around the issues. It is a tactic which is likely to be effective in achieving its goal because matters like teenage sexuality, abortion, and surrogacy are hotly contentious public issues in Kenya. Any calls for honest discussion based on facts are met with misinformation and strident moral arguments.

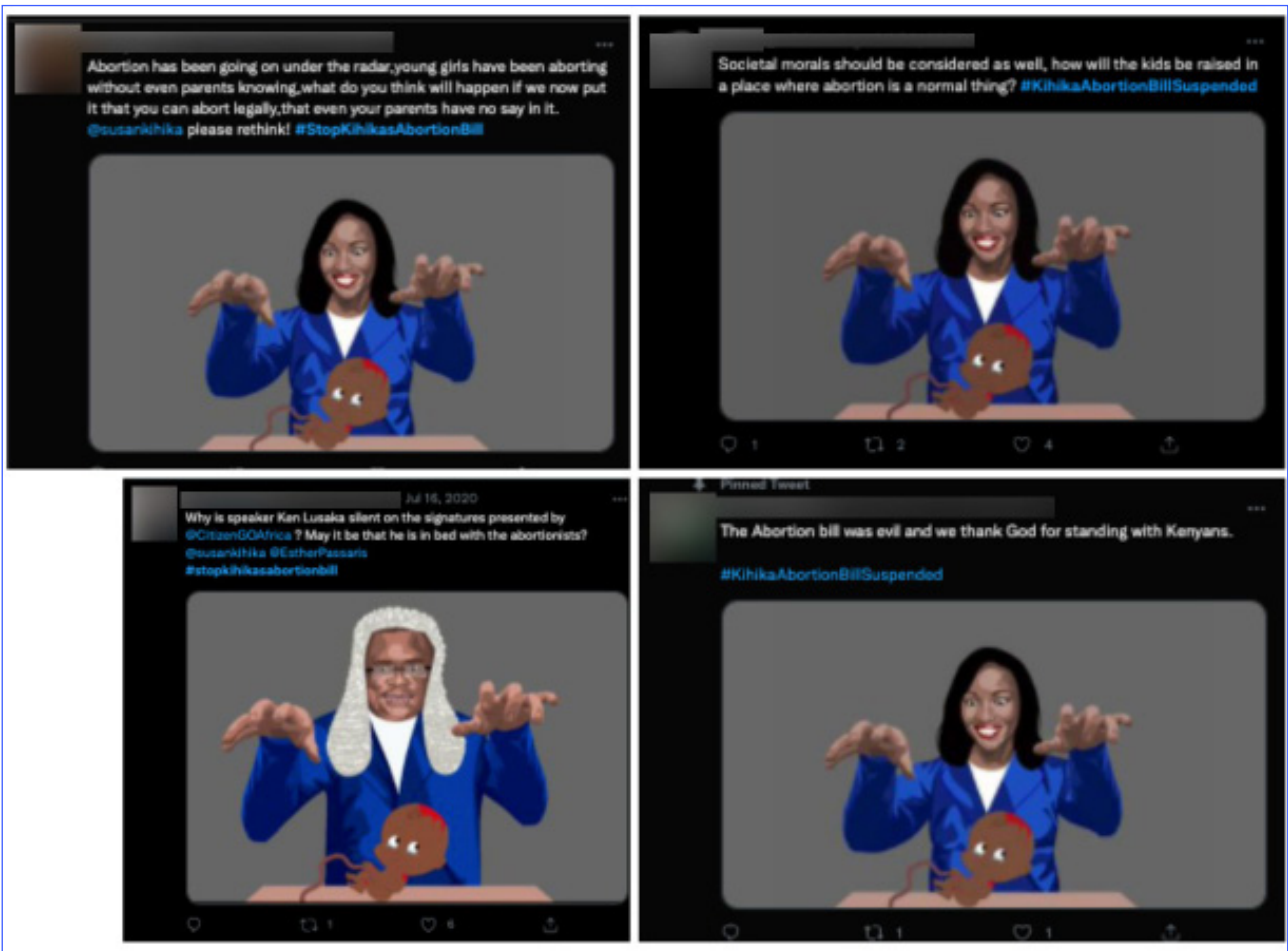
The Twitter campaigns appear to have inauthentically promoted anti-choice sentiment and attacked politicians and activists advocating for gender rights. These attacks on activists typically used cartoon caricatures and memes that echo messaging by American and European right-wing groups. For example, some of the tweets pushed George Soros conspiracy theories to audiences. He is often seen as the [Bogeyman](#) of the far right and some of the campaigns pushed the idea that he was funding attacks on Ann Kioko. This particular narrative was also evident in some of the briefs the influencers were given in their WhatsApp groups.



Once again, we were able to identify patterns that we've seen in previous inauthentic campaigns that make Twitter's trending topics so easy to manipulate, i.e.

Once again, we were able to identify patterns that we've seen in previous inauthentic campaigns that make Twitter's trending topics so easy to manipulate, i.e.

- **There was heavy repetition of content within the hashtags.** Hashtags and phrases were part of the media pack given to disinformation influencers via a brief in their WhatsApp groups. Accounts involved in the campaign used the same set of media assets in the tweets or repeated phrases within them. According to influencers we spoke to, this also happens when one person operates multiple accounts to push the campaign.

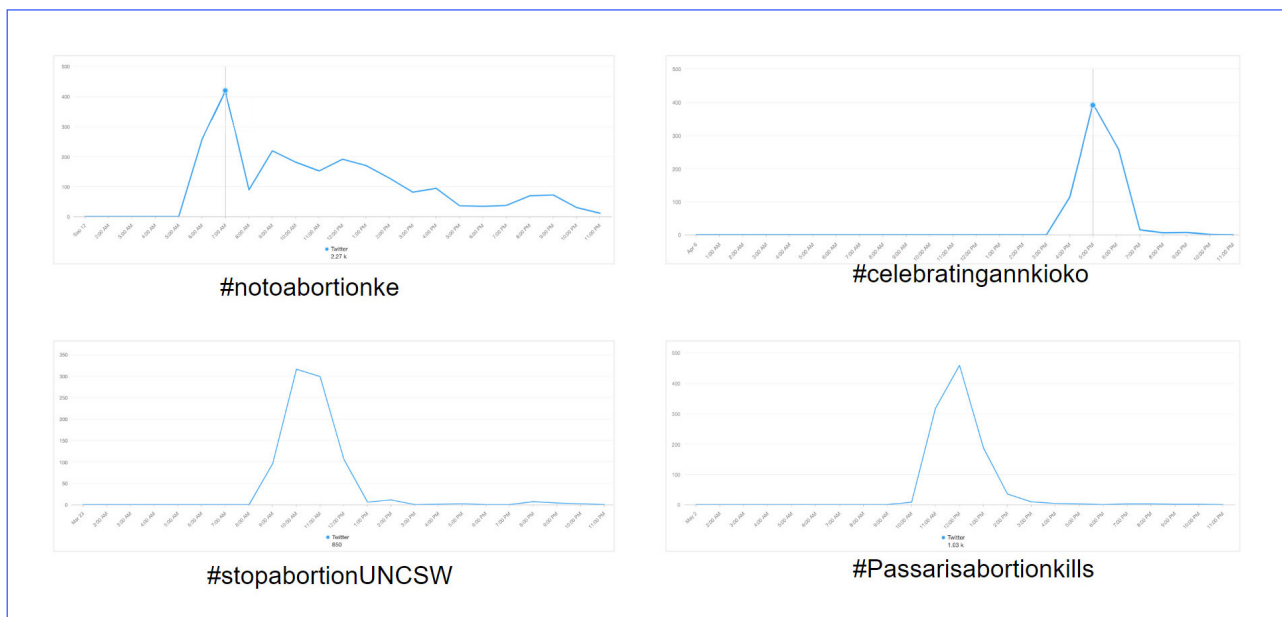


Tweets containing images depicting Kenyan Legislators



Tweets containing "stop surrogacy now" campaign material

- Several of the accounts that appear to have participated in these campaigns [tweeted](#) nothing but [political and marketing hashtags](#) for days on end. According to the influencers we spoke to, this is because these accounts build audiences with the aim of using them to sell marketing services to whoever comes calling.
- The hashtags had synchronized time of flight as well, with some producing their highest volume of content in the morning hours – a tactic which is [consistent with inauthentic behavior in previous pieces](#) of research we’ve published.



- In many cases, tweets using these hashtags shared false information about reproductive health and relevant bills in what appears to be an attempt to mislead the public, [which health officials say](#) frequently lacks information about these issues. These tweets labeled the Reproductive Health Bill as an abortion bill, despite several [researchers](#) and [rights observers](#) arguing that the bill was in keeping with the rules of the existing constitution. Additionally, the #stopsurrogacybillKE campaign that was run on 11/19/2020 (and was published on Twitter’s trending section) had several examples of health misinformation and false claims. [Tweets using the hashtag #stopsurrogacybillKE](#) made false claims, such as:
 - Children of surrogacy are more likely to [display behavioral and emotional](#) problems;
 - The surrogacy bill explored the “[experimentation of laboratory baby production](#);
 - Surrogate births are more likely to lead to preterm birth, stillbirth, low birth weight, fetal anomalies and high blood pressure; and
 - The practice of surrogacy is the [equivalent of buying and selling of children](#);

We ran all these claims by medical and public health experts at the [Africa Population and Health Research Center \(APHRC\)](#), and they confirmed to us that all of them are false. Kenneth Juma, an APHRC researcher, said: “These claims, especially around the Surrogacy Bill 2019, are FALSE and lack any scientific backing. Most of these misleading statements prey on the lack of awareness among the people.”

He went on further to say: “In detail, published research evidence has discounted the indicated claims. For instance, while there are certain risks that may stem from IVF births, researchers who have systematically reviewed results from multiple studies, find no evidence of an increase in behavioral or cognitive problems in children born after IVF. And where there are neurodevelopmental disorders, these may also stem from low birth weight or multiple fetuses.”

Where does this leave Twitter?

CitizenGO's case shows the kind of vulnerabilities that platforms like Twitter create in Kenya. Twitter [has often claimed](#) that it wants its platform to be a place where healthy conversations can happen. However, the problem with these kinds of disinformation campaigns is that they strip these conversations of the facts and create a moral panic, stifling any progress legislators seek to make on reproductive healthcare rights. According to researcher Kenneth Juma, disinformation campaigns take advantage of the lack of awareness the public has about these issues and attempts to legislate around them: "Limited knowledge of sexual and reproductive health and rights is a precursor for such misinformation. Most often, the intent for perpetrators of such distortions is not to pass credible information but rather to sensationalize the subject and make it untenable to have proper conversations."

Plenty of Kenyan women turn to social media platforms to seek out information on reproductive health, so it is critical that platforms improve how they identify and handle health disinformation. The rate, method and critical times at which these campaigns happened suggests that some of the messaging around reproductive health that people see on Twitter was the product of a manipulation campaign. As Peter Tatchel puts it, "The far right exploits platforms to get away with poisonous claims stirring fear and hatred that are then engineered into witch hunts and repressive legislation. Social media campaigns like these are cheaper to run than traditional media advertising. They get more bang for bucks if they can bypass local press to reach people."

These groups have made it very clear that [they are seeking to have a hand in electoral campaigning](#) and see Twitter as a vital tool for achieving their goals. As we've mentioned before, many groups on the right like CitizenGO recognize Kenya's influence in the region, and therefore have an incentive to influence its elections.



Once again, our research suggests that Twitter’s trending algorithm is simply too easy to manipulate. Despite the overwhelming evidence we’ve provided to Twitter about the abuse of the Trending feature, until recently no public statements have been made about what steps the social media giant will take in Kenya. This is especially troubling in light of the 2022 general election that draws nearer every day, and we anticipate that demand for disinformation-for-hire services will rise.

In response to our findings, Twitter through a spokesperson said:

“We are investigating the information shared with us by Mozilla Foundation and have permanently suspended more than 240 accounts under our [platform manipulation and spam policy](#). Our work to combat inauthentic virality and fake influencer accounts is long-running; since 2020, we have permanently suspended more than 600,000 accounts found to violate our rules in this manner.

“Twitter’s uniquely open nature empowers research such as this, and we remain vigilant about coordinated activity on our service. Using both technology and human review, we proactively and routinely tackle attempts at [platform manipulation](#) and mitigate them at scale by actioning millions of accounts each week for violating our policies in this area. We are constantly improving Twitter’s auto-detection technology to catch accounts engaging in rule-violating behavior as soon as they appear on the service.”

However, Twitter did not respond to questions about what partnerships they may have with civil society organizations and public health institutions in Kenya. They also didn’t respond to questions about whether their current public health policy is equipped to handle such problems.

It’s worth mentioning, however, that if the COVID pandemic has taught us anything, it’s that the internet is rife with health misinformation. Health is a political topic and sometimes the source of this problem is the medical establishment itself. This makes it a tremendously tough problem for platforms to referee, and social media companies traditionally have no subject matter expertise to tackle health misinformation. Twitter’s current approach to health misinformation is a patchwork of policies that are sometimes only invoked when [there is public attention or pressure](#). There is more that Twitter can be doing. Researcher Kenneth Juma of APHRC called upon platforms like Twitter to go beyond low hanging fruit solutions, stating: “Beyond labeling, these platforms can delete outrightly false information and ban superspreaders of false information.”

Twitter needs to exist in Kenya and it is in everyone’s interest that the platform become a vehicle for healthier discourse. The platform is a vital avenue of civic engagement and education and can be used to hold public officials accountable. As such, we recommend that Twitter take concrete steps to prevent its trending feature from being manipulated. We recommend that Twitter take the following steps in order to safeguard its platform against disinformation campaigns:

1. Undertake dedicated country-specific risks assessments that seek to identify, assess and mitigate the ways in which the trending algorithm can contribute to disinformation online.
2. Provide more clarity and information around their efforts to fix the trending algorithm. Simply notifying people when [they remove the feature](#) within certain territories is not enough.
3. Develop partnerships with fact checkers and civil society organizations to help them navigate the complexities of the region. Be transparent with the public about those partnerships.
4. Establish dedicated collaborations with in-region researchers, to facilitate greater oversight of and insight into the ways in which Twitter can be a vector for disinformation.

Conclusion

Bearing in mind our findings and the upcoming Kenyan election in 2022, we are concerned that Twitter's platform manipulation policies and even their enforcement have not been enough to address how paid political influence occurs on the platform. Even after sharing our evidence with Twitter, and Twitter suspending the accounts, [yet another campaign emerged](#) just before we published this report, using the same tactics. All this is happening despite the existence of Twitter's policies and, moreso, despite their enforcement. Our ask to the platform is that they consider our recommendations in order to safeguard against potential harm, especially because of Twitter's importance to political discourse in Kenya.

Further Reading

<https://foundation.mozilla.org/en/blog/fellow-research-inside-the-shadowy-world-of-disinformation-for-hire-in-kenya/>

<https://foundation.mozilla.org/en/blog/new-research-in-kenya-disinformation-campaigns-seek-to-discredit-pandora-papers/>

<https://www.opendemocracy.net/en/5050/revealed-the-trump-linked-super-pac-working-behind-the-scenes-to-drive-europes-voters-to-the-far-right/>

<https://bylinetimes.com/2021/06/21/the-money-men-behind-citizengos-anti-rights-agenda/>

