Hate is an emotion that has the potential to pave the way towards negative behaviors. Those behaviors can grow to more advance forms, starting from simpler acts such as: insensible remarks, hateful jokes, avoiding people of particular groups, rudeness, and meaningless bias. Growing in complexity, hatred takes advanced life-threatening forms that could result in political discrimination, burning of churches and mosques, rape and genocide.

In the United States, for example, each year a report by the FBI on hate crime statistics is released. It states the number of hate crime in the US, target groups, whether some hate crimes have increased or decreased through the years, and a comparison among groups in terms of the per capita rate of hate crimes committed against them. The 2016 issue of this report declared the increase of hate crimes. According to the report, motivations were: bias against a religion, gender, a particular sexual orientation, ethnicity/national origin. More general examples of actions driven by hate can be employers not hiring people whose native language is not English, the rare positive-photographing of overweight people in magazines, and many more that would not have happened without the underlying hate and stereotype.

None of the event mentioned above start suddenly; they all grow from hate. Therefore, it makes sense to begin with addressing and combating hate first, in order to take a more systematic approach of how complex negative behaviors are addressed.

The Pyramid of Hate

For this addressing of hate to be more efficient, the Pyramid of Hate is the referenced graphical diagram that is used to track the consequences of accepting and tolerating hate. Designed by the Anti-defamation league, the Pyramid of Hate is a classification of acts of humans in terms of prejudice. It is a good way of examining the escalating nature of hate. According to the pyramid, hatred is systematic. It shows biased behaviors growing in complexity from bottom to top. Upper levels are supported by lower levels. Thus, tolerating lower-level behavior leads to behavior in the next levels becoming more acceptable.

Discussions brought up by analyzing the pyramid can help in better understanding how hate can escalate to violence, different ways in which each level of the pyramid is contributing to the problem, and a clearer recognition of the roles of individuals in breaking the progressive cycle of hate.
To put it simply, hate can be looked at as the prerequisite for violence. It is common for an inflammatory public speech to precede the outbreak of a local or international mass conflict. In Rwanda for example, the conflict was aided significantly by hate speech being broadcast through radio stations, which has contributed to the massacre of 800,000 Tutsis and Hutus. Those stations went even further by broadcasting lists of people to be killed, in addition to where to find them. [1]

The situation in South Sudan does not differ much from Rwanda’s example, especially with the wide-spread use of social media among South Sudanese communities. Social media offers an open platform, which has contributed to perpetuating the call to violence and inner conflicts that erupted in South Sudan in December 2013, and again in July 2016. I

**South Sudan: Online versus Offline**

It was also notable that these actions were often driven by the South Sudanese diaspora. In one of the examples, certain groups were being referred to as terroris. Other examples involve using the word “MTN” to describe particular groups. MTN is a
telecom company that operates in South Sudan, and they use the slogan “Everywhere you go” in their campaigns.

Peach Tech Lab Africa researches came to emphasize on this specific point. An independent non-profit organization based in Nairobi, PeaceTech Lab Africa works towards reducing conflict using technology, media and data. Part of their research work involves monitoring and analyzing hate speech in South Sudan in order to mitigate the potential threat of violent language. Their published reports are a result of analyzing hate/dangerous speech of South Sudanese in social media during certain periods of time between August 2017 – November 2017. The initiative aims for developing online hate speech lexicon, providing data visualizations and social media monitoring, and conducting of sessions with South Sudanese groups to validate the context of the hate speech terms it identifies.

Their research also shows how online hate speech is contributing to the online and offline violence – especially the one coming from the South Sudanese diaspora in the US, Canada, UK, and Australia. A YouTube video by Peace News Network speaks of the perpetuating of fake news and hate speech in social media among South Sudanese. Besides the role nonprofit organizations has played in raising awareness and discussing the problem further.

There is, however, a debate of opinions.

South Sudan has limited access to the internet, which leads to questioning whether the former premise of online hate speech contributing to offline violence is actually true or not. The opposing side is basing their argument on how remote communities in South Sudan are, or at least the majority. The many tech-illiterate residents are far away from social media posts, and all the liking and sharing. According to IRIN News, only about 21 percent of South Sudanese have phones and around 17 percent can get online.

But regardless of the small number of population that is using online platforms, hate news and rumors can still be passed on to a wider offline community by word of mouth, phone calls, or group discussions. Lack of proper reporting from the ground, craving of information by users on the outside, and censorship of media are all factors that invite more spreading of fake news.

The UN Mission in South Sudan (UNMSS) report (published on 22\textsuperscript{nd} February 2018) on the right of freedom of opinions and expression in South Sudan since the July 2016 crisis, has stated facts about the sudden and alarming increase of inflammatory language by South Sudanese citizens.

As reported by this document, hate speech has spread through various forms of communication, including: text messages, phone calls, cartoons, private conversations, public speeches, and social media platforms. A staff member at a school in Yambio,
Western Equatoria, reported receiving text messages threatening her and 20 Dinka students with death, unless they left the school. In addition, according to the UN report, “South Sudanese internet users often resorted to social media platforms such as Facebook, Twitter and other blogs, as tools to spread derogatory content and inflammatory messages.

**An Intervention**

An intervention of several ways is urgently needed to be adopted to protect the audience against hate speech, making it less influential or dangerous. In addition, people need to be trained on how to be aware of their own negative actions, and which alternative approaches they should follow in order to make a safer and less dangerous online environment.

**Youth Driven Action**

Hence the #DefyHateNow project was launched, aiming to prevent and counter the spread of hatred on social media. Supported by the German zivik programme of ifa (Institute for Foreign Cultural Relations) and managed by the Berlin-based rog_agency for open culture and critical transformation, #DefyHateNow was established as an urgent community peace-building initiative aimed at combating online hate speech and mitigating incitement to offline violence in South Sudan.

The project organizes workshops and trainings on the impact hate speech has on fueling conflict, and offers training in how social media can be used in a constructive and peace-building manner.

These programs allow attendees of Sudanese and South Sudanese origins to explore their own attitudes and experiences toward hate-motivated actions, examining their roles and responsibilities, and to think of possible solutions and right actions to take amid the encountering of online and/or offline hatred.

On a recent workshop that was organized by #DefyHateNow in a collaboration with Andariya magazine, participants were trained on identifying hate speech, pyramid of hate, cycle of liberation, and more concepts that helped in better understanding of the rooting problem. A video documentation can be found here.

The tools and strategies provided by the #DefyHateNow project are not exclusive for online hate speech alone. They can be used to address and mitigate hate incidents that take different forms, such as: over the telephone, in person, at school, with family or at community events.
The Field Guide

Rog_agency’s work involves researching and exploring the ways in which open data, open source technologies and methodologies can be applied to a broad spectrum of activities. The aim at the end is more empowerment and positive transformation.

#DefyHateNow has provided trainings, workshops, campaigns, sport events, and music concerts to raise awareness on online hate speech and its offline effect. The outcome of this initiative has also come out in the forms of publications, posters and a mobile app.

One of the published documents is the #DefyHateNow Social Media Hate Speech Mitigation Field Guide; a detailed document that represents the guiding toolkit to achieve the project’s aims and goals.

The field guide includes: tools, strategies, and relevant grassroots projects regarding the spread of hate across social media. It is designed to be used as guiding material during workshops and training programs of the initiative.

This document is divided into sections, each of which handles a specific aspect of the issue. These aspects are: social media itself, hate speech, fake news, countering the dangerous speech, ethical reporting, peace activism, identity context, civil society initiatives and diaspora online. The field guide is also accompanied by concept cards that facilitate the flow of conversation among attendees, and the focusing on areas of interest of online hate speech.

Social Media Giants

Many private tech companies such as Facebook, Twitter, YouTube and Microsoft are expanding their policies on hate speech, violence and abuse. The expansion is in terms of what constitutes hateful and harmful behavior on their platforms, and stricter rules and conditions that are enforced.

Facebook’s “community standard” is a 25 pages documented translated into 40 language on rules of hate speech. The policy has been developed by people in 11 offices around the world, including experts and consultants on issues of expression, safety, rape and terrorism.

Twitter has also updated their rules around abuse, hateful conduct, violence and physical harm in November 2017. The changes were announced in anticipation of their enforcement starting from December 2018. Their policy updates, however, don’t apply to military or government entities. This would allow President Trump for example to continue his racist threats against target groups, promoting more hate. Moreover, the company was also occasionally criticized on their failure to punish harassers.
**Way Forward**

By interfering with the spread of hate across online platforms, the prevention of both online and offline violence will be more possible. The field guide in its detailed form is helping in clarifying the problem, highlighting concerns and educating on dangers and possible prevention mechanisms.

One of the big challenges that is facing the regulating of hate speech is not agreeing upon what exactly defines hate speech, and by whom is hate defined. There is a fine line between freedom of expression and the censorship of hate speech. On the other hand, freedom of expression can be taken advantage of by attackers as a veil to spread hate agendas. Thus, a more careful consideration to this when hate is being regulated should be taken care of.