MOXII AFRICA · MEDIA FREEDOM FESTIVAL 2026

What Happens When Fact-Checking Isn’t Enough and Trust Has Broken Down?

Spotting, reporting and responding to mis- and disinformation

MOXII AFRICA Presenter: Ntsako Manganyi · 14–16 October 2026

01

About Moxii Africa

Striving for open and trusted information

MOXII AFRICA

ABOUT

Moxii Africa (formerly Media Monitoring Africa)

Our mission — striving for an accessible, accountable and transparent information ecosystem we can trust.

Instill information integrity

Advocate for a healthy information ecosystem.

Promote safety

Safe, respectful, inclusive digital spaces where people can express themselves freely within the bounds of local and international law.

Open information

Universal, meaningful access to diverse views and information.

We do this through

Ensuring accountability · Empowering the public

MOXII AFRICA

02

What happens when fact-checking isn’t enough and trust has broken down?

When people no longer trust institutions, media, experts—or one another—correcting a false claim may not be enough to change minds.

WHY THIS CONVERSATION MATTERS

From correcting claims to rebuilding trust

Why trust breaks down — Disinformation and online harms rarely operate in isolation. Emotionally charged narratives, targeted harassment, hate speech and xenophobia can deepen suspicion, intimidate people and weaken social cohesion.

Look beyond the claim — Explore the emotions, identities and experiences that shape whether people accept or reject information.

Respond without amplifying harm — Challenge harmful narratives without giving them greater reach.

Build trust through action — Use credible messengers, community engagement, digital safety and clear escalation pathways.

Opening question — What makes you trust—or distrust—a correction, even when the evidence is clear?

A correct fact only works when the audience still trusts the source.

Opening question What makes you trust — or distrust — a correction, even when the evidence is clear?

MOXII AFRICA

“Misinformation, disinformation and hate speech now spread at unprecedented volume, velocity and virality.”

UN Global Principles for Information Integrity

MOXII AFRICA

Misinformation and disinformation

MISINFORMATION

False or misleading information shared without necessarily meaning to incite or cause harm.

Example: a satirical post taken literally, or a clickbait headline that does not match the story.

DISINFORMATION

False, inaccurate or misleading information deliberately spread to obscure the truth, covertly influence opinion or cause harm.

Example: a doctored video or fabricated post meant to deceive voters.

We avoid the term “fake news”. It is easily politicised and used as a weapon against the news industry.

Source: Spotters Network User Guide

THE TRUST PROBLEM

Disinformation and incitement often go together

Incitement online

Encouraging others, through digital and social media platforms, to commit a crime — in this case violent actions that may cause harm, damage or even death.

Why they travel together

Disinformation causes an emotional response, which makes incitement to violence more effective.

Intention versus potential

Counter-disinformation work often focuses on intention. Where it meets incitement, we must also weigh the potential for public harm.

The wider landscape of online harms

Astroturfing
Concern trolling
Cross platform harassment
Cyberbullying
Cyber-mob attacks
Cyberstalking
Deepfakes
Denial of access

Denial of service (DoS attacks)
Dog whistling
Doxing
Hacking
Hashtag poisoning
Hateful speech
Nonconsensual sharing of intimate images

Online impersonation
Online sexual harassment
Phishing
Swatting
Threats
Zoombombing

Source: PEN America

03

What 2026 showed

Real411 and MARS findings from the Local Government Elections

MOXII AFRICA

WHAT 2026 SHOWED

Real411 complaints in the election period

710

complaints received since election monitoring began

65%

recorded as election-related

81%

of those were false information (187 of 232)

Provisional figures. Election-related and category counts as at 29 Sep 2026. Source: Moxii Africa

MOXII AFRICA

WHAT 2026 SHOWED

Delegitimising the electoral commission

The most consistent and highest-volume pattern: four of the 16 cases reviewed in detail. Recurring narratives:

- Claims the IEC colludes with specific parties
- Unverified rigging and corruption allegations, including retroactive claims about 2024
- Fake IEC recruitment posts
- Legitimate scrutiny blended with speculation presented as fact

REAL411 VERDICT: DISINFORMATION

A post falsely alleged, without evidence, that the IEC rigs elections in favour of a party, presenting fraud as fact.

Why it matters: repeated claims like this can weaken public confidence in the IEC and in future outcomes.

Sources: Moxii Africa preliminary findings; Real411 monitoring brief

WHAT 2026 SHOWED

Attacks on journalists (MARS)

+35.9%

recorded attacks, from 1,205 to 1,637 (432 new incidents)

+289%

mis- and disinformation records, from 28 to 109

Common tactics: animal comparisons, anti-foreigner slurs and occult framing. They discredit the person instead of answering the reporting.

Provisional, before 4 Aug to 29 Sep 2026. Source: Moxii Africa preliminary findings

MOXII AFRICA

WHAT 2026 SHOWED

Four ways fact-checks fall short

Distrust is seeded early

Retroactive claims that 2024 was rigged were used to pre-seed distrust before 2026.

Speed outruns checking

A cross-border story spread through the DRC, Ghana and Nigeria faster than fact-checkers could respond.

Synthetic media looks real

An unlabelled AI poster showed a party leader endorsing a rival, and a cloned presidential voice circulated.

Emotion beats evidence

Algorithms amplify charged content, and fear and anger travel faster than corrections.

Sources: Real411 monitoring brief; Moxii Africa preliminary findings; SM4P

MOXII AFRICA

Example 1

Possible harm to discuss

- Hate speech: targets people because of their migration status
- Incitement: urges the community to confront and pressure them

DISCUSS

Under Section 16(2), where does angry expression end and incitement begin?

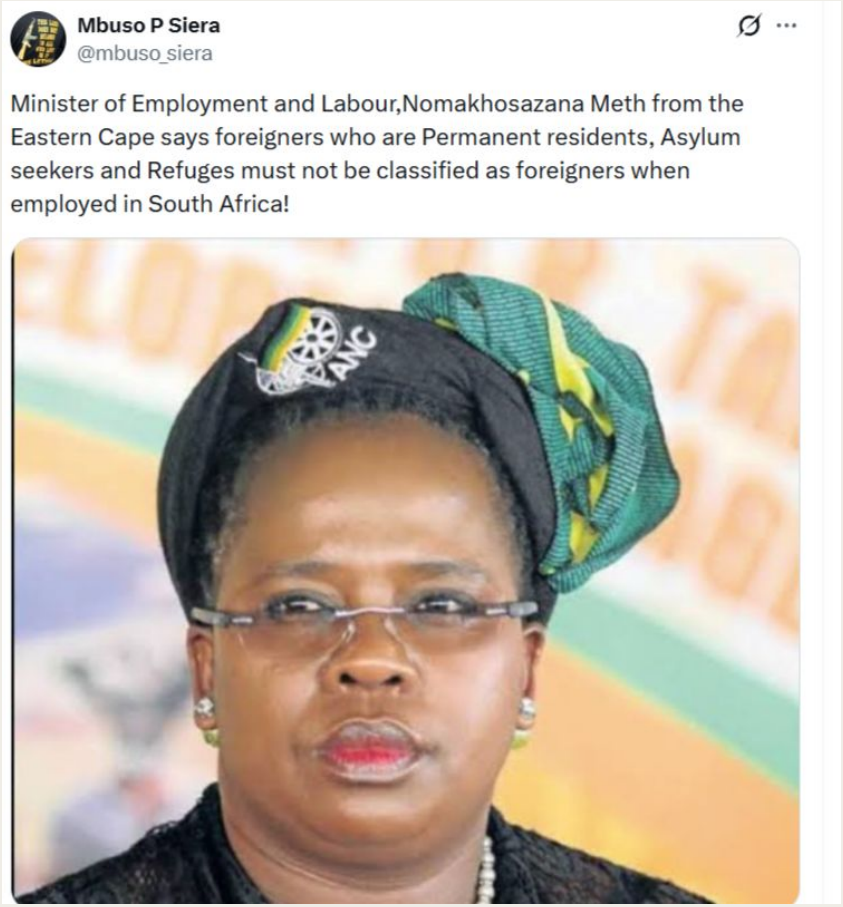


Examples reported to Real411. Possible harms are discussion prompts, not verdicts.

Example 2

Possible harm to discuss

- A claim attributed to a public figure that may be unverified
- Policy debate is legitimate; accuracy is the question



DISCUSS

What would you check first, and where would you check it?

Examples reported to Real411. Possible harms are discussion prompts, not verdicts.

Example 3

Possible harm to discuss

- An unsourced claim: “reportedly”, with no named source
- Framing foreign nationals as a threat to cultural identity

DISCUSS

The account has a blue tick. Does that make the claim more credible?



Examples reported to Real411. Possible harms are discussion prompts, not verdicts.

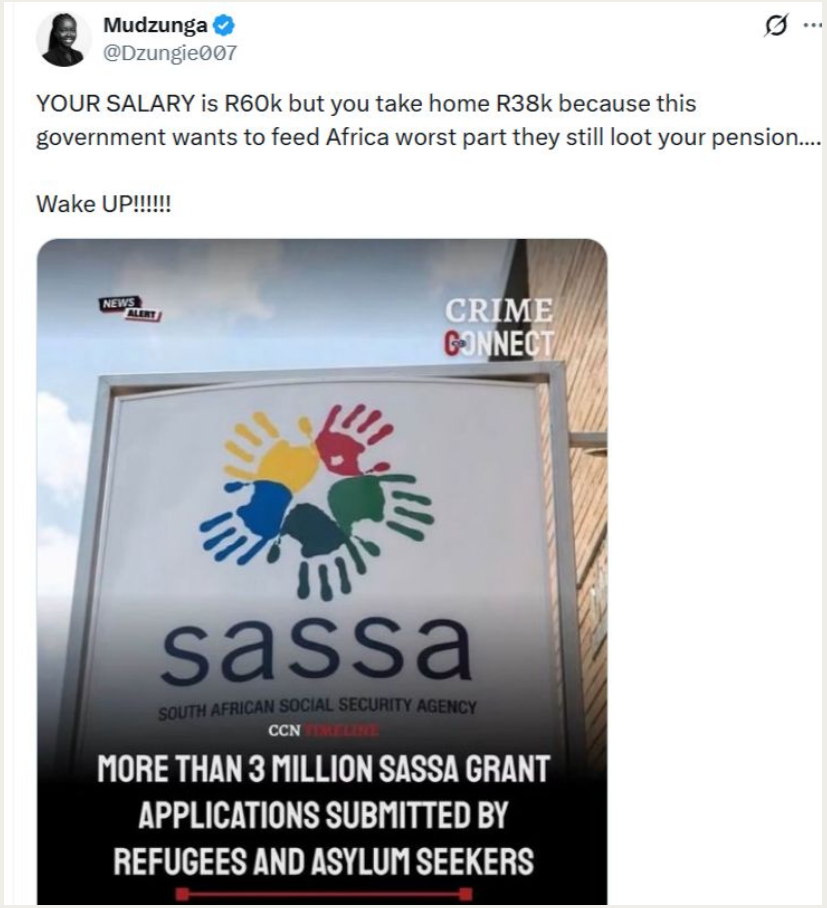
Example 4

Possible harm to discuss

- Misleading numbers: a pay claim paired with an unrelated grant headline
- Blaming refugees and asylum seekers for economic hardship

DISCUSS

Who produced the 3 million figure, what does it count, and does it support the claim?



Examples reported to Real411. Possible harms are discussion prompts, not verdicts.

04

Beyond fact-checking

A layered response: anticipate, report, assess, protect, coordinate

MOXII AFRICA

BEYOND FACT-CHECKING

A layered response

Anticipate

Track narratives early and flag undisclosed AI content.

Report

Give the public a channel to flag harmful content.

Assess

Review independently against legal principles and context.

Protect

Document attacks on journalists and support those targeted.

Coordinate

Share intelligence across media, civil society, platforms and the IEC.

Build literacy

Equip people with the skills to identify and counter misinformation.

Sources: Moxii Africa overview; Real411 monitoring brief

MOXII AFRICA

05



MOXII AFRICA

Real411: how it works

What it is

A secure, independent platform where the public and parties report online harms: disinformation, hate speech, incitement to violence and harassment of journalists.

Who runs it

Moxii Africa, with the IEC: the official complaints portal, offering a transparent, human-rights-compliant review process.

How it is reviewed

Each complaint goes to 3 trained reviewers. The secretariat is a qualified lawyer, and more legal firms and pro bono experts are being brought on.

How fast and how open

Submitting takes about 2 minutes. Unlike on the platforms, complaints are there for all to see.

THE SPOTTERS NETWORK

What happens after you submit



MOXII AFRICA

How a Real411 complaint is judged, and why it matters

What reviewers weigh

Freedom of expression, dignity, equality, political rights, context and likely harm.

Offensive, controversial or unpopular is not the same as harmful. Legitimate criticism and political debate are protected.

Why it matters

Proactive protection: an official channel to flag false information.

Speedy: urgent complaints dealt with within hours, to mitigate virality.

Accountability: bad actors held to the same ethical standards as traditional media.

Public trust: shows voters there is an active watchdog for the digital space.

Sources: Moxii Africa preliminary findings; Moxii Africa overview

BEYOND FACT-CHECKING · COORDINATION

Partnership on Information Integrity in Elections

A coalition that coordinates the work on election mis- and disinformation, with a single line of communication to the IEC through focal points.

FOUNDING STEERING COMMITTEE

Moxii Africa

Born from the IEC and MMA election review workshop in August 2024, and piloted in the Local Government Elections.

South African National Editors' Forum

Members include civil society, academia, experts, industry bodies, platforms, news media and Chapter 9 bodies.

Africa Check

Source: Moxii Africa overview

MOXII AFRICA

06

PADRE
Political Ad Repository

Verify credible political advertising



MOXII AFRICA

OVERVIEW

PADRE: How does it work?

- **What is it?** A public, transparent database of all digital political advertisements run during the election period across Meta and Google platforms. Tracks who paid for the ad, the amount spent, and the demographics/audience targeted.

- **Why it matters:**

  Levelling the Playing Field: Prevents "dark ads" (untraceable propaganda) from influencing voters without oversight.
  Ad Compare: Checker to quickly check if advert is official or not.
  Transparency: Ensures all participants adhere to campaign finance and disclosure rules.
  Compliance: Simplifies the process of auditing campaign activity for the IEC and relevant regulatory bodies.

07

The Spotters Network

How to spot, report and defy digital offences

MOXII AFRICA

THE SPOTTERS NETWORK

Introducing the Digital Spotters Network

An online network of people who monitor, spot and report digital offences on social media and the wider web. Over 200 people signed up since inception.

Spot

Report

Defy

Source: Digital Offences Spotters Network User Guide

MOXII AFRICA

What spotters look for

Disinformation

False or misleading content deliberately spread to obscure the truth or cause harm.

Incitement

Encouraging others online to commit violence that may cause harm or death.

Hate speech

Expression that advocates harm, discrimination or hostility against a group.

Journalist harassment

Persistent or serious conduct that demeans, humiliates or intimidates a journalist.

Source: Digital Offences Spotters Network User Guide

Six red flags

1 · Emotional framing

Outrage, fear or urgency, often with a deadline.

2 · No corroboration

No credible news source backs the claim.

3 · Out-of-context media

Old or unrelated photos and video reused for a new event.

4 · Synthetic artifacts

Deepfake glitches, or AI content with no label.

5 · Coordinated pushing

Bot-like accounts amplifying identical posts.

6 · Imposter sources

Real logos or names used for fake posts, such as fake IEC recruitment offers.

Sources: Truth Lab workshop; Real411 verdict criteria; monitoring brief

KEY TAKEAWAYS

What to take with you

Protect the messengers. Facts need trusted institutions and journalists.

Anticipate. Track narratives early instead of waiting for them to peak.

Respond without amplifying. Protect legitimate criticism.

Spot and report. Join the Spotters Network and use Real411.

THE SPOTTERS NETWORK

Online listening and monitoring tools

Graph.tips

graph.tips/beta

Trendsmap

trendsmap.com

Brandwatch

brandwatch.com/platform

Data Miner

dataminer.io

Tools surface what is trending. Human judgement decides what is harmful, and Real411 is where it gets reported.

Source: Digital Offences Spotters Network User Guide

MOXII AFRICA · STRIVING FOR OPEN AND TRUSTED INFORMATION

Thank you

Join the Spotters Network. Report harmful content on Real411.
Real411 app: Google Play and App Store
www.moxiiafrica.org · @moxiiafrica
Ntsako Manganyi · ntsakom@mma.org.za

MOXII AFRICA