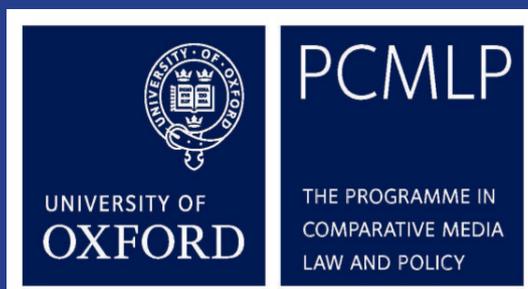


MECHACHAL:

ONLINE DEBATES AND ELECTIONS IN
ETHIOPIA - FROM HATE SPEECH
TO ENGAGEMENT IN SOCIAL MEDIA

MAIN FINDINGS





Claims that social media are increasingly being used to disseminate hate speech and incite violence often fail to build on comprehensive and publicly accessible empirical evidence. This makes it difficult to place the most extreme forms of expression into context, understand how pervasive they are, and determine how they spread.

This briefing paper summarizes the findings of a two-year study led by the University of Oxford and Addis Ababa University under the name of “Mechachal” (tolerance in Amharic) that has sought to address this gap and to understand both the nature and the prevalence of hate speech in social media. The study focuses on Ethiopia, and on the social networking platform, Facebook, but the tools that were developed and the evidence that was collected can contribute to larger debates on freedom of expression, extremism, and political contestation.

What is hate speech?

There is little consensus as to what qualifies as hate speech. As other empirically grounded research suggests,¹ most individuals, from ordinary Internet users to politicians, tend to adopt definitions that are broader than those included in international and regional

norms.² The framework we developed sought to accommodate this diversity and identified a continuum ranging from: statements that **go against** individuals or groups on any ground (e.g. politics, ethnicity, religion) by challenging, provoking, or explicitly threatening them; statements that are **offensive** on the base of ethnicity, nationality, religion, and gender (this definition thus excludes statement that go against purely on political grounds); **hate speech**, defined as speech that does not just offend, but **incites** others to discriminate or act against individuals or groups based on ethnicity, nationality, religion, and gender; and **dangerous speech**, text or images that build the bases for or directly call for widespread violence against a particular group (all definitions are discussed with greater detail in the main report).

In an effort to avoid the risk shared by most research on hate speech online to excessively emphasise the “darker side” of Internet communication because of the focus on extreme forms of expression, we also mapped statements that **go towards**. These are statements that help initiate, maintain, and/or build a communicative relationship. They can also contain strong criticism, but they offer at least a premise for recognizing adversaries as legitimate, rather than simply dismissing them. Taken together, these definitions allow

for the mapping of a broad spectrum of what is communicated on social media, placing different forms of expression, including the most extreme, into context.

What did we study?

Mechachal's team (which included researchers fluent in Amharic, Tigrigna, and Oromiffa) analyzed more than 13,000 statements posted on 1055 Facebook pages by Ethiopians in Ethiopia and in the Ethiopian diaspora. The findings presented here are based on

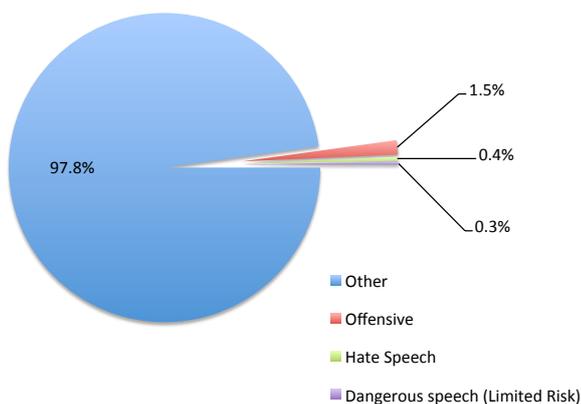
statements collected between 24 February and 24 June 2015, three months before and one month after the parliamentary elections that took place in Ethiopia on 24 May 2015.

Below we offer a list of our main findings, which are further discussed in the main **report** ("MECHACHAL: Online debates and elections in Ethiopia - From hate speech to engagement in social media. Final Report").

FINDINGS

FINDING 1 - HATE AND DANGEROUS SPEECH ARE MARGINAL FORMS OF SPEECH IN SOCIAL MEDIA

Only **0.4%** of statements in our sample have been classified as hate speech (i.e. speech that incites others to discriminate or act against individuals or groups based on their ethnicity, religion, or gender) and **0.3%** as dangerous speech (i.e. speech that builds the bases for or directly calls for widespread violence against a particular group). In addition, all statements in this latter category were associated with a limited possibility for the speakers (or the groups they appeal to) to actually carry out violence.



Only a minute portion of the statements analysed can be considered offensive, hate or dangerous speech (with limited risk) (FINDING 1)

FINDING 2 - THE 2015 ELECTIONS ON FACEBOOK WERE A "NON-EVENT" ON SOCIAL MEDIA

Most Ethiopian Facebook pages discussed the elections, but many statements either directly referred to, or seemed informed by, the perception that the outcome of the elections was already predetermined. This contributed to a general perception that the electoral contest was a "non-event". There is little evidence that new social or political tensions emerged, or was heightened, around

the elections. No single group was singled out as a target for antagonism. Antagonistic speech was mostly directed against the elections themselves, which were largely dismissed as a futile process.

FINDING 3 - DANGEROUS SPEECH IS A DISTINCTIVE AND MORE DELIBERATIVE FORM OF ATTACKING OTHER GROUPS OR INDIVIDUALS

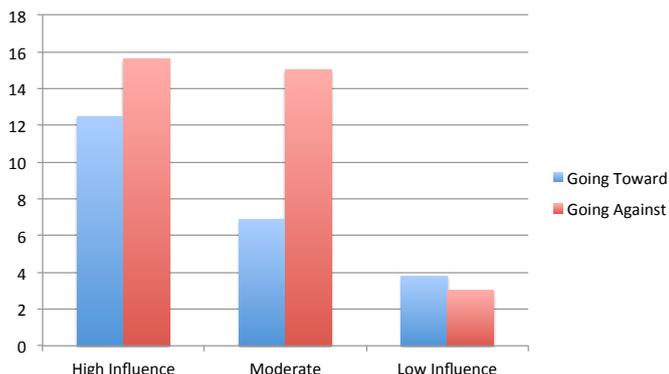
Dangerous speech reflects a more deliberate strategy to attack individuals and groups. Almost all dangerous statements in our sample are uttered by individuals seeking to hide their identity (92%). This proportion is significantly lower for statements classified as hate speech (33%) and offensive speech (31%). Shifting the focus from speakers to targets, dangerous statements appear to focus exclusively on ethnicity. The salience of ethnicity can be found also in other types of messages. 75% of hate speech and 58% of offensive speech have ethnic targets, but they also target individuals based on their religion, and, to a much smaller extent, their sexual identity.

FINDING 4 - THE POLITICAL, SOCIAL AND CULTURAL VIEWS REFLECTED IN SOCIAL MEDIA IN ETHIOPIA ARE LESS POLARIZED THAN MIGHT BE EXPECTED

The relative marginality of hate and dangerous speech is part of a broader trend that involves antagonistic statements more generally. If we zoom out from the sub-categories of offensive, hate, and dangerous speech to the broader category that includes all statements going against, the evidence indicates that, in this case too, they represent a minority on Facebook (only 16% of statements go against). This may not be surprising, especially when we consider that the majority of users do not log in to social media simply to attack others. This is significant, however, in the context of the political debate in Ethiopia, where media have historically been highly polarized.

FINDING 5 - WHILE STATEMENTS GOING AGAINST ARE FEW, THEY PRODUCE MORE REACTIONS AND TRAVEL FURTHER THAN OTHER STATEMENTS

Although statements going against are a minority, and are more likely to be uttered by individuals with little influence, they still have the power to shape debates. Once a certain threshold of visibility is passed, it appears that it is the content of the statements that matters, rather than the influence of the speakers. Framed in different terms, it can be argued that while statements going against are less frequent, they are more likely to promote reactions and travel further.



As the figure suggest, registering the comments each type of statements produce, once a certain threshold of visibility is passed, it appears that it is the content of the statements that matters, rather than the influence of the speakers (FINDING 5)

FINDING 6 - UNEQUAL POWER RELATIONSHIPS INCREASE ANTAGONISM BETWEEN SPEAKERS AND FOLLOWERS

Users with little or no influence tend to post more statements going against (18% of the total of statements analyzed) than highly influential speakers (11%). For example, if we consider some of most influential pages in Ethiopia’s online sphere, a large gap exists between the tone of the posts written by the owners of these pages and the tone of the comments written by their “fans/followers”. While posts tend to display a smaller proportion of antagonism, the percentage of comments going against tends to be significantly higher than in the general sample.

FINDING 7 - BROADCASTING VIEWS DOMINATES OVER DIALOGUE IN POLITICAL COMMUNICATION ON FACEBOOK

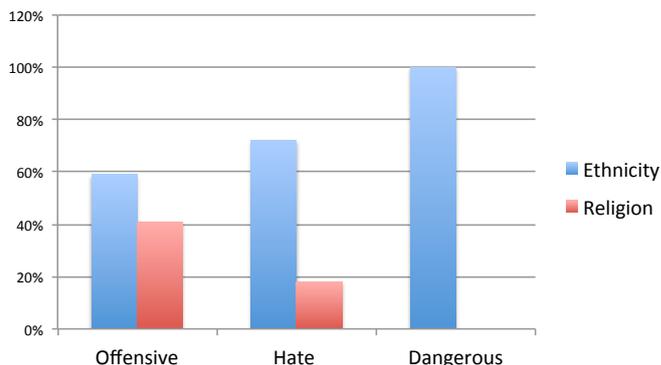
Most institutional pages tend to adopt a top-down broadcasting communication style, with very little interaction between page owners and page followers, and even less interaction with other Facebook pages, including those sharing the same political agenda. There are, however, exceptions such as the pages of some influential speakers that have become spaces for audiences of different political affiliations to engage in open discussions.

FINDING 8 - THE ETHIOPIAN FACEBOOK COMMUNITIES ARE NOT ECHO-CHAMBERS

Ethiopian Facebook users tend to form communities by coalescing around a few online pages of their choice, whether to express opinions in agreement with the page owner, or, in some cases, to dissent and criticize the page’s political line. Online political communities are not necessarily divided by political affiliation, and do not serve as “echo-chambers”, or spaces that only expose users to ideas with which they already sympathize with. While users may be attracted to specific spaces because of the issues they expect to be discussed, they also seem to steer these discussions into different directions if necessary.

FINDING 9 - DISCUSSIONS ABOUT RELIGION IN SOCIAL MEDIA ARE THE LEAST ANTAGONISTIC

If we divide the statements we sampled on Facebook into broad categories, such as politics, religion, and ethnicity, some relatively distinct patterns emerge. Discussions on religion are those displaying the lowest percentage of statements going against, with only 10.5% falling into this category. This finding is particularly relevant in the context of perceived growing divides among the major faiths in Ethiopia, and of a politicization of religious debates, including the forced closure of several papers that serve the Muslim community.



Statements focusing on religion contain the least amount of statements that could be considered offensive, hate or dangerous speech (FINDING 9)

FINDING 10 - NEW OPPORTUNITIES FOR ENGAGEMENT EXIST WHEN A MORE UNIVERSAL AGENDA IS EMBRACED

Over the past twenty years, a high level of polarization has characterized Ethiopia’s media. Conversations on social networking platforms have inherited some of the tones that characterized previous debates; but the strongest forms of antagonism have remained relatively marginal. The greatest opportunities for engagement seem to lie in politics that appeal to universal values. For example, despite attracting criticism around the status of freedom of expression in Ethiopia, both at the national and international level, the campaign for the release of the Zone9 bloggers was largely conducted without antagonism and divisions.

Conclusion

The findings suggest how Ethiopians' discussions on Facebook are much more nuanced than commonly perceived. Despite the polarized tones that characterize some political debates, social media also offer opportunities to experiment with new forms of engagement. As it is the case in many other societies, hate speech and dangerous speech do occur on these platforms, but they occupy a marginal proportion of the total online conversations.

While this is a new area and more research is needed, our project is the first academically grounded attempt to analyse online conversations focusing on Ethiopia. It also dispels the common claims and assumptions that social media are used by influential people to support or catalyze violence in the country.

If we shift the focus from anecdotal evidence, or individual cases, to a broader sample of Facebook pages, it becomes clear that the most violent and aggressive forms of speech come from individuals with little or no influence voicing their anger towards power.

These findings, which have been extensively debated with a variety of political actors representing both government and opposition parties, as well as academics, journalists and bloggers in Ethiopia, should be taken into account when political figures are accused of using social media in ways that are divisive or subversive. The research findings can also offer a base for developing a comparative and evidence-based understanding of how hate speech emerges and spreads in different national contexts in Africa, and beyond.

NOTES

1. For discussions of international and national norms regulating hate speech see, for example, James Banks, "Regulating Hate Speech Online," *International Review of Law, Computers & Technology* 24, no. 3 (2010): 233–39; Iginio Gagliardone et al., *Countering Online Hate Speech* (UNESCO Series on Internet Freedom. United Nations Educational, Scientific and Cultural Organisation, Paris, 2015).
2. For discussions of international and national norms regulating hate speech see, for example, James Banks, "Regulating Hate Speech Online," *International Review of Law, Computers & Technology* 24, no. 3 (2010): 233–39; Iginio Gagliardone et al., *Countering Online Hate Speech* (UNESCO Series on Internet Freedom. United Nations Educational, Scientific and Cultural Organisation, Paris, 2015).

The mechachal team

We are grateful to the British Embassy in Addis Ababa for their support in enabling us to envisage and realize this project. The views expressed in this report, however, are those of the authors and should not be attributed to the British Embassy or the British government.

For more information about the report, please contact the authors individually.

Iginio Gagliardone is Lecturer in Media and Communication at the University of the Witwatersrand and Associate Research Fellow in New Media and Human Rights at the University of Oxford. His research has focused on the relationship between new media, political change, and human development and on the emergence of distinctive models of the information society in the Global South. He holds a PhD from the London School of Economics and Political Science.

Matti Pohjonen is Research Fellow for VOX-POL, a European Union Framework Programme 7 (FP7)-funded academic research network focused on researching the prevalence, contours, functions, and impacts of Violent Online Political Extremism and responses to it. His work focuses on developing comparative and practice-based research approaches to understand digital cultures in the developing world. Previously he worked as an AHRC post-doctorate and a Teaching Fellow in Digital Culture at the School of Oriental and African Studies (SOAS).

Abdissa Zerai is Assistant Professor at the School of Journalism & Communication, Addis Ababa University, where he focuses on the political economy of the Ethiopian media and ICT in the context of a democratic developmental state. He has been working on issues related to conflict reporting, political economy of communication, and the nexus between media, democracy & civil society.

Zenebe Beyene is Assistant Professor of Journalism and Communication, Director of Office of External Relations, Partnerships and Communications at Addis Ababa University, Ethiopia and NAFSA Global Dialogue Fellow. Dr. Zenebe has taught in Ethiopia, Rwanda and the U.S.A. His publications include Media use and abuse in Ethiopia, the role of ICT in peace building, state building and governance in Africa (with Abdissa Zerai), and Satellites, Plasmas and Law (with Abdissa Zerai and Iginio Gagliardone).

Gerawork Aynekulu is reading for a MSc in computer science at University of Belgrade, where he focuses on data mining. He has been working on text analytics of online Amharic textual resources.

Jonathan Bright is Research Fellow at the Oxford Internet Institute, which is a department of the University of Oxford. He is also an editor of the journal Policy and Internet. He is a political scientist specializing in political communication, digital government and computational social science.

Mesfin Awoke Bekalu is Research Fellow at Harvard T.H. Chan School of Public Health, USA. Prior to his current post, he has been a postdoctoral fellow at the University of Leuven, Belgium and a lecturer in Journalism and Communications at Bahir Dar University, Ethiopia. His research interests include communication inequalities in the areas of health and development as well as media discourse analysis.

Michael Seifu is an independent researcher based in Ireland and has completed a PhD in politics from Dublin City University. He has been working on issues related to the politics of economic development.

Mulatu Alemayehu Moges is PhD candidate in the University of Oslo, at the Department of Media and Communication, where he focuses on conflict reporting in the Ethiopian media. He has been working as a Journalist in Ethiopian media, and as Lecturer and Researcher in Addis Ababa University, School of Journalism and Communication.

Nicole Stremlau is Head of the Programme in Comparative Media Law and Policy, Centre for Socio-Legal Studies, University of Oxford. Her research focuses on media and conflict in the Horn of Africa. She has worked extensively in Ethiopia, Somaliland/Somalia, Uganda and Kenya. As Head of PCMLP, she also directs the Price Media Law Moot Court Programme and co-directs the Annenberg-Oxford Media Policy Summer Institute.

Patricia Taflan is Research Assistant at the Programme in Comparative Media Law and Policy, University of Oxford. She completed an MSc in Criminology and Criminal Justice, also at the University of Oxford, where she focused on online hate crime.

Tewodros Makonnen Gebrewolde is PhD candidate at the University of Leicester, where he focuses on productivity growth and industrial policy. He has been working on issues related to economic growth and development of the Ethiopian Economy.

Zelalem Mogessie Teferra is PhD candidate in International Law at the Geneva Graduate Institute of International and Development Studies. His study focuses on the intersection between national security and the humanitarian norms of international law. He was previously an Instructor of Law at Jimma University (Ethiopia), Michigan Grotius Scholar in University of Michigan (U.S.A).