How African countries respond to fake news and hate speech

Garbe, Lisa; Selvik, Lisa-Marie; Lemaire, Pauline

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
Wissenschaftszentrum Berlin für Sozialforschung (WZB)

Routledge
Taylor & Francis Group

How African countries respond to fake news and hate speech

Lisa Garbe [a,b], Lisa-Marie Selvik [c] and Pauline Lemaire [c,d]

[a]WZB Berlin Social Science Center, Berlin, Germany; [b]Institute of Political Science, University of St.Gallen, St.Gallen, Switzerland; [c]Department of Comparative Politics, University of Bergen, Bergen, Norway; [d]Chr. Michelsen Institute, Bergen, Norway

ABSTRACT

While scholars have already identified and discussed some of the most urgent problems in content moderation in the Global North, fewer scholars have paid attention to content regulation in the Global South, and notably Africa. In the absence of content moderation by Western tech giants themselves, African countries appear to have shifted their focus towards state-centric approaches to regulating content. We argue that those approaches are largely informed by a regime's motivation to repress media freedom as well as institutional constraints on the executive. We use structural topic modelling on a corpus of news articles worldwide ($N = 7'787$) mentioning hate speech and fake news in 47 African countries to estimate the salience of discussions of legal and technological approaches to content regulation. We find that, in particular, discussions of technological strategies are more salient in regimes with little respect for media freedom and fewer legislative constraints. Overall, our findings suggest that the state is the dominant actor in shaping content regulation across African countries and point to the need for a better understanding of how regime-specific characteristics shape regulatory decisions.

1. Introduction

Online platforms gained enormous traction in political and social discourse, with platforms like Facebook evolving into transnational companies that are 'unmatched in their global reach and wealth' (Gorwa, 2019, p. 860). Social media platforms in particular have been blamed for poor efforts to moderate content in many instances around the world, failing to protect users from foreign influence during elections in the USA and France (Walker et al., 2019, p. 1532) or to adequately moderate hate speech in Ethiopia inciting violent ethnic protest (Gilbert, 2020). 'We take misinformation seriously,' Facebook CEO Mark Zuckerberg (2016) wrote just weeks after the 2016 elections in the USA. In the years since, the question of how to counteract the damage done by 'fake news' has become a pressing issue both for technology companies and governments across the

globe. Indeed, there is a growing debate about how to adequately regulate online content predominantly taking place in Europe and North America (Iosifidis & Andrews, 2020).

Yet, how are fake news and hate speech regulated across African countries? In this paper, we use news coverage of fake news and hate speech in Africa to analyse how regulatory strategies are framed, and how these frames are predicted by different regime characteristics. In essence, our analysis of 7'787 news articles covering 47 African countries suggests that in the absence of proactive content moderation by the platforms, discussions regarding the regulation of fake news and hate speech mostly centre on state-centric strategies.

Figure 1 underlines the salience of 'fake news', 'hate speech', 'misinformation', and 'disinformation' in news coverage of African countries. The trend over the last five years suggests that these issues have gained increasing importance in public discourse. The enormous spread of misinformation related to the COVID-19 pandemic on Facebook in South Africa and Nigeria (Africa Check, 2020; Ahinkorah et al., 2020, p. 2) further underlines this trend. The fact that both humans and bots are used in several African countries to spread government-propaganda and discredit public dissent online (Bradshaw & Howard, 2019) highlights the challenges related to limiting hate speech and fake news in African contexts. Indeed, in more authoritarian contexts, domestic governments themselves seek to manipulate both information and discourse to ensure their regime's survival.

How African countries respond to fake news and hate speech is a highly relevant question, especially in the absence of content moderation by Western tech giants. While platforms have started to engage in content moderation around the world, they appear comparatively inactive on the African continent. In 2019, upon request from governments, courts, civil society organizations, and 'members of the Facebook community' (Facebook, n.d.), Facebook removed content from its platform in several thousands of instances in countries like Pakistan (N = 7'960), Mexico (N = 6'946), Russia (N =

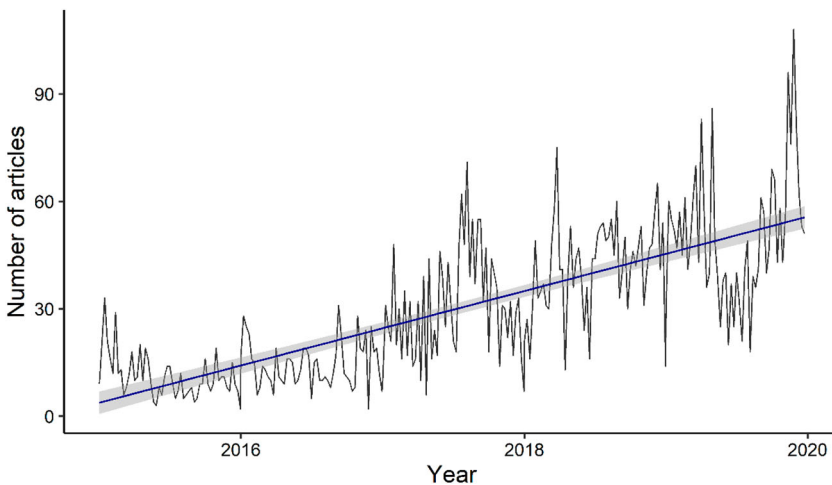**Figure 1.** Number of articles including the search terms, 2015-2019.

Note: Total number of news articles including the terms "fake news", "hate speech", "misinformation", or "disinformation", per week between 01.01.2015 and 31.12.2019 (N = 7'787).

2′958), or Germany ($N = 2′182$), but it hardly removed any content in Africa. In fact, Morocco had the highest number of content removals, with $N = 6$ (Facebook, n.d.). Twitter's transparency reports suggest similar figures for African countries (Twitter, n.d.). The lack of content moderation in African countries seems particularly counterintuitive given the vast use of social media platforms, such as Facebook, on the African continent (e.g., Bosch et al., 2020; Nothias, 2020).

Theoretically, we build on Lessig (1999) and Boas (2006) framework of the regulation of code as well as recent scholarship on regime survival to explain how regime-specific characteristics shape the prevalence of legal and technological discussions about online content regulation. Empirically, our results underline the importance of regime characteristics to understand debates about online content regulation. Technological approaches to content regulation such as blocking or censoring online content are more commonly discussed in regimes that rely on the repression of media and in which the executive's actions are less constrained by legislatures. Legal approaches to content regulation, such as legislation passed by parliament, in particular the criminalization of hate speech, are more commonly discussed in countries respecting media and press freedom, yet not necessarily in countries with higher institutional constraints on the executive. Overall, our findings suggest that regime-specific characteristics pave the way for different strategies to regulate content, some of which may have profound consequences for the freedom of expression online.

We proceed with a theoretical section in which we combine insights from internet governance and comparative politics to formulate expectations about the prevalence of technological and legal regulation in public discourse. Subsequently, we present our data and explain how we employ structural topic modelling to identify regulatory frames in our body of collected texts. We then present results from regression analyses and discuss these findings in light of recent regulatory trends in Africa.

2. Theory

Recent scholarly efforts seek to understand determinants and effects of governments' attempts to control online spaces using censorship (Hellmeier, 2016), internet shutdowns (Hassanpour, 2014; Freyburg & Garbe, 2018; Rydzak et al., 2020) or online surveillance (Michaelsen, 2018). However, these studies do not take into account the legitimate need for governments to address, prevent and punish the spread of hate speech and fake news. Crucially, the aim and motivation of such regulation can be legitimate as long as it addresses citizens' needs (Helm & Nasu, 2021). By connecting insights from scholarship on internet governance and regime survival, our aim is to explain how variation in (1) regimes' motivations to control Information and Communication Technologies (ICT) and in (2) institutional constraints on the adoption of regulation result in differences in the framing of regulatory strategies addressing fake news and hate speech.

We assume that regulation of hate speech and fake news covered in news reports can be seen as regulatory frames (Gilardi et al., 2021, p. 23) that represent different perspectives on regulation. Following DiMaggio et al. (2013), we consider that news reports offer a useful mirror of societal debates, both because they report on issues when these are under consideration by political institutions, and because they reflect debates among the informed public. Furthermore, by considering not only African but also global

news reports, we overcome potential biases in the way regulation is framed in more illiberal countries. We assess how well news reports reflect actual regulatory strategies using information from *Freedom on the Net* reports, annually released by Freedom House (2015, 2016, 2017, 2018, 2019, 2020) that cover legal and technological aspects of internet regulation in 16 African countries (see Methods section; Appendix D).

2.1. Legal versus technological approaches controlling online content

Lessig (1999) distinguishes, broadly speaking, between institutional and architectural means (or *legal* and *technological* means as we will call them during the remainder of this paper) to regulate online space (cf. Boas, 2006, p. 4f.). According to Lessig (2006, pp. 124–125), states can control the technological architecture of the internet through executive decisions, thus influencing or restricting the production of and access to specific content. He (2006, pp. 136–37) argues that technological approaches enable states to regulate online content without having to suffer any political consequence. Legal approaches to content regulation are the predominant institutional strategy to shape access to and production of online content (Lessig, 2006, p. 130), including both the formulation of legislation in the form of bills, laws, and acts as well as the judicial review of existing legislation by courts.

Applying this distinction to how governments seek to regulate online content, we argue that the main difference between technological and legal regulation is that technological strategies are an *ex-ante* approach to prevent the production of online content in the first place, while legal strategies are mostly *ex-post*, removing and/or punishing harmful content after it was produced or shared (Frieden, 2015). These two approaches are not mutually exclusive and are, in fact, often employed in combination with one another.

2.2. Motivations for controlling online content

We acknowledge that a differentiated understanding of regime type is needed when studying politics in Africa. One important and useful distinction with regards to a government's motivations for controlling the flow of information and communication, is the degree to which a political regime is relying on people's informed vote and a viable opposition as sources of its legitimacy. Most African regimes qualify as 'electoral regimes' (Schedler, 2002, p. 36), meaning they hold elections and tolerate some competition but also violate minimal democratic norms so severely and systematically that they cannot be classified as full-fledged democracies. In countries in which the ruler is not (re-)determined by means of free and fair elections, the government usually relies on a whole 'menu of manipulation' to stay in power (Schedler, 2002). This includes the control of media and civil society actors because a strong and well-informed civil society 'can hold governments accountable beyond elections' (Mechkova et al., 2019, p. 42). Traditionally, in order to control information and communication, authoritarian rulers rely on manipulation of public discourse through the control of media outlets (Kellam & Stein, 2016) or heavy restrictions on civil society (Christensen & Weinstein, 2013).

In the digital age, internet and social media provide both civil society and media actors with new means to access and share information (Breuer et al., 2015; Eltantawy & Wiest, 2011). Authoritarian rulers might therefore require new regulatory strategies to also

control the flow of internet-based information and communication. In particular, they need to overcome the challenge posed by some of the decentralized and low-cost features of the internet that facilitate the organization of collective action without formal organization (Bennett & Segerberg, 2012). From a regulatory perspective, authoritarian regimes should hence be inclined to use preventive measures to keep civil society and media actors from putting pressure on the incumbent by using ICT for mobilization purposes (Dresden & Howard, 2015; Goetz & Jenkins, 2005, p. 20). We therefore expect that those regimes that traditionally rely on the repression of media and press freedom are more likely to employ technological ex-ante strategies that prevent the production and sharing of content in the first place. This is likely to affect how regulating online hate speech and fake news is framed in media reports:

> **H1a**: With increasing levels of press and media freedom, the salience of <u>technological</u> regulatory frames <u>decreases</u>.

> **H1b**: With increasing levels of press and media freedom, the salience of <u>legal</u> regulatory frames <u>increases</u>.

2.3. Institutional constraints to controlling online content

The extent to which authoritarian regimes can impose means of regulation that prevent the creation of digital content should not only depend on their tendency to repress press and media freedom in general but also on institutional constraints. We argue that authoritarian regimes can apply more preventive measures of regulation without facing the need for approval by the legislature or the review by the judiciary. They should therefore be more inclined to use technological means of regulation. In turn, in regimes in which the executive faces more constraints by other branches of power, discussions about legal approaches to content regulation should be more prominent.

The separation of powers aims to prevent a government's abuse of power (Rose-Ackerman, 1996). In many authoritarian regimes, institutions such as legislatures or courts serve as a way to co-opt the opposition rather than provide de facto oversight (Gandhi, 2008; Rakner & van de Walle, 2009; Shen-Bayh, 2018). It is therefore important to focus on the *de facto* capacity of such institutions to constrain executive decisions and hence the government's capacity to regulate online content. Legislatures can challenge a government through non-confidence votes for example (Mechkova et al., 2019). This capacity might be even stronger when opposition actors are represented in the legislature (Herron & Boyko, 2015). Some African legislatures have become powerful institutions 'in terms of checking the executive, contributing to the processes of policy-making, and indeed as a monitor of policy implementation' (Bolarinwa, 2015, p. 20). Independent legislatures are important actors in Africa 'assessing proposed legislation, drafting amendments, […] asking questions, attending committee and plenary meetings, participating in debates or voting' (Nijzink et al., 2006, p. 315), all of which should be reflected in broader societal debates about different steps in the process of legislation. High courts have the possibility to sanction government actions. Examples from Africa highlight their capacity to challenge even fundamental government decisions such as amendments to the constitution to overcome presidential term limits (Vondoepp, 2005).

In a country with independent legislatures and high courts, which effectively constrain the government, the executive is thus more limited in its ability to regulate fake news and hate speech. Ad hoc technological regulation to prevent the circulation of fake news and hate speech appears to be more challenging in such an environment compared to contexts without institutional constraints, as highlighted by an example from Ethiopia. In response to violent protest and the circulation of fake news, the Ethiopian government repeatedly shut down internet access in part of the country. As outlined by Abraha (2017, p. 302) this strategy 'usually take[s] place in the absence of any specific legislative framework'. We hence argue that discussions about legal strategies to regulate content are more prevalent in regimes where the government is de facto constrained by legislatures and high courts:

> H2a: With increasing levels of constraint by legislatures and courts, the salience of <u>legal</u> regulatory frames <u>increases</u>.

> H2b: With increasing levels of constraint by legislatures and courts the salience of <u>techno-logical</u> regulatory frames <u>decreases</u>.

3. Methods

We assess legal and technological regulatory frames by analysing how regulation of hate speech and fake news are reported and discussed in news coverage of Africa. Importantly, news items come from both African and non-African publishers. We include news items from non-African publishers as reporting on politically contested issues like misinformation and hate speech might be scarce or biased in more authoritarian countries where news outlets are often owned by government authorities (Stier, 2015).

Still, domestic African media outlets are prominent in our sample (like Nigerian Vanguard, The Punch, and The Sun, with 17% of the news stories combined) or African reproduction of media content (like AllAfrica with 13% of the news stories). In contrast to analysing actual regulatory advances, news reporting can provide a sense of debates surrounding regulatory strategies pursued by governments and may provide an indication of regulation even before a law has been formally adopted (DiMaggio et al., 2013).

3.1. Corpus

Our data consists of 7'787 English-language news articles from a wide range of news outlets ($N = 380$), covering both digital and digitalized printed press, in 47 African countries. These articles are sourced from Factiva, containing the terms 'hate speech', 'fake news', 'misinformation' and/or 'disinformation' as well as terms related to online activity in the title or article published between 2015 and 2019. The Dow Jones Factiva database is a digital archive of global news content which is frequently used by scholars analysing media reporting on African countries (e.g., Bunce, 2016; Obijiofor & MacKinnon, 2016). Appendix A provides details of the full Factiva search query, which in total produced 22'457 news stories. To ensure that our corpus only consists of news stories discussing fake news and hate speech in online contexts, we subset the full corpus of news stories, only including articles that mention pre-defined words for online aspects. For each article in the final corpus, we only keep those paragraphs in which our key online terms are mentioned.[1] Table 1 provides an overview of the final corpus of 7'787 news stories.

3.2. Structural topic model

In order to analyse how regulatory strategies are framed, and to test our hypotheses about how these frames are predicted by different regime characteristics, we apply structural topic modelling (STM) (Roberts et al., 2019). We first estimate topic models ranging from 10 to 50 topics per model using the *stm* package in R (Roberts et al., 2019). We choose the 35-topic model as the most meaningful in terms of topic quality, based on quantitative measures for exclusivity and semantic coherence, and qualitative evaluation of the topics' interpretability (see Appendix B). Because the STM analysis relies on the probabilistic topic model technique Latent Dirichlet Allocation (see Blei, 2012), a technique which uses word counts and not the order of words, it is up to the researcher to infer meaning from the words and topics that appear, rather than assert it (Grimmer & Stewart, 2013, p. 272).

Based on the words in each topic and a close reading of the twenty most representative articles, we identify two topics as indicators for the framing of technological and legal approaches to content regulation. Representative articles can be found using the *findThoughts* function of the *stm* package, which provides documents highly associated with particular topics (Roberts et al., 2019, p. 14). To validate our interpretation and labelling of the selected topics, four human coders read and manually coded a sample of the most representative texts for each topic. The coders' agreement with the assignment of the structural topic model is around 70–75 percent (see Appendix C).

We further assess how well the identified topics for 'legal' and 'technological' approaches to content regulation in news coverage capture actual regulatory steps undertaken by African governments (see Appendix F). Specifically, we compare our country-year mean topic proportions with data from the *Freedom on the Net* reports (Freedom House, 2015, 2016, 2017, 2018, 2019, 2020), first through a t-test and then by investigating four cases more qualitatively. According to the results, news reports provide a fair indication of different legal and technological regulations by African governments For the remainder of this study, we use the expected proportion of each topic as dependent variable.

Table 1. Description of the textual corpus.

Pre-defined words for online aspects	'online', 'digital', 'Internet', 'web', 'social media', 'Facebook', 'Twitter', 'Google', 'YouTube', 'WhatsApp', 'Instagram'	
Number of news stories	7'787	
Countries in sample	47	
Texts per regime type	0: autocracy	339
	1: electoral autocracy	2'011
	2: electoral democracy	5'293
	3: democracy	144
Texts per year of publication	2019	2'632
	2018	2'157
	2017	1'645
	2016	762
	2015	591

Note. The corpus consisting of news stories on 'hate speech', 'fake news', 'misinformation' and/or 'disinformation' is subset to English news stories in which one or more paragraphs mention the pre-defined words for online aspects listed above. Regime type is coded by V-Dem's categorical regimes of the world measure (v2x_regime) following Lührmann et al. (2017).

3.3. Covariates

To predict the expected proportion of each topic, we use three different indicators from the Varieties of Democracy (V-Dem) project, version 10 (Coppedge et al., 2020). First, we use an aggregated index to assess media and press freedom (*v2x_freexp_altinf*) that ranges from 0 to 1 assessing the extent to which citizens are able to 'make an informed choice based on at least some minimal possibilities for collective deliberation' (Teorell et al., 2019). Second, following Mechkova et al. (2019), we use two different indicators to assess *de facto* rather than *de jure* accountability mechanisms through legislatures (*v2xlg_legcon*) and high courts (*v2x_jucon*) both of which range from 0 to 1. For both indicators, higher values indicate more freedom and more constraints on governments, respectively. In addition, we include a variable on state ownership of the telecom sector per country and year to control for a government's capacity to block internet access (Freyburg & Garbe, 2018). Specifically, the variable indicates the proportion of the telecom sector that is majority state-owned. Here, data comes from the Telecommunications Ownership and Control Dataset (Freyburg et al., 2021).

3.4. Methods

To estimate the effects of press and media freedom and institutional constraints on the proportion of the three selected topics, we use Linear Mixed Models (LMM; Baayen, 2008) and include country as random intercepts to acknowledge that the articles are nested in countries and time fixed effects. We use the logarithm of the proportion of topics as the distribution of these variables is right-skewed. All predictors are standardized. After fitting the model, we check whether the assumptions of normally distributed and homogeneous residuals are fulfilled. Appropriate tests indicate no substantial deviations from these assumptions. Finally, collinearity determined for a standard linear model without random effects, appeared to be no major issue (maximum Variance Inflation Factor: 5; Field, 2009).

4. Results

We identify two topics that reflect the two dominant state-centric regulatory strategies technological and legal approaches: Topic 31, which we label 'technological approaches', represents regulatory frames of governments using technological means to block, manipulate, or censor specific online content; and Topic 5, which we label 'legal approaches', reflects legislative strategies to regulate the production of fake news and hate speech. We illustrate how each of those topics reflects different types of regulatory strategies with excerpts from representative news articles from the corpus Table 2.

4.1. Technological approaches

Topic 31 appears to be related to more technological approaches to content regulation, with terms including 'shutdown', 'access', 'blackout', and 'block'. Both representative articles below point to the problem that fake news and hate speech might often be used as a pretence to prevent opposition actors from accessing specific content or sharing

Table 2. Topics related to state-centric online regulation in news coverage

Topic	Interpretation	Keywords
31	Technological approaches	Ethiopian, ethiopia, sudan, zimbabw, shutdown, addi, shut, mugab, protest, zimbabwean, ababa, burundi, mnangagwa, access, sudanes, uganda, prime, abiy, cut, blackout, unrest, block, mobil, diaspora, reform, harar, restrict, cpj, moyo, colour, activist, burundian, amid, fuel, congo, countri, govern, ahm, tax, disrupt
5	Legal approaches	Legisl, cyber, law, cybercrim, bill, draft, blogger, provis, protect, fine, regul, crimin, tanzania, bulli, amend, offenc, crime, act, kenya, propos, enforc, penalti, onlin, fraud, pornographi, kenyan, requir, comput, pass, provid, communic, appli, board, prohibit, legal, film, fee, enact, illeg, applic

Note: The selected topics presented by the 40 most frequent and exclusive words (*frex* terms). Appendix C provides an overview of all 35 topics and their labels.

information. One article exemplifies how a government, here the Ethiopian government, uses the blocking of specific websites as a means to prevent the spread of 'rumours':

> 'Amid reports of violent clashes that have led to at least 15 deaths, the Ethiopian government has partially blocked internet access […]. The government has justified such action in the past as a response to unverified reports and rumors, noting that social media become flooded with unconfirmed claims and misinformation when violence erupts.' (Solomon, 2017)

Another representative article focusing on government blocking in Cameroon underlines the preventive character of such measures:

> '[Cameroon] endured at least two Internet cuts since January last year with government saying the blackouts were among ways of preventing the spread of hate speech and fake news as the regime tried to control misinformation by separatists groups in the Northwest and Southwest.' (The Citizen, 2018)

4.2. Legal approaches

Topic 5 seems to be concerned mostly with legal processes, as shown by the combination of terms like 'legisl[ation]', 'bill', 'regul[ation]', 'fine' or 'prosecut[ion]'. A closer look at a representative article exemplifies that this topic embraces news coverage of specific legislation such as in Kenya:

> Kenyan President Uhuru Kenyatta signed a lengthy new Bill into law, criminalising cybercrimes including fake news […] The clause says if a person 'intentionally publishes false, misleading or fictitious data or misinforms with intent that the data shall be considered or acted upon as authentic,' they can be fined up to 5 000 000 shilling (nearly R620 000 [43′865 USD]) or imprisoned for up to two years. (Mail & Guardian, 2018)

Before Kenyatta signed the Bill, there were demands to have Parliament review the law to make sure that it does not violate the right to media freedom and expression. Similar laws or proposed laws are discussed in news articles covering Nigeria (Adegbo, 2019), Botswana (The Botswana Gazette, 2017), or Ethiopia (Ethiopian News Agency, 2019), all of which focus on criminalizing the publication of fake news or hate speech and holding to account the individual. In few cases, articles also point to legal approaches that would make it possible to hold internet service providers (ISPs) liable who fail to moderate content appropriately as indicated by a bill discussed in South Africa (Eloff, 2019).

4.3. The influence of media freedom and institutional constraints on regulation

We use our topics 'technological approaches' and 'legal approaches' as dependent variables and estimate the effect of a country's press and media freedom as well as institutional constraints on the expected proportion of each of these topics. Figure 2 depicts the results; more detailed results from the linear mixed models can be found in Appendix D.

First, results from the linear mixed models reveal a differential impact of a country's level of press and media freedom on the expected reporting of technological and legal approaches to regulation. Increasing levels of press and media freedom are associated with decreasing levels of technological regulatory frames (B = −0.55, SE = 0.09). This supports Hypothesis 1a that countries traditionally relying on the repression of press and media are more likely to appear in frames related to technological strategies of content regulation. Furthermore, press and media freedom is positively associated with legal frames (B = 0.17, SE = 0.09), supporting Hypothesis 1b that countries respecting press and media freedom are more likely to be associated with legal strategies of content regulation.

Second, institutional constraints vary in their effects on the expected proportion of technological and legal frames. Legislative constraints are associated negatively with the expected proportion of technological frames (B = −0.44, SE = 0.08) whereas judicial
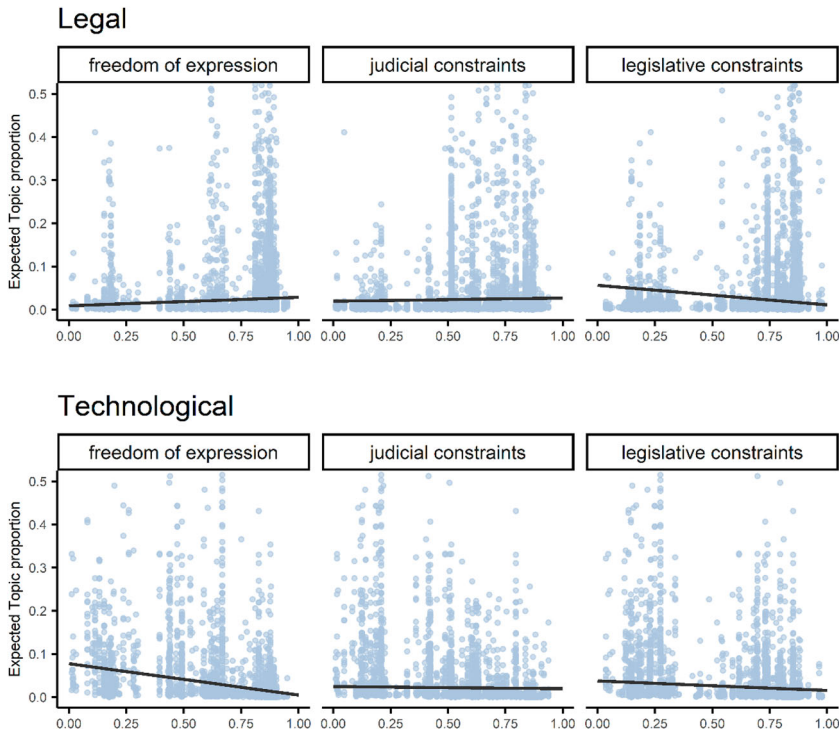

Figure 2. Marginal effects of covariates on expected topic proportions.

Note: Marginal effects are calculated using the ggpredict() function (Lüdecke, 2021); points indicate the expected topic proportion per text; in contrast to the statistical models (see Appendix D), the marginal effects are calculated using models in which the variables are not logarithmized and/or standardized to facilitate interpretation.

constraints are positively associated with the expected proportion of technological frames (B = 0.27, SE = 0.08). Overall, this suggests that only legislative but not judicial constraints are negatively associated with frames of technological regulation and lends mixed support for Hypothesis 2b. Furthermore, legislative constraints are negatively associated with legal frames (B = −0.37, SE = 0.09), whereas judicial constraints are positively associated with legal frames (B = 0.17, SE = 0.08). Overall, this provides mixed support for Hypothesis 2a, suggesting that in contrast to judicial constraints, legislative constraints are not necessarily associated with the reporting of legal strategies.

4.4. Discussion & limitations

Our investigation suggests that both traditional restrictions on press and media freedom as well as institutional constraints influence the salience of different regulatory strategies as reported in news outlets. Given that our indicators for different regulatory strategies are informed by computer-assisted text analysis rather than in-depth analysis of all articles in the corpus, we discuss our findings in light of country-specific examples.

The salience of regulatory frames of technological approaches to address fake news and hate speech appears to be higher in countries that traditionally restrict press and media freedom and that are less constrained by legislative institutions. This is also reflected by the growing trend across authoritarian African rulers to block internet access during elections (Freyburg & Garbe, 2018; Garbe, 2020). Indeed, our findings suggest that those rulers who have strong incentives to prevent the production and spread of content – often because it is considered potentially harmful to the regime – are more likely to use technological means to restrict access to and production of online content. Our findings further suggest that strong legislative constraints on the executive might prevent governments from using technological means of blocking. On the other hand, our findings indicate that regimes with strong(er) judicial constraints on the executive may still revert to technological strategies of content regulation. This is exemplified by the shutdown of social media in Zimbabwe amid protests in 2019 which was later challenged by Zimbabwe's high court (Asiedu, 2020). Given that governments who seek to fundamentally restrict access to and production of content online often do so in response to pressing political issues, *ex-post* judicial review of such measures might not deter governments from doing so. While some observers recognize the legitimate aim to contain the spread of fake news (Madebo, 2020), there is also widespread concern about the potential harm of such preventive measures in over-censoring potentially important information such as news related to Covid-19 (Nanfuka, 2019).

Our findings further indicate that legal approaches to regulating fake news or hate speech, i.e., media coverage of the introduction of bills, laws, and acts, are not more prevalent in those regimes with strong legislative constraints. This may reflect the increasing importance of law-making as a political tool of power consolidation and illiberal practices, also known as 'autocratic legalism' (Scheppele, 2018, p. 548). In fact, many African rulers started introducing legislation on the production and spread of content online. Kenya's Computer Misuse and Cybercrime Act, for instance, criminalizes the 'publication of false information in print, broadcast, data or over a computer system' (2018, Art 22, 23) and also explicitly refers to the publication of 'hate speech'. Digital human rights defenders have observed many of the changes in the

legal landscape in both authoritarian and democratic countries with worry. Regulations specifically criminalizing online content that is regarded as misinformation or hate speech are often 'inherently vague, and […] create a space for abuse of the law to censor speech' (Taye, 2020). While Helm and Nasu (2021) argue that criminal sanctions can be an effective way to counter hate speech, they underline that it is necessary to find an appropriate balance between censoring information and respecting freedom of expression. In addition to bills criminalizing the publication or spread of fake news, authoritarian regimes also seem to develop more indirect means of legislation that can be described as *ex ante* measures to prevent the production of fake news and hate speech. For instance, Tanzania, Lesotho, and Uganda, all introduced laws that indirectly prevent people from sharing content online either through fees on social media use itself or fees that are required from online bloggers (Karombo, 2020). Overall, our STM approach is limited in grasping more nuanced types of legislation and further qualitative work is needed to better understand how regimes differ in their legislative approach to regulating fake news and hate speech and to what extent legislatures affect this process. In addition, the increasing use of bots by African governments can also be seen as regulatory strategy in itself requiring more fine-grained approaches to investigate differences across countries (Nanfuka, 2019).

We acknowledge that there might be non-state solutions to regulation as, for instance, reflected in Topic 14 (see Appendix C). While governments appear to be the most prevalent actors emerging from our analysis of media coverage on hate speech and fake news in Africa, news reporting also points to other approaches, such as bottom-up initiatives to improve fact-checking skills, to regulate fake news and hate speech. This might reflect the fact that, facing increasing pressure on fundamental rights, civil society in Africa is advocating for online platforms to meaningfully invest in content moderation in Africa and collaborate with local civil society (Owono, 2020; Dube et al., 2020).

Finally, we want to highlight three limitations of our study. First, our approach using news coverage of African countries enabled us to assess the public discourse surrounding the regulation of fake news and hate speech. This has the advantage that we also include discussions about the regulation of fake news and hate speech, often before they translate into actual legislation. However, it is unclear to what extent news reports reflect actual regulation across African countries. As our validation highlights, news reports provide a good indication of technological approaches to regulate fake news (see Appendix D). Yet, it is less clear how well news reports reflect legal approaches to content regulation. The fact that most data sources on African legislation do not directly assess the extent to which legislation regulating the digital space is meant to address fake news, makes it difficult to validate the fit of news reports. Empirical studies comparing actual laws explicitly addressing fake news as well as technological manipulation of online activity beyond shutdowns are hence encouraged. Second, our sample is biased towards large African countries and countries with a high degree of digitalized press, like South Africa and Nigeria, which represent up to 20 and 40 percent in our sample respectively. As both are prominent and dominant countries on the continent, however, we can assume that they are important actors in both driving and shaping discussions on how online content should be regulated. Third, the salience of the two strategies is subject to variation over time (see Appendix F) and highlights that especially legal frames have only recently gained importance in the African context.

5. Conclusion

Our study contributes to the growing discussion on content regulation in two ways. Theoretically, we add to the understanding of online regulation by showing that regime-specific characteristics can alter a government's choice of different regulatory strategies. Empirically, we find that public discourse on online content regulation in Africa points to the relevance of technological and legal strategies pursued by governments. Discourse on technological approaches to content regulation is more prominent in coverage of countries with lower levels of media and press freedom and legislative constraints. Our analysis further suggests that legal frames are more dominant in coverage of countries with judicial constraints, but not in coverage of countries with more legislative constraints. More qualitative insights suggest that criminalization is among the dominant legal strategies. While criminal regulation can be an effective strategy to counter hate speech and fake news, Helm and Nasu (2021, p. 327) also warn about the potential for abuse of laws to supress dissent in more authoritarian regimes.

Overall, our analysis points to the central actors when it comes to content regulation in Africa: African governments. While theory has so far tended to either follow Lessig (1999, 2006) and focus on content regulation in democracies, or to focus on censorship in authoritarian regimes (Stoycheff et al., 2020; Keremoğlu & Weidmann, 2020), our analysis demonstrates that these issues cannot always be easily separated. News reporting on African countries underlines that all regimes face issues of fake news and hate speech and seek to find solutions to manage them. While technological strategies to address fake news and hate speech (including shutting down internet and blocking specific content) appear to be more prominent in regimes with low respect for media and press freedom and fewer institutional constraints, our results indicate that the same regimes may also revert to more legal means to regulate content. Not only technological but also legal strategies of content regulation may have severe implications for freedom of speech (Helm & Nasu, 2021), especially, but not only, in countries facing weak institutional constraints. Overall, our paper highlights that the regulation of fake news and hate speech are also pressing issues beyond the Western world. In turn, the current prevalence of state regulation addressing problems of fake news and hate speech points to a need to strive for multi-stakeholder approaches across continents.

Note

1. In preparing for the textual analysis, the body of textual data was properly pre-processed, white space, punctuation, and so-called 'stopwords' (the, is, are, etc.) were removed, as well as the search terms we used to delimit our body of texts (Benoit et al., 2018). Our full script for importing, preprocessing and analysing the corpus is available on GitHub: https://github.com/lisagarbe/ContentRegulationAfrica.

Acknowledgements

thank the research groups at the Comparative Politics Department at the University of St.Gallen and the Department of Administration and Organization Theory of the University of Bergen, as well as the two anonymous reviewers for their constructive and encouraging comments.

Author contributions

All authors contributed equally to the design and implementation of the research, to the analysis of the results and to the writing of the manuscript.

Data availability statement

The data that support the findings of this study are available from the Factiva Global News Monitoring & Search. Restrictions apply to the availability of these data, which were used under license for this study. R scripts for data preparation and analysis are available from the authors in the Github repository https://github.com/lisagarbe/ContentRegulationAfrica

Notes on contributors

Lisa Garbe is a postdoctoral researcher at the WZB Berlin Social Science Center. Her research focuses on inequalities in internet provision and use with a focus on authoritarian-developing contexts.

Lisa-Marie Selvik is a PhD Candidate at the Department of Comparative Politics, University of Bergen. Her work focuses on information rights, with a focus on contentious processes of advocacy on freedom of information laws in Africa.

Pauline Lemaire is a doctoral researcher at Chr. Michelsen Institute, and a PhD candidate at the Department of Comparative Politics, University of Bergen. Her research focuses on regime - youth interactions as mediated by social media.

ORCID

Lisa Garbe 🆔 http://orcid.org/0000-0001-7555-8007
Lisa-Marie Selvik 🆔 http://orcid.org/0000-0003-3442-2765
Pauline Lemaire 🆔 http://orcid.org/0000-0002-1785-151X

References

Abraha, H. H. (2017). Examining approaches to internet regulation in Ethiopia. *Information & Communications Technology Law*, *26*(3), 293–311. https://doi.org/10.1080/13600834.2017.1374057

Adegbo, E-O. (2019, November 5). The illegalities of social media regulation. *Nigerian Tribune*. Retrieved December 13, 2020 from https://tribuneonlineng.com/the-illegalities-of-social-media-regulation/

Africa Check. (2020). *Cures and prevention*. Retrieved November 9, 2020 from https://africacheck.org/spot-check/cures-and-prevention

Ahinkorah, B. O., Ameyaw, E. K., Hagan Jr, J. E., Seidu, A. A., & Schack, T. (2020). Rising above misinformation or fake news in Africa: Another strategy to control COVID-19 spread. *Frontiers in Communication*, 5, 45. https://doi.org/10.3389/fcomm.2020.00045

Asiedu, M. (2020, December 11). The Role of the Courts in Safeguarding Online Freedoms in Africa, *Democracy in Africa*. http://democracyinafrica.org/the-role-of-the-courts-in-safeguarding-online-freedoms-in-africa/

Baayen, R. H. (2008). *Analyzing linguistic data*. Cambridge University Press.

Bennett, W. L., & Segerberg, A. (2012). The logic of connective action. *Information, Communication & Society*, 15(5), 739–768. https://doi.org/10.1080/1369118X.2012.670661

Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S., & Matsuo, A. (2018). Quanteda: An R package for the quantitative analysis of textual data. *Journal of Open Source Software*, 3(30), 774. https://doi.org/10.21105/joss.00774

Blei, D. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77–84. https://doi.org/10.1145/2133806.2133826

Boas, T. C. (2006). Weaving the authoritarian Web: The control of Internet Use in nondemocratic regimes. In *J. Zysman & A. Newman (Eds.), How Revolutionary was the digital revolution? National responses, market transitions, and Global technology*. Stanford Business Books.

Bolarinwa, J. O. (2015). Emerging legislatures in Africa: Challenges and opportunities. *Developing Country Studies*, 5(5), 18–26.

Bosch, T. E., Admire, M., & Ncube, M. (2020). Facebook and politics in Africa: Zimbabwe and Kenya. *Media, Culture & Society*, 42(3), 349–364. https://doi.org/10.1177/0163443719895194

Bradshaw, S. & Howard, P. (2019). *The Global disinformation order: 2019 Global inventory of organised social media manipulation* (Working Paper 2019.2). Oxford, UK: Project on Computational Propaganda.

Breuer, A., Landman, T., & Farquhar, D. (2015). Social media and protest mobilization: Evidence from the Tunisian revolution. *Democratization*, 22(4), 764–792. https://doi.org/10.1080/13510347.2014.885505

Bunce, M. (2016). The international news coverage of Africa. In M. Bunce, S. Franks, & C. Paterson (Eds.), *Africa's media image in the twenty-first century: From the 'heart of darkness' to 'Africa rising* (pp. 17–29). Routledge.

Christensen, D., & Weinstein, J. M. (2013). Defunding dissent: Restrictions on Aid to NGOs. *Journal of Democracy*, 24(2), 77–91. https://doi.org/10.1353/jod.2013.0026

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Altman, D., Bernhard, M., Steven Fish, M., Glynn, A., Hicken, A., Luhrmann, A., Marquardt, K. L., McMann, K., Paxton, P., Pemstein, D., Seim, B., Sigman, R., Skaaning, S.-E., Staton, J., … Ziblatt, D. (2020). *V-Dem* [country–year Dataset v10]. *Varieties of Democracy (V-Dem) Project*, https://doi.org/10.23696/vdemds20

DiMaggio, P., Nag, M., & Blei, D. (2013). Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of U. S. Government arts funding. *Poetics*, 41(6), 570–606. https://doi.org/10.1016/j.poetic.2013.08.004

Dresden, J. R., & Howard, M. M. (2016). Authoritarian backsliding and the concentration of political power. *Democratization*, 23(7), 1122–1143. https://doi.org/10.1080/13510347.2015.1045884

Dube, H., Simiyu, M. A., & Ilori, T. (2020). *Civil society in the digital age in Africa: identifying threats and mounting pushbacks*. Centre for Human Rights, University of Pretoria and the Collaboration on International ICT Policy in East and Southern Africa (CIPESA).

Eloff, H. (2019, October 19). *ISPs could go to jail for failing to act against racism, hate speech and child pornography*. Randburg Sun. retrieved December 14, 2020 from https://web.archive.org/

web/20191019135116/https://randburgsun.co.za/377900/isps-go-jail-failing-act-racism-hate-speech-child-pornography/

Eltantawy, N., & Wiest, J. B. (2011). Social media in the Egyptian revolution: Reconsidering resource mobilization theory. *International Journal of Communication*, 5, 1207–1224. https://ijoc.org/index.php/ijoc/article/view/1242

Ethiopian News Agency. (2019). Draft Law Proposes Up to 5 Years Imprisonment for Hate Speech, Disinformation. Retrieved December 13, 2020 from https://www.ena.et/en/?p=8166

Facebook. (n.d). *Content Restrictions Based on Local Law*. Retrieved November 26, 2020, from https://transparency.facebook.com/content-restrictions

Field, A. (2009). *Discovering statistics using SPSS: And sex drugs and rock 'n' roll* (2nd ed.). Sage.

Freedom House. (2020). *Freedom on the Net: The Pandemic's Digital Shadow*.

Freedom House. (2015). *Freedom on the Net: Privatizing Censorship, Eroding Privacy*.

Freedom House. (2016). *Freedom on the Net: Silencing the Messenger: Communication Apps under Pressure*.

Freedom House. (2017). *Freedom on the Net: Manipulating Social Media to Undermine Democracy*.

Freedom House. (2018). *Freedom on the Net: The Rise of Digital Authoritarianism*.

Freedom House. (2019). *Freedom on the Net: The Crisis of Social Media*.

Freyburg, T., & Garbe, L. (2018). Blocking the bottleneck: Internet shutdowns and Ownership at election times in Sub-Saharan Africa. *International Journal of Communication*, 12, 3896–3916. https://ijoc.org/index.php/ijoc/article/view/8546

Freyburg, T., Garbe, L., & Wavre, V. (2021). *The political power of internet business: A comprehensive dataset of Telecommunications Ownership and Control (TOSCO)* [Unpublished Manuscript]. University of St.Gallen.

Frieden, R. (2015). Ex ante versus Ex post approaches to Network neutrality: A Comparative assessment. *Berkeley Technology Law Journal*, 30(2), 1561–1612. https://doi.org/10.15779/Z386Z81

Gandhi, J. (2008). *Political institutions under dictatorship*. Cambridge University Press.

Garbe, L. (2020, September 29). What we do (not) know about Internet shutdowns in Africa. *Democracy in Africa*. http://democracyinafrica.org/internet_shutdowns_in_africa/

Gilardi, F., Shipan, C. R., & Wüest, B. (2021). Policy diffusion: The issue-definition stage. *American Journal of Political Science*, 65(1), https://doi.org/10.1111/ajps.12521

Gilbert, D. (2020, September, 14). Hate Speech on Facebook Is Pushing Ethiopia Dangerously Close to a Genocide. *Vice*. Retrieved December 13, 2020 from https://www.vice.com/en/article/xg897a/hate-speech-on-facebook-is-pushing-ethiopia-dangerously-close-to-a-genocide

Goetz, A. M., & Jenkins, R. (2005). *Reinventing accountability: Making democracy work for human development*. Palgrave Macmillan.

Gorwa, R. (2019). What is platform governance? *Information, Communication & Society*, 22(6), 854–871. https://doi.org/10.1080/1369118X.2019.1573914

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis Methods for political texts. *Political Analysis*, 21(3), 267–297. https://doi.org/10.1093/pan/mps028

Hassanpour, N. (2014). Media disruption and Revolutionary unrest: Evidence from mubarak's quasi-experiment. *Political Communication*, 31(1), 1–24. https://doi.org/10.1080/10584609.2012.737439

Hellmeier, S. (2016). The dictator's digital toolkit: Explaining variation in internet filtering in authoritarian regimes. *Politics & Policy*, 44(6), 1158–1191. https://doi.org/10.1111/polp.12189

Helm, R. K., & Nasu, H. (2021). Regulatory responses to 'fake news' and freedom of expression: Normative and Empirical evaluation. *Human Rights Law Review*, 21(2), 302–328. https://doi.org/10.1093/hrlr/ngaa060

Herron, E. S., & Boyko, N. (2015). Horizontal accountability during political transition: The use of deputy requests in Ukraine, 2002–2006. *Party Politics*, 21(1), 131–142. https://doi.org/10.1177/1354068812472573

Iosifidis, P., & Andrews, L. (2020). Regulating the internet intermediaries in a post-truth world: Beyond media policy? *International Communication Gazette*, *82*(3), 211–230. https://doi.org/10.1177/1748048519828595

Karombo, T. (2020, October, 12). More African governments are quietly tightening rules and laws on social media. *Quartz Africa*. https://qz.com/africa/1915941/lesotho-uganda-tanzania-introduce-social-media-rules/

Kellam, M., & Stein, E. A. (2016). Silencing critics: Why and How presidents restrict media freedom in democracies. *Comparative Political Studies*, *49*(1), 36–77. https://doi.org/10.1177/0010414015592644

Kenya's Computer Misuse and Cybercrime Act. (2018). http://kenyalaw.org/kl/fileadmin/pdfdownloads/Acts/ComputerMisuseandCybercrimesActNo5of2018.pdf

Keremoğlu, E., & Weidmann, N. B. (2020). How dictators control the internet: A review essay. *Comparative Political Studies*, *53*(10–11), 1690–1703. https://doi.org/10.1177/0010414020912278

Lessig, L. (1999). *Code and other laws of cyberspace*. Basic Books.

Lessig, L. (2006). *Code. Version 2.0*. Basic Books.

Lüdecke, D. (2021). ggeffects: Marginal Effects and Adjusted Predictions of Regression Models. https://cran.r-project.org/web/packages/ggeffects/vignettes/ggeffects.html

Lührmann, A., Lindberg, S., & Tannenberg, M. (2017). *Regimes in the World (RIW): A Robust Regime Type Measure Based on V-Dem*. https://ssrn.com/abstract=2971869

Madebo, A. (2020, September 29). Social media, the Diaspora, and the Politics of Ethnicity in Ethiopia, *Democracy in Africa*. http://democracyinafrica.org/social-media-the-diaspora-and-the-politics-of-ethnicity-in-ethiopia/

Mail & Guardian. (2018, May, 16). Kenya signs bill criminalising fake news, retrieved December 13, 3030 from https://mg.co.za/article/2018-05-16-kenya-signs-bill-criminalising-fake-news/

Mechkova, V., Lührmann, A., & Lindberg, S. I. (2019). The accountability sequence: From De-jure to De-facto constraints on governments. *Studies in Comparative International Development*, *54*(1), 40–70. https://doi.org/10.1007/s12116-018-9262-5

Michaelsen, M. (2018). Exit and voice in a digital age: Iran's exiled activists and the authoritarian state. *Globalizations*, *15*(2), 248–264. https://doi.org/10.1080/14747731.2016.1263078

Nanfuka, J. (2019, January 31). Social Media Tax Cuts Ugandan Internet Users by Five Million, Penetration Down From 47% to 35%, *CIPESA*. https://cipesa.org/2019/01/%ef%bb%bfsocial-media-tax-cuts-ugandan-internet-users-by-five-million-penetration-down-from-47-to-35/

Nijzink, L., Mozaffar, S., & Azevedo, E. (2006). Parliaments and the enhancement of democracy on the African continent: An analysis of institutional capacity and public perceptions. *The Journal of Legislative Studies*, *12*(3–4), 311–335. https://doi.org/10.1080/13572330600875563

Nothias, T. (2020). Access granted: Facebook's free basics in Africa. *Media, Culture & Society*, *42*(3), 329–348. https://doi.org/10.1177/0163443719890530

Obijiofor, L., & MacKinnon, M. (2016). Africa in the Australian press: Does distance matter? *AfricaN Journalism Studies*, *37*(3), 41–60. https://doi.org/10.1080/23743670.2016.1210017

Owono, J. (2020, October 5). Why Silicon Valley needs to be more responsible in Africa. *Democracy in Africa*. http://democracyinafrica.org/why-silicon-valley-needs-to-be-more-responsible-in-africa/

Rakner, L., & van de Walle, N. (2009). Opposition parties and incumbent presidents: The new dynamics of electoral competition in sub-Saharan Africa. In S. Lindberg (Ed.), *Democratization by elections*. Johns Hopkins University Press.

Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). Stm: An R package for structural topic models. *Journal of Statistical Software*, *91*(1), 1–40. https://doi.org/10.18637/jss.v091.i02

Rose-Ackerman, S. (1996). Democracy and 'grand' corruption. *International Social Science Journal*, *48*(149), 365–380. https://doi.org/10.1111/1468-2451.00038

Rydzak, J., Karanja, M., & Opiyo, N. (2020). Dissent does Not Die in darkness: Network shutdowns and collective action in African countries. *International Journal of Communication*, *14*, 24. https://ijoc.org/index.php/ijoc/article/view/12770 https://doi.org/10.46300/9107.2020.14.5

Schedler, A. (2002). Elections without democracy: The menu of manipulation. *Journal of Democracy*, *13*(2), 36–50. https://doi.org/10.1353/jod.2002.0031

Scheppele, K. L. (2018). Autocratic legalism. *The University of Chicago Law Review*, *85*(2), 545–584.

Shen-Bayh, F. (2018). Strategies of repression: Judicial and extrajudicial Methods of Autocratic survival. *World Politics*, *70*(3), 321–357. https://doi.org/10.1017/S0043887118000047

Solomon, S. (2017). Ethiopia: As Violence Flares, Internet Goes Dark, *AllAfrica*. https://allafrica.com/stories/201712150414.html

Stier, S. (2015). Democracy, autocracy and the news: The impact of regime type on media freedom. *Democratization*, *22*(7), 1273–1295. https://doi.org/10.1080/13510347.2014.964643

Stoycheff, E., Burgess, G. S., & Martucci, M. C. (2020). Online censorship and digital surveillance: The relationship between suppression technologies and democratization across countries. *Information, Communication & Society*, *23*(4), 474–490. https://doi.org/10.1080/1369118X.2018.1518472

Taye, B. (2020, October 22). Internet censorship in Tanzania: the price of free expression online keeps getting higher, *Access Now*. https://www.accessnow.org/internet-censorship-in-tanzania/

Teorell, J., Coppedge, M., Lindberg, S., et al. (2019). Measuring Polyarchy Across the Globe, 1900–2017. *Studies in Comparative International Development*, *54*, 71–95. doi:10.1007/s12116-018-9268-z.

The Botswana Gazette. (2017, October 12). Kgathi's Cyber Bullying Law: Intrusion On Privacy Rights, Freedom Of Expression? Retrieved December 13, 2020 from https://www.thegazette.news/news/kgathis-cyber-bullying-law-intrusion-on-privacy-rights-freedom-of-expression/20861/

The Citizen. (2018, October 7). Internet shutdown fears as Cameroonians elect president. Retrieved December 12, 2020 from https://www.thecitizen.co.tz/tanzania/news/africa/internet-shutdown-fears-as-cameroonians-elect-president–2658018

Twitter. (n.d.). *Removal requests*. Retrieved November 26, 2020, from https://transparency.twitter.com/en/reports/removal-requests.html#2019-jul-dec

Vondoepp, P. (2005). The problem of judicial control in Africa's neopatrimonial democracies: Malawi and Zambia. *Political Science Quarterly*, *120*(2), 275–301. http://www.jstor.org/stable/20202519 https://doi.org/10.1002/j.1538-165X.2005.tb00548.x

Walker, S., Mercea, D., & Bastos, M. (2019). The disinformation landscape and the lockdown of social platforms. *Information, Communication & Society*, *22*(11), 1531–1543. https://doi.org/10.1080/1369118X.2019.1648536

Zuckerberg, M. (2016, November 18). A lot of you have asked what we're doing about misinformation, so I wanted to give an update. The bottom line ... Facebook. https://www.facebook.com/zuck/posts/10103269806149061