

ISSUE 144 • JUNE 2017

Using Digital and Social Media to Monitor and Reduce Violence in Kenya's Elections

Social media and digital technology offer immense potential for citizens, policymakers and practitioners to raise awareness of, monitor, and respond to violence. With Kenya's elections approaching, technology can help to raise awareness of insecurity, support early warning, combat incitement of violence and promote accountability. However, digital technology also carries a number of risks. To maximise effectiveness and inclusivity, 1) greater support must be given to locally legitimate peace messaging and counter-speech; 2) government, media and civil society should collaborate to improve transparency and accountability in the regulation of online activity; and 3) social media monitoring of violence should be undertaken in conjunction with other reporting systems that seek to overcome inequalities in digital access and use.

A new panacea?

Social media and digital technologies are transforming the way the public, civil society and policymakers engage with information about violence, and how that information is captured, analysed and acted upon.

Social media platforms allow members of the public to share information on insecurity, and widespread access to mobile phones enables individuals to produce and disseminate much more immediate and detailed audio, visual and textual reports of violence than previously possible. These new information and communication tools have been key in spreading awareness of insecurity, and supporting efforts to monitor, and respond to, violence in a range of contexts.

In Kenya, the use of WhatsApp, Twitter and Facebook to communicate has transformed the information landscape, and been used to raise awareness or promote peace messaging in the face of uncertainty, rumour and threat. Similarly, purpose-built digital systems for violence reporting, monitoring, and crisis response, have proliferated in recent years.

At the same time, after an initial period of uncritical optimism regarding the potential of technology to promote peace, there is now growing recognition of its limitations. Social media and digital technology can be used to incite violence and spread hate speech, features that have been well-documented in Kenya's past elections. Equally, even well-intentioned initiatives relying on social media and digital technologies to monitor and respond to violence can exclude or marginalise the voices of the vulnerable.

This briefing explores the contrasting roles of social media and digital technology in Kenya's elections; and proposes recommendations for policymakers and practitioners in related fields to inform and improve violence monitoring initiatives and ultimately reduce and better respond to violent crises.

Social media and digital technology in Kenya

Social media and digital technology are the latest frontier in political engagement, serving to facilitate communication between leaders and the public, bring political processes closer to the people, and mobilise political participation.

“Social media monitoring of violence should be undertaken in conjunction with other reporting systems that seek to overcome inequalities in digital access and use.”

This evolving role has been facilitated by the rapid spread of mobile phone and internet usage in Kenya over the past 15 years. Estimates suggest there are currently 39.6 million internet users and 31.9 million phone users in the country. Popular social media platforms include WhatsApp (with an estimated 10 million users), Facebook (6.1 million), and Twitter (2.2 million monthly active users). Informal coalitions of Twitter users, commonly referred to as 'Kenya on Twitter' #KOT, have been instrumental in influencing online public participation in policymaking, sometimes engaging in fierce critique of political processes and policymakers.

Threat conveyer and violence mobiliser

The rapid spread of these technologies also carries significant risks, including facilitating the rapid and widespread dissemination of rumours, threats, hate messages and calls to violence.

In Kenya, the 2007 presidential elections saw the dissemination of hateful language through multiple media outlets. However, these elections were the first in Kenya to see this occur via 'new' media (social media, mobile phones, and online platforms), alongside 'old' (radio, television and print). This included SMS bulk messaging through mobile phones, as well as e-mail, blogs, and social networking platforms. After the presidential election results were announced, the internal security minister announced a ban on live broadcasts. Given the news blackout, text messages became the primary means by which people updated each other on the election and its fallout. This was double-edged: while SMS was used to rally support for peace and raise awareness of insecurity, it was also used to spread hate and incite further violence.

As the 2017 elections approach, the National Cohesion and Integration Commission (NCIC) is mandated with fostering unity and tackling hate speech propagated on both 'old' and 'new' media platforms. The NCIC has actively pursued hate speech suspects, summoning politicians allegedly engaged in incitement. This has suppressed incitement in public political discourse, such as in politicians' and aspirants' speeches, and in official campaign material.

However, digital and social media provide new ways to spread inflammatory messages, sometimes discreetly or anonymously. Over the years, WhatsApp has become one of the most popular platforms in Kenya. This

closed messaging system is more private, and not open to online monitoring. Thus, people may be less cautious about what they post in groups of friends, neighbours or close networks. Hence, harmful content on WhatsApp is rarely reported to authorities, and is very difficult to monitor independently.

The NCIC also monitors hate speech online and collaborates with operators and the police to investigate alleged incidents. However, it has not yet successfully prosecuted any person on hate speech, due in part to challenges posed by technology and online speech.

Peace mobiliser and counter-messaging

At the same time, digital technology can play an important role in confronting and challenging hateful messages and incitement to violence.

In 2007/08, social media was used to promote peace and safeguard against the spread of hate speech. Re-tweets allowed users to voice their agreement with a news item or opinion, and popular tweets spread quickly through cascade effects. Hashtags also played a role in online peace promotion.

Counter-speech, defined as crowd-sourced responses to dangerous content, was also popular following the 2013 elections. Individual users as well as broadcasters used platforms such as Twitter and Facebook to spread positive messages and encourage people to remain calm. Messages promoting peace were also circulated through SMS.

Some regulators and service providers have taken further steps to tackle hate speech or incitement. In the lead up to the 2013 elections, Kenya's Communications Authority sought to regulate bulk SMS messaging by requiring providers to screen content for dangerous speech. Mobile service providers blocked an estimated 300,000 text messages per day, suspected to contain hate speech. However, this process of blocking content was not transparent and decisions were not subject to review by public bodies to allow for democratic scrutiny and to ensure citizens' rights to freedom of expression were not violated.

Documenter and forecaster of violence

Social media and digital technology can also be used to document and forecast violence. Digital technology has been used to document, model and predict violence using

“Digital technology can play an important role in confronting and challenging hateful messages and incitement to violence.”

crowdsourced platforms in contexts including Eastern DR Congo and Syria, where media and researchers typically have limited access.

In Kenya, Facebook pages such as 'Elections Iwitness Kenya' and 'Kenya Elections 2013' were established before the 2013 elections with the aim of allowing users to share accounts of electoral conduct. WhatsApp has also been instrumental in sharing videos, pictures and information during crises. Similarly, since 2007, custom-built initiatives have encouraged citizens to report and document violence using online platforms, including sharing videos and photos. This has enabled analysts to track and verify reported violence, identify flashpoints, and forward credible accounts to responders.

In addition to monitoring actual violence, digital technology can be used for early warning: in preparation for the 2017 elections, The Sentinel Project and iHub Research have created *Una Hakika* ('Are you sure?'), a mobile phone-based service that monitors the spread of rumours and acts to contain them. Subscribers can report rumours through text messages or by talking to trained volunteer community ambassadors. The well-known *Ushahidi* platform will also crowdsource reports of insecurity in the 2017 elections. In both systems, reports relating to actual or rumoured violence are filtered and referred to responders for action.

However, acting on reports often depends on the extent to which initiatives are joined up to, and have the support of, relevant responder groups.

Digital technology: Whose voice?

While social media and digital technology can play a variety of positive and negative roles in crises, they are also limited by issues of inclusion and representation.

In Kenya, despite relatively high levels of internet and smartphone use, the majority of people do not have regular internet access, and SMS is expensive for the poor. Unequal digital technology infrastructure across the country, and differential use according to age, income and education, can limit the representativeness and inclusiveness of these technologies.

Different interpretations of what constitutes violence may also mean certain forms of violence are systematically under-reported. Intimate partner and criminal violence may not

be routinely captured in systems designed for early warning in political crises, but evidence suggests these phenomena are inter-related, as patterns of gender-based violence, for example, are often correlated to wider unrest.

Equally, the political context matters: suspicion and distrust of security and intelligence services could lead to reluctance to share information on platforms that are actively monitored, due to concerns that it could be leaked or misused.

These factors hinder users' ability to engage with these platforms, including contributing to positive digital initiatives to promote peace. This can also produce biases in the data collected and used to inform early warning and crisis response.

Therefore, if policymakers and practitioners rely primarily on social media and digital platforms for violence reporting, many people experiencing violence may be silenced and typically affluent, urban, young, literate voices may be further amplified. Initiatives gathering data through digital platforms and social media should be used in conjunction with other reporting pathways that specifically seek to correct for potential exclusion effects.

“While social media and digital technology can play a variety of positive and negative roles in crises, they are also limited by issues of inclusion and representation.”

Digital technology – The 5 A's

Five factors effectively exclude many people from full use of digital technology:

- 1. Availability:** In many contexts, a large portion of the population live in remote areas where digital technology and social media are simply not available.
- 2. Affordability:** Even where available, it may not be affordable for many people.
- 3. Awareness:** Even when affordable, people may lack awareness of its presence or its use; or of particular initiatives (e.g., in violence reporting systems).
- 4. Ability:** Among groups that are aware, some people may lack the ability to effectively use technologies due to digital literacy skills.
- 5. Accessibility:** Even where those conditions are met, accessibility can still exclude people from participation: for example, platforms may exclude minority language groups.

As a consequence, marginalised groups may be systematically excluded from full digital participation, even though these groups are vulnerable to violence. The extent of this exclusion, and its effects on the accuracy of violence reporting in Kenya specifically, remain largely unknown.

Recommendations

- **Donors, government and civil society should support initiatives that promote locally legitimate peace messaging and counter-speech as means of confronting incitement to violence on open platforms, but increasingly popular closed messaging systems require more innovative and inclusive strategies:** Community-based counter-speech is an effective means of challenging and confronting hate speech and incitement to violence, leveraging popular opinion and promoting greater awareness of inclusive and democratic engagement. However, closed messaging systems, such as WhatsApp, are increasingly popular in Kenya, and are not subject to the same public oversight. Addressing incitement to violence and hate speech on these platforms requires more innovative and collaborative responses to generate more widespread public support, in contexts where official monitoring and response is limited. Options include efforts to encourage users to pro-actively counter hateful messages in closed messaging systems, and generating broad-based support for community-based counter-messaging.
- **Guidelines by government and the private sector aimed at monitoring or regulating online speech should be transparent and accountable in judging what constitutes dangerous speech or incitement, so as to secure wide popular support and buy-in:** Attempts to monitor, regulate and prosecute online hate or dangerous speech have had limited success to date. This is due in part to technical challenges of responding to online activity, but also related to issues of political will and public credibility. Despite the fact that the NCIC has developed a set of guidelines linked to hate speech legislation, there is a need for increased collaboration with media practitioners and civil society to improve transparency and accountability. It is vital for public support and wider civil society buy-in that any efforts to monitor, regulate, prosecute or prevent certain speech online be based on transparent, accountable and openly debated parameters for determining what constitutes an infringement, so freedom of speech and legitimate critique is not impeded.
- **Government and civil society initiatives to monitor and track violence using digital technologies should be used in conjunction with other monitoring initiatives and reporting pathways:** Digital platforms and social media monitoring can generate huge volumes of data at rapid speed, but the accuracy and reliability of that data is impeded in part by underlying inequalities in access, use, and representation. Policymakers relying on digital monitoring systems should invest in triangulation methods that effectively combine multiple and targeted reporting pathways to overcome geographic, demographic and economic differentials in digital access and use. Options include combining digital reporting with systematic media monitoring, networks of on-the-ground reporters, or targeted reporting initiatives such as those that document sexual- and gender-based violence, or those that actively engage digitally under-represented groups.

Further reading

This *IDS Policy Briefing* draws on the following papers:

Mutahi, P. and Kimari, B. (forthcoming) *Examining the Impact of Social Media and Digital Technology on Electoral Violence in Kenya*, IDS Working Paper, Brighton: IDS

Roberts, T., and Marchais, G. (forthcoming) *Assessing the Role of Social Media and Digital Technology in Violence Monitoring*, IDS Working Paper, Brighton: IDS

See also:

iHub Research and Ushahidi (2013) *Umati: Monitoring Online Dangerous Speech*, https://ihub.co.ke/ihubresearch/uploads/2013/february/1361013008_819_929.pdf, ihub.co.uk (accessed 7 June 2017)

Roberts, T. (2017) 'Digital technologies exclude', *Making All Voices Count blog*, 2 May 2017, <http://www.makingallvoicescount.org/blog/digital-technologies-exclude/> (accessed 7 June 2017)

Credits

This *IDS Policy Briefing* was written by **Caitriona Dowd** and edited by **Carol Smithyes**. The support of the Economic and Social Research Council (ESRC) (UK) is gratefully acknowledged.

The opinions expressed are those of the author and do not necessarily reflect the views of IDS or the ESRC.

Readers are encouraged to quote and reproduce material from the *IDS Policy Briefing* series. In return, IDS requests due acknowledgement and quotes to be referenced as above.

© Institute of Development Studies, 2017
ISSN 1479-974X



IDS Policy Briefings are published by the Institute of Development Studies and aim to provide high quality analysis and practical recommendations for policymakers on important development issues.

To subscribe: www.ids.ac.uk/idspolicybriefings

Institute of Development Studies, Brighton BN1 9RE UK
T +44 (0) 1273 606261 F + 44 (0) 1273 621202 E ids@ids.ac.uk W www.ids.ac.uk
twitter.com/IDS_UK #idspolicy facebook.com/idsuk