

TRUTH TELLERS

REPORT OF A TWO DAY
MIS/DISINFORMATION WEST AFRICA
CROSS-DISCIPLINARY CONVENING

ORGANIZED BY WITNESS
IN ABUJA, NIGERIA

FROM
SEPTEMBER 14-15, 2021





Table of Contents

Introduction and background	4
Executive Summary	4
1.0 Overview	7
1.1 Objectives of the convening	8
1.2 Participants' introductions and views	9
2.0 Misinformation and Disinformation and their impact	11
2.1 Which solutions are being prioritized?	12
2.2 An African continental landscape of Mis/disinformation	12
2.3 We are all susceptible	15
2.4 Effects of the problem	16
2.5 Exploring solutions that work	16
3.0 Analysis of Nigeria's Media Mis/disinformation Risk	17
3.3 Feedback on presentation	19
4.0 Breakout Group Session	19
4.1 Group 1 - Social Justice Movements	22
4.2 Group 2 - Democracy and good governance	22
4.3 Group 3 - Conflict	23
4.4 Discussions	24
5.0 Developing trustworthy videos - WITNESS	24
6.0 Mis/disinformation in the context of gender and inequalities	27
6.2 Discussions	28
7.0 Day 2 Preamble	29
7.1 Group Exercise	29
7.2 Panel discussion	29
8.0 Internet Shutdowns and Content Takedowns in Africa – WITNESS	33
8.1 Breakout Group Session	34
8.2 Group reports	35
9.0 Presentation by Facebook	39
9.1 Sharing of practices from across West Africa	41
9.2 Final Spectogram	42
10.0 Where do we go from here - Next Steps?	42
ANNEX	43

Introduction and background

Founded in 1992, WITNESS has been at the forefront of helping people use video and technology to protect and defend human rights. Over the years, our work on using video and technology has helped in advancing justice, ensuring accountability for atrocities, and promoting good governance. In its approach, WITNESS identifies critical situations and teaches those affected by them the basics of video production, safe and ethical filming techniques, and advocacy strategies. However, recent proliferation of mis/disinformation has made it increasingly difficult for videos to be trusted and relied upon for mobilizing and advancing human rights, justice and accountability. This has resulted in WITNESS prioritizing convenings that bring together stakeholders to deliberate over solutions for combating the negative impacts of misinformation and disinformation on the advancement of human rights. Findings from our convenings inform our advocacy to technology platforms and companies in order to drive systems change.

Executive Summary

The growing spread of mis/disinformation presents a real threat to democracies, digital safety and the enjoyment of human rights. Also, the advancement of technology developed to mislead makes it increasingly difficult to verify content. Furthermore, the emergence of deepfakes and other forms of synthetic media enable more seamless manipulation of audio and video content.

Following the increasing incidences of misinformation and disinformation on the continent, and its impact on creating more trustworthy videos particularly by grassroots communities, it became imperative for WITNESS to prioritize efforts to combat this threat. In 2019, WITNESS organised the first workshop on deep fakes and other forms of synthetic media in



sub-Saharan Africa which was held in South Africa. A key outcome from the workshop was the need to broaden and deepen the understanding of the nature and scope of mis/disinformation especially at regional levels.

The West Africa cross-disciplinary convening is WITNESS' pilot project on mis/disinformation. The aim of the project is to anchor a community led conversation and a solution-based approach to countering mis/disinformation. The two-day convening was held in Abuja, Nigeria from September 14-15, 2021 with participants drawn from leading human rights organizations, media organizations, technology platforms, and academia.

The convening identified governments' cynical exploitation of the proliferation of mis/disinformation to clamp down on freedom of expression and evade accountability. It also highlighted the diverse challenges misinfo/disinfo poses to human rights defenders, activists and civil society groups particularly in effectively documenting and exposing human rights violations. Internet shutdown was also highlighted as a key tactic which governments use to suppress the truth while using misinformation and disinformation as justification.

At the end of the convening, it was evident that low media literacy even amongst journalists, and civil society organizations, exacerbates the problem. Additionally, ineffective fact checking, unavailable detection and verification mechanisms, low level of media accountability and flawed content moderation by technology platforms were identified as key limitations to addressing the problem of mis/disinformation.

In conclusion, stakeholders agreed that the media, civil society organisations, fact checkers, human rights defenders, activists and technology platforms must synergize across geographical locations, thematic focus and organizational structures to combat the spread of mis/disinformation.

Some of the gaps identified for urgent intervention include:

1. Identify journalists and social media influencers that are concerned about mis/disinformation and build their capacity to spot, verify and fact check information, in the media space. This will also include building competency in effective video documentation to meet the highest evidential standard;
2. Conduct media and information literacy campaigns in collaboration with fact checkers, technology platforms, media organizations including community radios and local grassroots organizations and gatekeepers;
3. Support projects, initiatives that increase and improve media accountability, transparency, and adherence to ethical standards;
4. Create guidelines for reporting Sexual and Gender-Based Violence, insecurity, conflict, and elections, to reduce the spread of mis/disinformation and minimise its harm; and,
5. Create a network of actors that are committed to combating mis/disinformation. The network will be resourced with up-to-date information on strategies, tools, technologies necessary for defense against mis/disinformation.



Day 1

1.0 Overview

Digital technologies have opened the space for increased malicious spread of unchecked, manipulated, recycled, and mis-contextualized information, which challenges truth and reality. WITNESS is therefore engaged with preparing the information and human rights landscape for new forms of media manipulation especially synthetic media such as deepfakes. One of the strategies deployed by WITNESS includes organizing convenings, where threats are identified through detailed threat modelling, and relevant contextual solutions are developed with emphasis on Global South countries. This has helped in decentralizing conversations on mis/disinformation and its impact from the Global North to the Global South. Challenges created by mis/disinformation in the Global South have included:

1. The discrediting and labelling of images captured by human rights activists as fake news.
2. Providing a liar's dividend to bad actors who use mis/disinformation as cover-up to evade accountability.
3. Harmful, malicious targeting of journalists, civil society organizations and activists with disinformation campaigns.

All these contribute to the breakdown of the fabric of society through the exploitation of differences within communities.



1.1 Objectives of the convening

WITNESS continues to use the convening of stakeholders - information and media experts, institutions, academia, technology firms, human rights organizations and activists - to understand how communities can effectively push back against mis/disinformation. The West Africa convening was therefore organised to understand how mis/disinformation:

1. Amplifies and alters existing challenges;
2. Creates new challenges; and,
3. Reinforces other challenges.

The convening set out to achieve two core objectives:

1. Identify key threats of mis/disinformation to critical voices including journalists, human rights defenders and activists in the region.
2. Design collective solutions that would prioritize grassroots communities' response to mis/disinformation through the development of a contextual framework that would form the bedrock for a robust defense against mis/disinformation in the region.

1.2 Participants' introductions and views

Participants were drawn from several human rights organizations, media groups, as well as technology firms/platforms. See Annex for the list of participating organizations and independent experts.

Following participants' introductions, and opinion on their understanding of mis/disinformation, most participants agreed that misinformation is the unintentional dissemination of false, untrue, inaccurate, and unverified information. There was also a consensus that disinformation is the willful creation, dissemination and amplification of false information with the intention to mislead, distort facts and cause harm or damage. It was equally highlighted that sometimes mis/disinformation is carried out with the intention to

gain social capital/influence - increase social media visibility and followership.

Spectrogram of participants' thinking on mis/disinformation as a human or technological problem or both, showed that eight (8) participants thought mis/disinformation is purely a human problem, while 16 participants thought mis/disinformation is both a human and technological problem. The participants argued that while Artificial Intelligence is developed by humans, glitches can occur within the technology which can cause the technology to undertake autonomous tasks not intended by the designers. Some other participants shared the opinion that it is a multidimensional social problem which includes governance, and the private sector.



2.0 Misinformation and Disinformation and their impact

Around the world, mis/disinformation has become a huge concern as manipulated media is continuously being deployed by various actors including governments for propaganda purposes, to shape opinion and spark debates. When videos are manipulated, even to promote social justice without appropriate signaling, it could inadvertently make strong arguments for the discrediting of authentic videos that expose human rights violations. The case of the “Syria Hero Boy”¹ is a clear example of how the use of misleading media - staged or unreal footage though produced with genuine social and humanitarian motivations – can undermine the work of citizen journalists and civil society. The Syria Hero Boy case armed the Syrian government with a basis to label genuine video evidence of mass atrocities as false and manipulated. This is a serious challenge for activists, human rights movements, and campaigners as they grapple to further prove the authenticity of audio-visual evidence that expose human rights violations.

Other cases of viral mis/disinformation highlight include the manipulated picture by the Egyptian

State Media that showed Former Egyptian President, Hosni Mubarak leading four (4) other world leaders, including then US President, Barack Obama at the launch of Middle East Peace talks. Another is the deepfake Tiktok video of a golfer which had Tom Cruise’s face. These examples serve different agendas for the people that produced them, but they also have diverse impacts on both the persons impersonated and the larger society. The impact of mis/disinformation pervades the society, especially in societies where information is difficult to verify and governance is shrouded in secrecy. This sometimes leads to negative actions that could undermine democracies, such as the military coup that happened in Gabon following the belief that a video recording of President Ali Bongo was a deepfake whereas it was not².

These are real life situations that challenge the authenticity of video evidence of human rights violations around the world. The problem though is that while detection tools for authenticating and verifying videos are improving, technologies for media manipulation are also getting more sophisticated, widespread and accessible.

¹ In 2014, some Norwegian filmmakers produced a film that was shot in Malta. Their motivation was to draw attention and evoke debate on the challenges and suffering of children trapped in conflicts and wars. <https://www.bbc.com/news/blogs-trending-30057401>

² Following a televised speech by Ali Bongo, the people of Gabon overwhelmingly disbelieved the authenticity of the video. This was because of an ongoing controversy over the President’s health and the President’s long absence from public functions. The speech was dubbed a product of deepfake, because of the President’s gestures during the speech. With the increasing concern and confusion over a power vacuum, the military took over power. However, power was restored only after Ali Bongo surfaced 24hours after the military took over power to reassure Gabonese that he was indeed alive.

2.1 Which solutions are being prioritized?

Prioritizing solutions beyond deepfakes and shallow-fakes are critical to WITNESS. Deepfakes are created using advanced technology to make someone say or do something they never did. While shallow-fakes usually involve slight modifications/edits to videos and photos using non-sophisticated methods. This may also include mis-contextualization. Over the years WITNESS has made deliberate efforts toward addressing the problem of mis/disinformation through a multi-prong approach. WITNESS in collaboration with key partners including Partnership on AI has led an initiative that further helps people to document evidence - videos that are powerful enough to withstand attacks aimed at discrediting them.

Other solutions include: improving authenticity infrastructure; developing detection tools for manipulated

media; improving media literacy; and having a tool-maker responsibility that places the onus on developers to ensure their products cannot be easily used to cause harm.

A question was asked as to why producers of manipulated media should bear the responsibility of appropriate labeling. The Syria Hero Boy video was used as a clear example of how distinct and obvious labelling by the producers could have better illustrated the devastation in Syria and reduced the likelihood of recontextualisation or recycling. Essentially, when videos are properly labelled as manipulated it provides the viewers a useful context for taking action on such video.

2.2 An African continental landscape of Mis/disinformation

Mis/disinformation is a global problem that equally runs deep in Africa. Though the conversations are dominated in the Global North, Africa experiences its variant of the challenge. The growing use of digital technologies, particularly mobile phones that enable access to messaging platforms such as WhatsApp has exponentially increased the spread of mis/disinformation on the continent. In Nigeria, it is easy to recall the “Umaru are you dead or alive?” incident in 2007³. While there were no viral enabling social media platforms such as Twitter, WhatsApp, the nationwide mis/disinformation necessitated the intervention of former President Olusegun Obasanjo to dismiss the rumour. Other negative and harmful misinformation narratives include the narrative of the potency of body parts of persons living with albinism for money making rituals or for other medicinal value. This has led to the harmful targeting of persons with

albinism and even the desecration of their graves to harvest their body parts.

Recent negative impacts of mis/disinformation in the African landscape include the August 2021 viral image portraying the former South African President, Jacob Zuma in an orange prison jumpsuit. This mis/disinformation further escalated the violence that journalists have described as the most violent clash and demonstration in South Africa’s post-apartheid history. Another recent mis/disinformation incident is the circulation of a manipulated image of Ugandan presidential candidate - Robert Ssentamu (Bobi Wine) in front of a lavish home in San Francisco. This was done just to discredit him at the 2021 presidential polls. In other instances, manipulated, altered or mis-contextualized videos and images have caused or escalated violent ethnic and religious conflict.

³ During the 2007 Presidential Election campaign, the Presidential Candidate of the ruling Peoples’ Democratic Party, Alhaji Musa Umaru Yar’Adua wasn’t seen at several campaigns, sparking rumors of his death. At one of the campaign events, the then President, Olusegun Obasanjo at the campaign train had to call Umaru Yar’Adua on speaker telephone to enable people hear his voice and dispel the rumours.





These incidents have led to tens and hundreds of deaths, as well as destruction of properties. In the case of Kenya, a mis/disinformation incident caused reputational damage to a high-ranking Kenyan public officer and the Kenyan electoral commission. Instances of the dangerous impact of mis/disinformation abound around the continent from Ethiopia to Central Africa Republic to the Xenophobic attacks in South Africa.

Governments across the continent have responded to this by imposing severe limitations that restrict people's ability to mobilize for social justice. This is seen through the introduction of laws and regulations

that clamp down on freedom of expression that has led to a climate of fear and a shrinking digital civic space. For example, Cameroon, Egypt, Zimbabwe, Democratic Republic of Congo, Ethiopia, Rwanda, The Republic of Benin, and Senegal have imposed, to varying degrees, several forms of internet censorship. In Nigeria, on 4 June 2021, the federal government placed a ban on Twitter. This was imposed after several failed attempts to introduce anti-social media and hate speech bills.

2.3 We are all susceptible

Studies have shown that people are inclined to share false information for several reasons including:

- 1) Consistency - As humans we align with narratives that are consistent with our preexisting biases, ideologies, values. Hence people share content that is consistent with their beliefs, values, and ideologies.
- 2) Consensus - More people, for the fear of missing out, tend to align with the most popular opinion. Oftentimes, people intentionally share false information just to feel a sense of belonging.

3) Authority - People mostly share information from a perceived authority – an individual, group or organisation that is trustworthy, credible, or popular.

4) Civic duty - Some people see it as a civic duty to share information as they receive it, regardless of the authenticity of the content. Parents, family members, and faith groups are most prone to this.

2.4 Effects of the problem

Today, a lot of information that is in the public domain cannot be trusted. According to researchers at the London School of Economics and Political Science (2018) ⁴, there are five (5) evils of misinformation flowing from the perception that information cannot be trusted. These five evils are confusion, cynicism, fragmentation, irresponsibility, and apathy.

If we must defend the media space from the pervasive power of mis/disinformation, it is important that the war against current dominant and incumbent forces is won. Some of these include: bots, deep-fakes and shallowfakes, low digital literacy rate, limited detection capacity and so on.

2.5 Exploring solutions that work

At the core of WITNESS' work is the training of people to develop trustworthy videos that meet the highest evidential standards. The convening therefore explored solutions on how to:

- i. Proactively push back against deepfakes, shallowfakes and all sorts of AI manipulated media.
- ii. Counter negative narratives by mischievous and malicious actors.
- iii. Build journalistic capacity and coordination.
- iv. Build resistance and resilience in grassroots communities to combat mis/disinformation.
- v. Recognize existing harms that have manifested in gender-based violence and cyberbullying.
- vi. Ensure platform and tool makers responsibility.
- vii. Guarantee shared detection access and capacity.

⁴ How to tackle the spread of misinformation and the problems it causes <https://blogs.lse.ac.uk/businessreview/2018/12/01/how-to-tackle-the-spread-of-misinformation-and-the-problems-it-causes/>

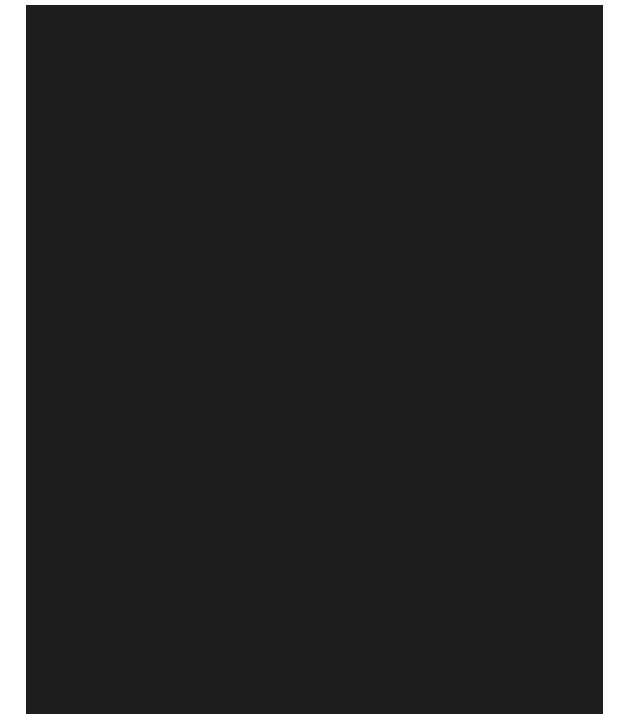


3.0 Analysis of Nigeria's Media Mis/disinformation Risk

The NGO – Paradigm Initiative ⁵ shared knowledge on practical ways communities and movements can disrupt the mis/disinformation system in Nigeria.

The internet has changed how news is funded, produced, consumed and shared. This has introduced new levels of risks for the news industry and the public who consume its content. The hysteria caused by mis/disinformation over the COVID-19 pandemic is a typical example.

The Global Disinformation Index along with Paradigm Initiative undertook a disinformation risk assessment of the media in Nigeria. They conducted their assessment by using the “content”, “operations” and “context” of the sampled news sites as indicators. This assessment was necessary to: reevaluate media affiliations; reevaluate news sources; measure strengths and weaknesses of the market with real figures; create a new journalism standard; and demand accountability.



⁵ Paradigm Initiative works to promote digital inclusion and digital rights - <https://paradigmhq.org/>

The key findings of the assessment are as follows:

- 1) More than fourth-fifths of the sites in the sample assessed had a medium risk rating;
- 2) Most sites in the Nigerian sample provide neutral and unbiased content;
- 3) Five news domains were classified with a high disinformation-risk rating;
- 4) Nigerian news domains in the sample have ample room to improve their operational transparency and accountability as a means of reducing their exposure to disinformation risks.

Fig 1: below shows the disinformation risk rating of the sample assessed.

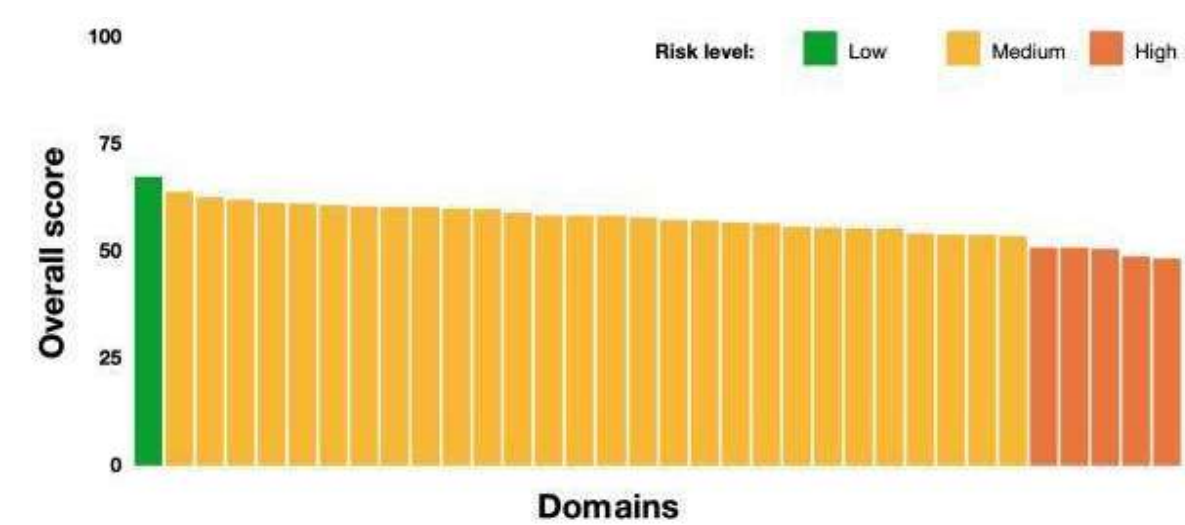


Fig 1: Disinformation Risk-Rating of New sites and Blogs

Sequel to this, the following recommendations were proffered for the news industry to implement:

- 1) Publish explicitly operational standards and ownership structures of the news organizations;
- 2) Make public the sources of funding for news, as this helps to build trust among users and serves as a check for conflicts of interest in reporting;
- 3) Adopt or develop a standard fact checking procedure, both prior to and after the publication of news content;
- 4) Ensure the adoption and publication of bylines/laws and/or related policies to promote transparency and accountability. In cases where there is concern for journalist's safety, sites could provide clear and justified policies explaining the need for author's anonymity;
- 5) Ensure the adoption and publication of comment reviews policies in order to restrict harmful content generated by users; and,
- 6) Make unequivocal editorial principles, especially with regards to independence from political affiliations, advertisers, government and any other external influences.

See paradigmhq.org/wp-content/uploads/2021/07/2021-06-15-Nigeria-Risk-Ratings-Report-Online.pdf for more details on the findings and recommendations.

3.3 Feedback on presentation

Questions were posed to understand how the news sites were selected; and why the methodology of the assessment merged mainstream news sites with blogs. Answers were provided that the selection of the sample was based on the traffic on the sites - the sites selected are the most visited in the country including mainstream news media and blogs.

The challenge of defunding news sites and blogs that disinform was raised considering the capitalist interest of the advertising world. A participant raised the concern of how difficult it may be to conduct a disinformation risk assessment for local/traditional news sources who are difficult to reach and sometimes situated in conflict locations. A disinformation risk assessment for this segment of the media is important as they are critical sources of information. In other comments, participants also agreed that increasing editorial independence, objectivity and attribution of news media, particularly blogs, would help improve the information environment.

4.0 Breakout Group Session

The breakout sessions were organized into three (3) groups that focused on different scenarios of mis/disinformation. The groups were tasked with identifying: potential harms; real and perceived impacts of mis/disinformation; and how to proactively and reactively push back against the impact of mis/disinformation. The last task was to identify tactics, techniques, tools and platforms that can be utilized in the fight against mis/disinformation.





4.1 Group 1 - Social Justice Movements

The scenario that this group analysed was the picture of a young man carrying another individual covered with a blood-stained Nigerian Flag. This image was circulated on Twitter as part of the pictures taken at the Lekki toll gate after the shooting of unarmed protesters by the Nigerian Army and Nigeria Police. However, investigations revealed that the photo was taken more than 2 weeks before the Lekki incident at a stage play in Akwa-Ibom, a location that is about 659km from Lekki.

The group's deliberations are broken down as follows:

Impact: Real and perceived impact of this mis/disinformation on the ENDSARS and social justice movements is that it created doubt about the demands and information provided by the ENDSARS movement. It also caused confusion and promoted cynicism as members of the public were unsure of what

to believe and what to disregard. This also discredited the movement to a large extent. It reduced the sympathy that the movement received, and further emboldened the government to discredit the genuine evidence of violations against the protesters. Lastly, it promoted the government's propaganda of the ENDSARS movement being a revolution aimed at scuttling the government.

Push back Approach: Establish a coordinated central communications system; engage in fact checking; record and document effectively with the aim of telling the authentic stories of the movement. This should be done using technologies and methods that ensure the documentation of trustworthy audiovisual materials that make manipulation difficult and verification easier.

4.2 Group 2 - Democracy and good governance

This group analysed the case of a viral voice note that was circulated during a Kenyan Presidential election. The voice note alleged gross electoral malpractices including rigging and voter intimidation which resulted in widespread violence, destruction of lives and property and the ultimate suspension of the elections. It was later realised that the voice note was first recorded 4 years prior and was merely recirculated in the aftermath of the election.

The analysis of the group are as follows:

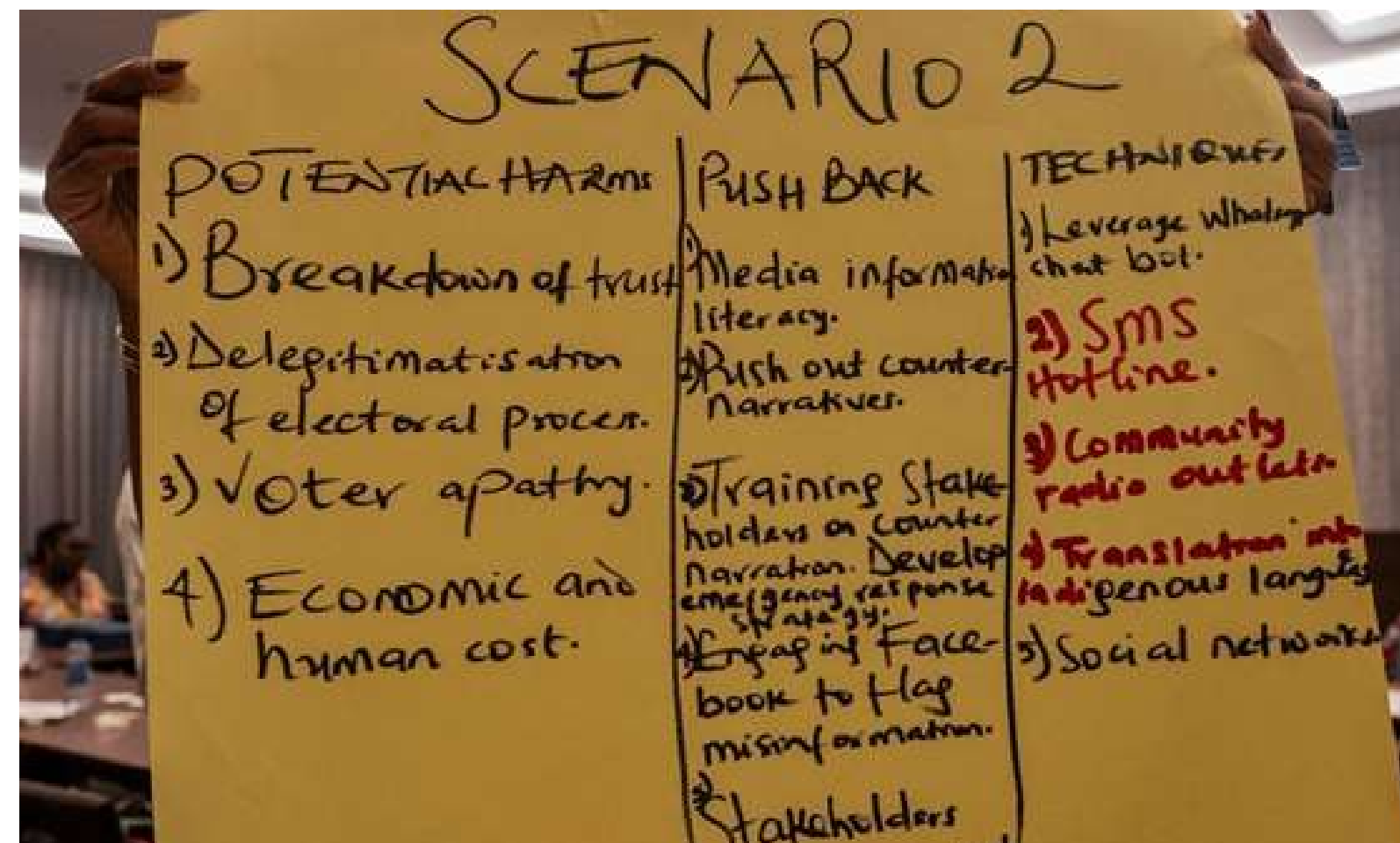
Potential harm/Impact: The viral video led to offline violence; loss of lives and property; loss of businesses and employment. It also led to loss of trust in the government and the electoral process; and increased voter apathy.

Breaking WhatsApp wildfires and Push backs: Pro-active approaches include increasing media literacy with training on detection techniques. Engaging with

social media platforms such as Facebook, Twitter, and WhatsApp on improving in-platform detection and flagging systems. Reactive measures include creating counter narratives.

Techniques, tools and platforms: Technological solutions include leveraging WhatsApp chat bots. Though this can be challenging, it can be deployed specifically for reporting fake news during elections and cases of sexual and gender-based violence (SGBV). Non-technological solutions include use of grassroots social networks, town criers amongst other grassroots channels to provide authentic news on sensitive issues.

Fig 2 below shows the report of group 2 as it was presented to the plenary of the convening.



4.3 Group 3 - Conflict

This group considered a viral video of a farmer who had a brawl with a herder over a football betting game. The video was shared on WhatsApp and Facebook, leading to a full-blown, deadly violent ethnic conflict between the farmer and herder communities in the state resulting in attacks and reprisal attacks. Unfortunately, the farmer chose to record a video in which he misrepresented the truth of how he sustained his injuries. His intention was to deliberately misinform and escalate the impact of his loss. The result was a full blown, deadly violent ethnic conflict between the farmer and herder communities in the state resulting in attacks and reprisal attacks.

The group's conversations generated the following:

Impact: The mis/disinformation escalated existing ethnic conflicts and divides. Lives and property were lost, resulting in displacement and forced migration.

Push back within local context: Measures include media literacy through opinion leaders, faith and

community leaders, women and youth leaders and other channels. Others include use of basic fact checking techniques; news consumption from trusted communication channels; and teaching simple video verification skills. Other measures are: formulating human rights centered laws and policies, and enforcing regulations.

Techniques, tools and platforms: Solutions include use of Information, Education and Communication (IEC) materials - fliers, handbills, videos; conducting town hall meetings and sensitization workshops; use of jingles on community radios; use of verification tools; teaching basic journalism skills; and, flagging of false content.

4.4 Discussions

Questions were raised regarding regulating the new media and whose responsibility it is or should be? Generality of the participants' agreed that:

- 1) Ideally, governments should be responsible for regulation. However, due to rising authoritarianism across the continent even within democracies, it has become imperative to proceed cautiously about granting government such far reaching responsibilities that are susceptible to abuse;
- 2) Governments are also guilty of mis/disinformation and therefore it will be challenging for them to hold themselves accountable;
- 3) Technology platforms should invest adequately in their community guidelines in order to strengthen their capacity to effectively fight mis/disinformation;
- 4) New laws are essentially to meet the emerging threats of technological, however there is the need for effective implementation of existing laws, respect of the rule of law and compliance with international human rights standards particularly the right to freedom of expression;
- 5) Often, Big tech companies have no legal personality in countries where they operate. This raises the question of judicial jurisdiction which leaves aggrieved parties with limited opportunities through which they can get justice in the court of law; and,
- 6) The need for a transnational solution where regional blocks can agree on regulatory frameworks that they would abide to. The European Union's General Data Protection Law for instance helps inform global regulation on data protection. However, regulatory policies like this can be over regulatory in some applications; and may not respond to context specific concerns. Therefore, a bottom up and top-down approach (a mix of national, transnational and global solutions) of problem solving is what might be most effective in this context.



5.0 Developing trustworthy videos - WITNESS

One of the earliest examples of video evidence being used to preserve the truth can be traced to General Dwight Eisenhower's request to the US Congress in 1945 to demand that members of the press and Congress be sent over to Europe to document the atrocities of the Holocaust. Eisenhower's fears were that if there was no proper documentation of the events, in the future, the evidence could be misrepresented to suit an entirely different agenda.

An excerpt from the letter is reproduced below:

"The things I saw beggar description...The visual evidence and the verbal testimony of starvation, cruelty and bestiality were so overpowering as to leave me a bit sick I made the visit deliberately in order to be in a position to give first-hand evidence of these things if ever, in the future, there develops a tendency to change these allegations merely to "propaganda".

This further demonstrates the power of video evidence in confronting misinformation/disinformation.

Unfortunately, today, there are several technological applications that are used to alter or manipulate pictures and videos. The following examples were created during the convening to demonstrate in real time what is possible using apps that are easily accessible, sometimes for free.

Adobe Lightroom: The output of this image editing application was demonstrated to the participants using a picture of the convening hall taken earlier in the day. Fig 3: shows an image of the convening room that was manipulated using Adobe Lightroom. On the right is the original image without any alterations. While on the left is the image with one participant erased from the photo.



Fig 3: Lightroom Editing Output

TweetGen: This is a web-based application that is used to generate fake tweets. An output of this application was shown to the participants. Fig 4 below, is the image of the fake tweet that was generated using the account of one of the participants – though with permission.

Other examples of video manipulation software applications that were demonstrated include Adobe's Content Aware Fill tool that is used to manipulate videos to make objects disappear without a trace; Reface app that enables users swap faces to create realistic alternatives; and Nvidia which makes it possible to alter seasons within videos.

Due to the ease with which videos can be manipulated using the foregoing examples, WITNESS therefore develops approaches and techniques that improve the reliability and trustworthiness of video evidence. This is done to make video evidence secure, difficult to undermine and less susceptible to manipulation. One of the ways to document more trustworthy video evidence is to capture the WHO, WHAT, HOW, WHERE and WHEN.

WHO: Document items such as: number of officers; supervisors; badges/IDs; uniforms; vehicle license plates and other distinct details. These serve as unique identifiers of persons and objects.

WHAT/HOW: Capturing what is happening and how it is happening is also important. For instance, proving that there was the use of excessive force will require capturing things such as deployment of tear gas, pepper spray or water cannons. You might also need to document an arrest, the use of racial slurs, threatening or coercive language. Context is also very important. So, endeavour to record for as long as possible.

WHERE/WHEN: It is important to capture landmarks and items that help to provide evidence of where and when an event might have occurred. Some of these will include clocks; newspapers; street signs; location of sun or where the shadows fall, etc.

6.0 Mis/disinformation in the context of gender and inequalities

Stand to End Rape Initiative ⁶ (STER) provided an insight on mis/disinformation in the context of sexual and gender-based violence (SGBV) in Nigeria. The problem of gendered mis/disinformation has long been biased against women and girls. This is evident from cases of sponsored advertisements peddling falsehoods against women and the use of social media bots that repetitively propagate dangerous narratives that expose women and girls to online and offline violent attacks. Others include unprofessional journalism targeting women in political positions and leading anti-SGBV advocacy; and statements in media by notable personalities reinforcing traditional gender stereotypes.

This gendered misinformation and disinformation selectively pushes false narratives to counter survivors' stories and a culture of shaming that propagates sex-based falsehoods against women. These create harmful information that results in polarizing debates against policies, strategies and legislation that are feminist-centered.

Gendered mis/disinformation is a serious challenge with very negative consequences for the female gender. Some clear examples include the false claim attributed to the former President of Tanzania, Late John Magufuli in which he was falsely quoted to have advocated for men to marry two or more women in order to reduce prostitution among ladies. The consequence of this is that it increases the objectification of women and promotes the belief that single women are promiscuous. Another is a sexist response by the Nigerian President, Muhammadu Buhari in which he said that the First Lady of Nigeria, Aisha Buhari belongs to his kitchen, his living room and "the other room". The impact of sexist mis/disinformation such as these especially from the highest political level are that: it reinforces the misinformed stereotype that women are only useful for sexual gratification, breeding children and house chores. It also wrongfully promotes the idea that women should only be seen and not heard.

The following were recommendations for how to combat gendered misinformation and disinformation:

- 1) App developers and tech platforms should take definitive steps to make their products safe for women, girls and other vulnerable groups.
- 2) Developing strategic partnerships that promote media literacy at the grassroots where there is higher susceptibility to believing false information;
- 3) Promoting access to education for women and girls which helps women and girls in deciphering facts from myths and falsehoods; and,
- 4) Legislation and policies that combat and criminalize gendered mis/disinformation.



Fig 4: Fake Tweet Generated Using TweetGen

⁶ Stand to End Rape Initiative <https://standtoendrape.org/>



6.2 Discussions

Following general discussions, there were comments that indicated that Premium Times, a leading independent newspaper in Nigeria has a program that is aimed at training journalists on reporting SGBV. They are developing a guideline for reporting on SGBV which will help solve the problem of how journalists have wrongly reported SGBV cases. There were also opinions that:

- 1) Highlighted the need to increase the accountability of media houses;
- 2) Raised the need to frame news stories in ways that empower the survivors and encourage others to speak up. Also, the need for increased thematic reporting of SGBV issues as against episodic reporting;

3) Highlighted the context of SGBV in the Northern Nigeria where survivors of sexual violence from Boko Haram and the military are invisible. Further dehumanization of survivors through sensational reporting of SGBV by journalists was also raised, and;

4) Highlighted the difficulty that women face traditionally particularly in Northern Nigeria in discussing issues of SGBV which is considered a taboo which is further reinforced and amplified by the lack of political and religious roles that women play in communities. This skews the narrative against the survivors because, the gatekeepers - the men, are responsible for shaping the narrative and telling stories of survivors if they ever get told.



Day 2

7.0 Day 2 Preamble

Participants shared their thoughts on how Day 1 sessions were useful for:

1. Appreciating that mis/disinformation is a regional problem and not just a national problem unique to their countries and that solving mis/disinformation requires an interdisciplinary approach of problem solving;

2. Understanding the impact of mis/disinformation on different actors including those that create and promote mis/disinformation and those that are targets of mis/disinformation;

3. Learning the nature of tools that are used in manipulating content and the need for more transparency in regulating the information space.

7.1 Group Exercise

The participants were divided into 5 groups of three participants per group. The groups competed against each other on lessons learnt from Day 1 sessions. Some of the questions were on the approaches to combating mis/disinformation and the impact of mis/disinformation. Fig 5 below, shows one of the questions that was asked as part of the group exercise.

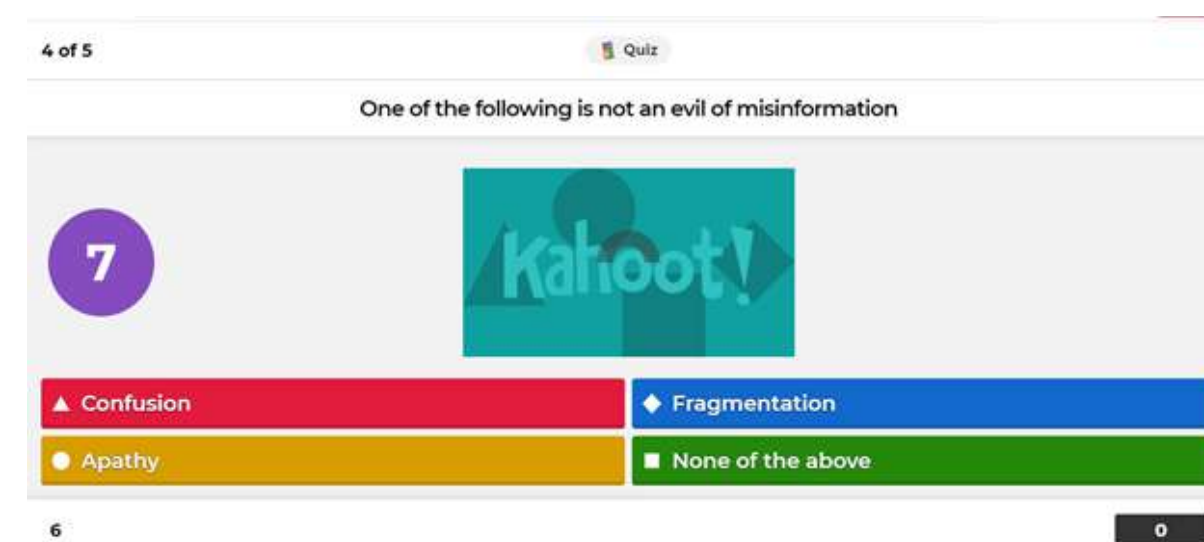


Fig 5: One of the Questions from the Group Exercise

7.2 Panel discussion

The panel discussion featured 5 panelists who discussed addressing mis/disinformation from a human rights perspective and the impact of mis/disinformation on social justice movements. Other topics they discussed are the role of the media in challenging mis/disinformation; and the impact of mis/disinformation on human rights defenders and activists.

The Panelists were representatives of human rights

and social justice organisations, journalists and activists including: Amnesty International, Global Rights, Foundation for Investigative Journalism, DF&co Law Firm, and Allamin Foundation for Peace and Development.

Fig 6 below shows Panelists in the panel discussion



Fig 6: Panel Discussion on Day 2 of Convening

In discussing how mis/information has introduced new challenges, the panelists expressed their views that:

- 1) Mis/disinformation has a huge impact on politics, governance and human rights in Nigeria, especially during elections, protests and in conflict regions.
- 2) Uncovering the truth is crucial to the advancement of human rights and governance. Human rights defenders and activists must anchor their work on unequivocal truth. This mitigates the threat of their work being discredited, refuted or controverted, though attempts may be made;
- 3) Mis/disinformation leads to breakdown of cohesion in the community. This increases prejudice and a breakdown of trust;
- 4) The depth and impact of mis/disinformation on civil society is staggering and affects the capacity of civil societies to do their work. Civil society organ-

isations have to continuously grapple with defending and countering narratives of government. These narratives are deliberately propagated to misinform and pitch human rights organizations against the State. A clear example is the constant labelling by the government that Amnesty International in Nigeria is a foreign agent. This is done to delegitimize their work and impact;

- 5) Mis/disinformation now makes it increasingly difficult to hold the government accountable and secure justice for citizens. Governments continue to explore loopholes using mis/disinformation as an excuse to evade accountability. This had an impact on the evidence presented by victims of the Lekki toll gate shooting.

On strategies to push back against coordinated strategic disinformation efforts by government, Panelists agreed that:

Systemic response is needed which requires synergy amongst civil society organizations irrespective of their thematic areas of focus since the problems that the sector faces are interconnected. This is important as the government's work to discredit is targeted and systematic.

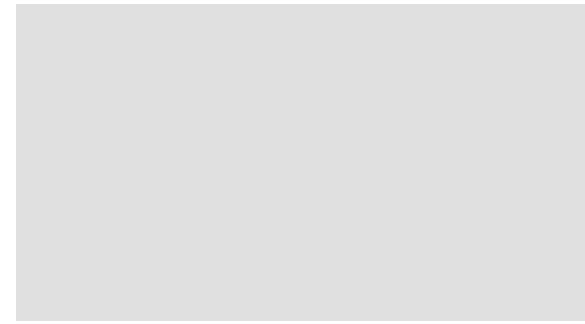
On how people can respond to the impact of information warfare in communities particularly in North East Nigeria, Panelists expressed the need to:

- 1) Empower survivors with skills, which will help them lead independent lives to tell their own stories and demand accountability from government and other actors.
- 2) Challenge the media to reevaluate their values and role in the society. This should include a change in certain practices that jeopardizes their independence such as accepting financial rewards for publishing news pieces. In essence, there is the need to build high ethical standards for journalists, provide them with requisite capacity for work in the age of technology and provide adequate remuneration and welfare.

- 3) Protect journalists from government intimidation and repression. This should include advocating for the repeal of repressive legislations that leave journalists and critical voices vulnerable

On solving the problem of censorship, repression and how to create alternative channels of communication, the Panelists were of the opinion that:

- 1) There is a need to identify authentic voices in the media and provide them with resources and training needed to uphold the truth. Also, the importance of long-term strategic planning to protect credible journalists from being poached by political actors was highlighted;
- 2) Telling the truth consistently is an effective way to fight government's repression and censorship which gives government the ultimate opportunity to discredit itself;



3) There is a need to encourage community ownership of media organizations since most media organizations are owned by the government, people in government and politicians. This way it becomes easier to counter deliberate misinformation and disinformation by the government and other actors.

4) Consistent mass citizen media literacy and education is very integral for citizen reawakening and understanding of how to engage with information in a critical and ethical way. The use of community radios will be key in this respect.

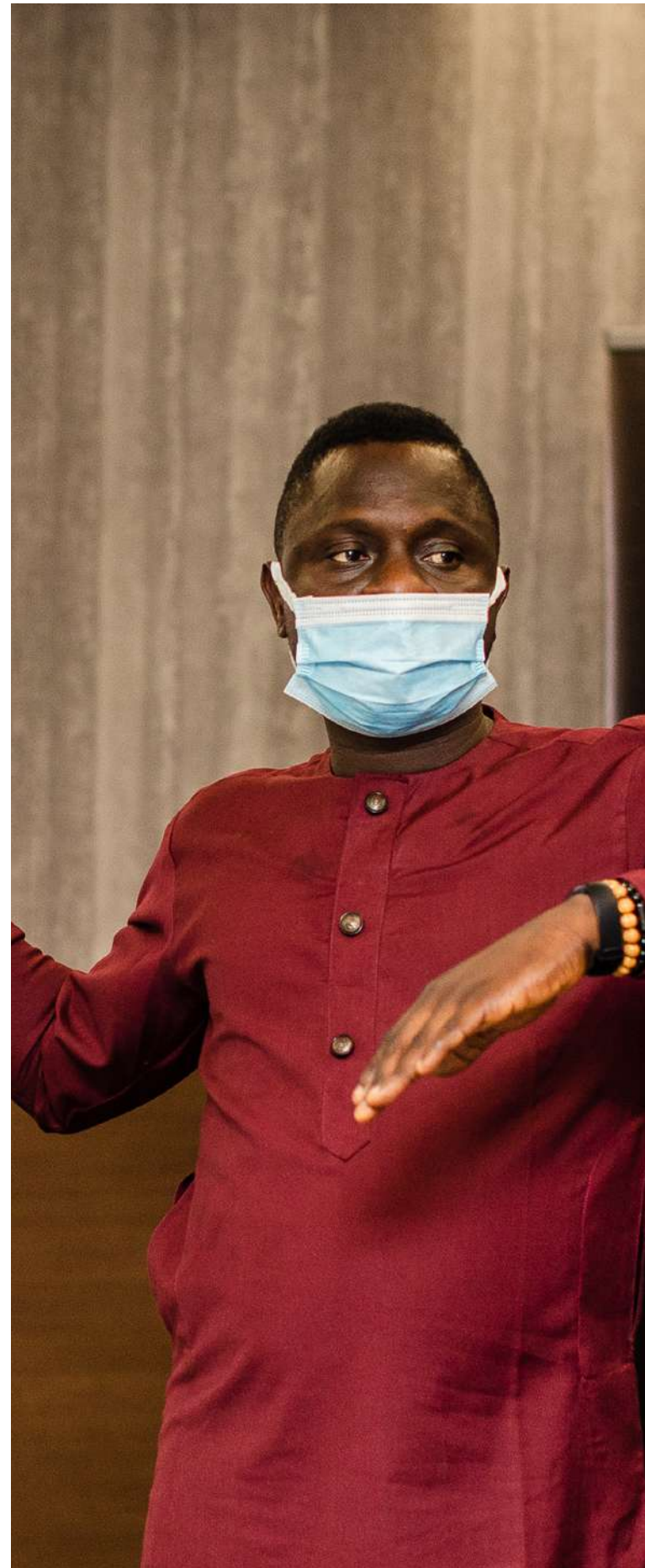
On how to counter government's narrative of mis/disinformation and still make impact within communities as CSOs, Panelists expressed the opinion that:

1) Integrity and credibility of an organization matters in the work of countering government's disinformation campaigns;

2) Verification of evidence - use of rigorous triangulation is very important in ensuring that only the facts are reported and amplified;

3) Serious efforts must be made to ensure that people are empowered to tell their stories from a human angle; and,

4) Civil societies need to be more strategic in investing and deploying digital tools and technologies in addressing mis/disinformation. It might consider crowdsourcing its response to mis/disinformation narratives by strategically creating verification platforms and networks within the civil society space and not necessarily relying on technology platforms to take the initiative. Some participants raised the need for civil society to conduct a media mapping exercise to identify credible media organizations and journalists and fund them effectively to enable them to do their work without fear.



8.0 Internet Shutdowns and Content Takedowns in Africa – WITNESS

Internet shutdown is the intentional disruption of internet or electronic communications rendering them inaccessible or effectively unusable for a specific population or within a location often to exert control of information.

Many African countries have engaged in the restriction of press freedom, free speech and human rights in the digital space. Recently, the governments of Uganda, Nigeria, Zambia, Eswatini, Ethiopia amongst others, have embarked on internet shutdowns, both total and partial shutdowns. All of these have been to use state media to perpetuate its own narratives and suppress counter narratives from the citizens and civil societies.

It is important to note that:

1. Fact checkers and journalists find it hard to do their work during an internet shutdown;

2. Hegemonic narratives can be easily spread; and,

3. AI moderation errors can fuel misinformation.

It was highlighted that technology platforms having oversight boards is useful in addressing some of the issues of infringement that take place on technology platforms. Facebook has an oversight board which is commendable but needs to be reviewed to improve its effectiveness and response.

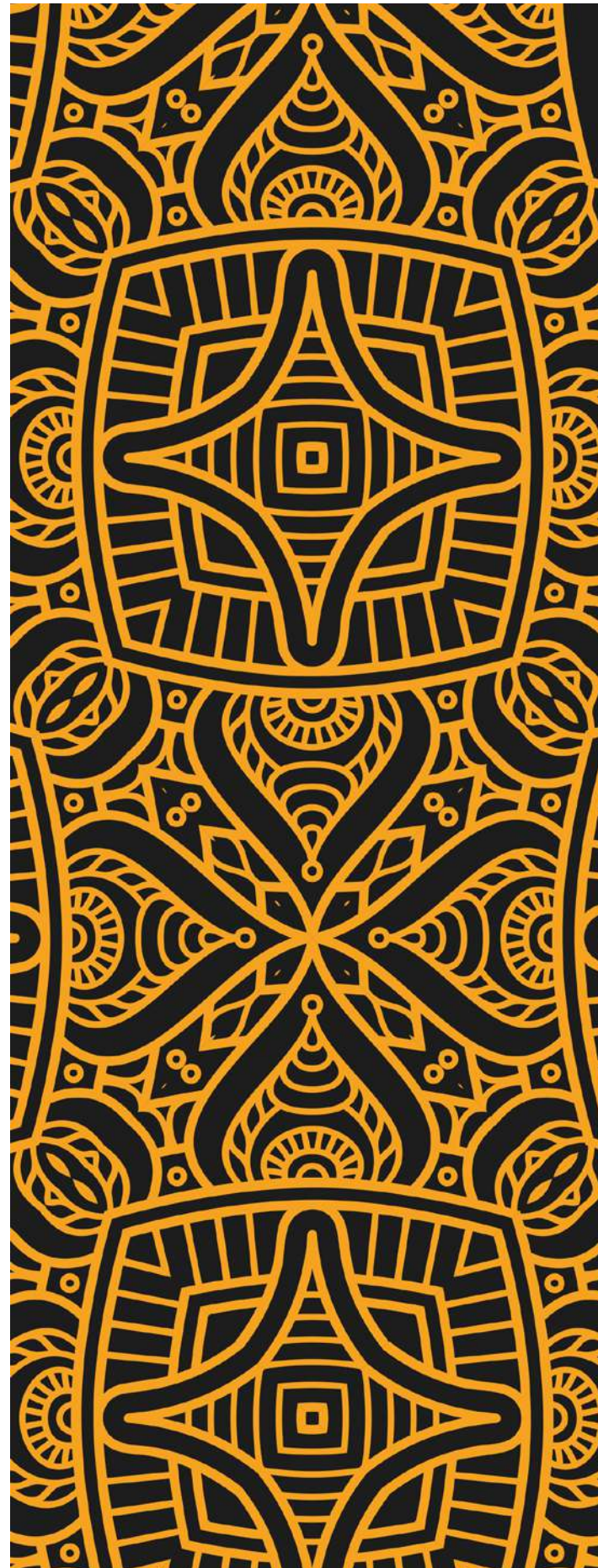
WITNESS disclosed that it has a campaign on internet shutdown called #EyesOnShutDowns. The campaign aims to help people better prepare for documenting and preserving trustworthy videos during an internet shutdown.

8.1 Breakout Group Session

Threats to human rights are emerging at the intersection of artificial intelligence, misinformation and disinformation. When it comes to systemic challenges WITNESS is on the watch for both existing problems – like how large amounts of critical human rights footage are lost because of the arbitrary decisions of platforms – but also emerging threats and opportunities, such as those posed by ‘deepfakes’, new ways to manipulate media to make it look like someone said or did something they never did.

WITNESS’s Technology Threats and Opportunities (TTO) program aims to identify and proactively address existing and emerging technologies’ that undermine truth, justice and accountability. Part of the successes of this program is the introduction of the blur tool by YouTube. The blur tool allows users to blur out unique features of the subject in a video they intend to upload. This was borne out of WITNESS’ advocacy efforts to ensure privacy protection of vulnerable groups on the platforms. One of the purposes of this session was to generate feedback that WITNESS can engage technology platforms with.

Hence, the participants were divided into three groups and were tasked with the objectives of: identifying solutions that help break internet echo-chambers and texting platforms wildfires and navigating privacy challenges related to messaging on the platform. Other objectives included exploring solutions on what WhatsApp can possibly do to reduce the spread of mis/disinformation on its platform; Identifying what CSOs and grassroots movements can also do; and, exploring which technology tools might help to combat these challenges?



8.2 Group reports

Table 1.0 shows the report of the three groups for the four different categories of actors in the mis/disinformation landscape.



Fig 7: Shows a Group during a Breakout Session on Day 1 of Convening



Table 1:0: Report of Group Session

Actor	Group 1	Group 2	Group 3
Individuals	<ul style="list-style-type: none">- Should verify before sharing- Flag suspicious content for institutions and others to verify- Be active participants in discussions- Have basic knowledge of verification tools	<ul style="list-style-type: none">- Individuals should be more selective with their sources of information- Reach out for verification of information before sharing- Report suspicious information on platforms – especially violations of community guidelines	<ul style="list-style-type: none">- Flag false messages on social media platforms- Choose not to share and refuse to propagate misinformation- Verify information via tools like google reverse image search
Media	<ul style="list-style-type: none">- Set up fact checking desks- Engage in self-regulation as an institution	<ul style="list-style-type: none">- Set up digital and non-digital means of verification- Train personnel on modern verification tools- Improve general ethics, fact checking, verification skills- Own up to errors and have clearly defined correction policy- Media literacy campaigns	<ul style="list-style-type: none">- Media investment in understanding the platforms and technologies in use- Have adequate knowledge on how to interface on platforms- Check and cross check information
CSOs, Grass-roots activists & Campaigners	<ul style="list-style-type: none">- Public sensitization and education – early warning signs, misinformation, how to sort, verify- Collaboration with stakeholders- Establishing & supporting fact checking mechanisms	<ul style="list-style-type: none">- Get involved with media literacy campaigns- Non-tech-based campaigns- Advocate for transparency and proactive disclosure of information	<ul style="list-style-type: none">- Define and drive policy direction- Work at protecting citizen data.
Tech tools and Policies	<ul style="list-style-type: none">- Verification and translation tools in local languages including sign languages- User friendly flagging tools across social media platforms- Setting up policies against disinformation that are enforceable locally.	<ul style="list-style-type: none">- Add number of forwards to WhatsApp messages- Verify if a forwarded message exceeds a particular number- Increasing media literacy- Continuous engagement with people, CSOs and grassroots activists	<ul style="list-style-type: none">- Set up identifiers of group of words that are termed as harmful- Increase verification of information shared in indigenous languages





9.0 Presentation by Facebook

The representative of Facebook commended the feedback from the groups as informative and helpful in understanding the gaps of what and how technology platforms such as Facebook approach misinformation.

Facebook has a three-prong approach to dealing with mis/disinformation on its platforms:

1. Identify and remove content that violate Facebook's policies - community standards and content policies which guide the types of content that are allowed on the platform;
2. Reduce distribution of content that doesn't necessarily violate community standards and content policies but have been determined to be false;
3. Inform people to help them decide what to read, what the truth is and what to share.

According to the representative, Facebook doesn't compromise on removing any content that is:

- 1) Clearly capable of causing offline harm, irreversible harm;
- 2) Misleading or manipulated. This is a recently instituted policy because of the growing threat of manipulated content on Facebook platforms;
- 3) Misrepresenting democratic process - elections, voting process, misinformation about dates, times, eligibility, methods of voting.

Facebook doesn't solely decide on content that will be removed, rather, it works closely with external experts and groups to take these decisions. These experts and groups work on policies, imminent harm, threats, language, sense of language, situations around the world in order to be able to respond promptly to context specific information.

On COVID-19 mis/disinformation, Facebook provides banners that lead people to genuine, credible and official information from agencies in different countries. It also flags misleading information and shares up to date information from national health care centers.

This year alone Facebook has removed over 7 billion misleading pieces of content that speaks to vaccines, transmission, health practices. The work of getting the platform free from mis/disinformation is a very challenging one as there are about 4 million users sharing content at any given time. Having a system that assesses all this information independently, and provides actionable feedback is really daunting. In this regard, Facebook works with fact checking organizations to provide language support in all countries where Facebook works.

A clarification was provided on the flagging of the Nigerian #ENDSARS campaign content by Facebook. What happened was that the use of an image of the deadly SARS virus was posted using the #ENDSARS hashtag causing the Facebook algorithm to flag that post and subsequent #ENDSARS posts as misinformation. This directly undermined the credibility of the movement and questioned the legitimacy of authentic videos from the protest.

Facebook is working to improve its mis/disinformation detection system by:

1. Setting up of the first content review team on the continent;
2. Increasing local language capacity; and,
3. Increasing the number of fact checkers and fact checking associations and groups.

Facebook is also invested in increasing and scaling up digital and media literacy. This is to improve on how the platform's response to mis/disinformation can increase the ability of users to detect and flag mis/disinformation.

In responding, participants highlighted the inefficiency of Facebook's algorithms to correctly flag or label information on its platform. Participants sighted instances where wrongful labelling of information on the platform have resulted in the delegitimization of authentic voices. Also, participants emphasised the need for in-app verification tools and media literacy.

The dangers of WhatsApp wildfires were equally discussed. The introduction of forwarding limits by WhatsApp was commended.

However, participants expressed the need for WhatsApp to invest more in tools that would check the autonomy of WhatsApp groups which have become a conduit for mis/disinformation. It was expressed that the ability of individuals to simultaneously administer numerous WhatsApp groups allows the platform to be a potent source of mis/disinformation in spite of the forwarding limit that was introduced. This is significant because group administrators have autonomous powers to control the information flow of the groups they create. They can consistently push out coordinated mis/disinformation contents across numerous groups they administer given that there are no caps to the number of groups an individual can administer at a given time.

The ease with which people can harvest contact details of individuals on WhatsApp groups is also worrying. This allows people to more easily include others

in groups and broadcast lists without their permission. To curb the spread of mis/disinformation on WhatsApp, participants made the following recommendations;

1. Introduce a limit to the number of groups an individual can administer at a given time.
2. Introduce in-app verification tools.
3. Introduce in-app media literacy content.
4. Better protect the data of individuals in WhatsApp groups from nefarious actors who could exploit such access.
5. Simplify the process of setting up chat bots particularly for CSOs who want to help fact check mis/disinformation in real time.

9.1 Sharing of practices from across West Africa

Three participants shared their experiences working in Nigeria and Ghana on combating mis/disinformation.

Center for Democracy and Development (Nigeria) shared revealing lessons from qualitative research work carried out in Kano and Oyo states on mis/disinformation. The lessons shared corroborated earlier findings that politicians strategically employ social influencers to help frame and drive often misleading narratives both online and offline. Some of them, which are referred to as data boys and Shekpe boys, operate a very structured command and control system in which they are able to infiltrate spaces and shape opinions. Some of the measures CDD has employed in combating mis/disinformation include fact checking, civic education and physical boots on the ground who directly debunk false narratives. Community radio was also cited as an effective means through which communities can push back against mis/disinformation. A case in point is the Chicoco radio station which is built by and for the residents of the waterfront slum communities of Port Harcourt in Nigeria. Through this radio station, the waterfront community is able to tell their story independent of external influences.

A Knight Fellow with the International Center for Journalism (Nigeria) shared insights of their work with Africa Check, Dubawa, and The Cable to improve dissemination of fact checked information. Digital strategy of working with partners using short videos is also conducted with the help of digital influencers who amplify fact checked information. Media literacy programs – webinars and skits are implemented. These approaches come with their challenges of commitment from the influencers who may later become polarizing figures or may demand payment for their participation in the campaign. However, engaging with influencers that have a certain connection with the issues helps secure a longer-term commitment. It was also expressed that getting influencers to sign an agreement of expectations can be useful in securing their credible commitment. Building the capacity of journalists such as Dubawa's fellowship

programme for journalists is useful for building movements of influencers committed to combating mis/disinformation. Finally, the use of credible community gatekeepers, grassroots networks and youth leaders was highlighted as useful in sharing fact checked information in local communities.

Joy News (Ghana) - Building public trust has been challenging especially at a time that different media organizations publish conflicting accounts of a specific story without verifying or fact checking. Fig 9, below shows how different captions from newspapers in Ghana on the same news subject. The discrepancy in the figures can deplete public trust in media organizations and engender mis/disinformation.



Fig 9: Ghanaian Dailies with different Captions on the Same Subject

To build public trust, it is a practice to develop a theme based on the public's interest and have such a topic discussed over a period of time. This has con-

9.2 Final Spectrogram

A spectrogram was conducted at the end of the convening to check if participants' opinions had changed on if humans or technology or both are responsible for mis/disinformation. The result didn't change significantly from the result at the commencement of the convening. Two participants indicated that humans were responsible while eleven (11) participants indicated that both humans and technology are re-

sistently helped in building the trust and confidence of the public in the media.

sponsible. These numbers are compared to the result of the Spectrogram at the commencement of the convening which indicated that eight (8) participants thought mis/disinformation is purely a human problem, while the remaining participants thought mis/disinformation is both a human and technological problem.

10.0 Where do we go from here - Next Steps?

These are immediate interventions agreed to be implemented by WITNESS and other stakeholders:

1. Journalists will be trained based on considerations and feedback from the convening. Specifically, journalists will be trained on digital verification and effective video documentation. The aim of this training is that the journalists will take the lessons into their news rooms and networks;
2. Training for activists and human rights defenders, especially those that work with physical geographical locations and thematic areas such as SGBV, governance, human rights and insecurity;

3. Media literacy campaigns will be rolled out in collaboration with other stakeholders where possible;
4. Embark on advocacy with policy makers, technology platforms, application developers and campaign efforts with other actors including civil societies, human rights activists and defenders; and,
5. Create a Truth Tellers Network where conversation will continue and resources will be shared.



ANNEX

List of participants

- Allamin Foundation for Peace and Development
- Africa Independent Television
- Amnesty International
- Centre for Democracy and Development (CDD)
- Centre for Environment, Human Rights and, Development (CEHRD)
- DF & Co Law Firm
- Dorothy Njemanze Foundation
- Dubawa
- Facebook
- Foundation for Investigative Journalism (FIJ)
- Global Rights Nigeria
- Human Rights Watch
- Joy News – Ghana
- Knight Fellow - International Centre for Journalism
- Media Reform Coordinating Group - Sierra Leone
- Nation Newspaper
- Paradigm Initiative
- Premium Times Centre for Investigative Journalism
- Stand to End Rape
- The Cable Nigeria
- WITNESS

Verify
before
sharing

africa@witness.org