



Report on a two-day workshop on:

Fortifying Community Truth in the Age of Generative AI and Synthetic Media in Africa

Nairobi, Kenya, March 22-23, 2023

Report by Gichunge Jackson and Jacobo Castellanos

Between March 22 and 23, WITNESS hosted journalists, fact-checkers, technologists, content creators, human rights defenders and other community activists from different parts of Africa in Nairobi, Kenya, to identify, discuss and prioritize threats, solutions and opportunities from synthetic media.

This workshop in Nairobi was part of ongoing efforts at WITNESS to engage with partners globally, including the 2019 workshops in Pretoria and Brazil, as well as South East Asia and the United States in 2020, and other in-person and online consultations on deepfakes and synthetic media since.

WITNESS's 'Prepare, Don't Panic' initiative aims to intervene early in the synthetic media ecosystem, focusing on its infrastructure, tools, and policy and legislative aspects. Our work seeks to ensure that the continuous development of synthetic media reflects an understanding of the key threats, solutions, and opportunities prioritized by individuals and communities globally who have experienced human rights violations similar to those emerging in the 'age of synthetic media.' These individuals are both most at risk and yet marginalized from the spaces that shape this emerging technology.

Contents of the report

Summary and analysis	4
Identified threats from synthetic media echo those of previous years	4
What has changed in the threat landscape	5
Addressing the bigger issues	5
Marginalized groups are vulnerable to threats from synthetic media.....	6
Decolonizing synthetic media and embedding diverse experiences.....	6
Establishing clear responsibilities across the pipeline.....	6
Leveraging parallel experiences to develop effective responses.....	7
Specific actions that we need to take	7
Media literacy will not be enough, but it is still as necessary as ever.....	7
Detection tools should be accessible to those that need it most.....	8
Verifiable provenance and watermarks can fortify the truth, but global human rights and accessibility concerns should shape the infrastructure and tools.....	9
Foundational model developers, tool developers, social media platforms and major news media outlets hold a key responsibility to guarantee disclosure.....	9
Synthetic media content moderation at scale still needs contextualized policies and local expertise to deal with threats from synthetic media.....	9
Civil society, academia and grassroot organizations can think more proactively about the creative potential of synthetic media.....	10
Non prescriptive policy templates can help newsrooms and archives develop resiliency against threats of synthetic media while leveraging their potential for creative and impactful expression.....	10
Collaborating and building local and regional networks to influence legislative efforts, technical infrastructure and platform policies.....	11
Report of Day 1	11
Session 1: Kick off	11
Session 2A: Fortifying the Truth - Introduction to Synthetic Media	11
Session 2B: Text-to-Image workshop	13
Session 3: Synthetic media creation in Africa	16
Session 4: Perspectives on synthetic media	17
Breakout group 1 - Gender.....	18
Breakout group 2 - Misinformation.....	19
Breakout group 3 - Journalism.....	21
Breakout group 4 - Digital rights.....	22

Session 5: Video for human rights, artistic creation and advocacy.....	23
Report of Day 2.....	24
Session 1: Threats Spectrogram and Analysis.....	24
Session 2: Current solutions and responses to synthetic media.....	30
Session 3: Mapping solutions and responses.....	31
Breakout group 1 - Journalism.....	31
Breakout group 2 - Human rights defenders.....	32
Breakout group 3 - Policy makers or advocates working on influencing policy.....	33
Breakout group 4 - Creators and technologists.....	34
Session 4: Leveraging synthetic media for advocacy and creativity.....	35
Session 5: Advocacy recommendations.....	37

Summary and analysis

An accelerated threat landscape for human rights that echoes across countries

Despite the hype surrounding synthetic media, it has not yet become a source of concern for urgent and widespread human rights violations. Consequently, many of the threats that were identified and discussed in relation to synthetic media still resemble those that emerged in previous workshops and discussions.

Among those that stand out is the threat that synthetic media, or the idea of it, could be misused by those in power to dismiss demonstrably true evidence by claiming it is false (also known as the **Liar’s Dividend**), to justify the enactment of laws that **limit the potential of synthetic media for free, creative, and impactful expression**, or to ‘**poison the well**’ – to discredit journalists, activists, civil society organizations, and the ‘truth’ they put out in the world. One example shared by a participant [shows a doctored audio clip](#) of former Nigerian presidential candidate of the Peoples Democratic Party (PDP) Atiku Abubakar, Ifeanyi Okowa, vice-presidential candidate of the PDP, and Aminu Tambuwal, the director-general of the PDP’s presidential campaign council supposedly plotting to rig elections in Nigeria.

Women were identified as being particularly vulnerable to threats from synthetic media, primarily because of how the technology has enabled new forms of **gender-based violence**, but also

because the models **perpetuate and reinforce biases and stereotypes**, which is something that affects women of color and intersecting identities more profoundly.

The threat of synthetic media to **spur misinformation and incite violence** was also echoed. Various participants noted that synthetic media could be leveraged to amplify ethnic, religious, and political divisions in different parts of the continent, disproportionately affecting communities in already vulnerable situations.

What has changed in the risk landscape

As synthetic media tools have become more accessible, more people have had the ability to engage with them. They have been able to imagine –or experience– how the tools could impact their lives. This shift has resulted in a reevaluation of the risks and potential harms that can be caused by synthetic media.

For example, the issue of volume has come up before, but upon experimenting with synthetic media tools and recognizing how easy they are to use, the matter of **an information ecosystem flooded with synthetic content** came up more regularly. One of the concerns that this raises, as noted by fact-checkers present in the room, is that **the scale of synthetic content needing verification may not be matched** by the community unless parallel responses are developed.

How synthetic media can be used to manipulate elections has also been discussed at prior convenings, but it is now, markedly, one of the major concerns voiced by participants. It was noted that, while those in power may have the same capacity to generate synthetic disinformation regardless of where they are, there is a higher risk of undermining democracy in historically vulnerable contexts. In these contexts, detection tools may not be as readily accessible, where civil society and independent media may have fewer resources to prepare for synthetic media, and where media literacy is comparably low.

In this workshop, participants also underscored **the impact that synthetic media could have on public health**. Although health mis- and disinformation was recognizable before, the Covid-19 pandemic showed the level of harm that can come from the inability to stop the spread of mis- and disinformation and to guarantee access to trustworthy information.

Addressing the bigger issues

In brainstorming synthetic media's risks and opportunities for human rights defenders, participants brought up underlying issues and common themes that need to be considered as strategies and actions to 'fortify the truth' are developed.

Marginalized groups are vulnerable to threats from synthetic media

Despite the hype of deepfakes and synthetic media, shallow fakes (or images that have been edited without AI-enabled tools) are still the main source of concern for communities at the frontlines of human rights. The main overarching concern echoed in Nairobi is no different from what we have heard over the past years across continents: threats from synthetic media will disproportionately impact those that are already at risk, because of their gender, sexual orientation, profession, ethnicity or belonging to a social group.

Decolonizing synthetic media and embedding diverse experiences into the technology

Throughout the workshop, participants highlighted the fact that synthetic media and the broader Generative AI ecosystem reflect and perpetuate the injustices and biases of the coloniality of power. It is not just that these systems and tools are not designed by, or with the input of, the people and communities that may be affected by its global deployment, but that they work against them by leveraging exploitative and extractive relationships.

This points to the need for the creation of a synthetic media ecosystem that regulates exploitative practices, promotes diversity, local expertise, and contextualization in the development and deployment of mainstream Generative AI tools. Such an approach will help to ensure that human rights defenders and civic journalists can benefit from synthetic media technologies and avoid being put at heightened risk.

Establishing clear responsibilities across the pipeline

There seemed to be a clear sense of agreement among participants that any strategy to respond to the threats and opportunities of synthetic media cannot rely on the actions taken on by end-users. For all the solutions discussed, there is a key role to be played by governments, social media platforms and other technology companies, and news organizations. These stakeholders should promote the development of mechanisms (such as regulation, policies,

functionalities, processes et al.) that proactively tackle threats without placing the responsibility on content creators or consumers alone.

This idea is also reflected in calls from different stakeholders, including the United Nations via its [Guiding Principles on Business and Human Rights](#), various civil society organizations¹, and more recently emerging regulation from legislative bodies², to require companies **to develop comprehensive human rights assessments prior to the deployment of AI models and tools**, including those that facilitate the creation of synthetic media or synthetically manipulated media.

Although this is something that has been recognized by leading companies and individuals in this space, and more and more so as well in emerging regulation and soft law, what is worth underscoring from the workshop is that the creation of these threat-mitigating mechanisms is not enough; ultimately, these upstream stakeholders should be accountable for the harm they cause or fail to prevent as synthetic media infrastructure and tools are deployed.

Leveraging parallel experiences to develop effective responses

Although synthetic media is still an emerging technology, the threats it poses are not new. There is a wealth of knowledge and experience from organizations and communities that have grappled with similar media-related issues. These experiences can inform responses to the threats and opportunities of synthetic media. For example, fact-checkers can build off of their experience addressing more traditional methods of video and audio manipulation, or ‘shallowfakes’; news organizations can draw from best-practices in journalism; and social media platforms can rely on the experience of local experts in content moderation to draft contextualized policies.

Specific actions that we need to take

These prioritized solutions reflect and build off of previous workshops. For all of these solutions, it is important to consider that the underlying issues and common themes mentioned above should shape the way that these solutions are analyzed and enacted.

¹ <https://www.accessnow.org/human-rights-impact-assessment-ai/>

² <https://www.euractiv.com/section/artificial-intelligence/news/ai-act-meps-want-fundamental-rights-assessments-obligations-for-high-risk-users/>

Media literacy will not be enough, but it is still as necessary as ever

Although participants agree that relying on media literacy to detect synthetic media is not a comprehensive, long-term solution, it can still serve as a first line of defense. This is especially true considering that we are in the early stages of the use of synthetic media, and that shallowfakes are still a bigger concern for the participants of the workshop and the communities they belong to, or center, serve and support.

Media literacy campaigns should also be expanded to include processes that lead towards a more informed public about what synthetic media is, as well as what is and isn't possible with new forms of multimedia manipulation. Similar to findings from previous workshops, these campaigns need to be framed within the broader context of mis- and disinformation in an effort to promote critical consumption of content online.

According to several participants, these media literacy campaigns could serve two purposes: They could help prepare the public to view and consume synthetic media more critically, while also empowering individuals and communities to engage with governments, civil society, and companies to develop responses and solutions that reflect their needs, circumstances, and aspirations. Media literacy campaigns as a precursor to public policy was a point of emphasis during one of the sessions in the workshop.

Participants did underscore that media literacy campaigns should be careful not to add to the existing hype around Generative AI and synthetic media. This is because it could cause unwarranted panic while also bolstering actors that may be more concerned about the growth of their products than justice and human rights.

Lastly, participants also emphasized the importance of media literacy campaigns highlighting the creative potential of synthetic media for content creators. However, they emphasized the need for proper guardrails to protect privacy and human rights, both for the content creators themselves and the individuals associated with the content.

Detection tools should be accessible to those that need it most

Detection tools are seen as another part of the strategy that, in and of itself, cannot be relied on to offer long-term solutions. However, they can still play a significant role, especially among

journalists and fact-checkers, as they look to debunk realistic forgeries or dismiss claims that genuine journalistic multimedia content is fake.

An important question raised focused on the ways by which we could ensure that these tools can be accessible to those who need them most. There were arguments that open-sourced tools facilitate broader access, but overall, there was a recognition that unrestricted access would rapidly lead to its obsolescence. Mindful of the limitations of a gatekeeping role, there seemed to be a preference towards identifying key stakeholders, as close to the ground as possible, that could leverage these tools to meet the need of marginalized communities globally, while also preventing it from being rapidly undermined.

It is worth noting that accessibility was used in this context in a broad sense to mean that marginalized communities should be able to *effectively use* the detection tool; that is, that they not be limited by financial or technical barriers to use it in a way that allows them to meet their needs.

As for the role of social media platforms, there were fewer discussions that specifically pointed to the need for detection tools to be readily available to users within each platform, although it can be implied by the conversations that emphasized the need for accessibility — which would make this a requirement in closed networks, where the metadata is often stripped before being shared. Additionally, a few participants did highlight the need for improved back-end moderation, including with the use of detection tools, while others pointed to a platform's responsibility to disclose to viewers when content is generated or manipulated synthetically.

Verifiable provenance and watermarks can fortify the truth, but global human rights and accessibility concerns should shape the infrastructure and tools

AND

Foundational model developers, tool developers, social media platforms, and major news media outlets hold a key responsibility to guarantee disclosure

Considering the threat of the information ecosystem being flooded with synthetic media, participants pointed to the need to have a preemptive 'flagging' system that could facilitate the process, especially for fact-checkers and journalists, of discerning 'real' from synthetic content, and truth from misleading or false content.

Buy-in from social media platforms, major news media outlets, and AI model and tool developers is essential. This is because they play a key role in guaranteeing that the verifiable provenance or watermarks are part of, or attached to, the audiovisual content from early on in its lifecycle, but also in making sure that their users and viewers are meaningfully informed of the nature of the content they are consuming.

The development of these solutions should also include input from global stakeholders, with an eye towards defending human rights, guaranteeing accessibility, and protecting privacy. As with synthetic media developers, the companies, organizations, and governments behind initiatives promoting the use of provenance and watermarking technologies have a responsibility to proactively assess these technologies for their potential to cause harm.

Synthetic media content moderation at scale still needs contextualized policies and local expertise to deal with threats from synthetic media

Content moderation at scale can be more complicated because synthetic media and Generative AI tools exponentially increase the amount of content created and shared online. Automating moderation, including by leveraging AI such as detection tools, or provenance and watermarks, will be necessary, especially in larger technology platforms. However, there is also a sense that their policies should be designed alongside impacted groups, and that there should be clear processes for local experts and experience to be involved in the moderation loop.

Although this gap has been repeatedly highlighted in broader discussions on content moderation, synthetic media could aggravate harms, in part because of the volume of content that can be created, but also because the increasing sophistication of these technologies will further blur the lines between what is real and what is synthetic.

Civil society, academia, and grassroots organizations can think more proactively about the creative potential of synthetic media

Participants in the workshop discussed how synthetic media could help enable creative content that furthers human rights while protecting their privacy and averting harassment or persecution. On several occasions, participants alluded to the use of avatars as a way to protect the identity of individuals while still creating content that can connect with their intended audience. Immersive experiences were also mentioned as a tool that can foster empathy and move people

to action, while the use of AI chatbots could help streamline the design and implementation of effective campaigns.

Non prescriptive policy templates can help newsrooms and archives develop resiliency against threats of synthetic media while leveraging their potential for creative and impactful expression

Research from organizations participating in the workshop show that newsrooms are unprepared to deal with synthetic media. Different experiences working with archives highlighted the same concern. This is especially true for smaller news outlets and community archives that have less resources to access tools and build the necessary capacities and processes.

Drafting non-prescriptive policy templates for newsrooms and archives could help these institutions think systematically about the threats and opportunities that synthetic media poses to them in order to inform the strategies and processes that they put in place to be ready to respond to, or use, synthetic media.

Collaborating and building local and regional networks to influence legislative efforts, technical infrastructure, and platform policies

Although there is an urgency to advance legislative efforts, technical infrastructure, and platform policies as they trail behind technological developments, it is necessary that these processes be carried out in consultation with all relevant stakeholders, in particular the individuals and communities that may be most affected by synthetic media.

There is a need for civil society organizations to form alliances that can help them ‘punch higher’. Participants discussed that, despite targeted efforts, they have not been able to influence legislation and policy, and that these networks can help them gain the credibility and the resources that are often required to get ‘into the room’. One specific strategy for these networks to influence these spaces is to fill in gaps in grounded research and to communicate findings effectively, for example via policy briefs.

There already are networks in place that could be leveraged to this end, though participants did highlight the need for African organizations to have a larger and more leading presence in these spaces. These regionally-led networks would be well-placed to check the propensity to copy legislation from Europe or the United States without proper consideration, or to counterbalance spillovers from the geopolitical disputes between the West and China.

Report of Day 1

Session 1: Kick off

WITNESS introduced the 'Fortifying the truth' framework, background to the regional workshops and the specific objectives of this event. This was followed by a 'two truths and a lie' activity for participants to introduce themselves.

Session 2A: Fortifying the Truth - Introduction to Synthetic Media

WITNESS presented what deepfakes are, how we prioritize threats and what our approach is. The objective of this session was to share some of the key information that would set the stage for the rest of the two-day workshop.

The speaker explained how deepfakes are being used in simple commercial apps. They also discussed how synthetic media can be created using mobile phones, highlighting the many possibilities that the technology offers.

The speaker then delved into the vacuum that exists in the middle ground of advanced machine learning and advanced research customization, which is part of the larger generative AI and synthetic media landscape. They also discussed the importance that training data has for the development of these advanced technologies.

Moving on, the speaker highlighted the many emerging possibilities with deepfake and synthetic media. These include altering video just like photographs, creating realistic voice manipulated videos with tools such as Google Pixel Prime, editing face features, and styling environments. He further gave examples of these possibilities, such as real-time face and expression mapping, motion transfer, and the usage of full 3D body deepfake for videos.

The speaker also discussed the use of text-to-image tools, such as Dall-E-2, which can create synthetic images, as well as their use in videos. They demonstrated the use of text prompts in text-to-video tools, including a feature that allows users to transfer one event from one video to another.

The speaker then turned their attention to voice and skin cloning, emphasizing the increasing use of voice cloning in synthetic media. They gave an example of how words can be used to clone someone's voice, and how tools like Dall-E 2 can capture the background noise of a person's voice, making it appear as though they are speaking from a telephone.

The speaker also touched on the availability of synthetic media tools for public use, such as Prime Voice AI, which can be used for good or for ill. They also discussed the trend of text-prompt to voice choice, which allows users to write a story and have another person broadcast it in their voice.

The speaker further elaborated on the use of human-like avatars in synthetic media, which can be used to say things that a human would say. He cautioned that such tools can be used for nefarious purposes, such as spreading fake news.

The speaker also highlighted the trend of video-to-video avatars, such as lip syncing, which can be used to match voices and generate videos that appear almost completely authentic. He mentioned the Dubbing Project, 2019, as an example of this technology.

Finally, the speaker concluded that while synthetic media and deepfakes are not inherently bad, the fake news generated using these technologies is a serious problem. They warned that these technologies are becoming more accessible and easier to use, and that the audios and videos generated using synthetic media are increasing and improving rapidly. They urged the audience to take these developments seriously and to consider the broader implications of AI-generated content.

Session 2B: Text-to-Image workshop

The speaker introduced Text-to-Image tools Dall-e 2 and Dreamstudio, including its inpainting and outpainting features. Participants tested out different prompts to emulate images as a way to discuss some of the intricacies of prompt engineering. Participants then broke out into six groups with the following instructions:

- Groups 1 and 2: Choose any image and recreate it, trying to get as close to the original as possible
- Groups 3 and 4: Generate or edit an image with an intent to create mis/disinformation involving themes around elections
- Groups 5 and 6: Explore the creative potential for human-rights advocacy of text-to-image by either editing or generating an image

Key results and conclusions

Group 1 generated an image of the army containing demonstrators by altering a single original image using text prompts.

Original



AI generated version



Group 2 used an original image to emulate an image of a middle-aged African woman at a political campaign at a stadium. The text prompt recognized unique features such as the African headgear word Gele and the bandana.

Original



AI generated version



Group 3 generated an image of people attempting to burn ballot boxes using text prompts. The resulting image depicted a young boy with ballot boxes and fire.



Group 4 attempted to generate an image of four teenagers running away from four military men. However, the group was unable to get the desired outcome, highlighting that text-to-image is not perfect.

Group 5 tried to create an image to show people protesting against police brutality with journalists filming the event. However, the resulting image had no female protesters, revealing apparent biases in AI-generated media.

Group 6 created a campaign inspired by recent protests in South Africa around rising costs of living. They used the hashtags **#NoFoodNoElection** and **#BlowTheWhistle**, alluding to allegations of corruption as a cause of rising costs of living. They used the prompt *'A big group of african men and women blowing whistles at a protest demanding justice. Ndebele artist style. Digital art'* in Dreamstudio to create the following image that would serve as a banner calling to action:



Session 3: Synthetic media creation in Africa

The first part of the session focused on the African landscape and how synthetic media is increasingly being used by governments, businesses, and individuals to reach a wider audience.

The speaker discussed the rise of synthetic media in Africa and highlighted statistics indicating that as of December 2022, there are 570 million internet users in Africa, over double the number recorded in 2015. She also noted that the rate of mobile penetration in Africa was 46% in 2021, expected to grow to 50% by 2025, and that 46% of this penetration were smartphones. The number of unique mobile subscribers was 515 million in 2021, projected to grow to 613 million by 2025.

The speaker provided case examples of synthetic media in Africa, such as virtual influencers Kim Zulu and iYANDA, synthetic videos, synthetic images used for NFT art and stock photo generation, synthetic art created through computer graphics, synthetic audios including text-to-speech and voice cloning, and synthetic music that is indistinguishable from human-created music.

Finally, they highlighted the sectors where synthetic media is being used, including entertainment, advertising, education, politics, and healthcare. In the entertainment sector, Usiku Games, a Kenyan-based gaming company, uses artificial intelligence to create games

reflecting African culture and experiences. In advertising, Safaricom, a Kenyan telecommunications company, uses synthetic media to create social media ads that reflect the diversity of its customers. In education, Eneza Education, a Kenyan edtech company, uses synthetic media to create animated educational videos that teach students in a more engaging and interactive way.

During the second part of this session, a second presenter highlighted the use of Synthetic Media in fashion designs. The speaker has extensive experience in AI-based designs and has collaborated with numerous studios to design attire for characters in movies using synthetic media.

Thanks to synthetic media, the speaker was able to create a synthetic fashion show for seniors that went viral. They also started a cloth-line designed by Artificial Intelligence and runs a tech-show that is entirely built on AI.

During the Q&A session, questions were raised about the legal complications of licensing AI-generated content, data sets, and who has the right to AI-generated content. The University of Pretoria was mentioned, where lecturers urge students to acknowledge the use of Chat GPT in generating academic text. It was also noted that laws to regulate synthetic media are playing catch-up and regulators need to level the playing field.

Session 4: Perspectives on synthetic media

The afternoon session's agenda focused on group breakout sessions which were divided into four categories: gender, misinformation, journalism, and digital rights. The groups aimed to delve further into Synthetic Media for audio and visual media, focusing on their respective areas of interest and connecting them to other areas.

In the gender group, the focus was on exploring how synthetic media can be used to promote gender equality and representation. This could include creating synthetic media that portrays women and other marginalized groups in a positive and empowering light.

The misinformation group focused on exploring how synthetic media can be used to detect and combat fake news and misinformation. Synthetic media can be used to create realistic but fake videos and images, which can be used to deceive people. By developing techniques to detect and combat such fake content, synthetic media can help to promote a more informed and accurate public discourse.

The journalism group explored how synthetic media can be used to enhance and augment traditional journalism. For example, synthetic media can be used to create interactive visualizations and immersive experiences that help readers to better understand complex stories.

The digital rights group focused on exploring the potential risks and benefits of synthetic media for online privacy and security. Synthetic media can be used to create highly realistic deepfake videos that can be used to harm individuals or spread misinformation. The group aimed to develop strategies to mitigate these risks while also harnessing the potential benefits of synthetic media.

Breakout group 1 - Gender

The presenter in one of the breakout groups discussed the role of women in AI and the hindrances they face in her presentation. They began by acknowledging the work of African women in AI such as Linda Bonyo from the Lawyers Hub in Kenya and Favor Borokini from Pollicy in Uganda. However, the speaker noted that women are largely locked out of AI in Africa.

They emphasized the need for a universally accepted definition of AI that is simplified for better understanding, especially among women. The speaker pointed out that AI is not designed in Africa, and there is a need to harmonize institutions that drive AI on the continent. They further stated that AI policies should be developed in Africa and be geared towards policies and strategic plans that drive tech and AI.

They identified the hindrances of AI, which include low trust in AI, the lack of education tools, a smaller number of universities offering AI courses in Africa, data deficits, ethical concerns, and colonial issues impacting AI. They concluded by stating that improvement in the uptake of AI is crucial towards the synthetic media uptake by women in Africa.

During the ensuing discussion, the group highlighted that gender plays a critical role in both misinformation and disinformation, two factors that can be amplified by the use of synthetic media. The group noted that these factors contribute significantly to Gender-Based Violence. Furthermore, female journalists often face discrimination and sexual harassment based on their gender.

A participant stressed the importance of harmonizing the definition of AI to increase the understanding of Synthetic Media and its potential to curb Gender-Based Violence. They also

emphasized the importance of research on AI and gender, and how this applies to synthetic media.

The group also noted that the intersectionality between poverty and gender is apparent in the uptake of synthetic media in Africa. Women in Africa have fewer opportunities to learn about technology, and the tech industry is often considered a male affair. In some African communities, girls are not allowed to take tech courses in school and are instead expected to leave such opportunities to their brothers. Additionally, technology customization in Africa is often based on men, which negatively affects women's participation in AI and Synthetic Media.

Finally, the group emphasized the need for government support to encourage women's participation in Synthetic Media. They stressed the importance of aligning government activities with this goal, and highlighted the need for education on Synthetic Media and its effects on women. The group also pointed out the need to design platforms that can curb the vice of non-consensual sex imagery. Finally, the group emphasized the importance of having designers of AI frame tech tools from a neutral perspective that is inclusive of women.

Breakout group 2 - Misinformation

The presenter in one of the breakout groups highlighted the growing concern over deepfakes in Africa. They noted that while many Africans previously viewed deepfakes as a less significant problem than shallow fakes, the situation is changing drastically. Today, deepfakes have become a major challenge in Africa, prompting the emergence of organizations dedicated to monitoring synthetic media and fact-checking.

Unfortunately, these organizations often face a lack of understanding from media stakeholders and other key players. The speaker emphasized that fact-checking is not only crucial for politicians but also for consumer protection. In areas of conflict, synthetic media has been used to fill gaps in information. However, during the COVID-19 pandemic, misinformation about the efficacy of vaccines was spread through synthetic media.

The rise of synthetic media has also led to a growing distrust of mainstream media. In Kenya, for example, the media has been dubbed the "Githeri Media" for publishing irrelevant narrations. Funders are also implicated in media funding, pushing their narratives and eroding the freedom of the media to report objectively. As a result, synthetic media is being used to spread misinformation rather than synthesize news for the public.

The speaker stressed the importance of fact-checking in establishing the truth and communicating it to the public. They pointed out that there is an emerging imbalance in social media, where misinformation can spread faster than mainstream media. To address this, we need to identify knowledge gaps, not only by asking people if they know about synthetic media but also if they can provide real-life examples.

The speaker called for legislators to be knowledgeable about the challenges posed by synthetic media. It is essential to communicate the risks of synthetic media and where we currently stand in terms of combating misinformation.

During the conversation that ensued, the team focused on the issue of misinformation and highlighted that it often thrives when there is an opportunity to push an agenda. To combat this, technology plays a critical role in fact-checking, and search engines need to be configured to prioritize and facilitate fact-checking efforts.

The group also discussed the importance of open-source tools for reliable fact-checking. These tools are crucial in ensuring that the public has access to accurate information, and they can help prevent the spread of misinformation. However, the team also noted that reliable tools for fact-checking archival data are not readily available on open-source platforms. There is a need for more access to such tools for fact-checking, and they should be made available to the public for free.

The importance of fact-checking was highlighted with recent events in Mali and Burkina Faso, where misinformation was spreading rapidly. The team emphasized that reliable fact-checking tools could have been critical in preventing the spread of false information and ensuring that the public had access to accurate information.

The team also reiterated the importance of fact-checking in combating misinformation and stressed the need for accessible and reliable fact-checking tools for both real-time and archival data. They stressed the crucial role of open-source tools and highlighted the need for governments and other organizations to prioritize the development and distribution of such tools.

Breakout group 3 - Journalism

The presenter in one of the breakout groups focused on the impact of synthetic media on journalism. As a journalist involved in fact-checking, they discussed the dangerous use of synthetic media by politicians to spread lies during elections. They cited examples from the

2023 Nigeria elections, where celebrities were manipulated to share synthetically altered audio, and use fake social media accounts and fake images were also used to spread disinformation about politicians.

According to the speaker, there is a lot of manipulation in Nigeria to discredit political issues. However, she praised the efforts of the Fact-checking Coalition of Nigeria, which was highly active during the 2023 elections and provided insight on credible news. They also emphasized the need for media literacy campaigns, particularly targeting social media influencers who are useful in fact-checking. In order to combat synthetic media in Africa, the speaker suggested the need for tools to detect fake content, which can be achieved through media literacy campaigns.

The speaker reiterated the importance of educating people on how to detect fake news and how to fact-check. They stressed the critical role of big tech companies in helping to combat synthetic media by developing and implementing tools to detect and flag fake content. Overall, the presentation highlighted the importance of media literacy, fact-checking, and the need for collaboration between media professionals and big tech companies to tackle the growing problem of synthetic media.

The discussion that followed highlighted the challenges faced by journalists due to the proliferation of synthetic media. These challenges included the creation of fake virtual journalists, photoshopped images used for shaming, and the spread of misinformation through social media accounts. One major hurdle is the lack of skills among journalists to conduct fact-checking. For example, in Ghana, journalists were unable to verify the authenticity of an audio recording by a politician, which fell under the category of the "liar's dividend" where the politician alleged that the audio was fake, leaving no way to determine its authenticity.

Another challenge highlighted was the unavailability of necessary resources and skills for fact-checking in many newsrooms, partly due to its high cost and low profitability. Additionally, traditional training for journalists has not kept up with emerging trends such as synthetic media in most African countries. The absence of media house policies on posting news was also seen as a significant challenge.

To address these challenges, the group proposed several recommendations. Firstly, there is a need to revise the journalists' training curriculum to incorporate synthetic media and fact-checking skills. Secondly, media houses should develop policies that address the dissemination of news, taking into account synthetic media. Thirdly, it was proposed that resources be mobilized to equip newsrooms with adequate fact-checking tools. Finally, the

group called for collaboration between journalists and media houses to enhance fact-checking efforts.

Breakout group 4 - Digital rights

The discussion began by acknowledging that digital rights are pegged on human rights. They stated that for a person to claim digital rights the starting point is the digital spaces. Countries who lack policies and laws on digital rights miss out on protection of some of the human rights.

Further in their reporting they applauded that not all synthetic media is bad and even deepfakes majorly associated with negativity can be used in a positive way. The use of synthetic media raises questions on digital rights which are not only human rights on the internet but also off the internet. Enjoyment of digital rights is pegged on the enjoyment of other human rights such as the freedom of the media and the freedom of expression. Sharing of information happening offline online outrightly compliments digital rights and human rights. Solidarity protests happening offline such as the Black Lives Matter movement was a perfect example of how digital rights are intertwined with human rights and applicable both online and offline.

Discussing emerging issues, the group noted that it is not possible to sustain civic engagements if we do not keep-up with technologies such as synthetic media. Ethics were hailed as important in synthetic media and especially for the players in the sector. It is important to evaluate how the government regulates synthetic media in connection with ethics.

The group also discussed the role of synthetic media in relation to the right to be forgotten as a result of the digital footprints because the digital presence demands more. We may enjoy being online but our rights to be forgotten are inhibited by our digital footprints.

On censorship, they discussed the tactics by the government to curtail digital rights such as the right to protest through internet shutdowns. The governments do these internet shutdowns and partial social media shutdowns with the intent to control a narrative. The governments create delayed feeds and more, which affects digital rights and ultimately human rights.

Digital rights have a legal aspect in the relation to synthetic media e.g., issues of copyright. It was commented that digital rights are universal but the enforcement of these rights are not universal. There is tension between exchange of ideas online and abuse. The right to privacy was featured, it was discussed that the internet is a networked space and how your interactions should remain private is not controlled. The provision of free internet by governments should be

examined on how it affects the right to privacy and especially in terms of government surveillance.

Session 5: Video for human rights, artistic creation and advocacy

Summary

The first speaker introduced their work as part of the Makadara Social Justice & Information Centre (Masjic) that focuses on land, water and housing rights, and youth education. They highlighted that they do face repression and harassment. They develop bottom-up media campaigns as part of their strategy.

The second speaker then shared about their work with the [Refugee Law Project](#) in Uganda. They use media as an important tool for change (and speaking to authority and power). They are concerned with amplifying voices of forced migrants in a grounded way while also connecting with broader audiences on issues they wouldn't normally hear about. They lead archiving efforts and boot camps for youth to train them on video advocacy.

This was then followed by an open discussion on the relationship between synthetic media and video for human rights.

Key discussion points

During the open discussion, several questions were raised regarding the strategies and projects mentioned in the presentations. It was asked what can threaten these initiatives and how synthetic media can be shaped creatively to advocate for human rights. One of the presenters responded by saying that synthetic media can be an enabler of rights and activism, impacting capabilities. They shared an example of a viral video they used for training that almost led to their organization being closed by the government.

However, there were concerns raised about the potential negative impact of synthetic media. Participants discussed how the power of social media can prompt the government to clamp down on digital rights and also alter information on social media. Governments can curtail the freedom of expression and the media as a result of the growth of synthetic media. To prevent this, safeguards should be put in place to ensure synthetic media is not used negatively, such as to spread misinformation. Inclusivity of synthetic media, especially in the local context, should also be taken into consideration.




Another issue that was highlighted was the potential stereotyping and bias in synthetic media. Participants emphasized the importance of responsible use of synthetic media and development of synthetic media to prevent misrepresentation of Africa. It was noted that AI and synthetic media should not be trained to only see the negative aspects of Africa. Research was also discussed as being important in informing how synthetic media operates, as it relies on data. The negative findings in research can impact how Africa is portrayed in synthetic media.

The need to protect indigenous people from over-sharing information that they would not like to appear in synthetic media was also discussed. Participants cited the example of the Black Panther Movie using information that was not voluntarily shared by the community, which led to them being sued. It is crucial to protect indigenous information in relation to synthetic media. Finally, the issue of research and how to protect identities was raised as a significant consideration in the use of synthetic media in Africa.

Report of Day 2

Session 1: Threats Spectrogram and Analysis

The first activity of the day was a wall walk-through, where participants used sticky notes in three different colors to indicate the degree of priority for identified human rights/issues thematic areas. The orange marker represented the highest priority, yellow indicated a middle priority, and green was used for the low priority. Participants were asked to place the markers on the mapped areas to express their opinion on the areas they considered high priority threat, middle, or low priority threat. They were also encouraged to use sticky notes to indicate the reasons for their prioritization.

	Threat	Comments				T*
General Verification Challenges	Missing or mis-contextualized date/time, context location, intention	<ul style="list-style-type: none"> Escalates conflict Media illiteracy will increase the threat. Most people cannot detect falsification Could be deepfake intel but traceable through IPs. Attracts potential threats after misinforming the public Fact-checking efforts already struggle with this. This could make it more difficult. 	-	5	5	25
	Content edited or manipulated	<ul style="list-style-type: none"> Opens to hate speech; human rights abuse It can create chaos in volatile states like Kenya Comedy videos can be cropped and altered 	-	4	10	38

		maliciously <ul style="list-style-type: none"> The “Breaking news syndrome” which makes everyone want to go viral will influence this 				
	Content staged	<ul style="list-style-type: none"> Fuels conflict and panic. Could lead to chaos Creates space for information manipulation 	5	3	6	29
Human rights issues / Thematic areas	Gender-based abuse	<ul style="list-style-type: none"> In extreme cases, could lead to death of female investigative journalists Young female journalists may be dissuaded to operate leading to further gender divide in field Reputational damage Lives in the internet ‘forever’ Removes voices from table out of fear of retaliation - SGBV as threat to hush ‘Revenge porn’ Targeted disinformation against leading female voices 	-	4	10	38
	War, conflict and violent events	<ul style="list-style-type: none"> Messaging platforms like Telegram can be used in conflict to spread misinformation 	-	4	7	29
	Elections	<ul style="list-style-type: none"> Ethnic discrimination influences decision to manipulate news Coordinated inauthentic behavior Compromising election integrity & fueling ethnic divide 	-	-	13	39
	Public health	<ul style="list-style-type: none"> Creates public health crises 	3	4	8	35
	Land rights and other ESC rights	<ul style="list-style-type: none"> Discrediting activists is a common tactic. This will increase threat at scale Land right defenders can be influential and be subject to targeted attacks May not be much an issue in Africa 	5	3	4	23
	Government propaganda	<ul style="list-style-type: none"> Allows spread of hate speech Low accountability (liar’s dividend?) Jeopardizes institutions and critical dissent 	4	3	8	34
	Attacks on civil society and the media	<ul style="list-style-type: none"> Targets leading CSO voices and groups Reduces credibility and authenticity of human rights defenders Freedom of expression accounts locked down Data privacy and surveillance technology Asymmetrical relationships of power benefit authoritative govts. Govt stifling live feeds of protest in Kenya’s protests (during workshop) ‘Anti-trust’ environments 	2	2	11	39

Information landscape	Scale and accessibility		4	2	4	20
	Volume	<ul style="list-style-type: none"> • ‘Flood the market’ so verification becomes impossible and misinformation ecosystem becomes an ‘echo’ • Needs media literacy at a parallel scale • Content creators could be tracked through IPs (note on connect with provenance) 	-	7	3	23
	Liars’ dividend	<ul style="list-style-type: none"> • #GandujeGate in Nigeria • Authoritarian govts will use this against activists • Governments trivialize critical issues 				
	Hallucinations	<ul style="list-style-type: none"> • Risk of media outlets picking up hallucinations, eroding trust • Very influential, but moderate threat • Adds to info noise 	7	3	2	19
	Multimodality		2	-	4	14
	Personalization and live interaction with deepfakes	<ul style="list-style-type: none"> • Public awareness / disclosure of use • Once it is in the internet, it is difficult to scrub from the internet • Once AI is used for cons, threats will increase for general public 	3	4	5	26

*Where T = (G*1)+(Y*2)+(O*3)

Elections

The election section received the highest number of high-priority threat markers with 15 participants indicating it as a top priority. All participants who walked through the election section identified it as a high-level threat. Comments made on the board suggested that the use of synthetic media to spread fake news in elections was rampant and was classified as a high threat. This is because it has the potential to spark ethnic conflicts and grow into full-blown chaos.

Gender-based Violence

According to the wall walk-through activity, GBV was classified as a high-level threat by 10 participants, while four participants considered it to be a middle-level threat. None of the participants classified GBV as a low-level threat. The comments made on the board indicated that non-consensual sex imagery was being targeted at female influencers, and there were attempts to campaign against leading female voices. Revenge porn was also highlighted as a

reason for marking GBV as a high priority threat. However, there were no comments on why some participants felt that GBV was a middle-level priority threat.

Attack on Civil Society

The participants categorized the attack on civil society and the media as a high priority threat, with 11 participants expressing this view. Two contributors felt that it was a middle-level threat while two members felt that it was a low priority threat. Participants cited various reasons for this prioritization. One of the reasons was the importance of freedom of the media, with social media shutdowns being a particular concern. Another reason was the intentional attack on the media and the potential for harm to journalists. Targeting of leading voices was also mentioned as a threat to civil society and the media. For example the Kenyan government issuing threats to shutting down media stations for covering opposition-led demonstrations was used as an illustration of the high priority threat posed by attacks on civil society and the media.

Content edited or manipulated

A high priority threat in the area of content manipulation was noted by ten participants, while four participants classified it as a middle-level threat. The reasons given for this classification were diverse and included concerns such as the potential of altered content leading to chaos, the creation of fake videos for fame and the malicious labeling of videos for disinformation. Other concerns were the possibility of leading to hate speech, occasioning human rights abuse, and the likelihood of creating deepfakes that can deceive people.

Personalization and live interaction with deepfakes

Personalization and live interaction with deepfakes was considered a high-level priority threat by five people, with four people classifying it as a middle-level priority threat and three indicating it as a low-level priority threat. The ease of creating deepfakes using a mobile phone was cited as a reason for the high classification, along with the lack of awareness among the public about the existence and danger of deepfakes.

Government propaganda

The spread of government propaganda was classified as a high-level priority threat by nine people, with three going for it as a middle-level priority threat and four classifying it as a low-level threat. The reasons for the classification of government propaganda as a high-level threat included the spread of hate speech, disinformation, and profiling.

Public Health

Creation of public health risk was one of the major concerns in classifying the public health sector as a high priority threat. Eight individuals classified it as a high-level priority threat, four people classified it as a middle-level priority threat, and three people claimed it to be a low-level priority threat.

War Conflict & Violent Events

Social media accounts and messaging platforms such as Telegram were regarded as enablers of spreading misinformation in conflicts, leading to the classification of war, conflict, and violent events as a high-level priority threat by seven people. Four people classified it as a middle-level threat, with no one indicating it as a low-level priority threat.

Content staged

Out of the fourteen markers presented, six participants classified the section on the balance between content manipulation and content verification as a high-level priority threat. The reasons given for the high priority threat included the likelihood of fueling panic and leading to conflict, ease of manipulation of information, and creation of space for information manipulation. Three people classified it as a middle-level threat, while five people viewed it as a low-level priority threat without providing any reasons.

Land rights and other ESC rights

Four people classified land rights and other ESC rights as a high-level priority threat, with three a piece classifying it as a middle-level and low-level priority threat. Those who regarded the topic as a high priority threat emphasized the use of technology as an enabler of activism and the resulting need to address synthetic media's potential impact on activism at a large scale. Another comment in support of high-level priority threat considered land rights defenders as influential and thus attracting high-level threats.

Volume

Seven people classified the issue of volume as a high-level priority threat, while three classified it as a middle-level priority threat. The reasons given included the need for media literacy, echo for misinformation, the use of lps to trace developers of content, and the ability to bring conflicts between the government and citizens.

Liars' dividend

Liars' dividend, the claim that content out there is deepfaked whereas it is not, was classified as a high-priority threat by six people, with one person indicating it as a middle-level priority threat and seven people regarding it as a low-level priority threat. The concerns for the high priority threat were mainly the risk of government authoritarian misuse against activists, the possibility of governments using it to ignore serious issues, and the threat to civil organizations.

Commercial growth

Commercial growth was classified as a low-level priority threat by nine people, with three people considering it a high-priority threat and four indicating it as a middle-level priority threat. The majority of the participants did not view synthetic media as a direct threat to society but rather an enabler of commercial growth and therefore not a threat on the face of it.

Scale and Accessibility

There were no comments provided regarding this section. Four participants classified it as a low-priority threat and another four participants categorized it as a high-priority threat, while only two participants marked it as a middle-level priority threat.

Multimodality

This topic did not receive much attention from the participants, with only six markers recorded. The markers were distributed equally among high, middle, and low priority threat classifications.

Hallucinations

Most participants did not perceive the capability of Chat GPT as a high-priority threat that requires immediate attention. However, two participants viewed it as a high-priority threat, stating that it may cause a loss of trust in organizations that rely on it. Seven participants considered it as a low-level priority threat, and three marked it as a middle-level priority threat.

Discussions

After the wall-walk through activity, the floor was opened for discussions where participants expressed their views on the thematic areas that were highlighted, the discussions commented on the following:

The issue of land and access to justice was identified as an emotional topic, especially when political figures are involved, which creates a high-level priority threat when using synthetic

media to advocate for human rights due to potential government surveillance. Satirical content creators were also considered a high-level priority threat as they may become targets of politicians. However, advocacy through synthetic media can be done anonymously, which is a positive aspect.

The concept of the liar's dividend was viewed as a difficult and high-level priority threat because it is challenging to win against the government and takes a long time. The categorization of liar's dividend as a low-priority threat was surprising since changing the narrative is a crucial aspect that government individuals will always want to do. The availability of witnesses to corroborate stories alleged as untrue was seen as a significant hindrance to liar's dividend, and this contributed to its categorization as a low-priority threat. The fact that the internet never forgets is also a critical factor in why politicians may not lean towards the liar's dividend.

Gender-based violence (GBV) was regarded as a high-level priority threat due to the potential for women to be sexually abused or harassed, even in journalism, and fail to report, which is also a threat to their livelihoods. The lack of security backup for women journalists was also a significant indicator of GBV as a high-level priority threat. Paid bloggers were also noted as facilitators of GBV, and witness protection for journalists was recommended.

Regarding civil society, it was noted that governments do not know the best way to deal with issues of synthetic media thus attacking civil society. Disinformation can lead people to underestimate the lengths governments are willing to go and what they are capable of doing to stick to their story.

Generative AI was praised as an enabler to speed up creativity but also highlighted as having the potential for negative use. Commercial growth indicators were seen as crucial because following the money would enable the discussion of all the other threats posed by synthetic media. As the commentator on commercial growth said, "follow the money".

Session 2: Current solutions and responses to synthetic media

Some of the potential solutions and responses to synthetic media included:

- Media Literacy
- Tools for detecting deepfakes, including for example identity based detection
- Legislation
- Labeling

- Authenticating content with trust signals

There were concerns that AI image detectors can fail and that creating the tools may not be the bigger problem; accessibility of these tools is the major issue. It was also expressed that there is a necessity for local journalists and fact-checkers to have access to forensic experts' input.

The need for identifying truthful content with provenance-capturing technologies and facilitating the process of authentication. Some of the questions raised were: how will consumers react to detection tools? How can we prevent synthetic media from being misused?

Hardline legislation on synthetic media was criticized. Example of hardline legislation was cited to be the law on synthetic media in China. Inputting deepfake into repressive fake news law was discussed. Having state-level laws and other national and regional laws against negative synthetic media content such as non-consensual synthetic sexual images was encouraged and an example of such laws in the EU was referenced.

Soft law based on Human Rights was highlighted and regarded as a step in the right direction but not to replace the need for enforceable laws.

Session 3: Mapping solutions and responses

Breakout group 1 - Journalism

The key points discussed by the group included:

- The need for media houses to buy into the need for solutions;
- The need to build capacity, including through peer-to-peer training;
- Develop specialized trainings for different journalists, recognizing the diversity of journalism, contexts, and needs
- The need to create toolkits for journalists
- To leverage existing responses and solutions in a globally equitable way
- Journalists need to collaborate on building mechanisms for authenticating content
- Going back to basics: Following traditional ethical guidelines is already a big step forward (for example, by fact-checking sources and avoiding click baiting)
- Develop mechanisms to effectively flag content without it leading to arbitrary take downs

Breakout group 2 - Human rights defenders

Key discussion points

The session began by referring back to the threats identified and discussed in the previous session. Some of the threats that were highlighted included the mechanisms by which the government stifles freedom of expression, including trolling (specially against women), poisoning the well and generating confusion, and the drafting of legislation that oversteps as an excuse to tackle mis- and disinformation or out of ignorance of local contexts. They also noted that platforms often play a complicit role in censoring human rights defenders (also highlighting the need for local expertise to inform moderation). For all of these threats, synthetic media is a tool that can be used against them, either directly by, for example, poisoning the well or trolling, or indirectly, by using it as an excuse to regulate to the point of censorship.

Participants then talked about potential solutions to these threats and ways by which synthetic media could be leveraged to raise their own voice. A few participants highlighted the need for more media literacy for human rights defenders – not to place the burden of identifying synthetic content on them, but as a practical recognition that this could enable them to tackle threats from the misuse of synthetic media. Another participant suggested the need to revise media literacy campaigns to reflect the age of synthetic media; ie: deepfake detection expertise and tools need to be part of media literacy campaigns.

Another participant emphasized on a matter that has been echoed in previous spaces: authentication tools can empower human rights defenders, but they need to be accessible and functional. The participant underscored the use of authentication tools for archiving, noting that it can facilitate the process of archiving and validating information, and then making it more readily accessible and verifiable. For this to be effective, the participant pointed to the need for more technical collaborations between human rights defenders and organizations like WITNESS in order to build the capacities necessary to enable these tools.

Lastly, there was a comment about thinking more deeply about the creative potential of synthetic media. One participant noted that the use of avatars (or face swapping) could, for example, help protect the identity of otherwise at-risk individuals while still sharing their stories in a way that connects with more people.

Breakout group 3 - Policy makers or advocates working on influencing policy

The group began by discussing some of the concerns/obstacles they face currently:

- A lot of discussions on legality/policy are happening in the US or Europe, and those conversations then trickle into Africa/elsewhere. I.e: plans were made in the US, not in Africa.
- In general, geopolitical issues affecting Africa, for example Chinese principles are having bigger role (in part because of rejection of US), but these principles also ignore needs, circumstances of Africa
- Datasets of AI models mainly come from global north (and English)
- In many countries, a lot of policies are very much a copy-paste of policies from the US, EU, and more recently China. There are some changes of course, but not in essence, and not with enough engagement from diverse stakeholders. Eg. Data protection act of Tanzania
- Civil society attempts to engage with governments on digital rights and transformation, but they end up having little impact – for a variety of reasons, including lack of resources. More resourced international organizations can often have a bigger impact.

This then led to a discussion about the possible pathways forward:

- Networks such as PAI are very global north driven, but they can still create a space for global south participation. This body can then serve to have an impact in legislation. It allows them to punch a little higher. I.e: networks, bodies that can have more impact, also on platforms. I.e: Partnership building
- The group discussed how to transform soft law into legislation. They discussed the value of creating policy briefs to include as many people as possible in these processes. Not just to make it more reflective, but also to push it forward as there is more participation/interest.
- We are often playing catch up with tech. How can we proactively engage on issues early on? For eg.: To have legislators reach out to civil society before having a draft law. Or how to be 'in the room' where the conversations are happening? One way to proactively engage is by doing research and putting it out there. This is useful for different stakeholders and one way for players to come to you. I.e: civil society fills in a gap of information necessary.
- How can we foster public participation in legislative processes? Law literacy is low so it is hard to get people to participate. And there are other more priority problems, that this does not garner enough attention. So literacy is important.

- Sometimes there are good laws, and good theoretical processes for getting public participation. But many times the voices of people are not actually heard. So people feel like there is no point in participating because they don't think that their voices will actually be heard, so there is a need for people to know how to participate and what impact their voice is having → Training of trainers to amplify impact.

To summarize:

- How to engage with platforms: Create foster alliances, partnerships (eg. PAI) with more global south to punch higher, with more impact.
- How to engage with legislation/gov: Policy briefs and fill in gaps in research to foster effective participation
- How to engage with the public: Media literacy and civic education and media literacy programs, promoted by CSOs as well.

Breakout group 4 - Creators and technologists

Some of the main points of discussion included:

- Media literacy: improve people's awareness when dealing with synthetic media → detect and debunk + keeping up with innovations in technology
- Should the tools being used be named or kept secret?
- How should we track the intention behind the creation or manipulation with synthetic media... The question above can be answered by this in part. Eg: if it is for marketing, it should always be mentioned.
- Platforms: Platforms do have a responsibility - platform discloses use of AI tools
- We need ethical parameters to help frame conversations
- Legislation is necessary

Session 4: Leveraging synthetic media for advocacy and creativity

The session started with WITNESS sharing a review of what [we have been hearing](#) over the past 5 years. Participants commented on these findings, and highlighted the need for local contextualization: afro-centered research (grounded research).

The speaker the presented the following questions to guide the conversation:

How to deal with the ease of creating volume of synthetic media? What technology, capacity demands do we have?

- Verification - access controls - can we reduce the volume? Feeding information back for AI training (eg. if uploading ID in the situation of access controls)
- Disconnect between builders and users (platforms building tools and their connection to academics; break up niches between disciplines)
- If price of tools is high this can bring barriers for HRDs
- Privacy concerns when implementing these controls
- Back end moderation ++
- Dissemination of information to the public has changed - use more media hubs
- What's a deterrent (to have less volume in the ecosystem)?
- Labeling / informing people about how a piece of information has been generated or original source - it could be a legal requirement for platforms
 - It works in forms → forms are used for misinformation (eg. a form for a fake job advert) So platforms need to have a responsibility of disclosing when an AI tool has been used. ie: individuals, users cannot be solely responsible.
- Redistribution of responsibility across the pipeline (creation to distribution)

What more should be demanded of tool makers and platforms?

- Afro-centered research
- Ruggie principles - when tool makers and platforms roll out new products we should demand that they meet existing standards like BHR and other intl HR standards

Importance of labeling and disclosure

- Eg: of covid misinformation. An 'i' on top right
- Different or discrete options embedded in platform design instead of asking for creators to do this voluntarily (e.g. icon platform user can click on to disclose more information; labels of how media has been created)
- Creators embedding it into the creation process
- Connection to media literacy - if option is not very visible or obvious, this may affect how some communities understand and access the information

Newsroom and archives resiliency

- Example of Code4Africa on support for newsrooms research - 21 countries research reveals that they don't have policies on synth media (neither much visible policies generally)
 - Template for newsrooms, without being descriptive, how to adapt and have an idea of what a good policy may look like in terms of synth media
 - It can help reinstall trust in the media
- Software that can generate and/or organize the metadata → helpful for communities, so important to make it accessible
- Closed vs open source software - suggestion to be open source + not heavy (to be used in older operating systems, phones, hardware)
- Detection tools - interest from the newsrooms but they may not be able to pay for it.
- More engagement from platforms, marrying commercial interests with research - give access for open research that is not necessarily for commercial purposes (APIs)

Government misinformation and media co-option

- Create a policy or mechanism in which the government is also accountable for the disinfo they share, the same they are accountable for mismanagement of public funds, for instance.
- Fact checking is a separate entity from newsrooms and they should be part of it - get platforms to sponsor it, or governments so it can potentially also be moved in house
 - Limitations of external entities for providing these services at scale
 - However... partnerships between media publishers and fact checkers
- Students' engagement so we are not always playing catch up
- Community / network to debunk gov propaganda
- Name and shame gov - requires integrity and ethics. Use of avatars (clone yourself) check everything (anonymous fact checkers)

Accessibility of protection tools for HRDs

- Making materials available in different formats to foster inclusivity + inclusion from/by staff
- Grassroots activists attached to organizing circles

Legislative responses - needs in current geopolitical context

- Lack of research -> clash of authorities, repetition of efforts, overlap of existing bodies doing regulation -> allows gov structures / regulators to gaslight or leave out marginalized or vulnerable communities who may not have the resources to go around. More research can help counteract this.
- Creation of databases and public participation - clarity on ownership and ethical hacking to understand data being captured
- Copy and paste of legislation without copying the institutions that are needed to uphold the legislation
- Regulation tends to be for punishing instead of to guide

Further responses to online sexual images

- Cyber security training for women protecting data
- Porn industry role
- Laws to takedown content of non consensual images - also clarity on their application

Preparation for specific scenarios - eg elections

- Set up groups to track misinfo and electoral results - coordination and capacity that includes public participation

Session 5: Advocacy recommendations

Group 1 group discussed the creation of a futuristic human rights advocacy using synthetic media tools and capabilities to protect the identities of activists at the frontline and minimize risks, they proposed creating an avatar influencer that would embody the characteristics of an everyday individual in an African community.

Individuals could have an immersive experience with the avatar, which would evoke emotions, consciousness, and a willingness to participate in the campaign. To ensure continued engagement beyond the virtual space, the group proposed creating a virtual room, a social media experience, and an interactive filter.

The avatar created would also be printed on products, which individuals could scan to access more information about the campaign or community. The group emphasized the need to do no harm and to provide adequate signaling before individuals enter the virtual space, allowing them to choose whether or not to participate.

The group also discussed the importance of education and how the level of education might affect police recruitment to ensure that policing officers are able to understand the use of synthetic media and what human rights entails. Finally, they talked about the potential for wearables, such as contact lenses, to be the means through which people experience digital content in the future and how to adapt content for such formats. The group stressed the importance of double-checking information as there is so much information available and it can easily be manipulated.

Group 2 recognized the challenges associated with verifying information and identifying the truth in the age of synthetic media. They stressed the need for a team of activists to work together to ensure the integrity of the election process.

The group discussed developing a solution for activism related to election integrity in the context of synthetic media. The conversation revolved around the need for verifying information and checking the authenticity of media shared on various platforms. The group emphasized the importance of forming strong bonds and working together to add information and fact-check claims. They suggested the use of a GPS system to identify media sources and share accurate information to counter misinformation.

The group also discussed the need for a tool to pinpoint the location of media and prevent the spread of false information. They emphasized the importance of coordinating with innovators to ensure that such a tool is in place.