

MODULE 3

DISINFORMATION DURING ELECTIONS

2026

MUNICIPAL ELECTIONS



Published by the South African National Editors' Forum

Prepared with the assistance of ALT Advisory

Funded by the Open Society Foundation for South Africa

June 2026

This work is licensed under a Creative Commons Attribution 4.0 International License. This means that you are free to share and adapt this work so long as you give appropriate credit, provide a link to the license, and indicate if changes were made. Full licence terms can be found at <https://creativecommons.org/licenses/by/4.0/legalcode>.



This module is part of a comprehensive set of online election resources which will be updated from time to time in the run-up to the 2026 municipal elections. For further information on the 2026 elections, please visit <https://elections.sanef.org.za>

Contents

1. Introduction to disinformation.....	4
2. What is disinformation?	4
3. Defining the terminology	7
4. Regional responses to disinformation	9
5. The challenge of regulating disinformation	11
6. Guidelines for dealing with election disinformation.....	13
7. Standard of care by the media	15
8. Guidelines for countering disinformation.....	17
9. Suggested resources.....	21

MODULE 3 | DISINFORMATION DURING ELECTIONS

Overview of this module:

- An introduction to disinformation and related concepts.
- Shortcomings of the term 'fake news'.
- The different regional and regulatory responses to address mis- and disinformation, and the measures being implemented in South Africa to address online mis- and disinformation during the elections.
- The standard of care required by the media and the importance of media credibility.
- Guidelines for countering disinformation

1. Introduction to disinformation

Disinformation is not a new phenomenon. The intentional sharing of false information to cause public harm, or for monetary gain, is not a modern concept. However, new technologies have expanded the available tools and methods for creating and sharing disinformation. Digital and social media platforms have enabled the perpetrators of disinformation to reach massive audiences almost instantaneously, and this challenge has often overwhelmed those trying to find solutions to combat disinformation.

Rapid developments in generative AI technology have introduced new layers of complexity to the disinformation landscape. For example, deepfakes, which are highly realistic but manipulated images, videos, or audio, can convincingly misrepresent people or events. As generative AI tools continue to evolve with increasingly hyper-realistic capabilities, news organisations in particular need to consider the novel ways these technologies can be harnessed to spread disinformation – including during election cycles.

2. What is disinformation?

DEFINING 'DISINFORMATION'

"[A]ll forms of false, inaccurate, or misleading information designed, presented and promoted to intentionally cause public harm or for profit. It does not cover issues arising from the creation and dissemination online of illegal content (notably defamation, hate speech, incitement to violence), which are subject to regulatory remedies under ... national laws. Nor does it cover other forms of deliberate but not misleading distortions of facts such as satire and parody."

Source: European Commission, 'A multi-dimensional approach to disinformation: Report of the independent High-Level Group on fake news and online disinformation', (2018) at 10 (accessible [here](#)).

The difference between malinformation, misinformation, and disinformation will be unpacked later in the module.

Disinformation during an election period is particularly dangerous. For elections to be free, fair, and credible, it is important that the electorate has access to accurate, credible, and reliable information. Disinformation can have the effect of creating confusion, doubt and mistrust in credible institutions (such as news media or independent electoral bodies) which can affect how and where the public chooses to access information. If those sources of information are not credible, it then becomes easier for false information to be disseminated. This directly impacts the nature of a free and fair election.¹

Disinformation, misinformation, and other forms of false news have increasingly been a factor in election cycles around the world. Following Brazil's 2022 presidential elections, which saw sustained disinformation about the credibility of the results circulating on messaging apps such as WhatsApp and Telegram, supporters of Jair Bolsonaro's unsuccessful re-election bid stormed government buildings to try overturn the results.² In 2023, Nigeria's elections were marred by reports showing that political parties had paid social media influencers to spread falsehoods about their opponents.³ During Kenya's 2022 elections, a manipulated image went viral which depicted one politician with a blood-stained shirt holding a knife to his neck, falsely branding him a murderer. Other false claims circulated widely, including premature declarations of election victories.⁴ In South Africa's 2024 elections, researchers found evidence of a growth in online campaigns to undermine faith in both electoral institutions and the voting process itself.⁵

Although not necessarily a new phenomenon, social media platforms have amplified the ease and reach of such disinformation, which has made it imperative for measures to be put in place to respond to such challenges in a timely and effective manner. False news and disinformation also adapt quickly to new technologies and new methodologies, making efforts to address it even more difficult. African countries are also sometimes used as 'testing grounds' for new disinformation techniques; for example, the company Cambridge Analytica, which became notorious for its role in the US 2016 elections, used African countries to test its model of extensive data collection through Facebook to target voters based on their perceived beliefs and to promote political polarisation.⁶

More recently, research conducted by a number of academic institutions and civil society organisations has looked at the impact of disinformation during an election period and the measures that can be put in place to counter it.⁷ One of the greatest challenges is managing the balance between upholding and promoting freedom of expression whilst combating and regulating the very real threat of disinformation.



¹ See UNESCO, 'Elections in Digital Times: A Guide for Electoral Practitioners', (2022). (Accessible [here](#)).

² See Rossini and others, 'Explaining beliefs in electoral misinformation in the 2022 Brazilian election: The role of ideology, political trust, social media, and messaging apps' *Harvard Kennedy School (HKS) Misinformation Review* (2023). (Accessible [here](#).)

³ Nkhowani, 'Tech Harms During Elections in Africa' (12 June 2024). (Accessible [here](#)).

⁴ *Ibid.*

⁵ Allen and le Roux, *Under the influence? Online mis/disinformation in South Africa's May 2024 election*, Institute for Security Studies (December 2025). (Accessible [here](#).)

⁶ Ekdale, 'African Elections as a Testing Ground: Comparing Coverage of Cambridge Analytica in Nigerian and Kenyan Newspapers', *African Journalism Studies* (28 November 2019). (Accessible [here](#)).

⁷ See UNESCO, 'Elections in Digital Times: A Guide for Electoral Practitioners', (2022). (Accessible [here](#)).



CASE STUDY: NAMIBIA ELECTIONS 2024

Namibia's 2024 parliamentary and presidential elections became a focal point for social media-led disinformation campaigns. For example:

- A forged letter was circulated widely, falsely claiming that a senior Electoral Commission of Namibia (ECN) official was colluding with the ruling SWAPO Party to secure victory in exchange for N\$2 million. The ECN swiftly dismissed the letter as malicious and baseless, stressing that it was a deliberate act of impersonation.
- Anonymous notes branded as 'Deep Intelligence' spread online, portraying SWAPO's presidential candidate Netumbo Nandi-Ndaitwah as a former apartheid agent and alleging that she and the party were selling the port of Walvis Bay to China for kickbacks.
- Reports of presidential candidate Nandi-Ndaitwah collapsing at a rally were amplified by an AI-generated image depicting the supposed incident, which, in turn, fuelled speculation about her health.

Source: Namibia Fact Check, 'Election fake news surge in October 2024', (2024) ([accessible here](#)).

The challenge of disinformation is truly a global one, extending beyond the political sphere to all aspects of information, including climate change, entertainment, and so on.⁸ Disinformation thrives during a crisis, as was the case with the [COVID-19 pandemic](#). This can be explained by the fact that disinformation relies heavily on emotions. When one comes across content that evokes particular emotions, be it anger, disgust, sadness etc, the reader is more inclined to engage with the content and to share it. During a crisis, when emotions and anxieties are already higher than normal, disinformation simply further heightens this, creating the perfect storm for disinformation to be spread.

However, the consequences have arguably been seen most starkly, and most concerningly, in its impact on elections, with there being a real risk that widespread campaigns driven by intentionally false information can affect the outcome of an election. As has been noted, "disinformation hurts democracy by undermining our faith in our institutions, weakening voter competence, and splintering the electorate."⁹

A recent UNESCO publication also notes that:

"Democracy requires free, periodic, transparent, and inclusive elections. Freedom of expression, freedom of the press, and the right to political participation are also critical to societies ruled by the respect of human rights. In today's rapidly evolving digital environment, opportunities for communication between citizens, politicians and political parties are unprecedented — with information related to elections flowing faster and easier than ever,

⁸ UNESCO, 'Journalism, 'fake news' and disinformation': Handbook for journalism education and training', (2018) at 20. ([Accessible here](#)).

⁹ Wood, Ravel, and Dykhne, 'Fool me once: Regulating 'fake news' and other online advertising' 91 *Southern California Law Review* Vol. 1 No. 6 (2018). ([Accessible here](#)) at 3.

coupled with expanded opportunities for its verification and correction by a growing number of stakeholders. However, with billions of human beings connected, and disinformation and misinformation circulating unhinged around the networks, democratic processes and access to reliable information are at risk.”¹⁰

The global picture is evolving daily, with a number of initiatives from various role-players being developed around the world, including from governments to social media platforms. What has emerged is that the purveyors of disinformation prey on the vulnerability or partisan potential of recipients to amplify the message through likes, sharing of posts, and retweets.¹¹ In addition to the impact that this can have on the outcome of an election, it can also have an impact on peace, security, and stability in a country during the election period and beyond. For instance, disinformation can breed distrust and a lack of confidence, or even ignite into violence when combined with existing tensions based on, for instance, race, ethnicity, or political affiliation; such tensions and conflict can, in turn, result in loss of life, internal displacement and despair, which have often left their mark in the aftermath of contested elections in other countries in the region.¹²

3. Defining the terminology

It is important to correctly define the different terms and the scope of what each term seeks to address. While the term ‘fake news’ has been popularised by politicians and particular media, there is an increasing trend to avoid using it as a synonym for disinformation. This is because current debates about so-called ‘fake news’ include a spectrum of different information types, ranging from relatively low-risk forms, such as honest mistakes made by reporters, partisan political discourse, and the use of clickbait headlines, to high-risk forms such as foreign states or domestic groups that try to undermine political processes through the use of various forms of malicious fabrications.¹³

As the European Commission argues,¹⁴ the term ‘fake news’ is both inadequate and misleading for the following reasons:

- The term falls short of the complexity of disinformation, which includes content that is not completely ‘fake’ but a mix of fabrications and facts, and practices that go well beyond anything resembling ‘news’.
- It has been appropriated by some politicians and their supporters to dismiss coverage that they find disagreeable.
- Citizens often associate the term ‘fake news’ with partisan political debate and poor journalism broadly, rather than more pernicious and precisely defined forms of disinformation.

The other argument against the term concerns news media credibility. Simply, if content is news it cannot be fake, and if it is fake, it cannot be called news. According to UNESCO: “‘fake news’ is an oxymoron which lends itself to undermining the credibility of information which does indeed meet the threshold of verifiability and public interest – i.e. real news.”¹⁵

¹⁰ See UNESCO, ‘Elections in Digital Times: A Guide for Electoral Practitioners’, (2022). (Accessible [here](#)).

¹¹ UNESCO, above n 7 at 7-8.

¹² UNESCO, above n 7 at 7-8.

¹³ European Commission, ‘A multi-dimensional approach to disinformation: Report of the independent High Level Group on fake news and online disinformation’, (12 March 2018). (Accessible [here](#)) at 10.

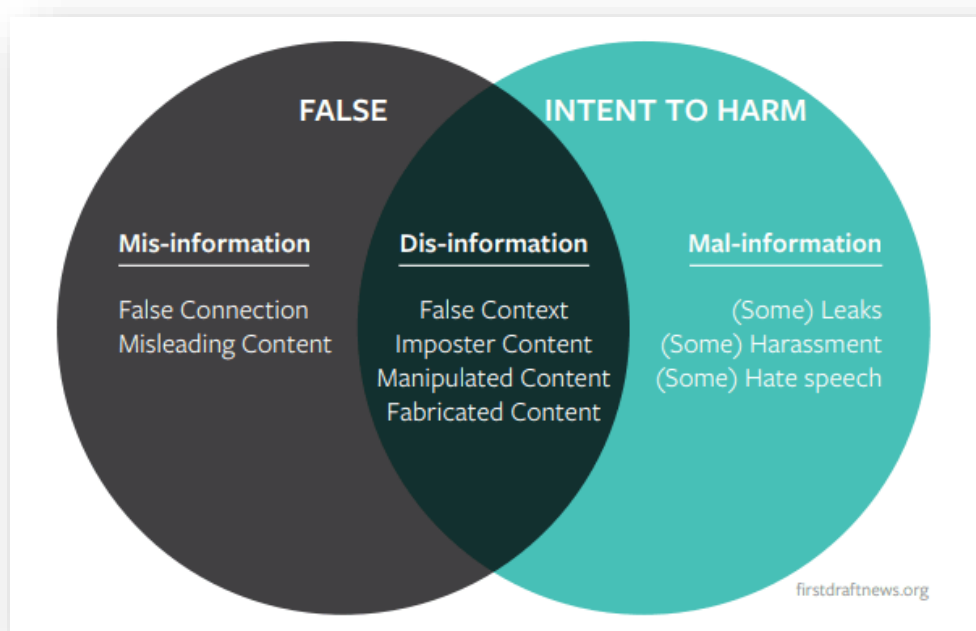
¹⁴ *Id.*

¹⁵ UNESCO, ‘Journalism, ‘fake news’ and disinformation’: Handbook for journalism education and training’, (2018). (Accessible [here](#)) at 20.

Various other terms have therefore been used in this regard, with subtle but meaningful differences:

- **Mis-information**, which is generally used to refer to misleading information created or disseminated without manipulative or malicious intent.
- **Dis-information**, which is generally used to refer to deliberate (often orchestrated) attempts to confuse or manipulate people through delivering dishonest information to them.
- **Mal-information** is sometimes used to refer to information that may be technically correct or based in fact, but is shared in an exaggerated or out-of-context form with the aim of misleading or doing harm.

The distinction in these terms has been illustrated as follows:¹⁶



Source: UNESCO, 'Journalism, 'fake news' and disinformation': Handbook for journalism education and training', (2018) accessible [here](#).

In general, the term 'disinformation' is preferred for regulatory purposes because of two key elements that narrow the scope:

- (i) it requires intention, and therefore takes into account that people may inadvertently share false information without meaning to do harm; and
- (ii) the required intention is targeted — for example, to cause harm or for profit — which both narrows the scope and focuses on the negative consequences that can arise.

¹⁶ UNESCO, above n 7 at 46.

As discussed in more detail in respect of section 89(2) of the **Electoral Act 73 of 1998** (“Electoral Act”) below, the term ‘disinformation’ fits best with the South African legal framework.

4. Regional responses to disinformation

Responses to the issue of disinformation have varied across the world. Some regions have targeted the regulation of social media platforms, while others have simply shut down communication channels. What is certain, however, is that there is no single approach to tackle disinformation; rather, it requires a collaborative, multistakeholder solution. In recognition of the widespread and multinational scope of disinformation, regional bodies such as the **African Commission on Human and Peoples’ Rights** (“ACHPR”) and the European Commission have sought to develop measures that can offer guidance to their member states.

For instance, in 2017 the **Joint Declaration on Freedom of Expression and ‘Fake News’, Disinformation and Propaganda** (“Joint Declaration”) was issued by the Special Rapporteurs for freedom of expression for the United Nations, the African Commission, the Organization for Security and Co-operation in Europe, and the Organisation of American States. The preamble to the Joint Declaration encompasses the competing concerns that arise when dealing with this issue:

- It is often designed and implemented to mislead a population.
- Some forms of disinformation and propaganda may harm individual reputations and privacy, or incite violence, discrimination, or hostility against identifiable groups in society.
- Some public authorities denigrate, intimidate, and threaten the media, including by stating that the media is “the opposition” or is “lying” and has a hidden political agenda, which increases the risk of threats and violence against journalists, undermines public trust and confidence in journalism as a public watchdog, and may mislead the public by blurring the lines between disinformation and media products containing independently verifiable facts.
- It stresses that the human right to impart information and ideas is not limited to correct statements, and that the right also protects information and ideas that may shock, offend, and disturb.
- It states further that “prohibitions on disinformation may violate international human rights standards, while, at the same time, this does not justify the dissemination of knowingly or recklessly false statements by official or State actors.”



KEY PRINCIPLES FROM THE JOINT DECLARATION

Three key principles can be distilled from the Joint Declaration:

- Proposals to address disinformation should avoid offering general prohibitions on speech as solutions, as this is unlikely to meet the test for a justifiable limitation of freedom of expression.
- State actors should not make, sponsor, encourage, or disseminate disinformation or propaganda.
- In addition to not disseminating disinformation or propaganda, state actors should also take positive steps to disseminate reliable and trustworthy information, including on matters of public interest.

In line with the above proposals, the European Commission's **Code of Conduct on Disinformation** contains voluntary, self-regulatory standards agreed to by industry stakeholders to address disinformation. Initially published in 2018 and then strengthened in 2022,¹⁷ many of the world's largest tech platforms had become signatories to the Code, including Meta, Google, Microsoft, and TikTok (Twitter was an initial signatory, but after its transition to X the company did not sign the updated version).¹⁸

These companies provide periodic reports on any actions taken to curb the spread of disinformation online, such as improving the scrutiny of ad placements, ensuring transparency of political and issue-based advertising, and to tackling fake accounts and malicious use of bots, which the Commission publishes together with its own assessments.¹⁹ In 2025, the Code was formally adopted into the framework of the European Union's Digital Services Act ("DSA"), which entered into force in August 2023.²⁰

In 2023, the **Principles and Guidelines for the Use of Digital and Social Media in Elections in Africa** were prepared by the Association of African Election Authorities ("AAEA") to help Election Management Bodies ("EMBs") and other electoral stakeholders to harness the advantages of social media and navigate the risks of emerging digital technologies.²¹ In doing so, the Principles include a series of recommended measures for news organisations to help tackle electoral disinformation, including:

- Putting in place measures and systems to enhance fact-checking and expose disinformation, and building the capacity of journalists and editors to conduct fact-checking.

¹⁷ European Commission, 'Code of Code on Disinformation', (2022). (Accessible [here](#)).

¹⁸ Code of Code Transparency Centre, List of Signatories. (Accessible [here](#)).

¹⁹ *Id.*

²⁰ European Commission, 'Code of Conduct on Disinformation: Signatories publish their latest reports in the Code's Transparency Centre' (2026). (Accessible [here](#)).

²¹ AAEA, 'Principles and Guidelines for the Use of Digital and Social Media in Elections in Africa', (2023). (Accessible [here](#)).

- Undertaking journalism to enhance the integrity of the entire electoral cycle, including coverage of any impact of digital platforms on elections, and any covert financing of content that causes harm.
- Raising public awareness of the risks of misuse of digital and social media.

5. The challenge of regulating disinformation

The advent of the internet and the exponential increase in popularity of social media platforms has led to the weaponisation of information on an unprecedented scale. New technology, such as generative AI, has made the manipulation and fabrication of content simple. Furthermore, social networks dramatically amplify falsehoods that are then shared by susceptible members of the public.²² Sophisticated disinformation is often persuasive because it looks credible.

However, regulating disinformation can be a complex question, especially in a democratic context where political speech is recognised as deserving of strong protections.²³ The complexity is exacerbated by the cross-border nature of social media platforms, with the potential for disinformation arising from both domestic and foreign sources. Given time and resource constraints, as well as a lack of digital literacy skills, voters may not dig deeper into false stories to verify the information.

Various countries have sought to impose regulatory measures aiming to tackle the dissemination of false information.²⁴ Unfortunately, many of the regulatory measures being proposed or implemented in various countries tend towards criminalising the dissemination of false information. This gives rise to serious concern, as there is always the risk that criminalising speech will have a chilling effect on the right to freedom of expression.

As noted in the Joint Declaration, the right to freedom of expression is not limited to the protection of correct statements and includes protection for information and ideas that may shock, offend, and disturb. Any regulatory measure that imposes restrictions on the dissemination of false information needs to be balanced against the right to freedom of expression contained in section 16 of the Constitution and needs to comply with the provisions of section 36 of the Constitution for it to be a justifiable limitation.

²² UNESCO, above n 7 at 15.

²³ The Guardian, 'Global crackdown on fake news raises censorship concerns', (24 April 2018). (Accessible [here](#)).

²⁴ Poynter, 'A guide to anti-misinformation actions around the world', (24 July 2018). (Accessible [here](#)).



EXAMPLES OF CRIMINALISING DISINFORMATION

According to one report, at least 78 countries enacted laws to limited the spread of false or misleading information²⁵ – including 64 which made provision for fines or imprisonment as penalties. The researchers noted that many of these were policies adopted to combat disinformation about the COVID-19 pandemic, suggesting that at least some (such as South Africa’s pandemic-era regulations) have subsequently lapsed.

Recent examples of laws which criminalise disinformation:

- The Criminal Code of Malta was updated in 2018 to make it an offence to “maliciously spread false news which is likely to alarm public opinion or disturb public good order or the public peace or to create a commotion among the public”.
- Singapore’s Online Falsehoods and Manipulation Act (2019) criminalises making a false statement which is detrimental to national security, public health, diplomatic relations, electoral integrity, or public confidence in the government.
- A 2021 amendment to Greece’s criminal code makes it an offence to “spread false news, capable of causing concern or fear to the public or undermining public confidence in the national economy, the country’s defense capacity or public health.”

When considering the legal framework around disinformation, and the regulatory responses thus far, it becomes clear that there is no silver bullet to the regulation of disinformation. On the one hand, the rights to freedom of expression and freedom of speech must be upheld, while tackling the very real danger of disinformation, and putting measures in place to minimise the impact. Across the African continent, there have been various attempts to tackle the issue of digital disinformation, while ensuring a human rights approach is upheld.

UN GLOBAL PRINCIPLES ON INFORMATION INTEGRITY

Accessible [here](#).

The UN Global Principles for Information Integrity aim to address disinformation by promoting a coordinated, human-rights-based approach that reduces the spread and impact of false, misleading and manipulative content while safeguarding freedom of expression. The principles focus on tackling the systemic drivers of disinformation such as opaque platform algorithms, unhealthy digital advertising incentives, declining trust in media, and gaps in digital and media literacy.

²⁵ Lim and Bradshaw ‘Chilling Legislation: Tracking the Impact of “Fake News” Laws on Press Freedom Internationally’ Center for International Media Assistance (2023). (Accessible [here](#).)

6. Guidelines for dealing with election disinformation

Although South Africa does not have a comprehensive law dealing with disinformation, section 89(2) of the Electoral Act contains an express prohibition against the publication of false statements with the intention of causing one of the three listed harms: (i) disrupting or preventing an election; (ii) creating hostility or fear in order to influence the conduct or outcome of an election; or (iii) influencing the conduct or outcome of an election. A person convicted of an offence can be sentenced to a fine or imprisonment for up to ten years.²⁶

The Electoral Code is also relevant to disinformation. Section 4(1)(a) states that every party and candidate must, among other things, publicly state that everyone has the right to freely express their political beliefs and opinions, to challenge and debate the political beliefs and opinions of others, and publish and distribute election and campaign materials. Section 4(1)(b) requires every registered party and candidate to “publicly condemn any action that may undermine the free and fair conduct of elections.”

Section 9 sets out the prohibited conduct in terms of the Electoral Code. Of particular relevance, section 9(1)(b) provides that no party or candidate may publish false or defamatory allegations in connection with an election in respect of a party, its candidates, representatives or members, or a ward candidate or that candidate’s representatives. Section 9 should be read with section 3(b) of the Electoral Code, which requires all parties and candidates to instruct candidates, representatives, members, and supporters to comply with this provision as well.

ELECTORAL LAWS REGARDING FALSE INFORMATION

Section 89 of the Electoral Act, 1998 prohibits intentional false statements:

- No person, when required in terms of the Act to make a statement, may make a statement knowing it is false or without believing on reasonable grounds that the statement is true.
- No person may publish any false information with the intention of disrupting or preventing an election; creating hostility or fear to influence an election; or influencing the conduct or outcome of an election.”

Section 9(1)(b) of the Electoral Code of Conduct includes false statements under its list of prohibited conduct for parties and candidates:

- No registered party or candidate may publish false or defamatory allegations, in connection with an election, about a party, its candidates or members, or an independent candidate or their representatives.

²⁶ Section 89(2) of the Electoral Act.



GUIDANCE NOTE ON COUNTERING MIS- AND DISINFORMATION DURING ELECTIONS

Accessible [here](#)

In 2024, the Information Regulator prepared a Guidance Note which proposed accountability-driven approaches to countering mis and disinformation during elections. It encourages political parties and independent candidates to develop and implement dedicated policies, strategies, and action plans aimed specifically at preventing, identifying, and responding to false or misleading information throughout the election period.

These measures include the active monitoring of digital and social media platforms, prompt investigation of incidents, and disciplinary action against members or officials who engage in conduct that undermines the right of access to information, infringes privacy, or causes online harm.

The Guidance also calls for formal statements of commitment by political actors to uphold electoral integrity, with particular emphasis on monitoring and managing high-risk practices such as data harvesting, micro-targeting, deepfakes, and AI-generated content.

In an effort to give effect to these provisions, Moxii Africa (formerly Media Monitoring Africa), supported by the IEC, SANEF and other stakeholders, has maintained an [online portal](#) and complaints mechanism for members of the public to lodge complaints regarding disinformation online that falls within the scope of the Electoral Act. As part of this process, a committee of experts considers complaints on disinformation and makes recommendations on the proposed recourse to take based on the specific type of disinformation and potential harm caused.

FINDINGS FROM 2024 GENERAL ELECTIONS

Accessible [here](#).

Moxii Africa has undertaken extensive research in respect of disinformation and possible strategies and approaches for elections. This includes a research report for the 2024 National and Provincial Elections. Notably, the report highlighted a marked increase of online disinformation compared to previous monitoring periods. It was noted that such content frequently aimed at confusing voters about electoral procedures, misleading the public about parties, candidates and results, and deliberately undermining trust in the IEC. While online platforms enabled broad participation and political engagement, they also amplified false and manipulative narratives at speed and scale, increasing the risk of voter confusion, polarisation, and post-election instability.

7. Standard of care by the media

Various codes of conduct — including the Press Code of Ethics and Conduct for South African Print and Online Media (“Press Code”), the Independent Communications Authority of South Africa (“ICASA”) Code of Ethical Conduct (“ICASA Code”) and the Broadcasting Complaints Commission of South Africa (“BCCSA”) Codes of Conduct — place requirements on members of the media to report news truthfully, accurately, and fairly. Different media organisations will likely put in place different measures to achieve this, to ensure that reasonable steps are taken to meet these requirements. However, there will inevitably be instances in which members of the media will make mistakes. Members of the media should at all times conduct themselves in a manner that is reasonable and compliant with the ethical standards expected of the profession.

National Media Limited and Others v Bogoshi

[1998] ZASCA 94

The case of *National Media Limited and Others v Bogoshi* deals with developing the reasonableness defence available to a media institution for publishing a false and defamatory statement. The Supreme Court of Appeal held that:

“In my judgment we must adopt this approach by stating that the publication in the press of false defamatory allegations of fact will not be regarded as unlawful if, upon a consideration of all the circumstances of the case, it is found to have been reasonable to publish the particular facts in the particular way and at the particular time.

In considering the reasonableness of the publication account must obviously be taken of the nature, extent and tone of the allegations. We know, for instance, that greater latitude is usually allowed in respect of political discussion ... and that the tone in which a newspaper article is written, or the way in which it is presented, sometimes provides additional, and perhaps unnecessary, sting. What will also figure prominently, is the nature of the information on which the allegations were based and the reliability of their source, as well as the steps taken to verify the information. Ultimately there can be no justification for the publication of untruths, and members of the press should not be left with the impression that they have a licence to lower the standards of care which must be observed before defamatory matter is published in a newspaper.”

Although the judgment pre-dates the current debates regarding the dissemination of false information online, the guidance is nevertheless of relevance, particularly with regard to the factors that ought to be considered when determining whether the publication of false information was justifiable.

It must also be remembered that denigrating legitimate news reports as ‘fake news’ has been used as a political tool to stifle criticism. In other parts of the continent, publishing so-called ‘false information’ is a criminal offence, which has historically been used as an intimidation tactic to silence journalists.²⁷ As noted in the preamble to the Joint Declaration, mentioned above, public

²⁷ See, for example, the judgment of the Community Court of Justice of the Economic Community of West African States in *Federation of African Journalists and Others v The Republic of the Gambia*, Judgment No. ECW/CCJ/JUD/04/18 (13 March 2018), in which it held that the false news provision contained in the Criminal Code did not conform with the

authorities claim that the media is lying and has a hidden political agenda, in an effort to undermine public trust and confidence in journalism. This may in turn mislead the public by blurring the lines between disinformation and media products containing independently verifiable facts.

However, this does not absolve the media of responsibility. As noted by UNESCO, “it is a time for news media to tack more closely to professional standards and ethics, to eschew the publishing of unchecked information, and to take a distance from information which may interest some of the public but which is not in the public interest.”²⁸ According to UNESCO, all news institutions and journalists, whatever their political leanings, should avoid inadvertently and uncritically spreading disinformation and misinformation.²⁹ Furthermore, journalists cannot leave it to fact-checking organisations alone to do the journalistic work of verifying questionable claims that are presented by sources. Journalism also needs to proactively detect and uncover new cases and forms of disinformation.³⁰

DIFFERENT FORMS OF FALSE CONTENT

Source: UNESCO, ‘Journalism, ‘fake news’ and disinformation’: Handbook for journalism education and training’, (2018) (accessible [here](#)).

- **False connection**, in which headlines, visuals, or captions do not support the content (commonly referred to as ‘clickbait’).
- **Misleading content**, in which there is a misleading use of information to frame issues or individuals in certain ways, such as by cropping photographs, or choosing quotes or statistics selectively.
- **False context**, in which genuine information is re-circulated outside of its original context.
- **Imposter content**, in which journalists have their by-lines used alongside articles they did not write or their organisations’ logos used in videos or images they did not create.
- **Manipulated content**, in which genuine content is manipulated to deceive, such as an image being manipulated to impute an improper relationship between two people.
- **Fabricated content**, which includes completely fabricated ‘news’ websites or fabricated images.

Referred to as ‘information disorder’, journalists need to separately examine the elements of information disorder: the agent, the messages, and the interpreter, as well as the different phases of information disorder: creation, production and distribution.

international law standards on freedom of expression contained in article 9 of the African Charter on Human and Peoples’ Rights and article 19 of the International Covenant on Civil and Political Rights.

²⁸ UNESCO, above n 6 at 11.

²⁹ *Id.* at 11.

³⁰ *Id.* at 11-12.

8. Guidelines for countering disinformation

Disinformation is a complex issue and therefore requires a multifaceted response. However, there are a number of practical ways in which journalists and the media are able to — indeed have the responsibility to — take steps in order to contribute to minimising the harm caused by disinformation.

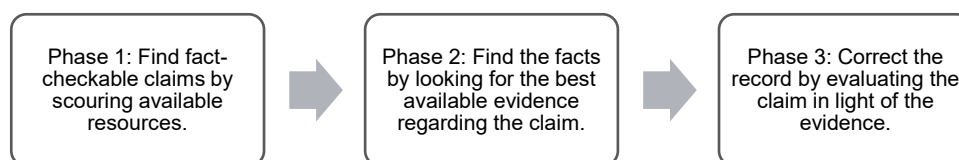
One of the key roles the media can play in addressing the challenge of disinformation is in publishing counter-narratives that provide corrections to false information and that highlight the work being done by fact-checking organisations. In essence, instead of killing a false story, this approach causes a story to be surrounded with related, credible articles to provide the reader with more context and alternative views.³¹ This invites easier access to alternative perspectives and information, including articles by third-party fact-checkers. While this approach raises questions about algorithmic transparency on social media platforms, it is an approach that some platforms have already begun to implement through suggested content appearing alongside a particular post.

This does require consideration for how the additional information can be made relatable to the audience to whom the false information was targeted, and can be readily accessed. The effect of misinformation can be very strong, and requires engagement and participation from audiences to be persuaded.

A 2017 study on misinformation recommends the following approaches for debunking misinformation:³²

- **Reduce arguments that support misinformation:** News accounts about misinformation should not inadvertently repeat or belabour detailed thoughts in support of the misinformation.
- **Engage audiences in scrutiny and counterarguing of information:** A state of healthy scepticism should be promoted. When trying to correct misinformation, it is beneficial to have the audience involved in generating counterarguments.
- **Introduce new information as part of the debunking message:** People are less likely to accept debunking when the initial message is just labelled as wrong rather than countered with new evidence.

Fact-checking of information has also been key to countering disinformation. In general, fact-checking comprises three phases:³³



³¹ Alemanno, 'Editorial: How to counter fake news? A taxonomy to anti-fake news approaches', in *European Journal of Risk Regulation* 9 (2018). (Accessible [here](#)) at 4.

³² Annenberg Public Policy Center of the University of Pennsylvania, 'Debunking study suggests ways to counter misinformation and correct "fake news"', (12 September 2017). (Accessible [here](#)).

³³ UNESCO, above n 7 at 89.

There are various online tools that can be used to assess the credibility of a particular online news resource. Moxii Africa, for example, has developed KnowNews, a browser extension for Google Chrome and Firefox that can be downloaded to help identify the credibility of a news website. The KnowNews database is a collection of news publishers, and once installed the KnowNews icon provides a colour-coded guide to how trustworthy a website is considered. People can also verify the existence of a news entity by going to <https://openanddisclose.org.za/>.

Moxii Africa has also built an app, RoveR, that enables users to fulfil three core functions:

- (i) upgrade their skills on spotting real versus false news by going through a few learning modules;
- (ii) testing their skills on spotting real versus false news through quizzes; and
- (iii) preventing sharing of false news by allowing users to check the credibility of the website from which it is obtained. RoveR can be downloaded from the Google Play Store or from here: <https://rover.directory/>.

Africa Check — a non-partisan organisation that works towards accuracy and honesty in public debate and the media in Africa — explains their approach to fact-checking in the following eight steps:³⁴

STEP	HOW IT WORKS
1. Select the claim to check	At Africa Check, the editors sift through the suggestions sent in by readers and raised by the team, based on criteria set out on the website: the importance of the topic, whether the claim was framed as a statement of fact or opinion, whether the claim matters, and whether it is a speaker Africa Check has focused on before.
2. Establish exactly what was said	Once the topic is selected, it is necessary to establish exactly what was said. The precise wording is needed. Consider the following questions: What exactly did they say? Was it as reported? And what was the context in which it was said?
3. Ask for their evidence	Having established the claim, try to contact the speaker or their office, and ask what evidence they have for their claim.
4. Check archives and other sources	Check through archives and other publicly available sources, both for evidence that supports and that contradicts the claim. It is advisable to cast the net as widely as possible. Africa Check's <u>InfoFinder</u> tool can be a useful resource.
5. Discuss the evidence with experts	Having secured the evidence, this should be discussed with specialist experts where necessary to help understand the data. According to Africa Check, they only discuss with experts willing to go on the record, and do not use anonymous sources.
6. Write up the report	The next step is to write up the report, setting out the following: (i) the claim that was made and the context in which it was delivered and reported; (ii) the evidence that supports the claim; (iii) any contrary evidence; and (iv) a balanced conclusion. For all evidence, a link or source should be provided.

³⁴ Africa Check, 'How we work', (undated). (Accessible [here](#)).

7. Have a colleague review the report and findings	To ensure that the report itself is accurate, a colleague should be asked to review the report, and independently assess the findings, before it is finalised.
8. Publish and monitor feedback	Finally, the report is published, and feedback is monitored. If or when a reader identifies an error, the report is updated openly.

When dealing with the spread of false information, verifying source content and visual content is also important. Verification tools can be used to establish where a source has posted from, but it is also possible to manually triangulate a source by analysing their social media history to check for clues that could indicate the feasibility of them being in a particular place at a particular time.³⁵ Examining the history of their interactions with other users and checking linked content within posts also assists in the manual verification process and can help eliminate information shared by bots.³⁶

CHECKLIST OF VERIFYING VISUAL CONTENT

Source: UNESCO, 'Journalism, 'fake news' and disinformation: Handbook for journalism education and training', (2018) (accessible [here](#)).

While it may not be possible to ascertain with full certainty the provenance of visual content, there are several indicators that can be uncovered through a verification process that asks:

- Is the content original, or has it been 'scraped' from previous reporting and re-appropriated misleadingly?
- Has the content been digitally manipulated in some way?
- Can we confirm the time and place of the photograph or video capture, using available metadata?
- Can we confirm the time and place of the photograph or video capture, using visual clues in the content?

In order to assess these indicators, it is also useful to understand the different types of false or misleading visual content that commonly arise.³⁷ This includes, for instance, wrong time/wrong place content that re-shares old visuals with new claims about what they show; manipulated content that has been digitally manipulated using editing software; or staged content that creates or shares original content with the intention of misleading.³⁸

It should be emphasised that in many instances, there will not be one single indicator that reveals whether the image or content is false information. Rather, the relevant conclusion is reached on a balance of all the information that is available. It is also important to remember that not all false information is intended to mislead or be malicious; this can sometimes arise from a genuine error of fact or judgment.

³⁵ UNESCO, above n 5 at 105.

³⁶ *Id.* at 105.

³⁷ *Id.* at 105-106.

³⁸ *Id.* at 105-106.

While the tools and resources available to assist with fact-checking and verification processes are invaluable, journalists should also remember that there will be times when one's instinct (in conjunction with discussions with editors and the broader news team) will be a key determining factor in assessing particular content.

TOOLS FOR VERIFICATION

- **Intel Techniques** undertakes Facebook account analysis and enables a journalist to find out more about a source by analysing their Facebook account (accessible here: <https://inteltechniques.com/osint/facebook.html>).
- **Google Reverse Image Search** enables a journalist to check if the image database contains an earlier version of that image, in order to ascertain whether the image is being recycled to support a new claim or event (accessible here: <https://support.google.com/websearch/answer/1325808?hl=en>).
- **YouTube Data Viewer** can detect video thumbnails for YouTube videos and facilitate a reverse image search on those thumbnails to check if earlier versions of the video have been uploaded (accessible here: <https://citizenevidence.amnestyusa.org/>).
- **Geolocation** is the process of determining where the video or image was captured. This can be obtained from metadata or by cross-referencing visual characteristics and landmarks from the content with satellite imagery, street-view imagery and content available from other sources.
- **Weather corroboration** relies on historical weather data to check if the weather observable in visual content is corroborated by the historic record.
- **Shadow analysis** examines the internal consistency of any shadows, such as where one would expect them to be and whether the visible shadows are consistent with the light sources.
- **Image forensics** seeks to detect inconsistencies in image metadata that suggest manipulation.
- **Microsoft's Video Authenticator** analyses still photos and video frames to provide a confidence score on whether the media has been artificially manipulated.
- **Reality Defender** offers real-time detection of synthetic media, including deepfakes, AI-generated images, and manipulated audio.

9. Suggested resources

- AAEA, 'Principles and Guidelines for the Use of Digital and Social Media in Elections in Africa', (2023) (accessible [here](#)).
- A. Alemanno, 'Editorial: How to counter fake news? A taxonomy to anti-fake news approaches', in *European Journal of Risk Regulation* 9 (2018) (accessible [here](#)).
- C. Silverman, 'Verification handbook: The ultimate guideline on digital age sourcing for emergency coverage', (undated) (accessible [here](#)).
- Data & Society, 'Reading metadata' in *Data craft: The manipulation of social media metadata*, (undated) (accessible [here](#)).
- European Commission, 'A multi-dimensional approach to disinformation: Report of the independent High Level Group on fake news and online disinformation', (2018) (accessible [here](#)).
- European Parliament, 'Societal costs of 'fake news' in the Digital Single Market', (2019) (accessible [here](#)).
- Global Investigative Journalism Network, 'Reporter's Guide to Detecting AI-Generated Content' (2025) (accessible [here](#)).
- UNESCO, 'Journalism, 'fake news' and disinformation': Handbook for journalism education and training', (2018) (accessible [here](#)).
- UNESCO, 'Elections in Digital Times: A Guide for Electoral Practitioners', (2022) (accessible [here](#)).
- UNESCO, 'Deepfakes and the crisis of knowing', (2025) (accessible [here](#)).

ENDS.

