IDS WORKING PAPER
Volume **2018** No **520**

Comparing 'New' and 'Old' Media for Violence Monitoring and Crisis Response in Kenya

Caitriona Dowd, Patricia Justino, Roudabeh Kishi and Gauthier Marchais

December 2018

Comparing 'New' and 'Old' Media for Violence Monitoring and Crisis Response in Kenya

Caitriona Dowd, Patricia Justino, Roudabeh Kishi and Gauthier Marchais

Summary

Social media and digital technologies are changing the way information about political violence is collected, disseminated, analysed and understood. Effective early warning and crisis response increasingly depends on the availability of timely, reliable reports of violence, and a growing body of research on violence relies on the availability of reliable violent event data to understand patterns, dynamics and trajectories of violence. While biases in traditional media – newspapers and print media – have been analysed and documented in the literature, there is relatively little information about biases in relation to new and emerging sources of data. This paper seeks to determine the comparative opportunities and limitations of 'new' and 'old' data sources for early warning, crisis response, and violence research. We compare the information set produced through social media violence reporting with conventional violence reporting around the August and October 2017 Kenyan elections. Specifically, we leverage data from a sample of social media reports of violence through public posts to Twitter. We then compare these reports with events coded from media and published sources coded by the Armed Conflict Location & Event Data Project (ACLED) along three dimensions: (1) the geography of violence reporting; (2) the temporality of reporting; and (3) the targeting of reporting. The paper has two main results. First, we find that Twitter has an urban and wealth bias. Twitter reports information most accurately in densely populated areas and areas that are more economically developed, but much less so in less populated areas. Second, Twitter reporting varies greatly over time. While reporting is high in the immediate period around elections, reporting during periods outside critical junctures is more limited. Reporting from 'old' media on violent events is much more consistent over time. We draw the implications of these results for policymakers, practitioners, and academic researchers.

Caitriona Dowd is a former Research Fellow in the Conflict and Violence cluster at the Institute of Development Studies (IDS), where her research focused on the politics of violent conflict and sub-national governance in sub-Saharan Africa. She currently works in a research role in the humanitarian sector; all views expressed are her own.

Patricia Justino is a development economist specialising in applied microeconomics. Her current research work focuses on the impact of violence and conflict on household welfare and local institutional structures, the micro-foundations of violent conflict and the implications of violence for economic development. She is the Director of MICROCON and co-founder and co-director of the Households in Conflict Network. Since June 2010, she convenes the Conflict and Violence cluster at IDS.

Roudabeh Kishi is Research Director of the Armed Conflict Location & Event Data Project (ACLED), where she oversees the quality, production, and coverage of all ACLED data across the globe; leads research and analysis across regional teams; aids in new partnerships with local sources and users; and supports the capacity building of NGOs and conflict observatories around the world.

Gauthier Marchais is a Research Fellow in the Conflict and Violence cluster at the Institute of Development Studies. He holds a PhD in Development Studies from the London School of Economics and Political Science. His research focuses on dynamics of armed mobilisation, social transformation and armed group governance in conflict affected areas.

Contents

Figures

Tables

Acknowledgements

Acronyms

ACLED	Armed Conflict Location & Event Data Project
API	Application Programming Interface
CHRIPS	Centre for Human Rights and Policy Studies
DRC	Democratic Republic of Congo
ESRC	Economic and Social Research Council
GDELT	Global Data on Events, Location and Tone
IDS	Institute of Development Studies
IEBC	Independent Electoral and Boundaries Commission
IED	improvised explosive device
KOT	Kenyans on Twitter
NGO	non-governmental organisation
OECD	Organisation for Economic Co-operation and Development
RSS	Type of web feed allowing users and applications to access updates to online content
SMDT	social media and digital technology

1   Introduction

Social media and digital technologies (SMDTs) are profoundly changing the way information about political violence is collected, disseminated, analysed and understood (Roberts and Marchais 2017; Steel 2015). SMDTs are creating new platforms for the horizontal sharing of information (Earl 2010); new sources of data for peacebuilding and conflict management (Reilly 2016); and new actors in the role of digital witnesses to violence and insecurity (Chouliaraki 2016; Gregory 2015). These violence monitoring systems can play a vital role in reducing, preventing and responding to insecurity. Effective early warning and crisis response increasingly depends on the availability of timely, reliable reports of violence, in order to react quickly, determine the scale and dimensions of crises, and target responses accordingly. Equally, a growing body of research on violence relies on the availability of reliable violent event data to understand patterns, dynamics and trajectories of violence (Justino *et al.* 2013; Salehyan 2015; Gleditsch, Metternich and Ruggeri 2014).

Reports of violence collected and coded from open access sources usually originate from 'old' media, such as newspapers or published print reports, or from 'new' social media, digital platforms, and crowdsourcing systems. Both approaches seek to produce timely, relevant and actionable data for analysis and response, but both encompass drawbacks that may potentially reduce their effectiveness. Systematic biases in selective reporting by traditional media sources have been extensively documented in the literature (*see* Baum and Zhukov 2015; Bocquier and Maupeau 2005; Desmaret and Langer 2018), and indicate ways in which systems relying on media monitoring alone may under-represent key segments of the population or areas of a country (Weidmann 2016). Initial optimism regarding digital technology suggested a new 'data revolution' in social media could address some of these concerns. A growing body of research has, however, indicated that digital technology may also be under-representing particular communities and groups (Lerman 2013; Read, Taithe and MacGinty 2016; Roberts and Marchais 2017). This research is, however, still in its infancy.

Sources used for reporting and documenting violence will unavoidably influence resulting narratives and perceptions about insecurity (Davenport and Ball 2002; Weidmann 2013). Yet, to date, we have relatively little information about the specificities and extent of this divergence in relation to new and emerging sources of data.[1] The implications of these differences are not only relevant to research, which draws on conflict event data to understand and theorise patterns of violence, but also to the policy and practitioner communities that create, support and use these systems to directly respond to and manage violence.

This paper seeks to determine the comparative opportunities and limitations of 'new' and 'old' data sources for early warning, crisis response, and violence research. We compare the information set produced through social media violence reporting with conventional violence reporting around the August and October 2017 Kenyan elections. Specifically, we leverage data from a sample of social media reports of violence through public posts to Twitter. We then compare these reports with events coded from media and published sources coded by the Armed Conflict Location & Event Data Project (ACLED) (Raleigh *et al.* 2010) along three dimensions: (1) the geography of violence reporting; (2) the temporality of reporting; and (3) the targeting of reporting. We focus on the case of Kenya, where elections have repeatedly been flashpoints of violence, and where crowdsourced violence reporting via SMDTs has been a feature of conflict management for several years (*see* Meier 2008; Mäkinen and Wangu Kuira 2008; Musila 2013).

---

[1]   *See* discussion on importance of this as a future direction in conflict research in Dafoe and Lyall (2015: 409).

The paper has two main results. First, we find that Twitter has an urban and wealth bias. Twitter reports information most accurately in densely populated areas (specifically Nairobi) and areas that are more economically developed, but much less so in less populated areas. Reporting from traditional media and published sources (hereafter referred to as 'old' media), has wider geographic coverage across the country, including reports from unique locations that Twitter reporting rarely covers, as well as across less densely populated and economically less developed areas. Reporting from traditional media and published sources is also generally more geographically precise than that which comes from Twitter.[2] We also find that reporting from traditional media and published sources tends to be more detailed, especially around reporting of fatalities.

Second, Twitter reporting varies greatly over time. While reporting is high (and more precise) in the immediate period around elections, reporting during periods outside critical junctures is more limited. Reporting from 'old' media on violent events is much more consistent over time. The timeliness of reporting – measured as the reporting lag, or the number of days between the date on which the event occurred and the date it is reported by the source – is generally similar across both sources, meaning that one is not considerably quicker than the other at large. However, Twitter tends to record events – specifically demonstrations – more quickly in the immediate period around elections, while traditional media and published sources report events – specifically organised violence[3] – more quickly outside of these periods. Reporting from 'old' media is also generally more temporally precise than that which comes from Twitter.[4]

The paper proceeds as follows: Section 2 presents key findings and gaps in existing research on violence monitoring, and derives a series of testable hypotheses to compare old and new media sources; Section 3 presents the research design; and Section 4 presents the main findings. Section 5 concludes the paper with a discussion of implications for future research and policy.

2   'New' and 'old' media for violence monitoring

There has been a growing wave of optimism regarding the potential of SMDTs to facilitate peacebuilding and transformative social and political change. The Organisation for Economic Cooperation and Development (OECD) 2009 report on *Preventing Violence, War and State Collapse* called on member states to recognise 'the critical importance of adopting innovative information and communication technologies for data collection, communication, visualisation and analysis' (OECD 2009: 19; Roberts and Marchais 2017). Similarly, the World Bank's 'Listening to Africa' initiative called for greater investment in cell phone-based data collection in crises, including forced displacement due to conflict, as a means of overcoming challenges of face-to-face data collection (World Bank 2017; Hoogeveen 2017).

Violence data generated through 'new' media provide a valuable opportunity to track, analyse and respond to crises. SMDTs have been deployed in a number of contexts for this purpose, including monitoring election-related violence in Nigeria through analysis of tweets (Bartlett *et al.* 2015); and a similar initiative in Ghana where civil society and journalists combed Twitter and WhatsApp to extract information using keywords such as 'violence',

---

[2]   For example, language such as "a specific neighbourhood of Nairobi" is more precise than "somewhere near Mombasa".

[3]   Organised violence refers to violence that is not spontaneous (the way violence that breaks out during a demonstration might be) and involves an armed agent and/or serious harm or death to at least one individual.

[4]   For example, language such as an event occurring "on August 10" is more precise than "sometime last month".

'gunshots', and 'killed' in order to produce incident reports of unrest (Moreno *et al.* 2017; Roberts and Marchais 2017). Similar technology has also been used to document and monitor violence against particular groups, such as election-related violence against women (Bardall 2013).

Previous studies have sought to test the reliability and actionability of similar data in Kenya specifically. For example, Sambuli *et al.* (2013) used the 2013 Kenyan elections as a case study to compare passive crowdsourcing from Twitter with active crowdsourcing through the Uchaguzi platform. They found that combing Twitter produced a significantly larger number of reports of critical incidents. However, it required more extensive human resources to filter out 'background noise' than the more targeted Uchaguzi platform, with its specific call to action. As a result, active crowdsourcing had a better signal-to-noise ratio and therefore lower human workloads to analyse the data, but it delivered fewer critical incidents (Roberts and Marchais 2017).

Aarvik (2015) also analysed the results of the 2013 Kenya elections, comparing Uchaguzi reports with reports generated through formal electoral monitoring missions, although only a subset of the analysis reports focused on physical violence and insecurity. The study found a high level of consistency between crowdsourced reports and the accounts of electoral observers, and suggested the utility of these systems for gathering detailed information on the conduct of the elections, although challenges included the need to have well-informed 'citizen reporters' in order to provide relevant and detailed accounts.

In spite of the growing number of initiatives using technology for monitoring violence and facilitating response, and an increasing number of studies analysing some dimensions of the resulting data, significant knowledge gaps remain. Chiefly, our understanding of how the nature of different reporting pathways influences the resulting data collected is still very limited; as is our understanding of the implications of this for research and policy responses that support or utilise these data.

These gaps have several implications for research and policy that undermine efforts to reduce violence through monitoring mechanisms. The main argument of this study is that, while new data sources on violence have the potential to considerably improve monitoring systems and resulting crisis response, they also differ in significant (and, to date, little understood) ways from other reporting systems. These differences, and the divergent information sets they produce, need to be assessed against the objectives and needs of researchers, policymakers and practitioners using these data. Below, we outline three key areas in which these gaps can potentially affect data usability and reliability: (1) geography, (2) temporality, and (3) targeting, in order to generate our research hypotheses.

2.1 Geography of violence

First, we consider the geography of violence reporting. Different sources of violence reporting can generate divergent pictures of the spatial patterns of violence, where it is concentrated, and where it spreads during a crisis. Geographic coverage at the subnational level has particular relevance in the case of Kenya's elections, as the process of devolution has created incentives for high-stakes electoral competition at the local-level (Lind 2018).

Systematic biases in the geography of traditional media coverage have been extensively documented. Geographic imbalances in coverage of humanitarian crises often reflect the economic and political power of urban elites over those in rural areas (Roberts and Marchais 2017). Media coverage may disproportionately focus on events that occur close to well-connected urban areas, where journalists tend to be posted, more so than in the rural periphery where networks are less developed (Davenport and Ball 2002; Wigmore-Shepherd 2015). As a result, the impact of crises on populations – and, in turn, policy and practitioner responses – may be shaped by these biases. For example, in the 2011 flooding in Thailand,

urban and more economically developed areas in the capital and central regions received more media coverage than in the poorer north-east, in spite of the latter experiencing higher levels of deaths and destruction as a result of the disaster (Chan 2017).

These same geographic biases may be even further exacerbated in cases of SMDTs, where digital technology 'layers' new geographical bias, relying as it does on a technical infrastructure that is unevenly distributed both internationally and within individual countries,' (Roberts and Marchais 2017: 13). Digital technology infrastructure is typically more developed in country capitals and urban centres than in rural peripheries, meaning regular and reliable internet access and telecommunications coverage are simply not available to many rural populations. Digital infrastructure inequalities map onto and often reinforce other existing social inequalities. For example, with high levels of poverty found in rural areas compared to economic centres, and a larger share of women in many low-income countries living in rural areas, further marginalising them from social media participation as both consumers and producers of content (Wilson and Gapsiso 2016; A4AI 2017; Roberts and Marchais 2017). These same biases can contribute to the precision of information being reported as well.

As a consequence, we hypothesise:

> *H1a: 'Old' media sources are more likely to capture violence in rural areas, and areas far from urban centres, than 'new' media sources.*

> *H1b: 'Old' media sources reflect the geographical distribution of violence more accurately than new media.*

2.2 Temporality of reporting

Next, we turn to the temporality of reporting, including coverage over time, timeliness, and temporal precision. Timeliness of reporting across 'old' media can be affected by a number of factors. First, the location of an event, with violence occurring in very rural areas potentially less likely to be reported immediately (or at all) due to delays in accessing and then transmitting relevant information. Second, the nature of violence, for instance, wherein high-profile actors, or violence considered particularly 'newsworthy', may be prioritised for timely reporting. Finally, the nature of reporting sources, with in-depth investigations typically requiring more time to unearth detailed accounts of insecurity than early-stage reporting (Chataing 2015).

By contrast, social media and crowdsourced reports often appeal to policymakers and practitioners because of the promise of producing near real-time accounts of violence. SMDT-generated data is generally considered more timely and immediate: in fact, the sheer pace at which real-time social media data are generated has been identified as a key challenge for processing and meaningfully analysing such large datasets (Andrews *et al.* 2016). However, it is important to consider whether this 'near real-time reporting' remains consistent over time, and how precise it is. This rapid pace of reporting can potentially enable timely action in the face of growing crises, though there is less evidence pointing to this rapid pace remaining sustained during non-crisis periods. There may also be reasons to anticipate delays in crowdsourced reports of violence. For example, crowdsourcing relies on the active participation of volunteers, who may be adversely affected by insecurity (for example, through displacement), limited access to communication infrastructure, and subsequent recall bias exacerbated in the face of trauma and stress.

To date, comparative studies have found inconsistent results concerning the extent to which the timeliness of events reported through SMDTs diverges from those generated through other monitoring systems. For example, in their comparison of crowdsourced data and non-social media-generated reports of violence in Syria, Price *et al.* (2013) and De Juan and

Bank (2015) both found a high degree of correlation on the dates and timing of reported events. However, De Juan and Bank (2015, Annexe 1: 3) and Masad (2013) both reported significant temporal discrepancies between violent events reported in the digital platform, Syria Tracker, and those captured in the Global Data on Events, Location and Tone (GDELT) data set, which is based on news reports from international news sources. According to De Juan and Bank (2015), this is most likely due to time lags between the reporting of violent events on crowdsourcing platforms and their reporting in international news outlets.

In light of these findings, we derive our next hypothesis:

> *H2a: 'Old' media sources are less likely to provide timely data in the immediate period surrounding insecurity.*

In considering temporality, we recognise that violence reporting entails more than the real-time information-sharing alone: it also concerns the comprehensiveness and consistency of reporting over an extended period. Pre- and post-election violence are separate conceptual categories with discrete conflict dynamics (Taylor, Pevehouse and Straus 2013). Pre-election violence is important to capture because violence is often deployed prior to elections to intimidate opposition or political supporters (Collier and Vicente 2012). Meanwhile, the post-crisis period can be accompanied by sporadic, but nevertheless significant, levels of violence long after the insecurity peaks (CIPEV 2008; Odhiambo 2017).

Social media platforms rely on user engagement to collect information on violence. Where such platforms are seen as ineffective, or exhaust users' willingness to report violence, these data may have a discrete temporal signature from violence reported in conventional media. Conflict and violence researchers know much more about conventional media's 'issue-attention cycle' (Jakobson 2000; Downs 1972) than about the nature and stamina of crowdsourcing efforts in reporting violence.

Moreover, the nature of SMDT oversight may also produce particular temporal signatures in violence reporting. Regime control over digital communication infrastructure may result in restrictions on the transmission of reports through SMDTs (Howard, Agarwal and Hussain 2011; Freedom House 2017). Blocking access to social media platforms has been a recurring feature of government crackdowns during periods of unrest in sub-Saharan Africa in recent years (Dahir 2016; CIPESA 2016), with examples such as Burundi in 2015, or Democratic Republic of Congo (DRC) in 2016, standing out (Roberts and Marchais 2017; Reuters 2016).

In the run-up to the 2017 elections in Kenya, there were concerns the government would institute an internet shutdown (Sang 2017). In past Kenyan elections, digital communication has also been restricted, in ways that may hinder reporting and monitoring of violence (IHRB 2013; Mutahi and Kimari 2017: 19). As in the case of Burundi, DRC and Kenya, communications crackdowns typically occur only after an increase in insecurity, suggesting a particular temporal profile (Robert and Marchais 2017).

As a result, we hypothesise:

> *H2b: 'New' media sources are less likely to provide more comprehensive account of violence over time.*
>
> *H2c: 'Old' media sources are more likely to capture the temporality of violence with higher precision than 'new' media sources.*

2.3 Targeting of violence

The last dimension on which reporting systems may differ in the accounts they capture concerns the nature and targeting of violence. Conventional media are affected by several forms of bias that may influence the types of violence and insecurity they report. 'Selection bias' (Earl *et al.* 2004) in the type and nature of insecurity that is reported can be driven by a tendency within news media to tailor coverage to events that will interest their audience and fit with their beliefs (Bocquier and Maupeau 2005), as well as documented examples of self-censorship among reporters under pressure from government (Campagna 1999). Triangulation of reports from multiple media sources, often with diverging editorial biases, can mitigate against the influence of individual outlets (Wigmore-Shepherd 2015), but the extent to which this affects reporting systems relying on media monitoring at an aggregate scale, is somewhat unclear.

SMDT system often promise to directly empower populations to document and disseminate information on violence, revealing patterns of insecurity which may be rendered invisible, or otherwise under-reported, in 'old' media (Weidmann 2015; Baum and Zhukov 2015). Examples include crowd-seeding systems for violence reporting in insecure contexts where journalist access is limited (Van der Windt and Humphreys 2016), or the use of social media in circumstances where more established media are banned from broadcasting (Fowler 2007; Mäkinen and Wangu Kuira 2008).

However, SMDTs alone do not entirely overcome the issue of selection bias. Where 'digital divides' and technology gaps in violence-affected contexts produce variable representation on social media platforms, users cannot be confident that these data truly reflect the underlying dynamics of violence, and the crisis response needs of a population. While existing research acknowledges the existence of this digital technology gap (Pew 2015; Blumenstock and Eagle 2012; May 2012), the precise nature and extent of these gaps and the resulting biases they produce in reports, remain under-specified.

Consequently, systematic biases within both 'old' and 'new' media reporting systems may result in the highlighting of certain forms of violence over others: this has implications for understanding the precise nature of violence as it unfolds, and the vulnerability of different population groups to that violence.

Traditional media may exhibit a preference for reporting on high-profile or high-intensity incidents of violence, based on their news value, or what is considered 'newsworthy' (Chataing 2015; Caple and Bednarek 2013). This phenomenon has been particularly extensively documented in relation to traditional media coverage of violence involving terror tactics (Asal and Hoffman 2015; Rohner and Frey 2007). However, newspaper coverage has also been shown to be related to event size (Barranco and Wisler 1999; McCarthy, McPhail and Smith 1996), intensity (Snyder and Kelly 1977; Earl *et al.* 2004; Pierskalla and Hollenbach 2013) and the likelihood of fatalities (Chojnacki *et al.* 2012).

'New' media reports, by contrast, may differ in terms of the type of events reported: while 'old' media may privilege reporting of high-intensity violence over other forms, or face greater challenges with respect to capturing and reporting small-scale violence such as riots and protests, or low-intensity attacks on civilians (Buhaug and Urdal 2013; Salehyan *et al.* 2012; Powell and Thyne 2011), diffuse social media users may have more information on lower-intensity events.

In these circumstances, high-intensity violence may be particularly difficult for individual (non-professional) reporters to capture through SMDT reporting systems: high-intensity violence may result in the destruction of telecommunications infrastructure and forced migration, as well as fear of reprisals, and pronounced physical and psychological distress, all of which may reduce the likelihood of volunteers reporting such incidents. *In extremis*, very high-

intensity violence may not be witnessed by volunteer reporters, as they may not remain *in situ* to observe it if violence is escalating considerably (De Juan and Bank 2015 Annexe 1: 1; Roberts and Marchais 2017).

In light of this, we hypothesise:

> *H3a: 'Old' media sources are more likely to capture high-intensity violence for which the media has a bias; than lower-intensity violence about which individuals may be more free to report.*

Socio-economic and political marginalisation may also play out in different ways in different reporting systems. Geographic discrepancies in coverage by traditional media, in particular the tendency to provide more coverage to urban areas, results in spatial biases in coverage that reflect existing spatial socio-economic inequalities. However, these biases are not necessarily corrected for in 'new' reporting systems. Technology gaps mean that social media usage is typically positively correlated to the level of human capital and technology available, and are not equally distributed across the country or segments of the population (Weidmann 2016). Crowdsourcing reports from a digitally-literate and networked population may be less likely to capture violence against marginalised groups, precisely because technology gaps map onto pre-existing social, economic and political exclusions. Through this, the voices of powerful groups may be amplified on social media, and those of more marginal groups, excluded (Lerman 2013; Perera 2015).

While these biases are by no means absent from traditional media coverage, they may be somewhat corrected for through the proliferation of media outlets, and less exaggerated than in those systems that rely on the volunteering of information from the digitally-literate and networked population alone. As such, we hypothesise:

> *H3b: 'New' media sources are less likely to capture violence against marginalised communities, owing to the overlap between poverty and digital technology gaps.*

3 Research design

3.1 Case selection

The paper takes as a case study the August and October 2017 elections in Kenya, for two reasons. First, the history and threat of violence in Kenya's elections; and second, the prevalence of social media and violence monitoring systems there.

3.1.1 Elections and electoral violence in Kenya

First, Kenya has experienced repeated cycles of violence surrounding elections over the past three decades, including in 1992, 1997, 2002, 2007/2008 and 2013 (Mutahi and Kimari 2017). Although there has been no repeat of the intensity of the 2007-2008 post-election violence to date, in the run-up to the August 2017 elections, there were worrying signs of tension and potentially violent mobilisation. Several observers predicted a high likelihood of political violence, characterising conditions around the country as a 'ticking time bomb' (Elder, Stigant and Claes 2014: 16; Claes 2016; Aling'o and Noor 2016; *Economist* 2017; *Africa Confidential* 2017), while others expressed concerns at the high, but sporadic, levels of violence that characterised some of the pre-election process, including the party primaries (Carter Center 2017). This risk of violence was particularly pronounced at the subnational level, owing to changes in the structure and distribution of power between Kenya's national and county governments following devolution (*see* Mutiga 2017; and Cheeseman, Lynch and Willis 2014 on the 2013 elections).

As the August 2017 election results were tallied, Presidential candidate and opposition leader, Raila Odinga, alleged that the Independent Electoral and Boundaries Commission (IEBC) system had been hacked, and that the preliminary results were fabricated to secure victory for the incumbent, President Uhuru Kenyatta (*Daily Nation* 2017b). The dispute plunged Kenya into a 'dangerous limbo' (Leithead 2017): rioting broke out in a number of areas, police clashed with demonstrators, and there were allegations of the use of live bullets, covert paramilitary violence, and excessive force, many of which were disputed by security forces (Burke 2017; Cherono 2017a, 2017b).

On 1 September 2017, the Kenyan Supreme Court invalidated the results and officially announced a re-run of the elections, citing irregularities and mismanagement of the elections. Jubilee party leaders and supporters accused the Supreme Court of being an 'opposition court' and issued threats to the judiciary and to individual judges (Worley 2017). Meanwhile, although the opposition initially welcomed the unprecedented decision, it soon lodged similar accusations of fraud in the run-up to the second elections, and eventually boycotted those, alleging that political manoeuvring by the incumbent president's party undermined their credibility. As the clock continued ticking and the opposition's demands in the form of '12 irreducible minimums' seemed to fall on deaf ears (IEBC 2017), the threat of violence in the repeat polls was very real.

The second election was held on the 26 October 2017 and led to Kenyatta's victory. While the opposition contested the credibility and legitimacy of the elections, the Kenyan Supreme Court upheld the results on 20 November 2017. Albeit on a smaller scale than in August, the October election also witnessed demonstrations and repression by state and paramilitary forces, which continued sporadically until Kenyatta's swearing in on November 28 (*Daily Nation* 2017a). A very tense period followed as Odinga refused to recognise the legitimacy of the new president, going so far as declaring himself president on 1 February 2018. While he eventually recognised the results and declared his support for President Kenyatta, the episode may have a lasting impact on the legitimacy of the electoral system in Kenya (Wadekar 2018; Gathigi 2018; Kegoro 2018).

3.1.2 'Old' and 'new' media in the Kenyan elections

Second, Kenya's digital landscape and the prevalence of violence monitoring systems there make it an interesting and valuable case study. Kenya has one of the highest levels of telecommunications and internet infrastructure development on the African continent. In 2017, there were 39.1 million registered mobile phone users in Kenya, and an estimated 40.5 million internet users, out of an estimated population of 44 million (CAK 2017: 8). Social media use is also extremely high: there are an estimated 2.2 million monthly active Twitter users in Kenya, in addition to over 10 million WhatsApp users (BAKE 2016: 3).

This is accompanied by a strong culture of social and political activism on social media. Historically, there is a strong record of SMDTs being used to report and monitor violence, including the Ushahidi platform, which was used to monitor the 2007/2008 election violence (Meier 2008; Mäkinen and Wangu Kuira 2008) and subsequent elections (Ushahidi 2010, 2013). This is also reflected in the emergence of a distinct community of 'Kenyans on Twitter' (#KOT) that have engaged in social media-based campaigns for political accountability and social action (Mutahi and Kimari 2017). The hashtags #ElectionsKE and #ElectionsKE2017 were used extensively to flag social media content related to the elections in the run-up to and during the election period (N'gang'a 2017).

Kenya also has a long history of an active press: the country is home to at least five daily newspapers, and several regional weekly newspapers, and international news media are both widely available and actively covered the elections (Freedom House 2016). In spite of recent efforts by the government to restrict press freedoms through the use of counterterrorism legislation – resulting in arrests, harassment and increased self-censorship

(Namwaya 2016; Kariuki 2016) – national media continues to regularly publish reports of violence, and media professionals see national media as playing an important role in shaping the public's perceptions of insecurity throughout the country.[5]

This combination of (1) high levels of digital penetration and activity, (2) a strong record of violence monitoring systems and response, and (3) an active press that reports on political violence, facilitates a study of the comparative benefits of different monitoring systems in a 'best case' scenario – one in which we can assume that both access to, and awareness of, these digital systems, is relatively high, and the media is relatively active. If the findings suggest that even in these circumstances, different reporting pathways exhibit comparative strengths or weaknesses, then this indicates that those same limitations may be even more pronounced in contexts where both the digital/social, and traditional, media landscape is less sophisticated.

3.2 Data collection

We collected data from two sources of violence monitoring for comparative analysis: 'old' and 'new' media sources. Further details of each are outlined below.

3.2.1 'Traditional' or 'old' media sources

Media monitoring involves researchers reviewing the contents of published media reports to identify and distil critical information on incidents. In this study, we refer to 'traditional' or 'old' media monitoring in contrast to 'new' or 'social' media monitoring, to distinguish between monitoring that draws on mainstream media such as newspapers, newswires, published reports and radio transcripts. These reports are drawn from the ACLED dataset (Raleigh *et al.* 2010).

ACLED is a crisis monitoring dataset that contains information on the date, location, actors and types of political violence and protest events across a number of countries. Data covering Kenya span from 1997 to the present, with on-going data collection published weekly. The unit of analysis is the individual event, disaggregated by date (day), event type, actor(s) involved, and location (region, country, administrative units, and geographic coordinates), with additional details on the temporal and geographic precision of each event, associated reported fatalities, and sourcing and its scale. This unit of analysis is constant across all years of the dataset, and facilitates meaningful comparative analysis with incident reports in other datasets over time and across territories. ACLED data have been used directly in several studies of conflict including electoral violence in Kenya (Linke 2013) and Africa more widely (Goldsmith 2015); urban violence in sub-Saharan Africa (Raleigh 2015); and resource politics and violent conflict in Kenya (Lind 2018).

ACLED relies on published media reports for information, drawn from international, regional and more local media (Wigmore-Shepherd 2015).[6] Media reports are analysed by a researcher and then coded according to disaggregated criteria. For the purposes of this study, the coding team expanded the metadata they usually record to include additional details, such as the date of reporting, to facilitate the comparative analysis for this study. This produced a set of events that were checked and cleaned for errors after going through an automated review tool to check for intra-coder reliability. Thereafter, the coding is checked by a second and third researcher for inter-coder reliability.[7]

---

5   Key informant interview with former journalist, Nairobi, 3 July 2017.

6   ACLED sourcing also includes reports from non-governmental organisations (NGOs) and international organisations (e.g. Amnesty International, Human Rights Watch, and UN) as well as through partnerships with local conflict observatories (though as of this writing, ACLED does not have a partnership with a local conflict observatory in the Kenyan context).

7   This three-step review process is the typical process used by ACLED to clean and review all data coded by researchers.

3.2.2 'New' media sources (e.g. crowdsourcing)

Crowdsourcing can involve both passive and active crowdsourcing. Passive crowdsourcing refers to the collection of data from a pool of information without a specific 'call to action,' (Roberts and Marchais 2017). In this study, we collect reports of insecurity in Kenya from public posts on Twitter through an automated search strategy utilising Method52. Below, we provide an overview of that data collection process, and a fuller discussion is contained in the Annex.

Method52 is a technology platform developed by the Informatics Department of the University of Sussex for collecting, processing, and analysing a large number of documents or texts, including social media data (Wibberley *et al.* 2014). The platform is based on components customisable by the user, each one with a different function. These components can be organised in pipelines so as to produce a tailor-made system for data collection and analysis. The components of the pipeline are able to perform standard operations, such as collecting data from social media platforms (e.g. Twitter and Facebook).

Despite its recent development (Wibberley *et al.* 2014), Method52 has been extensively employed by research organisations and cross-party think-tanks for monitoring social media during specific timeframes, such as during the 2015 Nigerian elections (Bartlett *et al.* 2015); or to assess EU sentiment during the Eurozone crisis of 2013 (Wibberley *et al.* 2014). The use of Method52 to automate Twitter monitoring has proved to be effective during electoral periods, when the number of tweets increases sharply and the amount of information is so large making other strategies of manual monitoring more difficult (Bartlett *et al.* 2015).

This study draws on the scarce but growing literature on crawling[8] tweets for event data and, more generally, on the literature around the use of social media during crises, which has grown in the last seven years (Imran *et al.* 2015). For example, Sambuli *et al.* (2013) use Datasift to focus on Kenya's 2013 elections to study differences in conflict events reported by Twitter and traditional media sources, and conclude that "Mining Kenyan Twitter data during an election scenario looks to be a very valuable and worthwhile technique when looking for timely, local information." (Sambuli *et al.* 2013: 47). Datasift and Method52 are similar, as they both belong to the category of search / streaming tools and are able to scan Twitter for all tweets published in the previous nine days. In our study, we extend the timeframe of the Sambuli *et al.* (2013) study, from two weeks to more than eight months of observations and focus on several other aspects of reporting. The use of Method52 is particularly suitable for this extended timeframe as it enables the searching and storing of information collected, as well as the creation of an online backup of data.

Moreno *et al.* (2017) provide another example of crowdsourcing data from social media platforms for violence monitoring during election, using as case study Ghana's 2016 election. The authors use Aggie, which is an open source tool that can track real-time events such as election violence from social media data. Similar to Method52, Aggie enables users to search, explore and filter millions of posts published on Facebook and Twitter, but also on WhatsApp, and as RSS. The appropriateness of Twitter for monitoring incidence data is confirmed by the fact that out of 584 incidents identified by the authors, more than 91.6 per cent are captured through Twitter.

This study limits its usage of Method52 to crawling tweets containing keywords related to political conflict during the Kenyan 2017 elections.[9] Table A4.1 in the Annexes lists the specific search terms used. The final set includes a total of 113 search terms, a set

---

[8]   'Crawling' refers to the scanning and collecting of tweets from Twitter.

[9]   The period of collection spans from March 2017, four months before the first 2017 election, through November 2017, one month after the second Kenyan election.

qualitatively and quantitatively comparable to the 98 terms employed by Sambuli *et al.* (2013) and those employed in other monitoring datasets (Raleigh *et al.* 2010).

After the automated search produced a list of tweets and associated meta-data, a research team member coded the resulting set of tweets into unique events according to coding criteria to enable comparative analysis. It is important to note that not all tweets containing search terms necessarily yielded 'relevant', or 'enough', information to be coded (i.e. 'false positives').

To meet the *relevant* criteria, tweets must include reports of political violence or protests, or information around strategic developments that would impact future conflict or protest dynamics. To meet the criteria for *enough* information, there must be information around 'what happened' (e.g. Was it a demonstration? Did armed groups engage?), where the event occurred (at minimum, what state did the event occur in, which could be coded with a lower precision rating), and who was involved (at minimum, were they armed or peaceful?). The date of the event is important as well, though in cases where an exact date was not noted, the date of the tweet itself was used instead (coded with a lower precision rating). This has a number of implications. For one, it means that analysis here is not around the sheer number of tweets, but rather around the number of relevant/sufficient tweets. This also means that tweets containing a link to an outside media report tended to have more information, by definition, and hence were more likely to be recorded; tweets without a link to an outside media report could only be recorded if they contained sufficient information to be coded.

From the coding, the set of events that were produced were then checked and cleaned for errors and inter-coder reliability by a second research team member, before being reviewed for inter-code reliability by a third research team member.

3.3 Description of datasets

The analysis is carried out on all events recorded by multiple monitoring systems surrounding Kenya's August and October 2017 elections, and specifically between 13 March 2017 and 30 November 2017.

In total, **852** discrete events were recorded. The total number of events reported in the ACLED dataset is **723,** while the total number of events reported the Twitter dataset is **412**. ACLED reported **440** unique events (events that did no feature in Twitter either because they were not reported on at all, or they were not reported on in sufficient detail to code consistently); while Twitter reported **129** unique events (using the same criteria as above). Finally, **283** events were recorded in both datasets (matched events). This can be seen on top left of Figure 3.1, below: ACLED's unique events are shown in orange; while Twitter's unique events are shown in blue. The teal events (matched events) were reported in both datasets.

Events across the two data sources were matched by hand. To do this, the narrative description ('notes') of each event in one dataset were read and matched to a corresponding event in the other dataset. Events were matched based on fuzzy temporal and geographic boundaries. This means that an event would be 'matched' to a similar event if it was reported by the other data source within a week and/or within the same province (administrative region) of the original event. This 'fuzziness' allowed for matching any events that were seemingly reporting the same violence, though with less granular temporal or geographic precision. Matched events need not necessarily match on all other criteria, as reporting of the details of events may vary across sources. For example, sometimes a peaceful protest coded in one dataset may be coded as violence against civilians in the other dataset if reports on which the latter was based gave reason to believe the peaceful protesters were met with excessive force; or, the actor in one dataset may be an 'unidentified armed group' while in the other it is noted as 'Al Shabaab'.

Figure 3.1 Understanding the information coming from different media sources
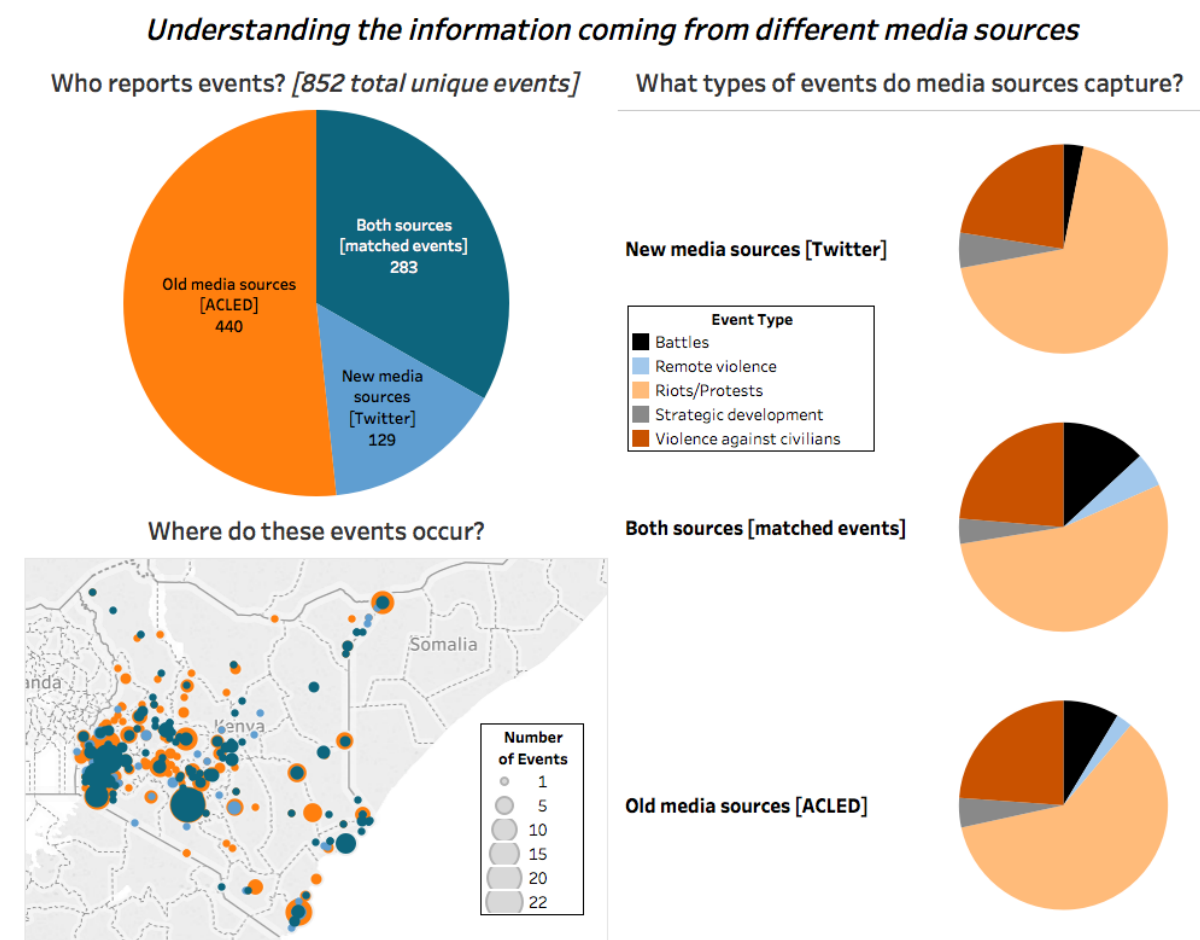
Source: Authors' own.

The following analysis uses three 'datasets': ACLED data to represent traditional or 'old' media reports, Twitter data to represent 'new' media, and the 'matched' events datasets.

4   Results

This section presents of the comparative analysis of 'old' and 'new' media reports, using the three datasets detailed above. We compare the datasets along the three dimensions of geography, temporality, and targeting/representativeness. It is important to recall, however, that neither 'old' nor 'new' media, nor the subset of reports contained in both, can present the true or full picture, as both sources carry a range of biases and limitations.

4.1 General descriptive

Figure 4.1 below depicts that both the August and October elections provoked peaks in events across all categories of events (battles, remote violence, violence against civilians, riots/protests, and strategic developments).[10]

---

[10]   These events types mirror event types used by the Armed Conflict Location & Event Data Project (ACLED) as of the writing of this piece. Event types are defined as follows:

A **battle** is an event "between two violent armed groups where control of the contested location does not change."

Figure 4.1 Political violence, protest and strategic development events during the 2017 election period

Political violence, protest, & strategic development events during the 2017 election period, including data from both old and new media sources
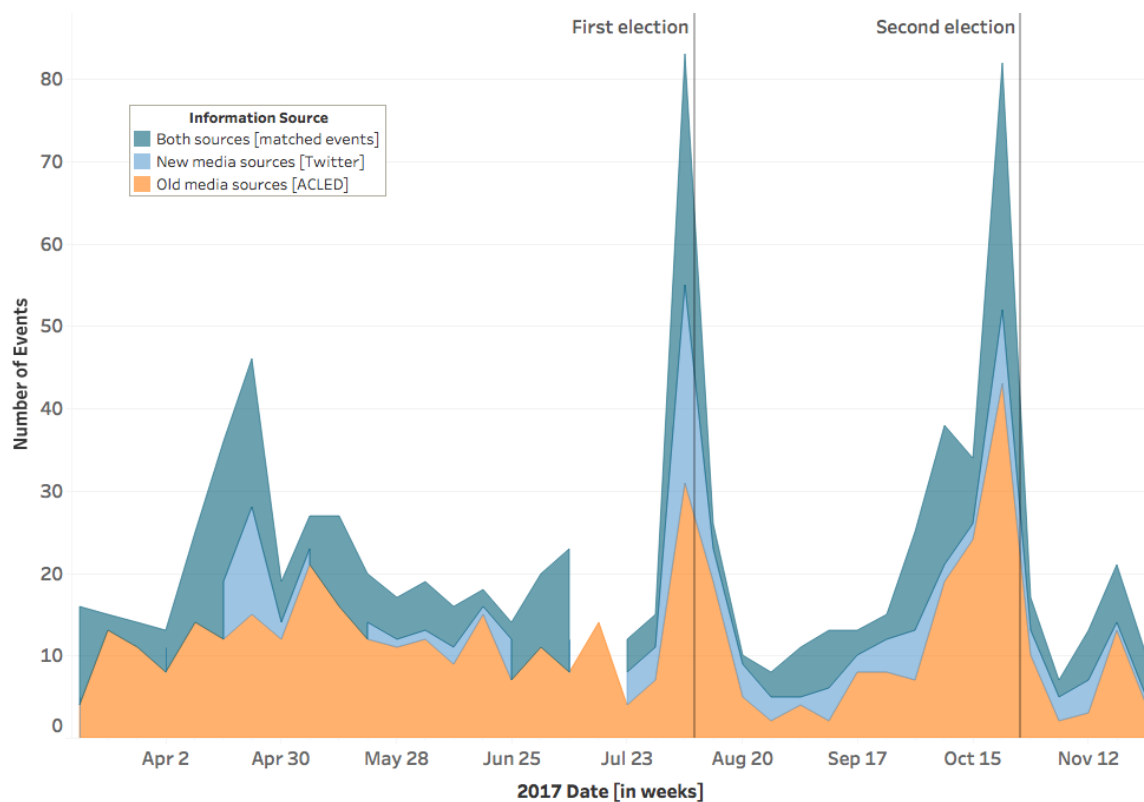
Source: Authors' own.

However, as can be seen in Figure 4.2 below, 'old' and 'new' media paint different pictures of the violence that occurred during and around the 2017 elections in Kenya.

The two graphs on the right-hand side of the infographic below show that the intensity of these peaks of violence varies according to whether 'old' or 'new' media sources are used. For reports aggregated through ACLED ('old' media) in the graph on the top right, the October election triggered a higher number of events relative to the August election, while the Twitter data in the graph on the bottom right suggest that more events occurred around

---

Remote violence refers to "events where engaging in conflict did not require the physical presence of the perpetrator. The main characteristic of this event is when a group determines the time, place, and victims of the attack, but is not directly present. These include bombings, IED attacks, mortar and missile attacks, etc. Remote violence can be waged on both armed agents (e.g. an active rebel group; a military garrison) and civilians (e.g. a roadside bombing)."

Violence against civilians is a violent act upon civilians by an armed, organised, and violent group. By definition, civilians are unarmed and not engaged in political violence." Any armed or violent group can inflict violence onto civilians. "Protesters are also civilians, and significant violence against protesters falls under this category."

Riots/Protests capture demonstrations. "A protest is a public demonstration in which the participants do not engage in violence, though violence may be used against them. Often – though not always – protests are against a government institution. Rioting is a violent form of demonstration where the participants engage in violent acts, including but not limited to rock throwing, property destruction, etc. Both of these can be coded as one-sided events. All rioters and protesters are noted by generic terms (e.g. "Rioters (Country)" or "Protesters (Country)"); if representing a group, the name of that group is recorded in the respective "ally actor" column."

Strategic developments provide "contextually important information regarding the activities of violent groups that is not itself political violence…. The inclusion of such events is limited, as its purpose is to capture pivotal events within campaigns of political violence."

the August election than around the October election. This demonstrates that 'old' and 'new' media sources present different and sometimes contrasting stories.

Figure 4.2 A tale of two elections
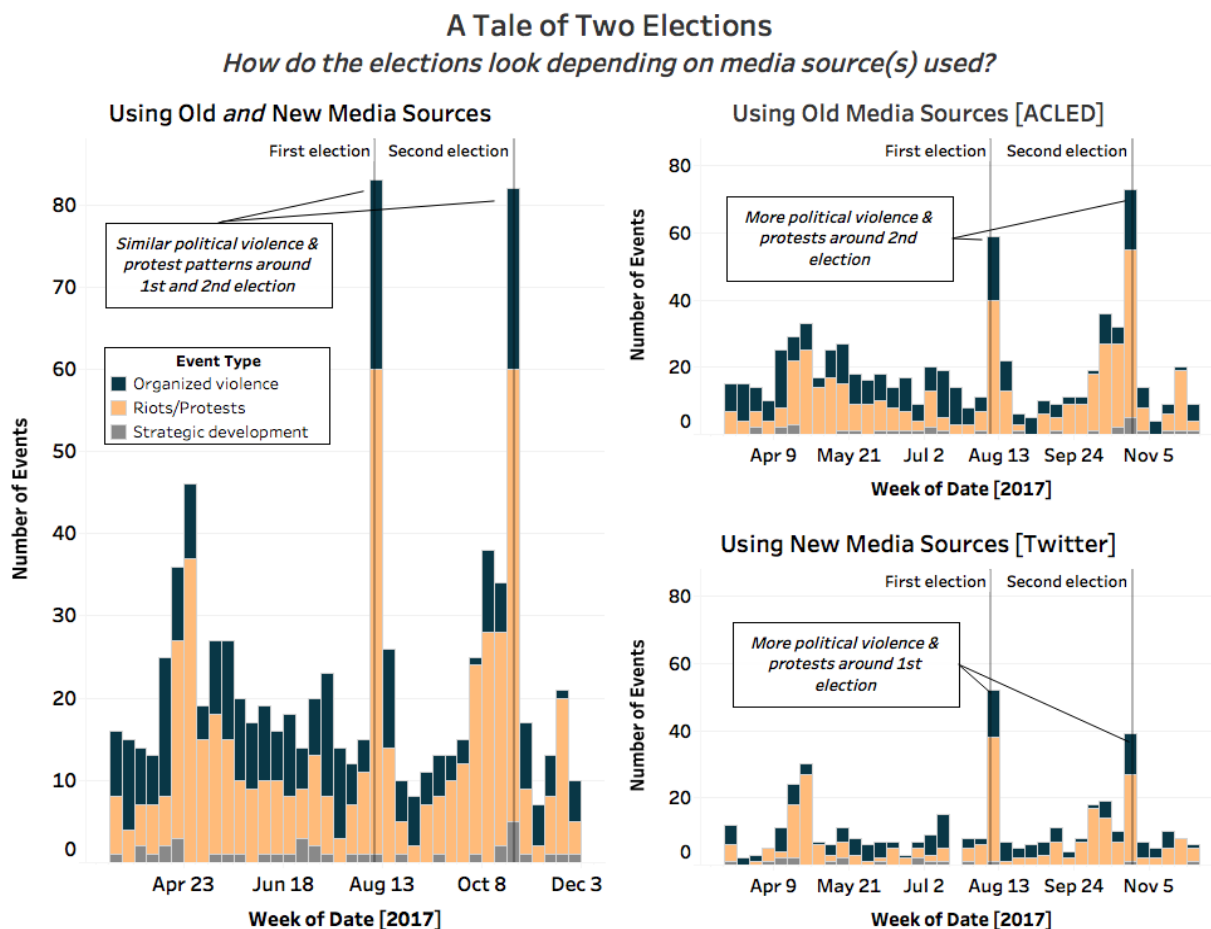
Source: Authors' own.

Figure 4.2 includes all 'matched' events (those events reported by both sources) as well as all unique events reported by both ACLED and Twitter. The infographic illustrates the fact that, rather than being exclusive or contradictory, 'old' and 'new' media sources should be considered as *complementary*, as both bring to bear unique and valuable information.

However, trends in unique reports from both sources vary over time. In the run up to the election – from March to July 2017 – ACLED ('old' media) reports significantly higher numbers of unique events than Twitter ('new' media). Given that the nature and trends in violence before the election is highly differentiated, this suggests that 'old' media sources provide more information on violence that is not related to a 'high profile' event such as the elections. This is explored in more detail below in the discussion around temporality. Also interesting is that while there is a sharp increase in the number of events coded from Twitter around the first election, this increase is much smaller around the second election – despite the fact that ACLED provides *more* unique events around the second than the first election. This may be the result of the underlying behaviour of Twitter users. Qualitative evidence collected during a Nairobi Stakeholder workshop held in January 2018 suggests that the second election captured less attention by the broader public, as a result of a form of 'election fatigue'. This was the result of a prolonged electoral cycle, and a less competitive one following the opposition boycott in October 2017. The presence of only one contender removed the impetus for continued online debate, and left a resigned electorate focusing on

their individual lives and on recovering from the hit on the economy generated by the unrest around the elections.[11] These results suggest that, while 'new' media sources can and do provide valuable information for critical episodes, this might be conditional on the existing circumstances and the attitude and behaviour of users, particularly their level of attention, and their level of fatigue, with regard to the events of interest.

4.1.1 Geography: urban and rural divides, economic geography and geographic precision

We then analyse the geographical coverage of 'new' and 'old' media sources. What regions seem more violent varies depending on which source is relied on. The maps in Figure 4.3 present where reported events occurred; the maps in Figure 4.4 present where reported fatalities occurred. The maps on the left of each figure are based on information collected by ACLED ('old' media) while those on right of each pair come from Twitter ('new' media). The spatial distribution of events suggests that Twitter is more effective at reporting violent events that occur in densely populated areas (such as urban areas), while ACLED has wider geographic coverage. Table 4.1 (below) confirms this: on average, the population density of areas of events ACLED reports is 743.4 people per square km; while it is 851.6 for Twitter events; and 770.4 for those events both datasets captured (matched events).

Figure 4.3 How does the conflict landscape vary by media source?

Source: Authors' own.

---
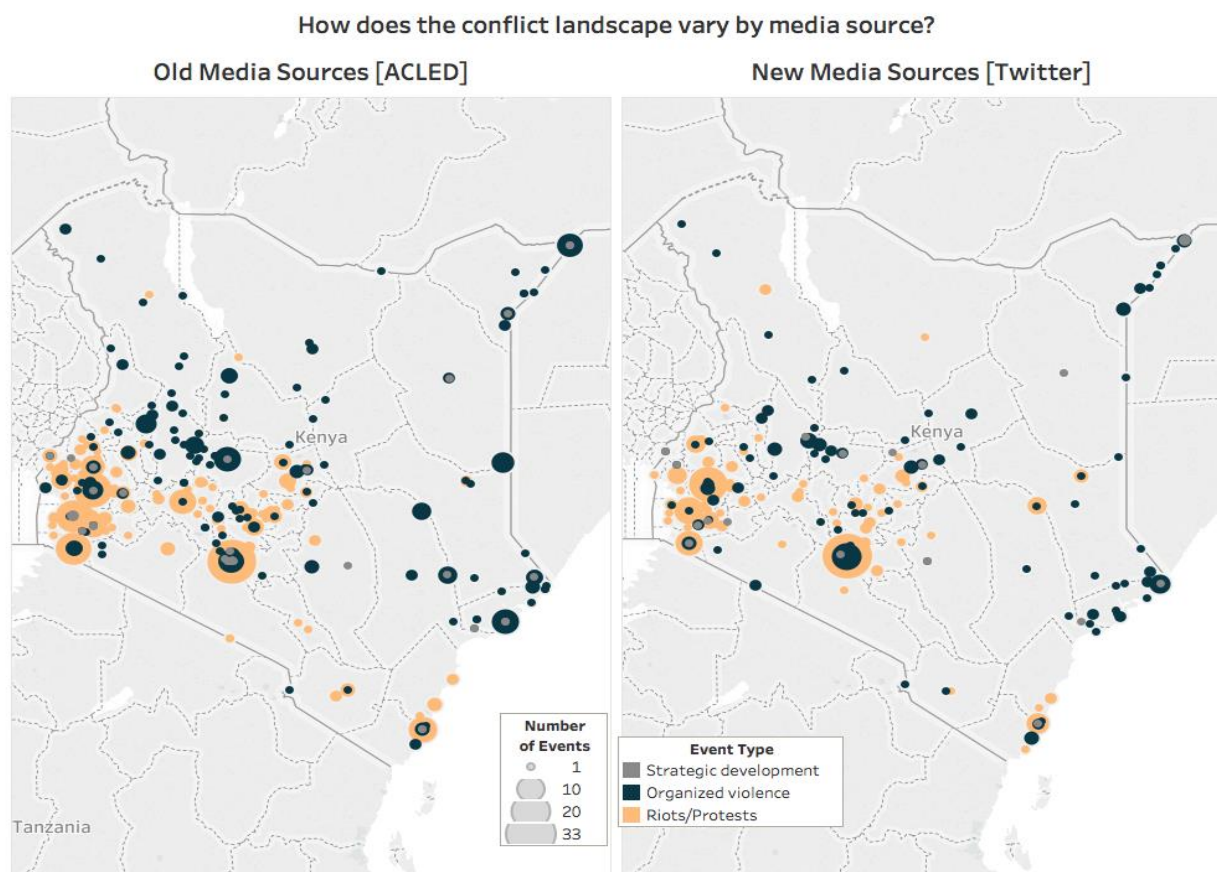
[11]   Interviews with Centre for Human Rights and Policy Studies, Nairobi (29 January 2018) and London (24 May 2018).

Figure 4.4 How does the conflict landscape vary by media source?

How does the conflict landscape (based on reported fatalities) vary by media source?

Old Media Sources [ACLED]   New Media Sources [Twitter]

Reported Fatalities: 0, 10, 20, 30

Event Type
Riots/Protests
Organized violence
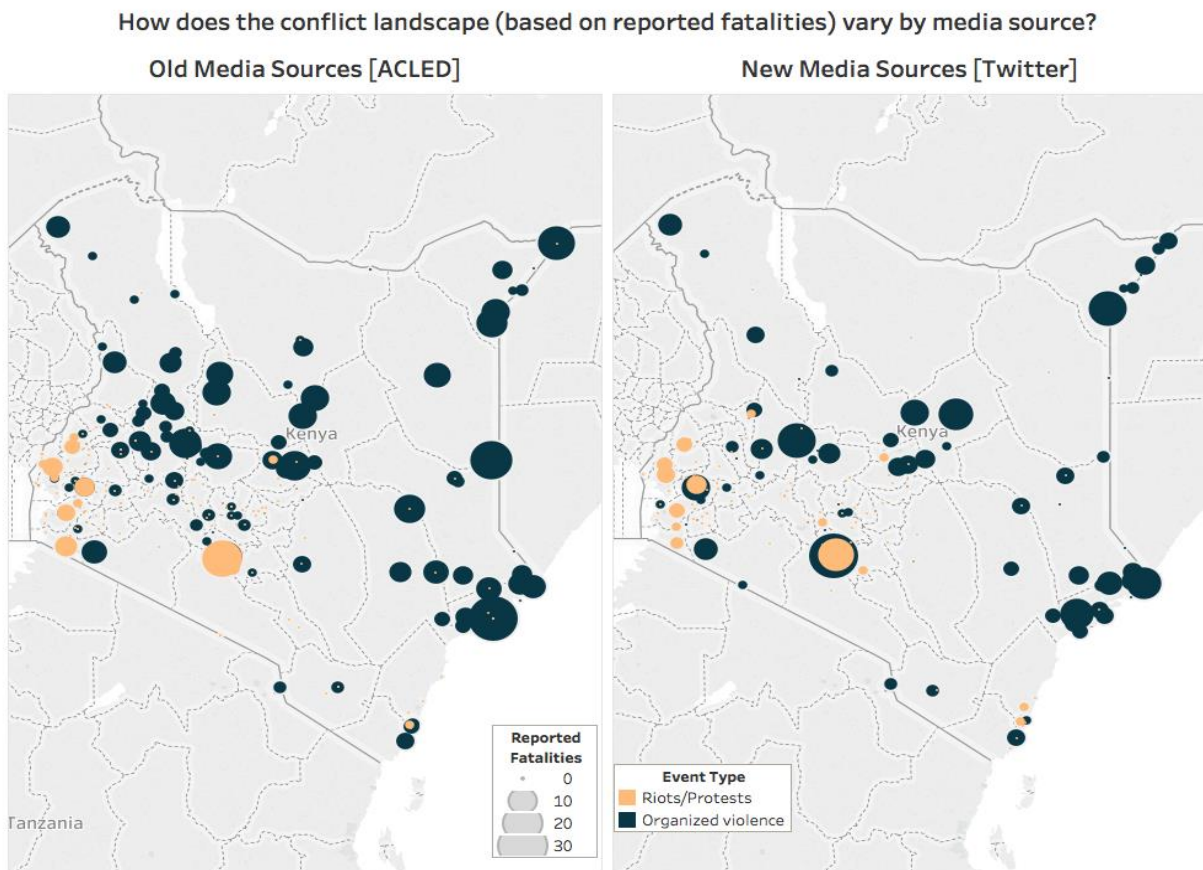
Source: Authors' own.

Again, when assessing geographic scope, reports from ACLED cover a wider area than those from Twitter. Specifically, ACLED reports events from 279 uniquely named locations; Twitter reports events from 137 uniquely named locations – or about half as many unique locations as ACLED. There are no locations that events coded from Twitter capture that events coded from ACLED fail to also cover. There are, however, a number of locations that events coded from ACLED capture in which those from Twitter fail to. In Figure 4.5 below, the map on the left depicts all events that ACLED reports that Twitter also captures; the map on the right depicts all events that ACLED reports that Twitter misses. The wider array of locations covered in this latter map point to the larger geographic range of ACLED ('old' media).

Figure 4.5 Locations of events captured by 'old' media sources

Locations of Events Captured by Old Media Sources [ACLED]

Where do events occur that are unique to new media sources?
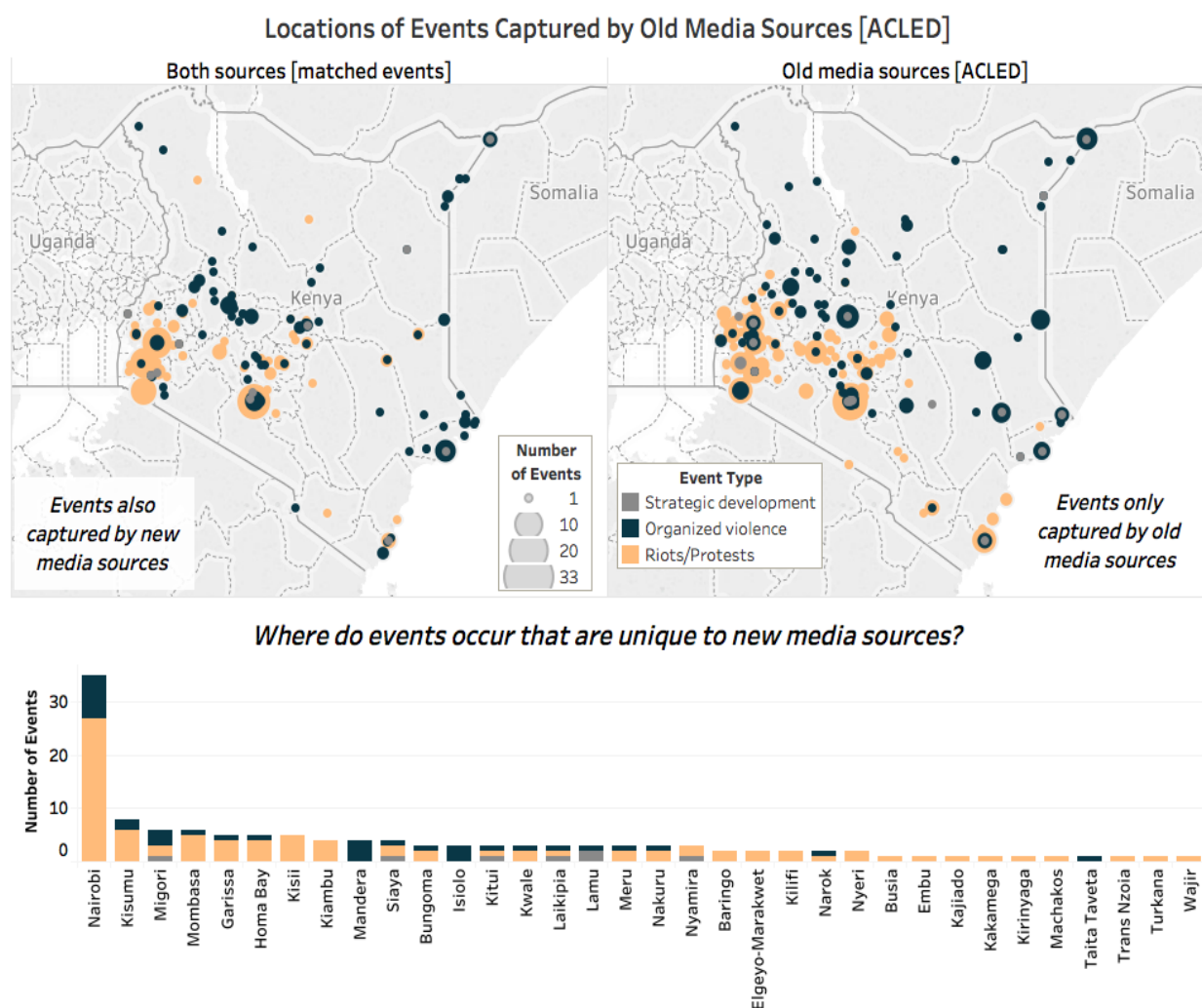
Source: Authors' own.

When looking at the locations of events that Twitter capture that ACLED does not (on the graph at the bottom of the Figure 4.5 infographic), it is clear that the vast majority of these events occurred in Nairobi. This fits with the trend that Twitter ('new' media) is more effective in dense urban areas where Twitter users are largely based. As an illustration of this, over 22 per cent of all events reported by Twitter occurred in the greater Nairobi region; about 17 per cent of all events reported by ACLED occurred in this same area.

This is in line with Hypothesis 1a: ACLED ('old' media) is more likely to capture violence in rural areas, and areas far from urban centres, than Twitter ('new' media).

Beyond the breadth of physical coverage alone, it is also important to take into account the underlying economic geography of the conflict landscape being reported, as this can impact where crisis and emergency response is needed spatially.

Table 4.1 depicts the economic makeup of the areas where events are reported. The table suggests that Twitter tends to report more events in areas with a higher GDP per capita, a higher population density, and higher level of development (as proxied by the nightlights index). In real terms, on average, residents in areas where Twitter reported violence make over US$100.00 more than those in areas where events were reported by 'old' media. This is in line with our understanding of where Twitter users tend to be based. This is also in line

with Hypothesis 3b: Twitter is less likely to capture violence against marginalised communities, owing to the overlap between poverty and digital technology gaps.

This finding is particularly significant in light of the (sometimes implicit) claims that social media contributes to a democratisation of the reporting and information landscape: while that may be true in some cases, evidence that, on the whole, more economically developed areas are better covered by social media reporting suggests that significant biases and technology gaps remain.

Table 4.1 The economic makeup of the areas where violent events are reported

	'Old' Media Reports (ACLED)	'New' Media Reports (Twitter)	Common Events (Matched)
GDP per capita (US dollars PPP)	1446	1548	1469
Population density (pop per km^2)	743.4	851.6	770.4
Nightlights index (0-63)	10.14	11.25	10.49

Source: Authors' own, based on GDP data from Nordhaus (2006): http://sedac.ciesin.columbia.edu/data/set/spatialecon-gecon-v4; 2000 population density data from the WorldPop Project: www.worldpop.org.uk/; and 2013 Average Visible Stable light nightlight data from the Earth Observatory Group: https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html.

Different data sources are able to offer varied levels of accuracy in understanding where events occurred. This geographic precision is a crucial determinant of both the actionability and usefulness of different media sources on crisis contexts. Figure 4.6 presents information around the 'geographic precision' of reported events. This depicts the level of specificity in reporting the location. For example, if a report says that an event occurred "in the Westlands neighbourhood of Nairobi", this reflects higher level of specificity than if a report says that an event occurred "somewhere near Mombasa." Differences in precision are coded on a deceasing scale of precision, from 1 (highly precise, usually a specific neighbourhood or area) to 3 (reference to a sub-national unit like a county).

The top graph depicts the geographic precision of events reported by ACLED, while the bottom graph depicts the same for events from Twitter. To allow for more accurate comparison, the events that are compared here are those events which both sources report, 'matched' events. The shorter the bar of the graph, the less fuzziness there is (or the more accurate the location reported is). In general, ACLED reports event location with more precision (i.e. on average shorter bars) than Twitter. On average, events reported by 'old' media had a geographic precision score of 1.2 out of 3; while those on 'new' media have a score of 1.4 out of 3. [12]

---

[12]   It is important to note that the geographic precision here is in regard to *reporting* and not in regard to the "truth". The precision denotes how precise the reporter (whether media or Twitter) is in their reporting with the assumption that the reporter has not misconstrued this information. Events were not specifically corroborated on the ground by authors.

Figure 4.6 How does reporting precision of the event location vary by media source?

How does reporting precision of the event location vary by media source?
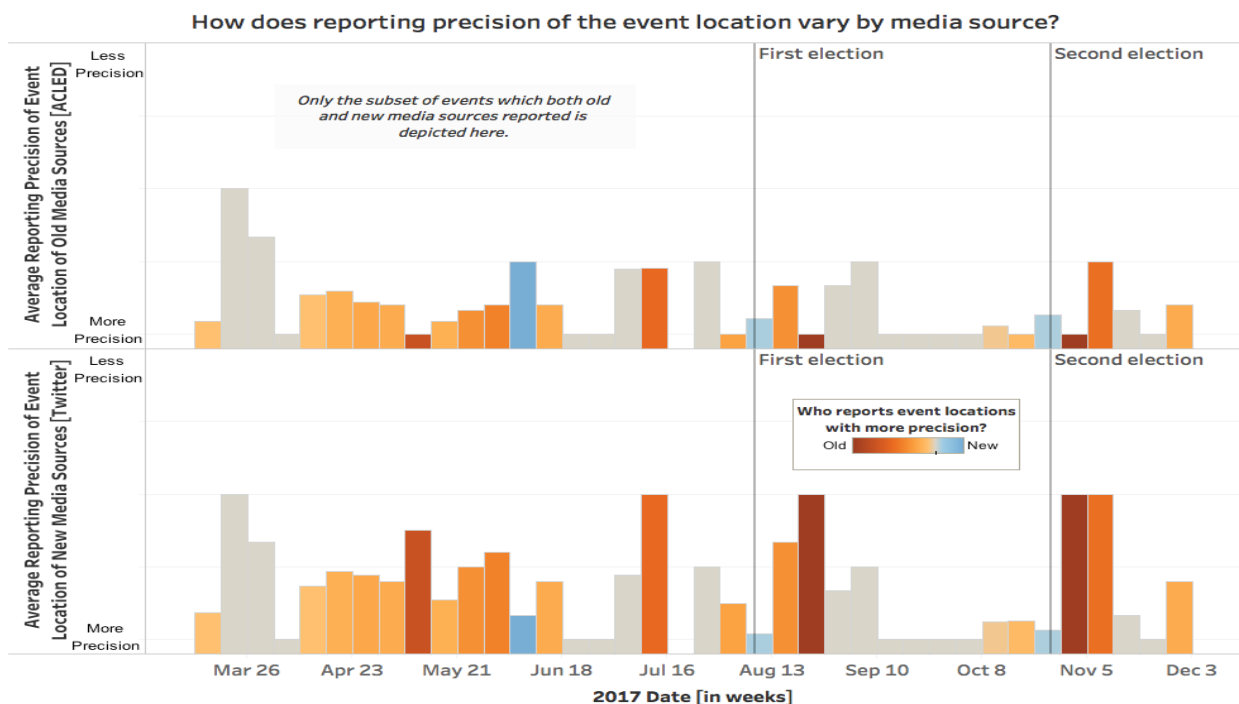
Source: Authors' own.

The shading of the graph aids in further comparison, and points once more to variability over time and the greater degree of precision and detailed captured by 'new' media in the period immediately surrounding critical junctures. In weeks where the colouring across the two graphs is more blue, Twitter reported with higher geographic precision, while weeks where the colouring is more orange, ACLED had higher precision. There is more blue around the first election, where events coded from Twitter report locations with higher precision, while outside of the immediate election period there is more orange, where events coded from ACLED report locations with higher precision.

This is in line with Hypothesis 1b: ACLED ('old' media) is more likely to capture the geography of violence with higher precision than Twitter ('new' media).

4.1.2 Temporality: coverage, timeliness and temporal precision

Next, we examine the temporality of 'new' and 'old' media sources, specifically the timeliness of reporting, the coverage of events over time, and the temporal precision.

It is important to measure how long it takes different media sources to report on events in order to determine the timeliness of reporting. We refer to this as the reporting lag, and measure it in units of days between when the event is reported to have occurred, and when the actual report of that event surfaced. For example, a report stating that an incident of violence occurred earlier that same day, would have a reporting lag of 0; while one citing violence the preceding day would have reporting lag of 1. We compare the timeliness of ACLED ('old' media) to Twitter ('new' media) by comparing the reporting lag between the date that an event occurs to the date that the event is reported.

As in the preceding section, we limit the comparative analysis in Figure 4.7 to the subset of matched events, to enable comparison of the two sources for the same event. It is also important to note that, for the same reason, only those events with the highest time precision (a score of 1 out of 3) are examined here. If a source did not specifically report the date of an

event, the default practice was to record the report with a fuzzier time precision. While this allows us to get *close* to the 'true' date of the event, it means that we would not be able to meaningfully test the timeliness of the reporting of those events with lower precision scores.[13]

The top graph depicts the reporting lag of all events reported by ACLED, while the bottom graph depicts the same for events from Twitter. The shorter the bar, the shorter the reporting lag (or the quicker the event is reported). In general terms, the two sources have similar reporting lags – neither is considerably quicker than the other over all. On average, across the matched dataset, the reporting lag for ACLED is approximately 1.83 days, compared to 2.21 days for Twitter.

This varies significantly, however, by time period. In the weeks immediately around the elections, ACLED has a reporting lag of about 4.37 days on average. This is a result of also drawing on more in-depth coverage (such as reports from organisations like Human Rights Watch) which often do not come out until much later, when detailed investigations have been completed, thus driving up the average reporting lag in that time. By contrast, the average reporting lag for Twitter events during this time period is 1.15 days. However, outside of the immediate election period, the average reporting lag for ACLED is 1.79 days, while for those same events the lag for Twitter increases to approximately 2.87 days.

Figure 4.7 How does reporting lag vary by media source?
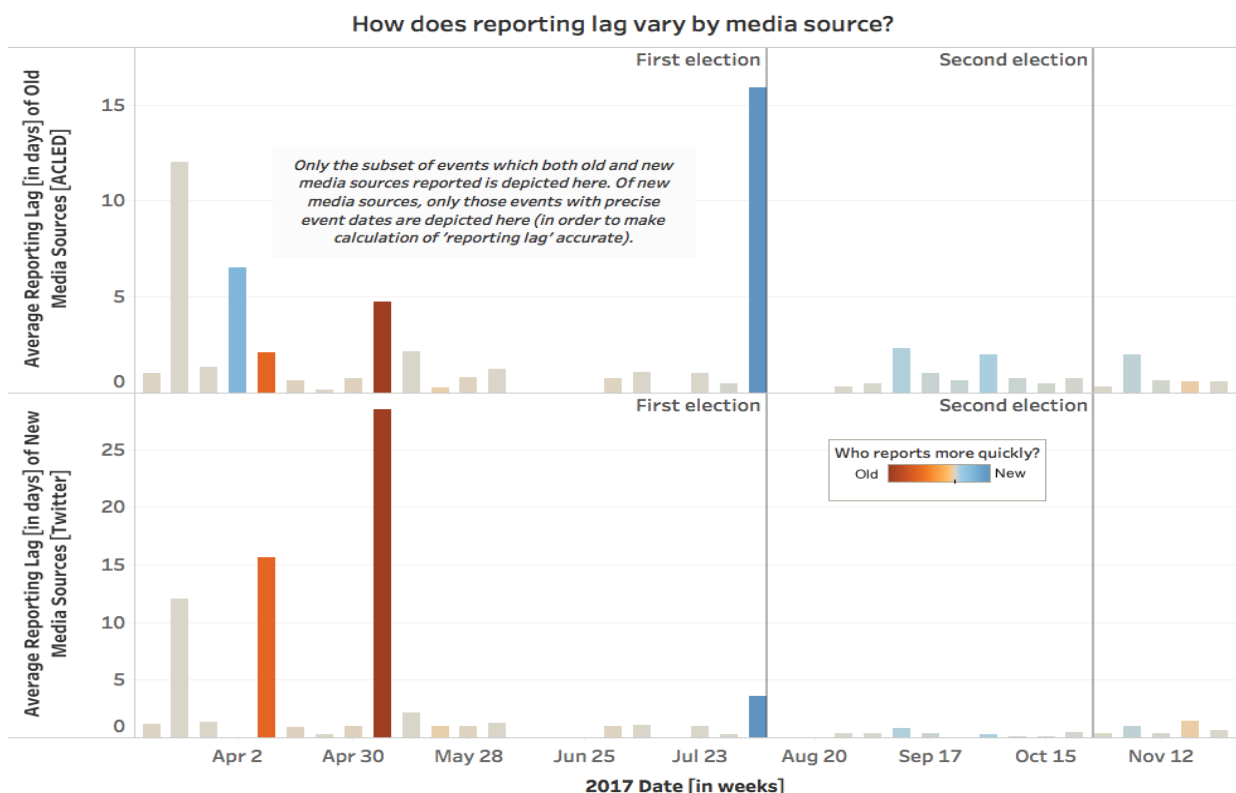
Source: Authors' own.

In weeks where the colouring across the two graphs is more blue, Twitter reported events more quickly, while in weeks where the colouring is more orange, ACLED reported events more quickly. There is more blue, for example, around the first election, where Twitter reports events more quickly – and considerably so (the bar on the bottom graph is considerably

---

[13]   It is important to remember is that neither traditional/old nor new media sources necessarily report the 'true' date and time of an event.

shorter than the one on the top graph). This is in line with Hypothesis 2a: ACLED ('old' media) is less likely to provide timely data in the immediate period surrounding insecurity.

In other words, not only does Twitter provide relatively more unique reports around the first election than during non-election periods and around the second election, but these reports may also be timelier. However, outside of the immediate election period there is more orange, where ACLED reports more quickly – and again, considerably so (the bars on the bottom graph are considerably larger than those on the top graph during these periods). This provides further evidence of the relative contribution of 'new' media being conditional on the underlying usage of these media, and particularly the level of attention and engagement of Twitter users.
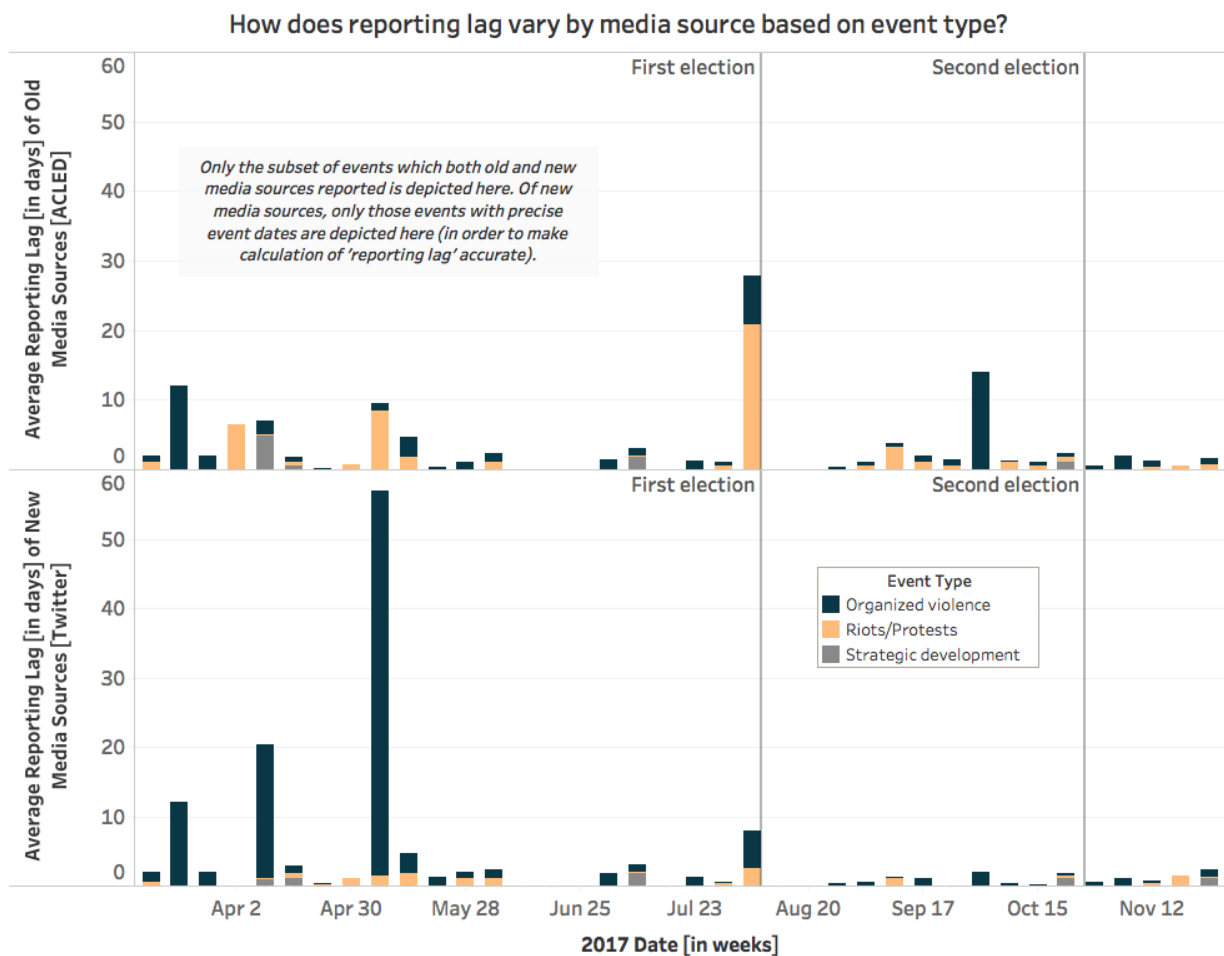
The timeliest tweets seem to be 'live tweets' – those in which an individual reports an event without a link to an outside media report. This makes intuitive sense as these correspond to people tweeting as the event is ongoing. However, live tweets do not always provide sufficient information – the user does not always take the time to offer sufficient information around all necessary dimensions.

Aside from live tweets, a significant number of tweets include links to traditional media sources. When restricting the analysis to these, we find that these are unsurprisingly less timely than traditional media sources, as they include a 'built-in' re-tweeting lag. Among these, the timeliest are those posted by media outlets themselves on their own Twitter accounts. This shows that, rather than being separate, 'old' and 'new' media sources are deeply interlinked. These patterns also highlight the extent to which social media contains – and reflects the same biases as – 'traditional' media.

We next disaggregate the analysis by event type. As previously noted, current evidence points to the fact that 'new' and 'old' media report different types of violent events differently (Roberts and Marchais 2017). Here, we ascertain whether there is evidence that different types of events generate variations in reporting lag between the two sources.

Figure 4.8 presents two graphs that are the same as those in Figure 4.7, but with the shading reflecting the types of events being reported. This helps to shed light on how this timeliness of coverage might be related to event types. We see, for example, that the events that Twitter reported much more quickly around the first election (see bottom graph) are largely riots and protests (i.e. more 'spontaneous' violence and action) while the events that they reported much more slowly outside of the immediate election period are largely organised violence (e.g. battles, remote violence, violence against civilians). For example, on average, when comparing the subset of events that both ACLED and Twitter reported, when it comes to riots and protests, ACLED reports events with a lag of about 1.58 days, while Twitter reports with a lag of about 0.46 days. In contrast, when it comes to organised violence – such as battles, remote violence, and violence against civilians – ACLED reports events with a lag of about 1.9 days while Twitter reports with a lag of about 3.71 days.

Figure 4.8 How does reporting lag vary by media source based on event type?

How does reporting lag vary by media source based on event type?

Only the subset of events which both old and new media sources reported is depicted here. Of new media sources, only those events with precise event dates are depicted here (in order to make calculation of 'reporting lag' accurate).

Event Type
- Organized violence
- Riots/Protests
- Strategic development

Source: Authors' own.

Beyond timeliness alone, different sources may also differ in the consistency of their reporting and coverage over time and when different sources may be more or less comprehensive depending on the stage in an election period. For example, Figure 4.9 below presents graphs depicting events for the time period between March and July 2017, prior to the first 2017 election in August. It is apparent that events coded from Twitter reports (the bottom graph) are far fewer during this pre-election period relative to those coded from ACLED reports (the top graph) – 181 events reported by Twitter relative to 365 by ACLED, or about half as many – with a spike seen around the primary elections of April 2017. ACLED reports, however, offer much more continuous coverage during this period.

Figure 4.9 Coverage of events outside of immediate election periods
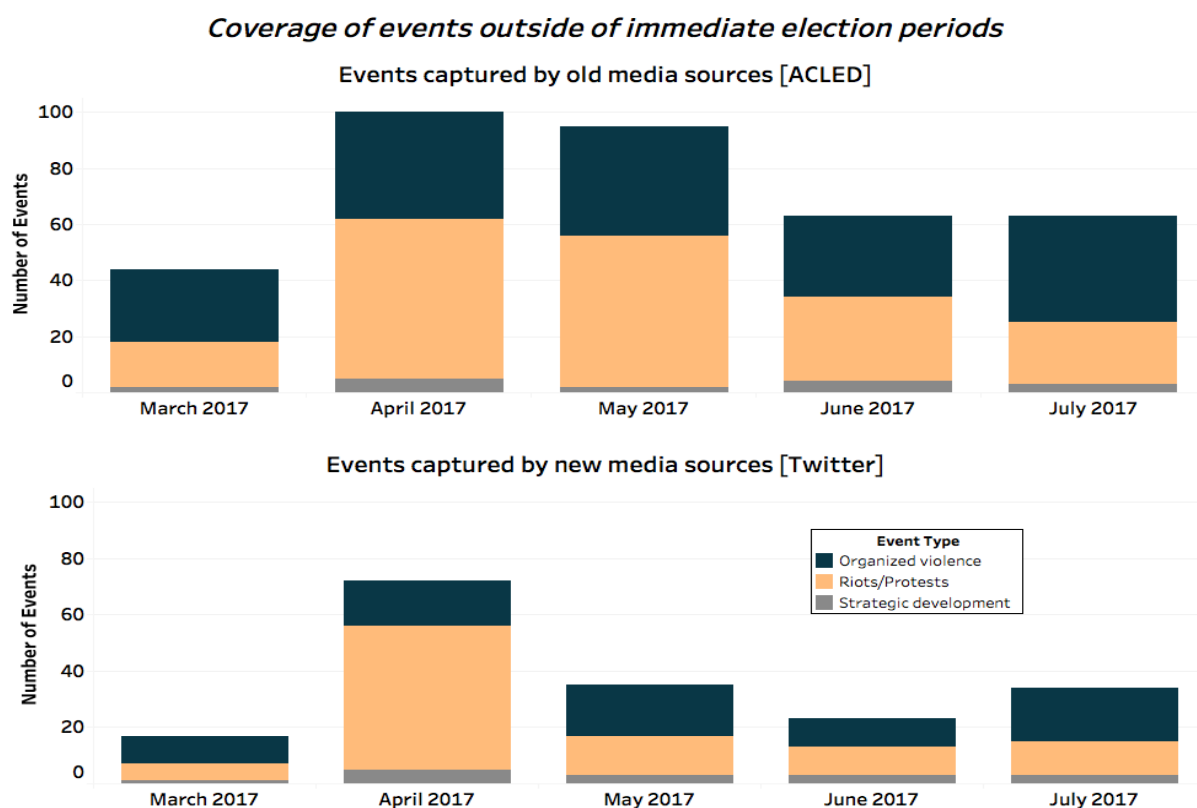
Source: Authors' own.

While Twitter may be a helpful tool in the period immediately surrounding critical junctures such as elections, ACLED reporting ('old' media sources) can be much more consistent over time if interested in monitoring over a longer longitudinal period. This is in line with Hypothesis 2b: Twitter is less likely to provide more comprehensive account of violence over time.

A separate but related issue is that of time precision, briefly addressed above, as it may also be important to understand that different data sources offer varied levels of accuracy around when events occurred. [14]

Figure 4.10 presents a set of graphs that present further information on time precision. As with geographic precision, time precision depicts the level of specificity in reporting, but in this case, referring to the date of the event. For example, if a report says that an event occurred "on August 10", this is a higher level of precision than if a report says that an event occurred "sometime last month." Differences in precision are coded on a deceasing scale of precision, from 1 (highly precise, corresponding to a specific day) to 2 (corresponding to a specific week), and 3 (corresponding to a specific month). The graph on the top depicts the 'time precision' of events reported by ACLED, while the bottom graph depicts the same for events from Twitter. And again, the events compared here are those events which both sources reported.

The shorter the bar, the less fuzziness there is – or the more accurate the event date reported is. In general, ACLED reports event dates with slightly higher precision (on average

---

[14]   It is important to note that the time precision here is in regard to *reporting* and not in regard to the "truth". The precision denotes how precise the reporter (whether media or Twitter) is in their reporting with the assumption that the reporter has not misconstrued this information. Events were not specifically corroborated on the ground by authors.

shorter bars), though there is not quite as much discrepancy between the two sources here; ACLED reports these events with a precision of about 1.1 out of 3, while Twitter reports with about 1.2 out of 3. The shading aids in the comparison again here. In weeks where the colouring is more blue, Twitter reported events with higher time precision; in weeks where the colouring is more orange, ACLED had higher precision. Again, outside of the immediate election period there is more orange, where ACLED reports event dates with higher precision. This is in line with Hypothesis 2c: 'Old' media is more likely to capture the temporality of violence with higher precision than Twitter.

Figure 4.10 How does reporting precision of event data vary by media source?
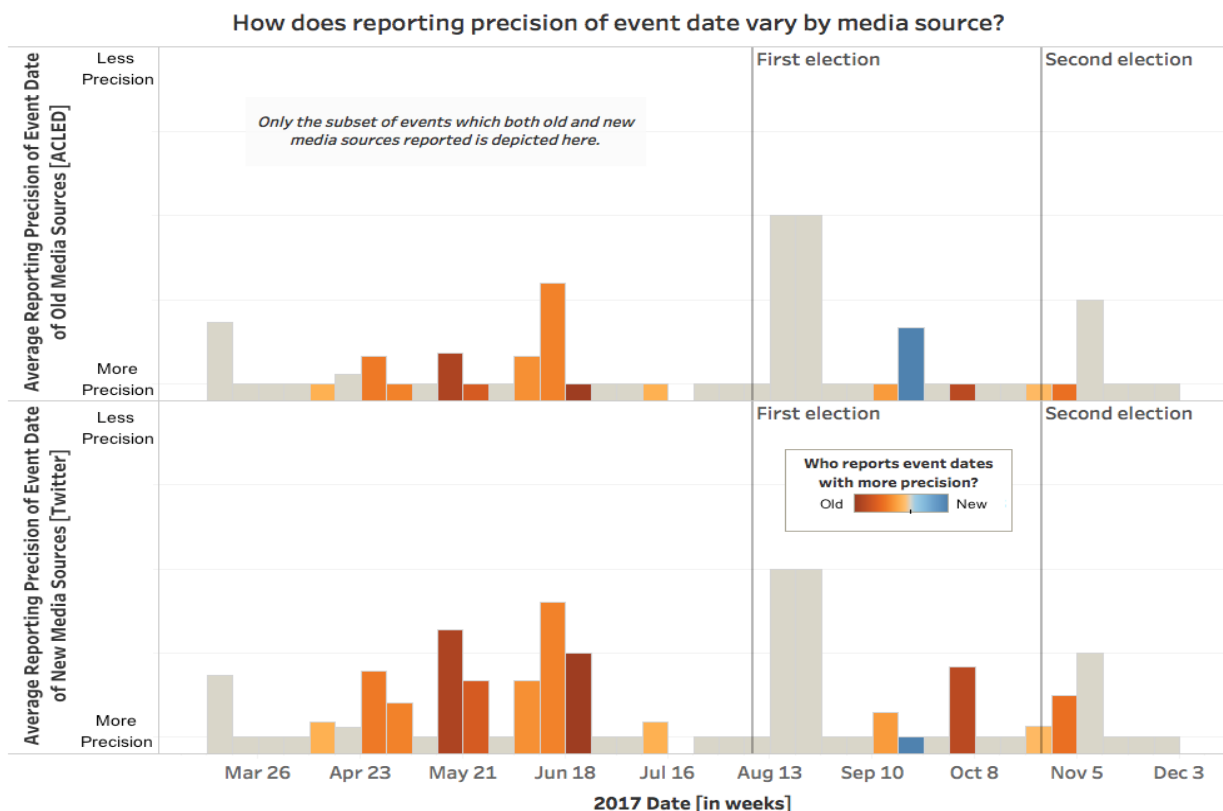
Source: Authors' own.

4.1.3 Targeting/representativeness: event intensity and economic geography

Lastly, we turn to the targeting and representativeness of information across sources – specifically, the question of how different sources of data inform different perceptions of crisis and emergency response needs.

Different data sources offer varied information around the same event. This includes information around the location or date of an event (as discussed above), but also what groups are reported to have been involved, and whether demonstrations turned violent or not. Another facet that is comparable across datasets is the number of fatalities reported. Here, we defer to the most conservative fatality estimate reported by each source as to not overinflate the number of reported fatalities.[15] What this means in practice is that if event reports did not note anything about fatalities, then we must assume 'no fatalities' – and hence code zero fatalities. Thus, while in general when comparing datasets, more fatalities

---

[15]   The exception to this is if a report is released later offering additional information and hence reports a higher fatality count, that higher (more accurate) count is coded instead.

being reported does not necessarily mean that a source is 'more accurate' than another, in this specific case it could indeed point to this in that most of the 'zeroes' stemming from reports could be attributed to no fatalities *being explicitly mentioned* and not necessarily that there were truly no fatalities. This is particularly the case as fatalities were often mentioned as a part of those same events by 'old' media sources.

Figure 4.11 is a set of graphs which depict information about the average number of fatalities reported. The graph on the top depicts this as reported by ACLED, while the bottom graph depicts the same for Twitter. Again, the events compared here are those events which both sources reported. A shorter bar means fewer fatalities on average have been reported by the source.

Figure 4.11 How does the number of reported fatalities vary by media source?
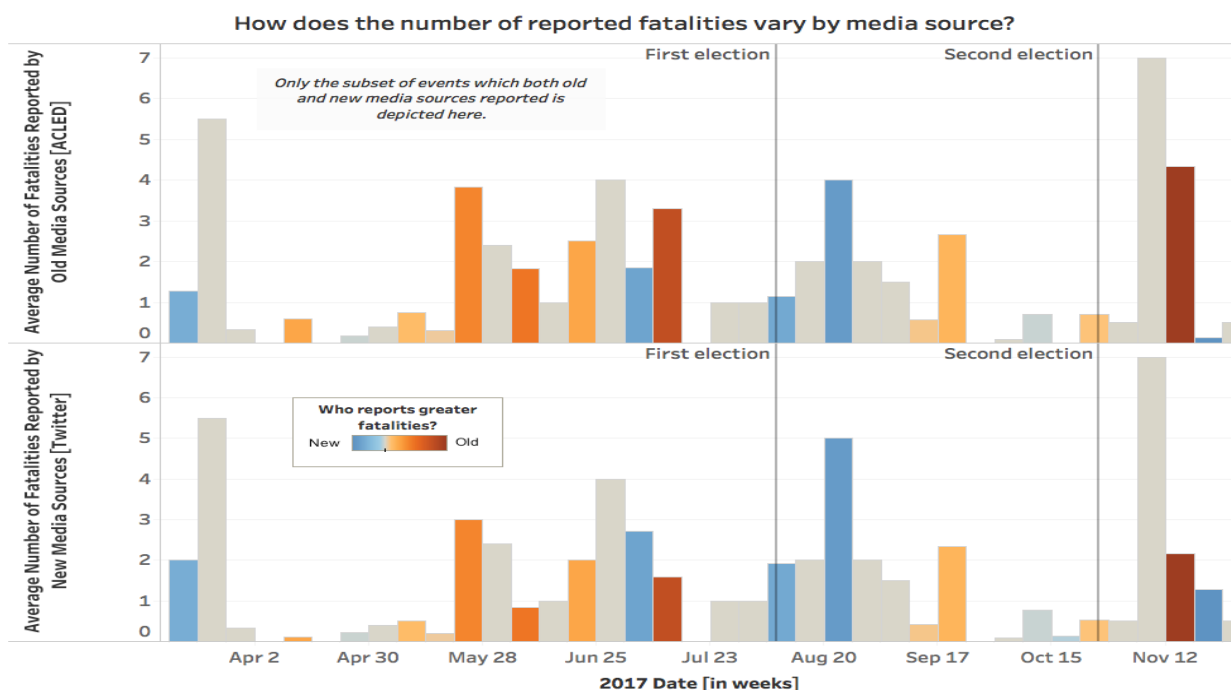
Source: Authors' own.

Again, the shading aids in our comparison. In weeks where the colouring is more blue, Twitter reported more fatalities on average, while in weeks where the colouring is more orange, ACLED reported more fatalities on average. Around the first election, for example, there is more blue, where Twitter reports a greater number of fatalities on average: during this time, Twitter reports an average of 1.6 fatalities per event, while ACLED reports 1.4 fatalities. Meanwhile, outside of the immediate election period, there is more orange, where ACLED reports slightly more fatalities on average (an average of 1.9 fatalities per event while Twitter reports 1.8).

What is also important to note is that more fatalities on average are reported by ACLED following the second election, compared to Twitter. This further highlights that relying on 'new' media sources alone might result in making only certain periods seem 'in crisis' based on fatality counts.

5   Discussion and conclusion

These findings indicate that both 'old' and 'new' media sources provide valuable information on the violence that unfolded before, during and after the Kenyan elections of 2017. Rather than being seen as alternative or competing sources, they should be conceived as complementary as they both provide unique information that, together, paints a fuller picture of political violence. The results contribute to the growing literature on the use of SMDT for violence reporting and policy in several ways.

First, the findings show that traditional media remain extremely valuable for reporting violence. We found that on average, they report more unique events than Twitter, are more temporally consistent in their coverage, have a wider geographical coverage as well as better coverage of violence in less economically developed areas, and provide significantly more reports on specific categories of violent events, in particular organised violence and remote violence.

However, these results are conditioned by the fact that Kenya has a comparatively large traditional media sector, with a relatively high level of freedom of the press (Geopoll and Portland 2017). These results would not necessarily hold in other countries and contexts. The quality, consistency and timeliness of traditional media reporting on violence is dependent on the size, quality and diversity of the traditional media sector, which can vary significantly across and within countries, and over time. These in turn are tied to a host of factors, from regime censorship and control over media (Freedom House 2017), to structures of ownership and levels of concentration of traditional media (Njoroge *et al.* 2011), to the education and training of journalists. In a context where dis-accreditation of traditional media is happening at a global scale, eroding their legitimacy and provoking significant disinvestment from the sector, this point has clear policy implications. For donors and policy actors committed to supporting an environment in which reporting on violent events is done as consistently and accurately as possible, and transparently brought into the public debate, supporting national, regional and local media organisations remains a key entry point with significant value-added in terms of monitoring and enabling response to a wide variety of violence dynamics.

Second, the findings indicate that 'new' media - and more specifically, Twitter - can have considerable value in reporting on violence, but paint a more nuanced picture about their role than what is often assumed. 'New' media generate a vast amount of data on a particular event, but only a fraction of that data is sufficiently detailed to be 'actionable' - i.e. to be used in a reliable, timely and efficient manner for rapid response and de-escalation interventions. While live tweets can be extremely timely and geographically precise, they often lack sufficiently detailed and precise information on other dimensions to be actionable. Apart from live tweets, we have found that the most precise tweets in terms of time precision, geographic precision and event detail are tweets that contain links to media sources. In these cases, the added value of Twitter as compared to 'old' media is limited as they incorporate the existing biases and limitations of the latter.

These results point to a fundamental trade-off between timeliness on one hand and accuracy and precision on the other in social media generated data on violent events. While live tweeting is the timeliest form of reporting, it is rarely sufficiently accurate, detailed and precise along other key dimensions. This can partly be mitigated by increasing the precision and accuracy of the information reported in live tweets. While this might be due to variations in the level of training of citizen reporters, we contend that it is also due to the nature and characteristics of violent events. More than other types of events, information on violent events is difficult to collect and verify because of security and access related restrictions, but also because it is particularly prone to silencing, distortion and manipulation. Generating

accurate information on violent events requires careful compilation, verification and triangulation of different sources, which is time consuming.

The fundamental trade-off between timeliness and accuracy does not mean that data that are timelier and less accurate along other dimensions are not useful. Usefulness depends on who is collecting the data, and for what purposes. For early warning and de-escalation platforms and initiatives, the precision of reports may be less important than their timeliness, as their main role is to draw attention to escalating events and trigger rapid responses and interventions. Conversely, platforms whose central objective is to document violence can decide to focus on accuracy and precision over timeliness. Recent initiatives, such as the Kivu Security Tracker, have explicitly chosen to focus on accuracy rather than timeliness, recognising that despite their capacity to deploy teams to sites of violent events in a timely manner, the compilation and verification of information might can take several weeks (Kivu Security Tracker 2017). Moreover, while time and geographic precision of reports are key for actionability, this does not mean that reports that fail to meet the criteria cannot be used for other purposes such as research, and particularly research on the magnitude and characteristics of social media activity around violent events.

Third, a central result of this study is that the density, timeliness, precision, and geographic spread of Twitter reports on violence are dependent on the behaviour of Twitter users. We have found that new media generate significantly fewer unique reports outside the election period, and report less on organised violence and remote violence, which suggests that a crucial underlying factor is the 'attention' of Twitter users to different types of events. The fact that the second election generated fewer unique reports than the first election also points in that direction. Qualitative evidence suggests this might be due to a form of 'fatigue' which increased as the political crisis dragged on. These results provide backing to the idea that 'issue-attention cycles' are also key conditioning factor in 'new' media coverage, as has been known to be the case for 'old' media (Jakobson 2000; Downs 1972). This provides further evidence to recent contentions that 'new' media attention cycles are dependent on the underlying characteristics of 'new' media users, and their behaviour (Dwyer and Molony 2018).

While limited information on the profile and behaviour of Twitter users prevented a deeper analysis in this study, the academic literature points to the fact that these have strong socio-economic, class, gender and rural/urban dimensions (Dwyer and Molony 2018; Roberts and Marchais 2017). This is consistent with the results of this study, which indicates that these dimensions may also influence the reporting priorities of these users. These results call for caution in relying solely on Twitter and other 'new' media for violence reporting, as this could lead to an amplification of underlying biases in attention of social media users toward different types of events, and generate new dynamics of (in)visibility of violent events.

As many of the earlier studies discussed have demonstrated, 'old' media are not free from systematic biases that may silence or under-represent marginalised, rural or low-income voices. However, these biases are typically well-documented and often at least partially taken into account in studies that rely on media monitoring as a pathway for violence reporting. Moreover, in spite of the promise of empowering everyday citizens with the tools to record and disseminate their own experiences, we find that these biases are often even more pronounced in SMDT systems, than in existing media counterparts. Together, this suggests that both sources of data should be considered with caution, and that particular measures might be taken to triangulate and supplement reporting systems with a particular focus on amplifying the voices and experiences of those 'left behind' by traditional and social media alike. Lastly, it is worth noting that the comparative analysis carried out in this paper required that 'old' and 'new' media be identified as separate entities. Yet our study has shown that, rather than being autonomous entities functioning in separate spheres, 'old' and 'new' media are closely intertwined and interdependent. Traditional media such as newspapers and TV

stations usually have social media accounts which they use to promote their material, and social media users then relay these through their networks. This has several implications.

Chief among these, is that further studies of violence reporting should focus on understanding these two media sources as part of a broader ecosystem of information through which information on violence is reported. While the promises of SMDTs to abolish the hierarchies and biases of coverage inherent to traditional media have largely been dispelled, little is known about how biases of coverage and inequalities of access and usage cut across different types of media. A 'horizontal' analysis of cross-cutting biases in coverage, incorporating the existing understanding of the two should be carried out to understand not only which violent events get reported, but also which ones are picked up and amplified by different media, and which ones are ignored or silenced by the combined effects of 'old' and 'new' media. Such an analysis can also shed light on the complex relationship between media sources and dynamics of violence. As dynamics of violent contention are deeply shaped by information and diffusion technologies (Dafoe and Lyall 2015), understanding the diffusion of reporting of violence, and which events are reported and amplified through channels can allow understanding of dynamics of mobilisation and counter-mobilisation.

Annexes: Passive crowdsourcing data strategy

A1   Scanning Twitter: an overview of the methods and tools

In this study, we collect reports of insecurity in Kenya from public posts on Twitter through an automated search strategy utilising Method52. Below, we describe that data collection process.

According to Burghardt (2015), Twitter has become one of the most popular social media platforms of academic study as it provides 1) a relatively quick and straightforward method for collecting and analysing information; 2) a large sample size with several million tweets published daily; 3) a sizeable amount of metadata; and 4) accessibility of data through predefined Application Programming Interfaces (APIs). The Twitter API is the only means that scholars can employ to generate and download their own collections of tweets (McCreadie *et al.* 2012). Twitter provides two different types of API: a search API, which is used to search for tweets according to certain criteria (e.g. keywords, username, location), and a streaming API, that can be employed to collect a real-time continuous stream of the published tweets.

There is a vast number of online and offline software that provide a graphical user interface for the Twitter API, which can be categorised by means of their analytic focus. A first group is composed of tools focusing on the social network of Twitter users (i.e. user-centric tool). These include Twittercounter, Twittonomy, TweetReach, which are all tools using parameters such as follower counts, re-tweet counts, or like counts, to infer the spatial linkages of the user within the Twitter universe.

The second class of tools are designed to obtain and analyse text and metadata from actual tweets (Twitter-centric tools), such as Aggie, Datasift, Gnip, Topsy, TweetArchivist, among others. Some of these tools, for example, Gnip and Topsy, are certified reseller companies of Twitter, meaning that they have access to, and can deliver, all tweets ever published on the platform. Others, such as TwitterArchivist, are Search or Streaming Tools, which provide access only to tweets published in the preceding nine-day period.[16] This characteristic influences our data collection strategy, which is based on several weekly searches rather than a unique search at the end of the period.

A2   Overview of Method52

Method52 is among this second class of tools. It is a software developed by the Informatics Department of the University of Sussex for collecting, processing, and analysing a large number of documents or texts, including social media data (Wibberley *et al.* 2014). The software is based on components customisable by the users, each one with a different function. These components can be organised in pipelines so as to produce a tailor-made system for data collection and analysis. The analysis can be both numerical and visual, and the results can be visualised using a graphical user interface accessible from any web browser. The components of the pipeline are able to perform standard operations such as collecting data from social media platforms (e.g. Twitter and Facebook).

In addition, Method52 enables the use of active learning to split data into different categories according to static or dynamic classifiers (Wibberley *et al.* 2014). Among Method52's innovative functions is the dynamic classifier, which allows the user to process data into user-defined categories. The dynamicity come from the ability of the system of achieving this by learning to imitate user's own decisions. For example, the user can create a sentiment classifier categorising a subset of observations according to certain terms. Method52 can

---

[16]   Access to older data is possible, but there are accompanying costs to accessing this archived data.

classify the remaining set of observations simulating the user's categorisation. Once this first step is complete, the software will ask for confirmation of its classification to the user, and will reclassify the observations according to this new indication. Looping this process, Method52 facilitates the classification of a large volume of observations with minimal effort and high precision (Wibberley *et al.* 2014). The capability is broad and the classifier can be based on any available set of categories.

In addition, Method52 can filter information, or other types of processing by dividing the data down into multiple layers. This speeds up analysis by helping the user to quickly understand the underlying structure of the data. Method52's technology draws on two strands of work on data mining and analysis: a first set of literature related to automatic data collection and analysis, using supervised machine-learning approaches (Carvalho *et al.* 2011; Meraz and Papacharissi 2013; Hopkins and King 2010); and a second strand of computer science literature dedicated to help the users in processing its own search pipeline, so to help improving semantic investigation on social networks (Black *et al.* 2012; Burnap *et al.* 2013).

A3   Method52 and crawling Twitter for incident data

Despite its recent development (Wibberley *et al.* 2014), Method52 has been extensively employed by research organisations and cross-party think tanks such as Demos[17] for monitoring social media during specific timeframes, such as during the 2015 Nigerian elections (Bartlett *et al.* 2015); to assess the EU sentiment during the Eurozone crisis of 2013 (Wibberley *et al.* 2014); or to study mental health-related discussion on public online forums (Buck *et al.* 2017).

The use of Method52 to automate Twitter monitoring, has proved to be effective during electoral periods, when the number of tweets increases sharply and the amount of information is so large as to other strategies of manual monitoring more difficult (Bartlett *et al.* 2015). A previous study sought to use Method52 to track references to violence during election periods: Bartlett *et al.* (2015) used the system to identify 408,000 violence-related tweets during the Nigerian elections in 2015. However, to the best of our knowledge, this research is the first one using Method52 for collecting and coding unique incidence data.

This study draws on of the scarce but growing literature on the crawling[18] tweets for incidence data and, more generally, on the literature on the usage of social media during crises which has grown in the last seven years (Imran *et al.* 2015).

For example, Sambuli *et al.* (2013) use Datasift to focus on Kenya's 2013 elections, in order to study differences in conflict events reported by Twitter and traditional media sources. The study finds that Twitter reports incidents as fast or faster than traditional media sources. In terms of the volume of events, Sambuli *et al.* (2013) find 75 incidence events on Twitter, compared to 46 on traditional media sources. The authors conclude that 'Mining Kenyan Twitter data during an election scenario looks to be a very valuable and worthwhile technique when looking for timely, local information.' (Sambuli *et al.* 2013: 47). Datasift and Method52 are similar, as they both belong to the category of Search/Streaming tools and are able scan Twitter for all tweets published in the previous nine days. In our study, we extend the timeframe of the Sambuli *et al.* (2013) study, from two weeks to more than six months of observations and focus on several other aspects of reporting. The use of Method52 is particularly suitable for this extended timeframe, as it enables the searching and storing of information collected, as well as the creation of an online backup of data.

---

[17]   See www.demos.co.uk/ for more information on Demos.

[18]   Crawling is a term indicating the scanning and collecting of tweets from Twitter.

Moreno *et al.* (2017) provide another example of crowdsourcing data from social media platforms for violence monitoring during election, using as case study Ghana's 2016 election. The authors use Aggie, which is an open source tool that can track real-time events such as election violence from social media data. Similar to Method52, Aggie enables users to search, explore and filter millions of posts published on Facebook and Twitter, but also on WhatsApp, and as RSS. The appropriateness of Twitter for monitoring incidence data is confirmed by the fact that out of 584 incidents identified by the authors, more than 91.6 per cent are captured through Twitter.

A4 Method52 and Kenya's 2017 elections

This study limits its usage of Method52 to crawling tweets containing keywords related to political conflict during Kenyan 2017 elections.[19] Table A4.1 lists the terms used in. The final set includes a total of 113 search terms, a set qualitatively and quantitatively comparable to the 98 terms employed by Sambuli *et al.* (2013). As Method52 allows a limited number of keywords per search, we therefore divide the search terms in seven different searches. Most of the terms are composed by a word denoting political conflict matched with a second one indicating the country of interest (Kenya). Words such as attack, ambush, battle, bomb, clash, displace, injure, kill, violence, among others, have been selected as denoting the occurrence of a violent event related to political conflict or suggesting situational awareness. In addition, these terms are similar to the ones employed by other sources (Sambuli *et al.* 2013; Raleigh *et al.* 2010).

We have used singular and plural forms of selected words, and present and past tense for some verbs, as Method52 is only able to collect tweets containing words matching precisely the search terms. In addition to general terms connoting violence or insecurity, we have included specific hashtags such as #ElectionKE and #KenyaDecides to monitor electoral-related tweets that may contain information on violence (Sambuli *et al.* 2013). Method52 allows users to crawl tweets posted up to nine days before the day of the search. To guarantee coverage, we have conducted searches approximately every seven days, setting as the starting period the day and the time of the previous search. Once the search terms are set up, we added a tool to filter out tweets containing words identifying replicates or irrelevant tweets. In this step, we filtered out retweets, which is a replicate of the original tweet.

Finally, the tool presents the set of tweets collected in tables of data. In the resulting tables, each column contains different information about the tweet, such as the tweeter's username, user description, tweet's text, time of creation, whether media or pictures were included in the tweet, an online link to the original tweet, latitude and longitude of the tweeter, [20] number of friends, and number of retweets. Once each of the seven searches are appended in a single table, we generate the categorical variable "SearchTermId" ranging between one to seven and allowing to reconstruct the search that has returned a given tweet.

Figure A4.1 shows the pipeline for one of the searches and the three steps explained above.[21] When the seven searches have been run and the seven tables produced, an "export" tool is used to export the .csv tables from the server to the local computer, where the set of information (number of columns) is reduced, and the duplicates are dropped.

---

[19]   The period of collection spans from four months before to three months after Kenyan election (9 August 2017).

[20]   This information is almost always absent for tweets from Kenya.

[21]   A similar pipeline is employed to export the tweets from the server to the local computer (not showed).

Figure A1 Example of the Method52 diagram for searching and collecting tweets
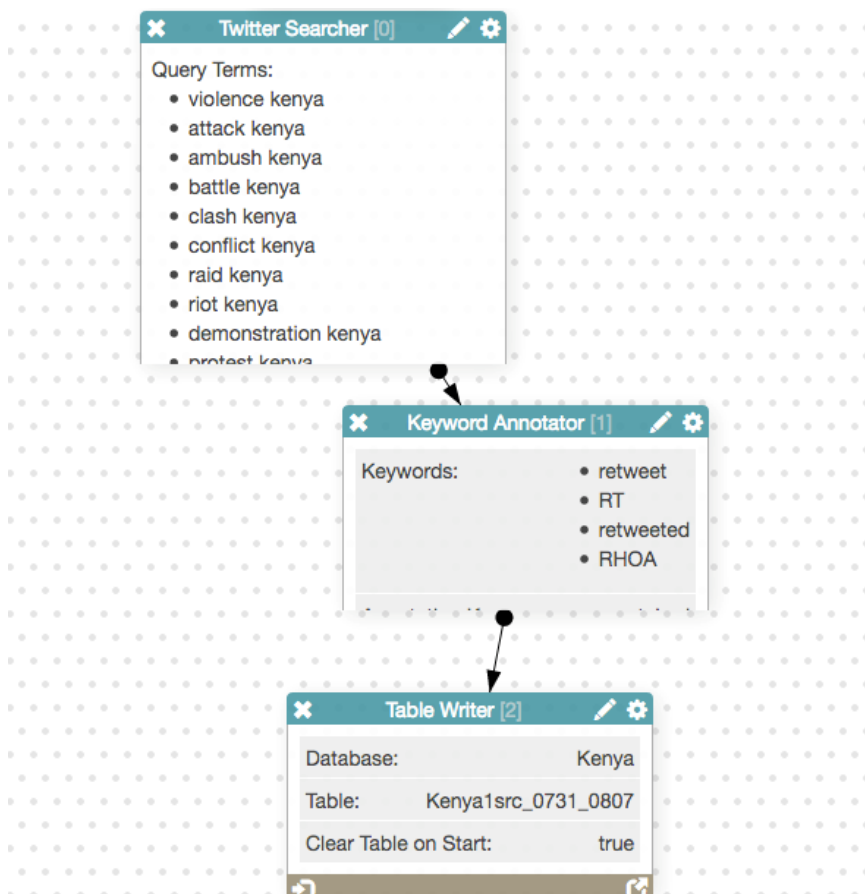
Source: Authors' own.

In a final step, a research team member (coder) codes the resulting set of tweets into unique events according to coding criteria similar to ACLED's, and produces a set of events that are checked and cleaned for errors and inter-coder reliability by a second research team member (cleaner).

Table A1 Search terms employed within Method52

#ElectionsKE	demonstrations Kenya	riot Kenya
#ElectionsKE2017	displace Kenya	rioted Kenya
#KenyaDecides	displaced Kenya	rioting Kenya
#KEPoliceBrutality	displaces Kenya	riots Kenya
#PwaniDecides	displacing Kenya	scuffle Kenya
#SecuringElections2017	exploded Kenya	scuffles Kenya
ambush Kenya	exploding Kenya	shoot Kenya
ambushed Kenya	explosion Kenya	shooting Kenya
ambushes Kenya	explosions Kenya	shoots Kenya
ambushing Kenya	gunshot Kenya	shot Kenya
attack Kenya	gunshots Kenya	skirmish Kenya
attacked Kenya	harass Kenya	skirmishes Kenya
attacking Kenya	harassed Kenya	tensions Kenya
attacks Kenya	harasses Kenya	unrest Kenya
battle Kenya	harassing Kenya	victim Kenya
battled Kenya	injure Kenya	violence Baringo
battles Kenya	injured Kenya	violence Busia
battling Kenya	injures Kenya	violence Coast
beat Kenya	injuring Kenya	violence Garissa
beaten Kenya	insecurity Kenya	violence IEBC
beating Kenya	inter-community Kenya	violence Isiolo
beats Kenya	kill Kenya	violence Kenya
bomb Kenya	killed Kenya	violence Laikipia
bombed Kenya	killing Kenya	violence Mandera
bombing Kenya	kills Kenya	violence Marsabit
bombs Kenya	land dispute Kenya	violence Migori
casualty Kenya	land disputes Kenya	violence Mombasa
clash Kenya	loot Kenya	violence Nairobi
clashed Kenya	looted Kenya	violence PEV
clashes Kenya	looting Kenya	violence Pokot
clashing Kenya	loots Kenya	violence Rift Valley
conflict Kenya	protest Kenya	violence Siaya
conflicting Kenya	protested Kenya	wound Kenya
conflicts Kenya	protesting Kenya	wounded Kenya
demonstrated Kenya	protests Kenya	wounding Kenya
demonstrating Kenya	raid Kenya	wounds Kenya
demonstration Kenya	raids Kenya	

Source: Authors' own.

References

A4AI (2017) *2017 Affordability Report*, Washington: Alliance for Affordable Internet

Aarvik, Per. (2015) *Uchaguzi: An Analysis of the Crowdsourced Election Monitoring in Kenya 2013*, www.academia.edu/15135872/Uchaguzi_An_analysis_of_the_crowdsourced_election_monitoring_in_Kenya_2013 (accessed 31 July 2017)

Africa Confidential (2017) 'Kenya: Pressure-Cooker Polls', *Africa Confidential* 58.1: 15-16

Aling'o, P. and Noor, H. (2016) *Signs of Violence Ahead of Kenya's 2017 Elections*, Institute for Security Studies, 17 June 2016, www.issafrica.org/iss-today/signs-of-violence-ahead-of-kenyas-2017-elections (accessed 1 November)

Andrews, S.; Gibson, H.; Domdouzis, K. and Akhgar, B. (2016) 'Creating Corroborated Crisis Reports from Social Media Data through Formal Concept Analysis', *Journal of Intelligent Information Systems* 47.2: 287-312

Asal, V. and Hoffman, A.M. (2015) 'Media Effects: Do Terrorist Organizations Launch Foreign Attacks in Response to Levels of Press Freedom or Press Attention?', *Conflict Management and Peace Science* 33.4: 381-399

BAKE (2016) *State of the Internet in Kenya*, Bloggers Association of Kenya, https://ifree.co.ke/wp-content/uploads/2016/11/State-of-Internet-Report-Kenya-2016.pdf (accessed 3 August 2017)

Bardall, G. (2013) 'Gender-Specific Election Violence: The Role of Information and Communication Technologies', *Stability: International Journal of Security & Development* 2.3: 1-11

Barranco, J. and Wisler, D. (1999) 'Validity and Systematicity of Newspaper Data in Event Analysis', *European Sociological Review* 15.3: 301-322

Bartlett, J.; Krasodomski-Jones, A.; Daniel, N.; Fisher, A. and Jesperson, S. (2015) *Social Media for Election Communication and Monitoring in Nigeria,* London: DEMOS, www.demos.co.uk/wp-content/uploads/2015/11/Social-Media-in-Nigerian-Election.pdf (accessed 17 August 2017)

Baum, M.A. and Zhukov, Y.M. (2015) 'Filtering Revolution: Reporting Bias in International Newspaper Coverage of the Libyan Civil War', *Journal of Peace Research* 52.3: 384-400

Black, W.J.; Procter, R.; Gray, S. and Ananiadou, S. (2012) 'A Data and Analysis Resource for an Experiment in Text Mining a Collection of Micro-Blogs on a Political Topic', paper presented at the 8th International Conference on Language Resources and Evaluation, Istanbul, Turkey, 21-27 May 2012

Blumenstock, J.E. and Eagle, N. (2012) 'Divided we Call: Disparities in Access and Use of Mobile Phones in Rwanda', *Information Technologies and International Development* 8.2: 1-16

Bocquier, P. and Maupeau, H. (2005) 'Analysing Low Intensity Conflict in Africa using Press Reports', *European Journal of Population* 21: 321-345

Buck, D.; Honeyman, M.; Bartlett, J. and Smith, J. (2017), *Online Support: Investigating the Role of Public Online Forums on Mental Health*, www.kingsfund.org.uk/publications/articles/online-support-public-forums-mental-health (accessed 17 August 2017)

Buhaug, H. and Urdal, H. (2013) 'An Urbanization Bomb? Population Growth and Social Disorder in Cities', *Global Environmental Change* 23.1: 1-10

Burghardt, M. (2015) 'Introduction to Tools and Methods for the Analysis of Twitter Data', *10plus1: Living Linguistics* 1: 74-91

Burke, J. (2017) 'Clashes in Kenya after Opposition Leader's Election Fraud Claim', *The Guardian*, 9 August, www.theguardian.com/world/2017/aug/09/kenya-presidential-election-raila-odinga-early-results-uhuru-kenyatta (accessed 17 August 2017)

Burnap, P.; Avis, N.J. and Rana, O.F. (2013) 'Making Sense of Self-Reported Socially Significant Data using Computational Methods', *International Journal of Social Research Methodology* 16.3: 215-230

CAK (2017) *Third Quarter Sector Statistics Report for the Financial Year 2016/2017 (January – March 2017)*, Communications Authority of Kenya, http://ca.go.ke/images/downloads/STATISTICS/SECTOR%20STATISTICS%20REPORT%20Q3%20FY%202016-2017.pdf (accessed 31 July 2017)

Campagna, J. (1999) *Siege Mentality: Press Freedom and the Algerian Conflict*, Committee to Protect Journalists, February 1999, www.refworld.org/docid/47c567c232.html (accessed 3 August 2017)

Caple, H. and Bednarek, M. (2013) *Delving into the Discourse: Approaches to News Values in Journalism Studies and Beyond*, Reuters Institute for the Study of Journalism Working Paper, https://reutersinstitute.politics.ox.ac.uk/sites/default/files/Delving%20into%20the%20Discourse_0.pdf (accessed 3 August 2017)

Carter Center (2017) *Carter Center Statement on Kenya's Pre-Election Period*, 27 July, http://reliefweb.int/report/kenya/carter-center-statement-kenya-s-pre-election-period (accessed 7 August 2017)

Carvalho, J.C. *et al.* (2011) Validity of an Information and Communication Technology System for Data Capture in Epidemiological Studies, *Caries Research* 45(3): 287–93

Chan, E. (2017) *Public Health Humanitarian Response to Natural Disasters*, London: Routledge

Chataing, P.A. (2015) *Reporting Lags in Real-time Data Collection*, ACLED Working Paper Series No. 8, www.acleddata.com/wp-content/uploads/2015/09/ACLED-Working-Paper-No.-8_Reporting-Lags_2015.pdf (accessed 31 July 2017)

Cheeseman, N.; Lynch, G. and Willis, J. (2014) 'Democracy and its Discontents: Understanding Kenya's 2013 Elections', *Journal of Eastern African Studies* 8.1: 2-24

Cherono, S. (2017a) 'Mystery over Shooting of Mathare Girl Deepens', *Daily Nation*, 14 August, www.nation.co.ke/news/Stephanie-Moraa-Mathare-shootings/1056-4056736-su4l89/index.html (accessed 17 August 2017)

Cherono, S. (2017b) 'Kenya Police Deny *Mungiki* Attacks in Mathare, Dandora', *Daily Nation*, 14 August, www.nation.co.ke/news/Mathare-Dandora-attacks-Munguki-police/1056-4056870-cjb3es/index.html (accessed 17 August 2017)

Chojnacki, S.; Ickler, C.; Spies, M. and Wiesel, J. (2012) 'Event Data on Armed Conflict and Security: New Perspectives, Old Challenges, and Some Solutions', *International Interactions* 38.4: 382-401

Chouliaraki, L. (2016) 'Digital Witnessing in Conflict Zones: The Politics of Remediation', *Information, Communication & Society* 18.1: 1362-1377

CIPESA (2016) *State of Internet Freedom in Africa 2016: Case Studies from Select Countries on Strategies African Governments Use to Stifle Citizens' Digital Rights*, Collaboration on International ICT Policy for East and Southern Africa, https://cipesa.org/?wpfb_dl=225 (accessed 31 July 2017)

CIPEV (2008) *Kenya: Commission of Inquiry into the Post Election Violence Final Report,* Commission of Inquiry into the Post-Election Violence, http://reliefweb.int/report/kenya/kenya-commission-inquiry-post-election-violence-cipev-final-report (accessed 3 August 2017)

Claes, J. (2016) 'Tension Rising ahead of Kenya and Liberia 2017 Elections', United States Institute for Peace, 19 September, www.usip.org/olivebranch/2016/09/19/tension-rising-ahead-of-kenya-liberia-2017-elections (accessed 1 November 2016)

Collier, P. and Vicente, P.C. (2012) 'Violence, Bribery, and Fraud: The Political Economy of Elections in Sub-Saharan Africa', *Public Choice* 153.1: 117-147

Dafoe, A. and Lyall, J. (2015) 'From Cell Phones to Conflict? Reflections on the Emerging ICT-Political Conflict Research Agenda', *Journal of Peace Research* 52.3: 401-413

Dahir, A.L. (2016) 'More African Governments Blocked the Internet to Silence Dissent in 2016', *Quartz Africa*, 31 December 2016, https://qz.com/875729/how-african-governments-blocked-the-internet-to-silence-dissent-in-2016/ (accessed 31 July 2017)

Daily Nation (2017a) "Kenya's repeat presidential poll: The incidents," *Daily Nation*, 26 October, www.nation.co.ke/news/Incidents-rock-Kenya-repeat-presidential-poll/1056-4156348-425v7gz/index.html (Accessed 9 July 2018)

Daily Nation (2017b) 'Raila Odinga Says IEBC Computer System Hacked', 9 August, www.nation.co.ke/news/politics/Raila-Odinga2017--Elections-hacked/1064-4050980-dsmk7r/index.html (accessed 17 August 2017)

Davenport, C. and Ball, P. (2002) 'Views to a Kill: Exploring the Implications of Source Selection in the Case of Guatemalan State Terror, 1977-1995', *Journal of Conflict Resolution* 46.3: 427-450

De Juan, A. and Bank, A. (2015) 'The Ba'athist Blackout? Selective Goods Provision and Political Violence in the Syrian Civil War', *Journal of Peace Research* 52.1: 91-104

Desmaret, L. and Langer, A. (2018) ''Peace journalism on a shoestring? Conflict reporting in Nigeria's national news media', *Journalism,* Published online at: https://journals.sagepub.com/doi/abs/10.1177/1464884918797611?journalCode=joua

Downs, A. (1972) 'Up and Down with Ecology: The "Issue-Attention Cycle"', *Public Interest* 28: 38-50

Dwyer, M. and Molony, T. (2018). *Social Media and Politics in Africa: Democracy, Security, and Surveillance*. London: Zed Books

Earl, J. (2010) 'The Dynamics of Protest-Related Diffusion on the Web', *Information, Communication & Society* 13.2: 209-225

Earl, J.; Martin, A.; McCarthy, J.D. and Soule, S.A. (2004) 'The Use of Newspaper Data in the Study of Collective Action', *Annual Review of Sociology* 30: 65-80

Economist (2017) 'I Burned A Farm in Africa: Land Invasion in Kenya Portend Election Violence', 16 February 2017, www.economist.com/news/middle-east-and-africa/21717067-politicians-are-urging-their-supporters-graze-their-cows-other-peoples (accessed 26 July 2017)

Elder, C.; Stigant, S. and Claes, J. (2014) *Elections and Violent Conflict in Kenya: Making Prevention Stick*, United States Institute for Peace Peaceworks Publication, www.usip.org/sites/default/files/PW101-Elections-and-Violent-Conflict-in-Kenya-Making-Prevention-Stick.pdf (accessed 1 November 2016)

Fowler, G.A. (2017) '"Citizen Journalists" Evade Blackout on Myanmar News: Blogs and Shaky Videos Find Way into Mainstream; Photo of Bloody Sandals', *Wall Street Journal*, 28 September, www.wsj.com/articles/SB119090803430841433 (accessed 3 August 2017)

Freedom House (2017) *Freedom on the net 2017 – Kenya*, www.freedomonthenet.org (Accessed 10 July 2018)

Freedom House (2016) *Freedom of the Press 2016 – Kenya*, https://freedomhouse.org/report/freedom-press/2016/kenya (accessed 3 August 2017)

Gathigi, G. (2018) "How closure of TV stations has narrowed democratic space in Kenya," *The Conversation*, February 6, https://theconversation.com/how-closure-of-tv-stations-has-narrowed-democratic-space-in-kenya-91164 (accessed 9 July 2018)

Geopoll and Portland Communications (2017) 'The Reality of Fake News in Kenya', Report: https://portland-communications.com/pdf/The-Reality-of-Fake-News-in-Kenya.pdf

Gleditsch, K.S.; Metternich, N.W. and Ruggeri, A. (2014) 'Data and Progress in Peace and Conflict Research', *Journal of Peace Research* 51.2: 301-314

Goldsmith, A.A. (2015) 'Elections and Civil Violence in New Multiparty Regimes: Evidence from Africa', *Journal of Peace Research* 52.2: 607-621

Gregory, S (2015) 'Ubiquitous Witnesses: Who Creates the Evidence and the Live(d) Experience of Human Rights Violations?', *Information, Communication & Society* 18.11: 1378-1392

Hoogeveen, J. (2017) 'Let's Invest in Mobile Phone Surveys to Monitor Crises', *Africa Can End Poverty*, 27 June, http://blogs.worldbank.org/africacan/lets-invest-in-mobile-phone-surveys-to-monitor-crises (accessed 3 August 2017)

Hopkins, D.J., and King, G. (2010) 'A Method of Automated Nonparametric Content Analysis for Social Science', *American Journal of Political Science* 54.1: 229-247

Howard, P.N.; Agarwal, S.D.; and Hussain, M.M. (2011) 'When Do States Disconnect Their Digital Networks? Regime Responses to the Political Uses of Social Media', *The Communication Review* 14.3: 216-232

IEBC (2017) 'Nasa Irreducible Minimums On Electoral Reforms Before The Fresh Presidential Elections Are Held', Independent Electoral and Boundaries Commission, www.iebc.or.ke/uploads/resources/qWgsGfYCKO.pdf (Accessed on 9 July 2018)

Imran, M.; Castillo, C.; Diaz, F. and Vieweg, S. (2015) 'Processing Social Media Messages in Mass Emergency: A Survey', *ACM Computing Surveys (CSUR)* 47.4: 67

Institute for Human Rights and Business (2013) 'Corporate Responses to Hate Speech in the 2013 Kenya Elections: Case Study Safaricom', Report, Institute for Human Rights and Business www.ihrb.org/uploads/reports/DD-Safaricom-Case-Study.pdf

Jakobson, P.V. (2000) 'Focus on the CNN Effect Misses the Point: The Real Media Impact on Conflict Management is Invisible and Indirect', *Journal of Peace Research* 37.2: 131-143

Justino, P.; Brück, T. and Verwimp, P. (2013). *A Micro-Level Perspective on the Dynamics of Conflict, Violence and Development.* Oxford University Press

Kariuki, N. (2016) 'Freedom of Press "Still Elusive" in Kenya', *Daily Nation*, 4 May, www.nation.co.ke/news/Freedom-of-Press-still-elusive/1056-3188008-pfpxewz/index.html (accessed 3 August 2017)

Kegoro, G. (2018) "Raila's swearing-in a boon and bane at the same time," *Daily Nation,* 4 February, www.nation.co.ke/oped/opinion/Raila-s-swearing-in-a-boon-and-bane-at-the-same-time-/440808-4290594-fy7y4iz/index.html (Accessed 9 July 2018)

Kivu Security Tracker (2017) 'Kivu Security Tracker Methodology', 2 Nov 2017, https://kivusecurity.org/static/KST_Methodology_Nov2017.pdf , (Accessed 12 July 2018)

Leithead, A. (2017) 'Kenya Election: "Dangerous Limbo" ahead of Official Result', *BBC News*, 9 August, www.bbc.co.uk/news/world-africa-40880601 (accessed 17 August 2017)

Lerman, J. (2013) 'Big Data and its Exclusions', *Stanford Law Review Online* 66:55-63

Lind, J. (2018) 'Devolution, shifting centre-periphery relationships and conflict in northern Kenya', *Political Geography* 63: 135-147

Linke, A.M. (2013) 'The Aftermath of an Election Crisis: Kenyan Attitudes and the Influence of Individual-level and Locality Violence', *Political Geography* 37: 5-17

Mäkinen, M. and Wangu Kuira, M. (2008) 'Social Media and Postelection Crisis in Kenya', *The International Journal of Press / Politics* 13.3: 328-335

Masad, D. (2013) 'How Computers Can Help Us Track Violent Conflicts - Including Right Now in Syria', *The Monkeys Cage*, 9 July 2013, http://themonkeycage.org/2013/07/09/how-computers-can-help-us-track-violent-conflicts-including-right-now-in-syria/ (accessed 5 May 2017)

May, J.D. (2012) 'Digital and other poverties: Exploring the connection in four East African countries', *Information Technologies and International Development*, 8.2: 33-50

McCarthy, J.D.; McPhail, C. and Smith, J. (1996) 'Images of Protest: Dimensions of Selection Bias in Media Coverage of Washington Demonstrations, 1982 and 1991', *American Sociological Review* 61.3: 478-499

McCreadie, R.; Soboroff, I.; Lin, J.; Macdonald, C.; Ounis, I. and McCullough, D. (2012) 'On Building a Reusable Twitter Corpus', *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*: 1113-1114

Meier, P. (2008) 'Crisis mapping Kenya's election violence', *iRevolutions*, 23 October, https://irevolutions.org/2008/10/23/mapping-kenyas-election-violence/ (accessed 1 November 2016)

Meraz, S. and Papacharissi, Z. (2013) Networked Gatekeeping and Networked Framing on #Egypt, *The International Journal of Press/Politics* 18(2): 138–66

Moreno, A.; Garrison, P. and Bhat, K. (2017) 'WhatsApp for Monitoring and Response during Critical Events: Aggie in the Ghana 2016 Election,' *WiPe Paper – Social Media Studies*, http://collections.unu.edu/eserv/UNU:6189/iscram_unu_cs.pdf (accessed 3 August 2017)

Musila, G. (2013) 'Early Warning and the Role of New Technologies in Kenya', in Mancini, Francesco (ed.) *New Technology and the Prevention of Violence and Conflict*, New York, NY: International Peace Institute, 42-56

Mutahi, P. and Kimari, B. (2017) *The Impact of Social Media and Digital Technology on Electoral Violence in Kenya*, IDS Working Paper 493, Brighton: IDS, https://opendocs.ids.ac.uk/opendocs/bitstream/handle/123456789/13159/Wp493_Online.pdf?sequence=287

Mutiga, M. (2017). 'Essential Guide to Kenya's High Stakes Election on 8 August', *International Crisis Group*, 1 August, www.crisisgroup.org/africa/horn-africa/kenya/essential-guide-kenyas-high-stakes-election-8-august (accessed 7 August 2017)

N'gang'a, D. (2017) 'Twitter Just Gave Kenya its Own Hashflags ahead of Election in August', 25 July 2017, www.plive.co.ke/bi/tech/tech-twitter-just-gave-kenya-its-own-hashflags-for-ahead-of-election-in-august-id7047150.html (accessed 28 July 2017)

Namwaya, O. (2016) 'Dispatches – Media Freedom under Threat in Kenya', *Human Rights Watch*, 23 March, www.hrw.org/news/2016/03/23/dispatches-media-freedom-under-threat-kenya (accessed 3 August 2017)

Njoroge, M.; Kimani, P. and Kikech, B. (2011) 'New Media in Kenya: Putting Ethnicity in Perspective', in Wachanga, D. (ed.), *Cultural Identities and New Communication Technologies: Political, Ethnic, and Ideological Implications*, Hershey PA: IGI Global

Nordhaus, W.D. (2006) 'Geography and Macroeconomics: New Data and New Findings', *Proceedings of the National Academy of Sciences of the United States of America* 103.10: 3510–17, https://doi.org/10.1073/pnas.0509842103 and http://sedac.ciesin.columbia.edu/data/set/spatialecon-gecon-v4 (accessed 19 December 2018)

Odhiambo, A. (2017). 'Justice Lacking for Victims of Kenya's Post-Election Violence: As Election nears, Victims Demand Reparation and Justice', *Human Rights Watch*, 11 July, www.hrw.org/news/2017/07/11/justice-lacking-victims-kenyas-post-election-violence (accessed 3 August 2017)

OECD (2009) 'Preventing Violence, War, and State Collapse: The Future of Conflict early Warning and Response', Organisation for Economic Co-Operation and Development, www.oecd.org/dac/conflict-fragility-resilience/docs/preventing%20violence%20war%20and%20state%20collapse.pdf

Perera, Suda. (2015) *Accessing the Inaccessible in Difficult Environments: The Uses and Abuses of Crowdsourcing*, Development Leadership Program, Research Paper No. 34, www.dlprog.org/publications/accessing-the-inaccessible-in-difficult-environments-the-uses-and-abuses-of-crowdsourcing.php (accessed 1 November 2016)

Pew Research Center (2015) 'Cell Phones in Africa: Communication Lifeline', www.pewglobal.org/2015/04/15/cell-phones-in-africa-communication-lifeline/ (accessed 25 November)

Pierskalla, J.H. and Hollenbach, F.M. (2013) 'Technology and Collective Action: The Effect of Cell Phone Coverage on Political Violence in Africa', *American Political Science Review* 107.2: 207-224

Powell, J.M., and Thyne, C.L. (2011) 'Global Instances of Coups from 1950 to 2010: A New Dataset', *Journal of Peace Research* 48.2: 249-259

Price, M.; Klinger, J. and Ball, P. (2013) *Preliminary Statistical Analysis of Documentation of Killings in the Syrian Arab Republic*, Benetech Human Rights Program, www.ohchr.org/Documents/Countries/SY/PreliminaryStatAnalysisKillingsInSyria.pdf (accessed 5 May 2017)

Raleigh, C. (2015) 'Urban Violence Patterns across African States', *International Studies Review* 17.1: 90-106

Raleigh, C.; Linke, A. and Hegre, H. (2010) 'Introducing ACLED: An Armed Conflict Location & Event Dataset', *Journal of Peace Research* 47.5: 651-660

Read, R.; Taithe, B. and MacGinty, R. (2016) 'Data Hubris? Humanitarian Information Systems and the Mirage of Technology', *Third World Quarterly* 37.8: 1314-1331

Reilly, P. (2016) 'Tweeting for Peace? Twitter and the Ardoyne Parade Dispute in Belfast, July 2014', *First Monday* 21.11: 1-17

Reuters (2016) 'Congo asks Companies to Block Social Media before Anti-Kabila Protests', 15 December, www.reuters.com/article/us-congo-politics-telecoms-idUSKBN14420M (accessed 5 May 2017)

Roberts, T. and Marchais, G. (2017) *Assessing the Role of Social Media and Digital Technology in Violence Reporting*, IDS Working Paper 492, Brighton: IDS

Rohner, D. and Frey, B.S. (2007) 'Blood and Ink! The Common-Interest-Game between Terrorists and the Media', *Public Choice* 133.1-2: 129-145

Salehyan, I. (2015) 'Best Practices in the Collection of Conflict Data', *Journal of Peace Research* 52.1: 105-109

Salehyan, I.; Hendrix, C.S.; Hamner, J.; Case, C.; Linebarger, C.; Stull, E. and Williams, J. (2012) 'Social Conflict in Africa: A New Dataset', *International Interactions* 38.4: 503-511

Sambuli, N.; Crandall, A.; Costello, P. and Orwa, C. (2013) *Viability, Verification, Validity: 3 Vs of Crowdsourcing*, iHub Research, https://ihub.co.ke/ihubresearch/jb_VsReportpdf2013-8-29-07-38-56.pdf (accessed 3 August 2017)

Sang, A. (2017) 'What To Do in Case There's Internet Shutdown', *Nairobi News*, 7 August, http://nairobinews.nation.co.ke/news/blog-case-theres-internet-shutdown/ (accessed 7 August)

Snyder, D. and Kelly, W.R. (1977) 'Conflict Intensity, Media Sensitivity and the Validity of Newspaper Data', *American Sociological Review* 42: 105-123

Steel, H. (2015) 'Streets to Screens: Mediating Conflict through Digital Networks', *Information, Communication & Society* 11.18: 1269-1274

Taylor, C.; Pevehouse, J. and Straus, S. (2013) *Perils of Pluralism: Electoral Violence and Competitive Authoritarianism in sub-Saharan Africa*, Simons Working Paper Series in Security and Development No. 23, http://summit.sfu.ca/item/14904 (accessed 1 November)

Ushahidi (2013) 'Case Study: Uchaguzi', *Ushahidi Blog*, www.ushahidi.com/case-studies/uchaguzi (accessed 17 August 2017).

Ushahidi (2010) 'Uchaguzi: An Ushahidi Deployment for Kenya's 2010 Constitutional Referendum', *Ushahidi Blog*, www.ushahidi.com/blog/2010/08/02/uchaguzi-an-ushahidi-deployment-for-kenyas-2010-constitutional-referendum (accessed 17 August 2017)

Van der Windt, P. and Humphreys, M. (2016) 'Crowdseeding in Eastern Congo: Using Cell Phones to Collect Conflict Events Data in Real Time', *Journal of Conflict Resolution* 60.4: 748-781

Wadekar, N. (2018) "Kenya's Dangerous Path Toward Authoritarianism," *The New Yorker,* February 9, www.newyorker.com/news/news-desk/kenyas-dangerous-path-toward-authoritarianism (Accessed 9 July 2018)

Weidmann, N.B. (2016) 'A Closer Look at Reporting Bias in Conflict Event Data', *American Journal of Political Science* 60.1: 206-218

Weidmann, N.B. (2015) 'On the Accuracy of Media-Based Conflict Event Data', *The Journal of Conflict Resolution* 59.6: 1129-1149

Weidmann, N.B. (2013) 'The Higher the Better? The Limits of Analytical Resolution in Conflict Event Datasets', *Cooperation and Conflict* 48.4: 567-576

Wibberley, S.; Reffin, J. and Weir, D. (2014) 'Method51 for Mining Insight from Social Media Datasets', www.aclweb.org/anthology/C14-2025 (accessed 17 August 2017)

Wigmore-Shepherd, D. (2015) *Reporting Sources*, ACLED Working Paper Series No. 5, www.acleddata.com/wp-content/uploads/2015/04/ACLED_Reporting-Sources-Working-Paper-No.-5_2015.pdf (accessed 1 November)

Wilson, J. and Gapsiso, N. (2016) *Overcoming Gender Inequalities Through Technology Integration*, Hersey: IGI Global

World Bank (2017) 'The Tech Revolution That's Changing How We Measure Poverty', *World Bank Blog*, 27 July, www.worldbank.org/en/news/immersive-story/2017/07/26/the-tech-revolution-thats-changing-how-we-measure-poverty?CID=ECR_TT_PUBS_EN_EXT (accessed 3 August 2017)

Worley, W. (2017) "Kenyan President Uhuru Kenyatta vows to 'fix' judiciary after Supreme Court election annulment," *The Independent,* 2 September, www.independent.co.uk/news/world/africa/kenyan-president-uhuru-kenyatta-vows-fix-judiciary-supreme-court-election-annulment-a7925586.html (accessed 9 July 2018)