
Monitoring online Hate Speech in South Sudan (Oct 2015 – July 2016)

Hastings Lemi Surur¹, Moses Bul² & Isaac chool³

¹Lead Researcher, ^{2,3}Data Coder

iHub Research – Umati Project Kenya
Search for Common Ground South Sudan

Abstract: *The ethnic and political rivalries and conflicts in South Sudan are extremely complex and have spilled over to the social media. The hostilities/violent communication on social media have poisoned the social and political environment. These unhealthy online conversations/hostilities that are mostly initiated by South Sudanese internet users in Diaspora sometimes are reported to be spilling over to real life situations in South Sudan. Search for Common Ground (SFCG) in its pursuit to build sustainable peace in South Sudan at all levels funded this study through iHub Research Umati with the view of building stronger national platform for diverse, constructive and non-violent dialogues in South Sudan. Thus, the study as part of SFCG's "I Love My Country": Strategic Communications for Peace building in South Sudan project under the research leadership of iHub Research Umati involved monitoring hate speech in the Republic of South Sudan cyber space from Oct 2015 to July 2016. In a nutshell, the study has two-fold aims. First, to identify, monitor and map speech that could catalyze violence in the South Sudan social media space. Second, it is to disseminate the research findings and discuss response strategies for constructive peace messages against hate speech in South Sudan within the project.*

The research was guided by three paradigms. The first paradigm is the concept of 'hate speech'. What exactly constitutes hate speech? The first paradigm – which is within Susan Benesch's (2012) framework, perceives hate speech; one, as insults targeted at a group of people and not a single person; two, as a speech that compares a group of people with animals, insects or vermin; and finally, as a speech that calls the audience to a violent action. A very important point to note is that within this 'hate speech definition framework', there are three categories of hate speech that the research team used to code the texts. Texts that are merely insults against a particular group (tribe, race, sex, religion etc.) fall under the 'offensive category'. Texts that compare people with animals and invoke fear in the audience fall under the

'moderately dangerous category'. Texts that call audience to a violent action fall under the 'extremely dangerous category'. To determine the dangerousness of the speech, Benesch's (2012) five-point analytical tool guided the coding process; the 5 elements considered were the level of a speaker's influence; the grievances or fears of the audience, that is, whether or not the speech act is understood as a call to violence; the social and historical context; the language used, and the medium of dissemination. The second paradigm is the context in which a meaning is formed. The third paradigm is the platform for mapping hate speech, which in the case of this project are the Internet platforms (online newspapers, Facebook, YouTube etc.).

Data was collected from 1088 internet users and couple of social networking sites during the 10-month research period and coded and arranged in themes and then analyzed using spreadsheet.

Some keywords of Hate Speech online: Nyagat¹, Nyamnyam², Aryan Jenge³, stupid cockroaches of Equatoria, Jenge Council of Awirin (Awirin is Arabic word for fools)⁴, Dinkacracy⁵, Animal kingdom, Cow kingdom, MTN⁶

1.0 Background to the Study

The current crisis in South Sudan was triggered by power struggle within the Sudan People's Liberation Movement (SPLM) and indeed represents fundamental political and ethnic tensions that date back to the Sudan People's Liberation Movement and Sudan People's Liberation Army (SPLA) split in the early nineties. In August 1991, a political upheaval within the SPLM/A led to the formation of a breakaway Nasir faction⁷ under the leadership of Dr. Riak Machar. Since then the relationship between members of the SPLM/A mainstream and the members of SPLM/A Nasir faction - both who got united and lived under one political house prior to the signing of the Comprehensive Peace Agreement - has been plagued by mutual mistrust. The political disputes, interpersonal rivalries and political stage

management ahead of 2015 elections were initially not based on ethnic identity. However, all these underlying elements soon culminated into an ethnic conflict when the crisis erupted on 15 December, 2013 with Dinka-on-Nuer violence reported in Juba and Nuer-on-Dinka violence reported in Jonglei, Unity and Upper Nile States. The African Union Commission of Inquiry Report on South Sudan (AUCISS) released on October 27, 2015 provides a more detailed account of the horrifying war crimes. For instance, the report of people being burnt in places of worship and hospitals, the mass burials, women of all ages being gang raped, and captured people being forced to eat human flesh is indeed disturbing (African Union, 2014).

Majority of SPLM politicians are either military personnel or people who have served in the military. As cited by Community Empowerment for Progress Organization (CEPO):

“[The] majority of Sudan People’s Liberation Movement members are militarized, political minded leaders who militarized politics and politicized military at the same time. This is a key reform that we expect to be addressed by the agreement for reunification of Sudan People’s Liberation Movement”⁸

Just like the hate propaganda broadcast in Rwanda, some radio stations were used to incite violence in South Sudan. For instance, radio Bentiu FM Station was abused by some soldiers from Sudan People’s Liberation Army in Opposition (SPLA-IO) to escalate ethnic violence where hundreds of people were killed on 16-17 April 2013. Keith Somerville (2014) in his article **“South Sudan: how hate radio was used to incite Bentiu massacres”** cited the United Nations Mission in the Republic of South Sudan (UNMISS) commenting that the rebels at times *“broadcast hate messages declaring that certain ethnic groups should not stay in Bentiu and even calling on men from one community to commit vengeful sexual violence against women from another community”*. Social media are reported by CEPO to be *“instrumental in facilitating violence”*. A study by CEPO suggests that false information may have led to violence:

“.....Facebook users looked at in the survey spread false information among communities, which had the effect of creating fear and panic, and may have led to violence. Three quarters of the accounts that were researched were found to be deeply involved in misinforming the public through fabricating facts.

Media was also manipulated to paint a false picture of events, for example an image of a Ugandan soldier believed to have been killed in Somalia was instead said to have been killed in Bor. A photo demonstrating the Lord’s Resistance Army killings in the Democratic Republic of Congo was fabricated to be a killing in Juba residential area”⁹

UNMISS (2014) highlights that the crisis displaced more than 500,000 people within South Sudan, an estimated 74,300 people crossed into neighboring countries and 78,477 Internally Displaced Persons (IDPs) having taken refuge at their PoC (Protection of Civilians) sites as of April 2014. The ethnic and political rivalries and conflicts in South Sudan are extremely complex and have spilled over to social media. The hostile and violent communication on social media has poisoned the social and political environment. These unhealthy online conversations/hostilities, which are mostly initiated by South Sudanese internet users in the diaspora, are apparently conceptualized to be aggravating the situation on the ground in South Sudan. The reaction by ‘Big Boy’ to a piece of news bulletin that read *‘Nuer community claims 17,613 killed in South Sudan violence’¹⁰* as reflected below is one of the examples of unhealthy conversations that sometimes take place on the South Sudan cyber space:

*“Dear Poor’s Nuer and Food Lover
How many dinka did you killed during the violence when you talked about you peoples being killed in Juba? Don’t cry out first you guys are cowards and food=lover so we have no doubt for that. You had killed double figures from dinka but we’re not crying like you guys”. (April 2014 08:40, by Big Boy)*

Search for Common Ground (SFCG), in its pursuit to build sustainable peace in South Sudan at all levels carried out a pilot study on online hate speech through iHub Research Umati Project. The study, as part of SFCG’s [“I Love My Country”: Strategic Communications for Peace building in South Sudan](#) project, involved *‘mapping and continual analysis of conflict drivers in social media, and a developing rapid response’* component for the project.

1.1 Objectives

The objective of this study are as follows:

- a. To identify, monitor and map speech that could catalyze violence in the South Sudan social media space

- b. To disseminate the research findings and discuss response strategies for constructive peace messages against hate speech in South Sudan

To achieve the objectives of the study, this project attempted to answer the following three questions:

1. To what extent are the conversations/messages/texts on the internet inflammatory?
2. What are the channels/techniques/tactics used by influential internet users, when engaging their audience in inflammatory speech?
3. What techniques, channels and reactions are applied by the audience when engaged/provoked by inflammatory speakers?

This study ran for 10 months from October 2015 to July 2016 and involved examining hate messages and conversations on South Sudan Internet platforms. The components that this study examined were around the five Benesch (2012) Dangerous Speech Framework that offers five key variables for assessment: the speaker, the audience, the speech act itself, the social and historical context, and the mode of dissemination.

1.2 Defining Hate Speech

Gagliardone et al (2014) in their working paper suggest that hate speech “refers to words of incitement and hatred against individuals based upon their identification with a certain social or demographic group. It may include, but is not limited to, speech that advocates, threatens, or encourages violent acts against a particular group, or expressions that foster a climate of prejudice and intolerance”. Cohen-Almagor (2011) defines it as biased-motivated, hostile, malicious speech aimed at a person or a group of people because of some of their actual or perceived innate characteristics. Precisely, the hate speech expresses discrimination on the bases of gender, race, religion, ethnicity, color, national origin, disability or sexual orientation.

<http://dictionary.reference.com/browse/hate-speech> define hate speech as a ‘speech that attacks, threatens, or insults a person or group on the basis of national origin, ethnicity, color, religion, gender, gender identity, sexual orientation, or disability’. Article 20 (2) of the International Covenant on Civil and Political Rights (ICCPR) states that any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law. Though there is no provision in

the African Charter on Human and Peoples’ Rights for banning hate speech, both Articles 2 and 9 (2) of the Charter guarantee every individual’s freedom and right to express and disseminate his opinions within the law. However, South Sudan has not yet signed off on these key human rights instruments such as the ICCPR, the International Covenant on Economic Social and Cultural Rights (ICESCR), nor the African Charter on Human and Peoples’ Rights (ACHPR).

The South Sudan Transitional Constitution of 2011 guarantees freedom of expression in Article 24, and the Bill of Rights upholds general rights and freedoms but there is no mention of hate speech. There are no laws restricting the use of mobile and online Internet content and platforms by individuals or organizations, whether in terms of accessing or publishing content; and users are not subject to sanctions for accessing or publishing sensitive or detrimental content on the Internet. Most online media enjoy a relatively high level of press freedom, particularly since most Internet websites are hosted outside of South Sudan (UNESCO, 2015). However, security personnel often intimidate and arrest journalists without arrest warrants as reflected below:

The media environment in South Sudan is often challenged with harassment, arrests, and detention of journalists by security personnel without arrest warrant, as well as direct and indirect censorship of the press and broadcast media for expression of opposition opinions.

It is within this nexus of apparent press freedom, unmerited censorship and the lack of a constitutional recognition and definition of hate speech, that this study lies. The study was to identify, monitor and map inflammatory speech in South Sudan’s online space with a view of understanding the motivations and reactions and drivers of such speech.

This study adopted the Umati Project’s narrower definition of hate speech, which is based on a three-point analytical tool. Umati’s definition provides the project with a workable framework for detecting, monitoring and responding to hate speech.

Most of the websites monitored (in addition to some in the guide of [South Sudanese credible news sources](#)) are run by South Sudanese nationals in the diaspora and provide online spaces to debate on political, social and economic issues. The sites’ comment sections are largely un-moderated, which could indicate why hate speech continues to be

found on the platforms. A key point to note is that some websites have their own control mechanism such as not publishing any content that is vulgar or obscene, as evidenced by the complaints from some commenters to the editorial team for not publishing their contributions.

2.0 Literature Review

This review focuses on having a clear understanding of forms of violent communication risks/hate speech online. An attempt has been made to cover some concepts, themes and interconnecting threads on *models and methodologies for monitoring and mapping of hate speech online*.

The project reviewed research conducted by the *Institute of Human Rights and Prevention of Extremism and Xenophobia* on hate speech in the Ukrainian cyber space. Their analysis revealed that:

“33% of respondents, who came across the displays of cyber hate, react on its presence in the same way (for instance, write a hatred comment in answer). At the same time, 21% of respondents said that at least once they were in the situation when a comment or a phrase in the Internet was so offensive that it made him/her to respond in the same way.

As for the streams of hatred themselves, we can see that hatred comments had to deal with politicians most of all, very often the authors of the comments expressed themselves against political views of their opponents. Comments that had to deal with nationality comprise around 35% of the general amount of hatred comments, as for the region of living there were 27%, as for confession – 19%”¹¹.

The study revealed that internet users from one political camp can become objects of hatred when conversing online with internet users from a different political camp. Appeals for the annihilation of the adversary group or politicians dominate these discussions.

Gagliardone et al (2014) provide insight into three key paradigms that can be gateways for understanding hate speech, and developing or adopting a methodology for monitoring hate speech online. The first paradigm is the concept of ‘hate speech’ and the circumstances under which a speech becomes dangerous and prone to inciting violence.

The second paradigm is the *context* in which a meaning is formed. For example, the chants by Wildcats in the sports game between the Notre Dame (The Fighting Irish) and the North-Western (Wildcats) that ‘kill the fighting Irish’ may suggest violence, but the context does not imply violence of any kind. However, the context in which Kenyan radio journalist Joshua Arap Sang allegedly made some statements such as ‘the war has begun’ and ‘the people of the milk’ should ‘cut the grass’- colloquial terms referring to the cattle-rearing Kalenjin (Sang’s ethnic group) and the agricultural Kikuyu (the ethnic group that was targeted by Sang’s audience) - may affect the meaning and amplify violence. The social-historical context also favored conflict, as Kenya has experienced violence during every [election](#) since 1992, and the Kalenjin and Kikuyu have had long-term disputes over land¹².

The third paradigm is the platform for mapping hate speech, which in the case of this project is the Internet (online newspapers, Facebook etc.). Understanding the concept and context empowers one to make informed decisions on methodological approaches that seek to identify, analyze and draw conclusions from the various instances of hate speech in the media. Brian Lucas (2014) classifies mapping of online hate speech into three groups based on their purpose. The first category is ‘Real-time monitoring and mapping’ where the purpose is to provide continuous monitoring of online media as reflected in the Umati project in Kenya (Sambuli & Awori, 2014). The second category is ‘Retrospective monitoring and mapping’ where monitors go through and analyze archived messages. The third category is the ‘Discourse and content analyses’ where monitors examine potential hate messages within their historical, political and social contexts to form meanings.

There is a thin line between free speech and hate speech in that freedom of expression, embedded in most countries’ constitutions, can be used by people to incite hatred and violence and sometimes does not have limits when it crosses the border of what is perceived as hate speech. Scholars have not yet found a middle-ground between maintaining freedom of expression and preventing incendiary speech that sparks overwhelming violence¹³. For example, Dutch MP Geert Wilders was cleared of all charges on charges of defamation and hatred against Muslims in June 2011 as the Judge ruled that, though sometimes “hurtful and humiliating”, Wilders’ statements had remained within the boundaries of the law¹⁴.

Some messages and conversations on hate speech can cause considerable offense, but they do not

always directly incite violence. Scholars of ‘free speech’ have made efforts to differentiate between ‘hate speech’ and ‘dangerous speech’ and to illuminate how ‘hate speech’ can become ‘dangerous speech’ - speech that has a reasonable chance of amplifying or catalyzing one group to carry out violence on another¹⁵. Benesch’s (2012) five-point analytical tool is a useful framework for determining the dangerousness of the speech and consists of: the level of a speaker’s influence; the grievances or fears of the audience, that is, whether or not the speech act is understood as a call to violence; the social and historical context; and the way in which the speech is disseminated. The Rwanda genocide is an important case study where there is correlation between speech and action i.e. the government radio’s chilling messages for vengeance against the minority Tutsi population indeed resulted in massacres:

‘You have to kill [the Tutsis], they are cockroaches..... Fight with the weapons you have at your disposal, those of you who have arrows, with arrows, those of you who have spears with spears... Take your traditional tools... We must all fight [the Tutsis]; we must finish with them, exterminate them, sweep them from the whole country... There must be no refuge for them, none at all.’¹⁶

The Umati project, in its pursuit to monitor hate speech online during the 2013 Kenyan elections, espoused a feasible definition of hate speech that also included other forms of hate speech not embedded in the National Cohesion and Integration Act of 2008. These forms are hate-based on religion, gender, nationality, sexual preference or political affiliation¹⁷. Given the broad yet partial

definition of hate speech in Kenya’s laws, Umati adopted the dangerous speech definition proposed by Benesch (2012), which is narrower and less ambiguous. Umati then developed a dangerous speech collection, classification and analysis method that was based on Benesch’s Dangerous Speech Guidelines.

Monitors scanned online platforms (such as Twitter and Facebook) for occurrences of hate speech and then saved and coded them into an online database. Umati used both quantitative approaches (mapping the number of incidences where they occurred) and qualitative approaches (analyzing discourse). Following translation of texts from some Kenyan languages into English, they then categorized the texts into offensive speech, moderately dangerous speech and extremely dangerous speech. The project evaluated the level of influence the speaker had over the audience on a scale of 1 to 3, with 1 being little influence and 3 being a lot of influence. It also evaluated the level of inflammatory remarks in the texts on a scale of 1 to 3, with 1 being barely inflammatory and 3 being extremely inflammatory.

Umati highlights that not all the five factors of Benesch’s framework (the speaker and his/her influence over an audience, the susceptibility of the audience, the content of the speech, the social and historical context of the speech and the means of spreading the speech) have to be present for a speech statement to amount to dangerous speech. It therefore cut down Benesch’s definition of dangerous speech into a three-point analytical tool suitable to the Kenyan context. Dangerous speech was therefore defined by the Umati Project as speech that:

Table 1: The Benesch Dangerous Speech Framework

Speaker’s Influence	Audience receptiveness	Speech Content	Medium of dissemination	Social/historical speech context
e.g. political or religious leader will likely have influence over a crowd	Subject to incitement by speaker	Content that may be taken inflammatory by the audience and may be understood as a call to violence	This includes language used and the medium for dissemination	e.g. previous clashes or competition between groups that can make them more prone to incitement

NB: not all variables must be present for speech to qualify as dangerous speech.

Table 2: Factors for identifying online inflammatory speech (from Benesch’s framework)

Targets a group of people	May contain one hallmark of dangerous speech	Contains a call to action
---------------------------	--	---------------------------

Dangerous speech towards a group can occur across different lines, including religion, tribe/ethnicity, gender, sexuality, political affiliation and race	Comparing a group of people with animals, insects or vermin	Encourages a particular audience to commit acts of violence towards a group of people
	Suggests that the audience faces a serious threat or violence from another group, specifically the same group that is a target of inflammatory (“accusation in a mirror speech”)	
NB: A hate comment about an individual is not necessarily a hate speech UNLESS it targets the individual as part of a group	Suggesting that some people from another group are spoiling the purity or the integrity of the speaker’s group	These can include call to kill, beat/injure, loot, riot, and forcefully evict

3.0 Research Methodology

The purpose of this chapter is to give an account of the research and provide the reason for the methodology used. This chapter describes the conceptual framework of the research e.g. purpose of the research design, theoretical underpinnings vis-à-vis data gathered, the unit of analysis and criteria for interpreting findings. It was therefore important to design the research in such way that the methodological tools identified answer the research questions as well as test and elaborate on these theories. In essence, this research design attempted to derive answers to the research questions. Yin (1994) argues that:

‘The main purpose of the design is to help to avoid the situation in which the evidence does not address the initial research questions. In this sense, a research design deal with a logical problem and not a logistical problem’.

Elsewhere Yin (1994) suggests that the overall framework of research design should include items such as the purpose of the study (series of questions that have to be answered), unit of analysis (the actual source of information), preposition (suggesting possible links between theoretical underpinnings and data gathered/analysed) and criteria for interpreting the findings (iteration between preposition and data etc). Yin (1994) emphasizes that such a conceptual framework links the data to be gathered and the conclusions to be drawn to the initial questions.

The purpose of this chapter was to answer the main research questions during the research process related to teaching and learning practices that affect

student learning and performance in science. For instance, this chapter attempted to answer the following generic questions: -

1. To what extent are the conversations/messages/texts on the internet inflammatory?
2. What are the channels/techniques/tactics used by influential internet users, when engaging their audience in inflammatory speech?
3. What techniques, channels and reactions are applied by the audience when engaged/provoked by inflammatory speakers?

The research design was therefore considerably influenced by the theoretical foundations in Chapter 2. Since this research is basically a study of recorded human communication online and given that iHub Research has carried out research of a similar nature, the project decided to adopt the [Umati Project’s](#) methodology for the following two reasons. First, because its three-point tool, as guided by Benesch’s Dangerous Speech Framework, is very comprehensive. For instance, the study sees its *methodological framework* of quantitative and qualitative method of content analysis, discourse and contextual analysis as the most suitable for studying the online conversations in South Sudan and analyzing who said what to whom, why, how, and with what effect. Second, because it perceives its three-point tool as practical in arriving at grounded conclusion on hate speech in South Sudan.

The methodology involved identifying and monitoring hate speech conversations/messages on

the Internet (blogs, Facebook, online newspapers etc.) for the last 10 months. 2 data coders and the lead researcher scanned online platforms for incidents of dangerous speech and recorded the speech acts they perceived to be hateful in an online database. The Umati project's three-point tool was used to derive quantitative measurements to questions such as the extent to which the conversations/messages/texts collected from the Internet are inflammatory.

The unit of analysis primarily was the online sites identified in https://docs.google.com/spreadsheets/d/1BT_JLi04_HtTZTs7F3WAXPkevjrFpWfdirR1UEyy7oV4/edit#gid=144847878. Unfortunately, the research team has not been able to capture inflammatory remarks on some of these identified sites. The methodology of data analysis is a combination of qualitative and quantitative approaches. Through the Umati Project's categorization formula, the project arrived at conclusions whether the speech is offensive, moderately dangerous speech or extremely dangerous speech. The coding sheet is provided in file:///C:/Users/Hasting/Desktop/Lemi%20Desktop%202017/IHUB/My%20reports_UMATI/Sample%20Coding%20Sheet.html. A guide adopted from Umati project on the definition of hate speech categories is provided in https://docs.google.com/document/d/1ZFn_PVFUPcsXCi71fUj9sXFpual-ai0QQOTD2VtMEFQ/edit

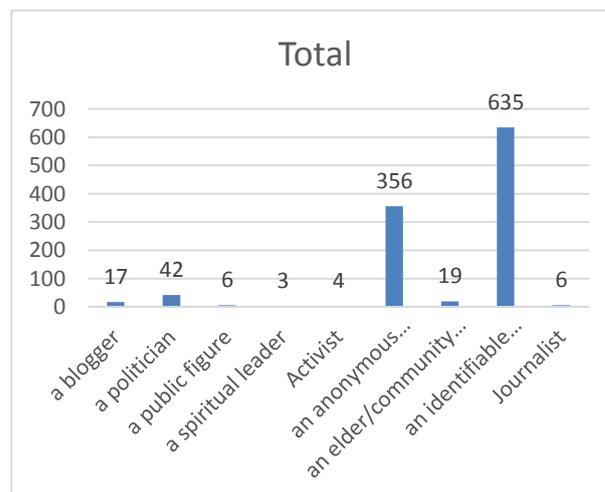
4.0 Research Findings

This chapter provides an account of the main findings of the research that the researchers established during the 10-month period of the research. The findings are premised on the transcripts analysed (data from the online sources). Thus, the findings are in conformity with the study's methodology framework (e.g. quantitative and qualitative method of content analysis, discourse and contextual analysis). 1088 transcripts (raw data) went through a thorough coding process prior to compiling this report. Indeed, our qualitative content analysis yielded significant incidences of hate speech that is akin to the concept of hate speech in the three paradigms discussed. The metaphors expressed in the transcripts characterize members of an ethnic group in non-human terms such as cancer, a virulent disease, Ebola, incurable disease etc. Words such as 'nyagat', nyamnyam, Aryan Jeinge, MTN, Cow kingdom appeared several times in the transcripts than any other inflammatory words. Words or phrases alone were not sufficient in some instances for coders to categorize the text; thus, coders in

such instances had to analyse sentences to arrive at the underlying meaning of the communication to code accordingly.

Blogs on the economic, political and social issues in South Sudan are normally initiated by intellectuals in the diaspora. Most reactions to posts on these blogs are unhealthy online conversations/ethnicity-based arguments, mainly from South Sudanese nationals in the diaspora in reaction to what they perceive as social, political and economic injustices in the system. 33% of these hate speech incidents (offensive, moderately dangerous and extremely dangerous) come from anonymous commenters (people who use pseudonyms names), while 58% come from identifiable commenters (people who use real names).

Figure 1: Hate speech propagators per commenters category during the research N= 1088



The identifiable commenters accounted for 58% of all hate speech while the anonymous commenters accounted for 37% of all hate speech. These 2 categories were the leading disseminators of the hate speech captured during the research. Interestingly politicians accounted for 4% of the hate speech. On the whole, majority of the identifiable and anonymous commenters are in diaspora as demonstrated by their own pronouncements during the heated conversations.

One key issue in relation to hate speech in South Sudan is that hate speech speakers often make sweeping statements against a tribe in lieu of an individual from that particular tribe that they have picked a bone with. For example, commenters end up stereotyping the whole tribe as cockroaches, Ebola, children of Satan etc. just because of one individual they hate i.e. castigating, insulting and calling for the annihilation of a whole tribe that particular individual comes from. Each community

presupposes itself as mightier and superior than the other, and often you would find one particular tribe being castigated as women and another being scolded as wild with limited reasoning capacities. For instance, some commenters made provocative comments that Equatorians are *Nusuwan* of Dinka (Arabic word for women) on one hand, while other commenters made inciting comments that the *foolish majority* or rather *Lugoro* (Bari language word that refers to the cultural marks or tattooed heads of the Dinkas) are illiterate and primitive on the other hand. Other commenters vented their anger on the government of the day and made inflammatory speech such as ‘Dinkacracy’. Apparently, this coined word suggest that the Dinka have monopolized the army, security, police and the economy

There were also commenters (especially Riak’s supporters) that targeted their hate speech at the Ugandans because of the involvement of Uganda in the conflict. The social networking sites monitored by the research included Facebook, YouTube channels and un-moderated online newspapers run by South Sudanese in Diaspora who are pro and against the government. As Figure 2 shows, 43% of the inflammatory speech captured in this research were observed on the Facebook. 82% of the inflammatory speech monitored and captured were in English language and 18% in Arabic language.

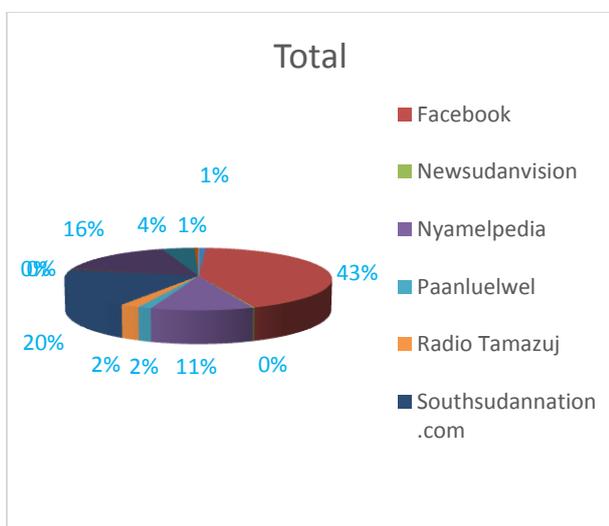


Figure 2: Source of hate speech statements during the research (n=1088).

As cited early on the Ukrainian cyber space that internet users from one political camp can become objects of hatred when conversing with internet users from a different political camp, the same hypothesis applies in the South Sudan cyber space. Often respondents were driven to respond to inflammatory texts pertaining to political, social

and economic issues in the same offensive way a commenter, blogger etc. has raised an issue. By and large, most of the conversations centered around politics and the economy and prompted users from both camps to generate inflammatory texts. The issues most common in the inflammatory texts are:

1. related to land
2. related to federalism
3. related to poor governance

4.1 Land-based hate speech

Land injustices are the precursor of hate speech. The land issue has been a predominant theme in the hate speech instances collected and analyzed so far. Speech tensions arise between commenters who support host communities whose lands have supposedly been grabbed by cattle keepers, and those who support the cattle keepers and feel that it is their constitutional right to settle in any part of South Sudan. Some identifiable and anonymous commenters have gone to the extent of emphasizing that the land grabbers will get buckets of their own human blood in exchange for their ancestral lands on one hand, while on the other hand supporters of the land grabbers have gone as far as stressing that they will occupy the lands forcefully and rule over the host communities through thick and thin.

Some hate speech commenters cluster tribes from Greater Equatoria as one ethnic group and labelled them as ‘cowards, weakest tribe and least powerless’ from which one can grab their lands. They promote the myth of ‘we are born to rule and will rule you’ forever. Commenters from the other political divide call their people to stand up and defend their lands and even go as far as calling their people to turn these lands into ‘mass grave yards’ for the grabbers. Some have even encouraged the picking of MTNS¹⁸ or Bu-tha Mumnu¹⁹ from public vehicles and their annihilation.

Of all the hate speech captured during the research, land-based hate speech accounted for 28%, federalism-based accounted for 14% and poor governance-based accounted for 58%. 17% of the land-based hate speech was extremely dangerous...commenters pitting communities against each other to the extent that they call their audience to a violent action (i.e. annihilation of the challenging ethnic group).

Basically, the inflammatory texts captured among others emerged as a result of disputes over land between communities (i.e. Dinka Padang vs Shilluk, Anyuak/Murle vs Lou-Nuer/Jikany Nuer, Bari vs Dinka, Moru vs Dinka and so on and so

forth).

4.2 Federalism based hate speech

Federalism has been a popular demand in South Sudan, predating the southern Sudan secession calls (from present day Sudan), and even after South Sudan became a sovereign state. Again, reactions to blogs on federalism and what is taking place currently on the ground have evoked heated discussions sometimes with others theorizing the country's 28 states through ethnic lens.

- Dangerous prelude to breaking up the country
- Fosters ethnicity and corruption
- Plot by Jeing Council of Elders to grab land
- Dividing the people (Shilluk and Nuer)
- The 28 states have been created for the purpose of grabbing lands from other tribes and not per se for development
- A way of keeping one ethnic group at bay of terrorizing the Dinka
- As contest between the president and designate first vice president (Dr. Riak Machar)
- Taking oil resources from Nuer and Shilluk communities
- Allow states to control their resources rather than the current system whereby a few ethnic leaders squander the national wealth.

4.3 Poor governance-based hate speech

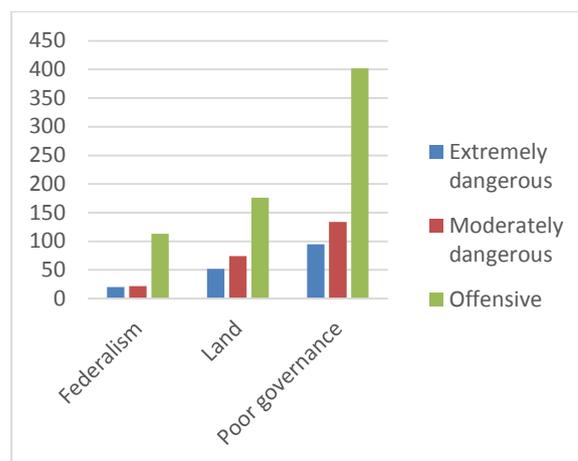
Majority of the commenters in our sample were unhappy with the leadership in the country because they believe that the government is controlled by JCE²⁰. The conversation on poor governance centered around human rights violations, corruption, political rivalry between the President and designate first vice president, reckless embezzlement, misappropriation of public funds and the national cake being accessible to only to few. As supporters of the political divide engaged in these conversations hate speech incidences directed at the JCE and Dinka, Nuer and people of Greater Ekuatoria were happening left, right and center.

Poor governance was cited as the cause of the conflict in the Republic of South Sudan and both ethnic groups (Nuers and Dinka) felt victims of 'each other's political machination' e.g. the massacres that happened in the country. The various ethnic groups subscribe more to local identities rather than to the national identity. The government referred to as foolish majority. There

wasn't religion-based hate speech captured during the research

Some commenters viewed *Mathiang Nyor* (pro-government militias) specifically trained to exterminate Nuer ethnic group in South Sudan and encouraged their respective communities to exterminate the Dinka Aweil.

Figure 3: Profiles of federalism-based, land-based and poor governance-based hate speech analyzed for each category (offensive, moderately dangerous and extremely dangerous speech) n = 1088



5.0 Conclusion and Recommendations

The social networking sites are seen as informal spaces for venting out anger against the system, ethnic group etc., contributing to dangerous speech. Most of the websites monitored (in addition to some in the guide of [South Sudanese credible news sources](#)) are run by South Sudanese nationals in the diaspora and provide online spaces to debate on political, social and economic issues. The sites' comment sections are largely un-moderated, which could indicate why hate speech continues to be found on the platforms. A key point to note is that some websites have their own control mechanism such as not publishing any content that is vulgar or obscene, as evidenced by the complaints from some commenters to the editorial team for not publishing their contributions.

The websites that most of our samples were drawn from are both rebel and government-affiliated. We have learned from the research that *among the 'Internet warriors', they also are voices of peace that continually advocate for peaceful co-existence among the country's 64 tribes*. A key point to note is that some of these 'Internet warriors' are teenagers who spew vitriolic convictions simply because of a conditioned hatred for another tribe.

Majority of the commenters (identifiable and anonymous) especially the ones making inflammatory remarks against the system are in diaspora as demonstrated by their own pronouncements where they are writing from. Even the ones defending the system and making counter inflammatory remarks are mostly residing in diaspora. 90% of the inflammatory remarks are reactions to the unfolding developments in South Sudan or what has been reported online regarding political developments. For instance, unfolding of events to do with political and social issues and the conflict per se on the ground are the ones that feed the streams of hatred on the cyber space

The research is yet to find instances of online hate speech that catalyze events offline. For instance, empirically establishing whether an extremely dangerous speech (i.e. commenters making extremely dangerous speech that they will turn their lands into mass graveyards for the land grabbers/occupiers/settlers etc.) spills over onto the ground. The research needs to capture in the future instances of the kind (linkages/correlation) where internet warriors fight their wars on the internet and relatedly things happen on the ground.

As part of the project, a response framework has been developed to mitigate issues to do with hate speech on online. Most of these websites we browsed for our data are run by South Sudanese in Diaspora and provide online spaces for South Sudanese to debate on political, social and economic issues. And since they are not regulated by the government, they do provide a space for publication and posting of viewpoints that contain hate speech transcripts. A mini conference is recommended here this mini conference will provide a forum for a plethora of key players (activists, anti-discrimination bodies, South Sudanese website owners in diaspora, some internet users/warriors, journalists, government etc.) to exchange views, concerns and practical experiences regarding hate speech. The conference will also provide a forum for the players to come up with tangible recommendations and solutions.

Secondly, some 'internet warriors' and users are adolescents and write unimaginable horrible stuff that they have not even committed simply because of an inbuilt hatred for another tribe. Users on the social media and the online forum are not aware of impact of their hate speech messages on the communities. Increase users' and the public awareness that some of the texts are viral and to some degree have negative impact on peaceful coexistence of the communities in South Sudan.

Lastly, as mentioned early - the research hasn't established any evidence linking online hate speech with what is happening on the ground e.g. the indiscriminate killing of civilians in the villages as a result of the online hate speech etc. In addition, this project doesn't have any information of what is taking place offline in terms of hate speech. However, the project assumes that there is much hate speech offline. Awareness-raising in the communities is very important for inhibiting and preventing offline hate speech. Communities need to learn that not only are these hateful offline statements detrimental to peaceful co-existence, but that some of these statements are indeed criminal.

6.0 Acknowledgement

This research was funded by Search for Common Ground, and indeed we are deeply grateful for their support. We would also like to thank IHuB Research and particularly their 2 lead researchers (Nanjira Sambuli and Kagonya Awori) who guided us at the beginning of the research. Without their valuable guidance and input, the research could not have been successfully conducted.

7.0 References

- [1] African Union (2014, Oct 15), Final Report of the African Union Commission of Inquiry on South Sudan, Addis Ababa; pp. 23
- [2] Benesch, S. (2012, January 12). Dangerous Speech: A Proposal to Prevent Group Violence; pp. 2
- [3] Gagliardone, I., Patel, A., and Pohjonen, M. (2014) *Mapping and Analysing Hate Speech Online: Opportunities and Challenges for Ethiopia*. University of Oxford and Addis Ababa University; pp. 5-6
- [4] Cohen-Almagor, Raphael (2011), *Fighting Hate and Bigotry on the Internet*, Published in: *Policy & Internet*; Vol. 3: Issue. 3, Article 6 pp. 1
- [5] Lucas, B (2014, July 14) *Methods for monitoring and mapping online hate speech*, Helpdesk Research report; pp. 2.
- [6] Sambuli, N. and Awori, K. *Monitoring Online Dangerous Speech in Kenya: Insights from the Umati Project*, In *Communications technology and humanitarian delivery: challenges and opportunities for security risk management*. European Interagency Security Forum (EISF), (2014),
- [7] Somerville, K (2014, April 24) *South Sudan: how hate radio was used to incite Bentiu massacres*
- [8] United Nations Educational, Scientific and Cultural Organization (2015) 'Interim Assessment of Media Development in South Sudan' 7 place de Fontenoy, 75352 Paris 07 SP, France and UNESCO Regional office

for Eastern Africa, United Nations Avenue, Gigiri, Nairobi 00100, Kenya and UNESCO Juba Office, Thompington, Juba, Republic of South Sudan. pp. 15

[9] United Nation Mission in the republic of South Sudan (2014, May 8), Conflict in South Sudan, A human Rights Report, pp 51

[10] Yin, Robert (1994) *Case study research: Design and methods*. California, USA: Sage Publications. pp.134

¹ A derogatory term for a 'traitor' and 'rebel'

² [Wikipedia](#) - The name Niam-Niam (or Nyam-Nyam) was frequently used by foreigners to refer to the Azande in the 18th and early 19th century. This name is probably of [Dinka](#) origin, and means *great eaters* in that language (as well as being an [onomatopoeia](#)), supposedly referring to cannibalistic propensities. This name for the Azande was in use by other tribes in South Sudan, and later adopted by westerners. Today the name Niam-Niam is considered [pejorative](#). Today it is used to refer to people of Greater Equatoria as great eaters

³ عريان (Aryan) Arabic word for 'naked'. 'Aryan Jienge – naked Dinka' is a derogatory term used by tribes from Equatoria region against Dinkas

⁴ Jeing Council of Elders being referred to as Jeing Council of fools

⁵ Coined word suggesting that the Dinkas have monopolized the army, security, police and economy

⁶ a euphemism for the Dinka that they are everywhere you go (MTN business logo of "MTN everywhere you go")

7

<http://www.southsudanhumanitarianproject.com/background/docb-13/>

⁸ <http://www.insightonconflict.org/2015/03/south-sudan-peace-talks-role-igad/>

⁹ <http://www.insightonconflict.org/2014/10/role-social-media-south-sudan-crisis/>

¹⁰ [Nuer community claims 17613 killed in South Sudan violence](#)

www.sudantribune.com/spip.php?article50571 Sudan Tribune

¹¹ The Phenomenon of Cyber hate in the Ukrainian Cyberspace (May 30, 2011)-

<http://www.ihrpex.org/en/docs/>

¹² http://freespeechdebate.com/fa/discuss_fa_ir/when-does-hate-speech-become-dangerous-the-link-between-words-and-violence-in-rwanda-and-kenya/

¹³ <http://www.worldpolicy.org/blog/2013/02/06/lets-talk-about-genocide>

¹⁴ <http://freespeechdebate.com/en/case/geert-wilders-on-trial/>

¹⁵ <http://freespeechdebate.com/en/discuss/when-does-hate-speech-become-dangerous-the-link-between-words-and-violence-in-rwanda-and-kenya/>

¹⁶ <http://www.counterfire.org/history/17141-the-rwandan-genocide-from-imperialism-to-barbarism>

¹⁷ <http://www.ihub.co.ke/blog/wp-content/uploads/2014/06/2013-report-1.pdf>

¹⁸ a euphemism for the Dinka that they are everywhere you go (MTN business logo of "MTN everywhere you go")

¹⁹ Arabic word referring to the Dinkas as prohibited goods

²⁰ Jieng Council of elders