Legal Response to Social Media Disinformation on National Level

MM Watney
University of Johannesburg, South Africa

mwatney@uj.ac.za

Abstract: Social media has an enormous impact on the manner in which society communicates and shares information. Digital is no longer a supplementary channel, but is the first place most people go to for news, information and communication. The transmission of social media disinformation has increased dramatically across the world and it necessitates a response. The discussion focuses on the response to social media disinformation on a national level. The discussion does not focus on foreign state or state-sponsored actors of misinformation. The focus and publicity may - within the context of cybersecurity - predominantly have been on cyberattacks, such as ransomware attacks. However, recent incidents - unrelated to foreign state interference and cyberattacks - illustrate that cybersecurity law must encompass the threat of disinformation. The 2020 COVID-19 pandemic, 2021 Washington, DC, United States, and South African as well as the 2023 Brazil riots illustrate the harmfulness of social media disinformation. Cognisance should be taken of the lessons learnt from the examples of social media disinformation as it may assist in determining a response to disinformation. There are various responses to national social media disinformation, such as legislative social media platform regulation, censorship, and criminalisation of the disinformation by itself. The response within the context of a cybersecurity threat landscape necessitates scrutiny as the response may impact on human rights. The trade-off between security and human right protection may be the violation of human rights to prevent harm from disinformation.

Keywords: Disinformation; misinformation; disinformation as a cybersecurity threat; social media platform regulation of disinformation; response to social media disinformation on national level.

1. Introduction

Facebook sees approximately 300 million new photos uploaded daily whereas six thousand tweets are published every second. The most popular YouTube channels receive over 14 billions views weekly, whilst the messaging app Telegram boasts over 500 million users (Hook and Verdeja, 2022). These statistics illustrate the enormous and unmistakable impact social media has on the manner in which society communicates and shares information and how deeply woven social media is into the fabric of society.

Social media has many societal benefits, such as providing a voice to many who have been voiceless in the past. However, social media communication and information-sharing can also be abused and exploited, as seen with the creation and distribution of disinformation. The concept of disinformation refers to false, inaccurate, or misleading information designed, presented, and promoted intentionally to cause public harm or make a profit (Colomina, Margalef, and Youngs, 2021).

Anyone with a social media account can create and spread disinformation, including governments, companies, other interest groups, or individuals. There are different actors responsible for disinformation (Colomina, Margalef, and Youngs, 2021), namely

- those who fabricate the disinformation, referred to as the instigators; and
- those who spread the content, referred to as the agents ('influencers', individuals, officials, groups, companies, and institutions).

Disinformation is aimed at driving a specific agenda. The motivation behind disinformation can be (Watney, 2018; Colomina, Margalef, and Youngs, 2021):

- personal (damage to a person or reputation of a company);
- political (to influence the public's viewpoint or ideology); or
- financial (to deceive).

Hook and Vereja (2022) indicates that researchers believe that organised social media disinformation campaigns have operated in at least 81 countries, a trend that continues to grow yearly.

Although the discussion does not focus on foreign interference with elections, Lasiello (2021) makes a valid point when he indicates that in 2016 the US government focussed predominantly on the threat of cyberattacks and "may have missed seeing the forest for the trees". Not only did the 2016 election reveal how ill-prepared the country was to address the soft-power tenets of information warfare (IW), but it also highlighted that social

media disinformation should form part of the cybersecurity threat landscape. Since 2016 there has been many incidents of social media disinformation unrelated to foreign interference and the attention has shifted to the cybersecurity threat that social media disinformation presents and the response thereto.

The aim of this discussion is to consider the manner in which parties, such as governments, social media users, private companies and service providers, may respond to social media disinformation. The main question will be discussed with reference to the following interlinked issues:

- Key terms will be conceptualised from a legal perspective such as misinformation, disinformation, fake news, propaganda, and cybersecurity.
- Reference will be made to examples of disinformation, such as the 2020 COVID-19 pandemic, the 6 January 2021 Washington, DC riot and the July 2021 South African riots.
- Disinformation and the impact of it will be highlighted within a cybersecurity threat landscape.
- The importance of a human rights centred approach to the mitigation of social media disinformation will be emphasised.
- The response to disinformation as a cybersecurity threat on a national level will be scrutinised.

2. Conceptualising terminology

It is important that terminology relevant to the discussion is conceptualised as it will serve as a point of reference with regards to the manner in which cybersecurity law may deal with social media disinformation nationally from a legal perspective. It appears that there are not universal definitions for the key concepts which is problematic as it may result in legal uncertainty.

The discussion focusses on the distribution of inaccurate or misleading communication and information by means of social media. “Social media” refers to computer-based platforms that enable people to communicate and share information across virtual networks in real-time. This includes community apps such as Facebook and Twitter, media-sharing channels like YouTube and Instagram, messaging apps such as WhatsApp, and hybrid platforms like Telegram (Hook and Verdeja, 2022).

A distinction is drawn between social media misinformation (SMM) — involving false or misleading information shared on these platforms — and social media disinformation (SMD), the subset of intentionally shared misinformation. The difference between misinformation and disinformation is that the latter is intentionally shared to cause harm. Malinformation on the other hand is true information which is shared to cause harm (Colomina, Margalef, and Youngs, 2021).

Understanding the difference between propaganda and disinformation is important as these terms are sometimes used interchangeably. Propaganda is information that might be true, but is often exaggerated, false, or misleading. It is used to influence public opinion to help a person, organization, or government. Propaganda creation is often associated with governments, but individuals and organizations outside of politics may use propaganda as well. The purpose of propaganda is to encourage people to promote a political cause or to think in a particular way and turn people away from a political cause (Bentzen, 2022).

Hook and Verdeja (2022) use the term *“social media misinformation*” and refer to SMD when there is a plausible indication of an intent to deceive such as with influence operations. In the discussion, the term “disinformation” is used as the discussion focuses on misleading or inaccurate information intentionally created and distributed on social media within a country and how the parties affected by such a cybersecurity threat should respond.

After the 2016 United States (US) elections, the term “fake news” was used, but experts have called for this term not to be used as it is too vague and ambiguous to capture the essence of disinformation (Colomina, Margalef, and Youngs, 2021). The term ‘fake news’ is commonly used to discredit the media. Since 2016 a clearer understanding of disinformation, has developed and for the purpose of legal certainty, the use of a more precise term such as disinformation is preferred. Baptista and Gradim (2022) propose the following working definition for “fake news”: “(1) a type of online disinformation, with (2) misleading and/or false statements that may or may not be associated with real events, (3) intentionally created to mislead and/or manipulate a public (4) specific or imagined, (5) through the appearance of a news format with an opportunistic structure (title, image, content) to attract the reader’s attention, in order to obtain more clicks and shares and, therefore, greater advertising revenue and/or ideological gain”. The latter definition will be applicable to the use of “fake news” in the discussion.

Disinformation should be seen within the context of cybersecurity. Cybersecurity is generally defined as the measures, policies, tools, and laws aimed at protecting hardware, software, and data in cyberspace against threats. Disinformation should be seen as part of the cybersecurity threat and risk landscape.

3. Examples of disinformation

The examples hereafter show that disinformation undermines human rights and many elements of good democratic practice (Colomina, Margalef and Youngs; 2021). Disinformation threatens freedom of thought, the right to privacy, and the right to democratic participation, as well as endangering a range of economic, social, and cultural rights. It also diminishes broader indicators of democratic quality, unsettling citizens' faith in democratic institutions by encouraging digital violence and suppression. On a global level, countries should take note of the lessons learnt from the social media disinformation examples in order to respond effectively to the threat and risk of disinformation.

Example 1: 2020 COVID-19 pandemic

Colomina, Margalef and Youngs (2021) indicates that during the pandemic, social media acted as a double-edged sword. On the one hand, digital platforms were useful in promoting debate within the scientific community and disseminating valuable information as well as investigative results. On the other hand, there was a great deal of disinformation. The disinformation was harmful as it resulted in mistrust in public institutions and put people's health at risk. The pandemic has also been referred to as an "infodemic" of misinformation. An infodemic refers to a large increase in the volume of information associated with a specific topic and whose growth can occur exponentially in a short period of time due to a specific incident. It should be noted that any big event may coincide with disinformation.

Example 2: Washington, DC riot

The 6 January 2021 Washington, DC, riot clearly illustrates the impact of disinformation during national elections (Almany *et al*, 2021; Jurecic, 2022). The riot followed a months-long disinformation campaign by former President Donald Trump and his allies, who claimed, without evidence, that the election had been stolen through fraud. The disinformation caused confusion and manipulation of social media users receptive to this rhetoric. The riot lasted seven hours, during which approximately 10,000 people came onto Capitol grounds, with many engaging in violent clashes with police officers trying to protect the building and lawmakers inside. At least 2,000 made it inside the Capitol building (Rubin, Mallin and Steakin, 2022).

Globally governments must be mindful of the impact of disinformation prior to and after an election. Increasingly it appears that elections on a national level may coincide with the use of social media disinformation aimed at causing political instability if a certain political party is not elected.

Example 3: Brazil

On 9 January 2023, Brazil experienced rioters storming democratic institutions, such as the Congress, Supreme Court, and Presidential Place. Brazilian analysts have long warned of the risk in Brazil of an incident similar to the 2021 insurrection at the US Capitol (Dwoskin, 2023). In the months leading up to the country's presidential election in October 2022, social media channels were flooded with disinformation, along with calls in Portuguese to "Stop the Steal" and cries for a military coup should the right wing incumbent, Bolsonaro, lose the election.

Example 4: 2021 South African riots

In July 2021, the province KwaZulu-Natal in South Africa (SA) ,experienced one of the worst uprisings since 1994. In 2022, **Media Monitoring Africa (MMA) director, William Bird told the SA Human Rights Commission that the disinformation posts shared during the July unrest sowed anger and distrust and contributed to the widespread violence and looting (**Pijoos, 2022)**.** While some social media messages served a genuine purpose to alert people to dangerous situations and places or roads to avoid, and areas to evacuate, many messages were misleading with out-of-context images or videos repurposed purely to be alarmist and deepen the sense of panic and anxiety (Malinga, 2021).

4. Challenges in mitigating social media disinformation

The characteristics of social media exacerbate the distribution and spread of disinformation (Watney, 2018). It provides for:

- instant and interactive communication;
- easy and cheap access;
- quick dissemination of disinformation to many; and
- a sense of interconnectedness.

Social media disinformation presents various inter-linked challenges, such as:

1. The rapid pace of social media and the technology culture associated with its rise accelerate the spread of misinformation, pushing it faster and further (Hook and Verdeja, 2022). The famous quote "A lie can travel around the world and back again while the truth is still lacing up its boots" resonates more than ever (Lapping; 2021). While professional journalists are trying to verify the facts before filing their reports, a deliberate falsehood created to stoke panic or anger can spread rapidly across social media. During COVID the term infodemic was used as this neologism (information + epidemic = infodemic) refers precisely to the speed with which disinformation spreads, similar to a virus (Baptista and Gradim, 2022).
2. Many actors of misinformation believe what they are sharing and are not intentionally spreading false information (Hook and Verdeja, 2022). There is a consensus that misinformation happens unintentionally (Colomina, Margalef and Youngs; 2021).
3. Not all harmful social media posts are wholly false. They may contain some truth, be largely true but misleading or out of context, or may advocate claims that are problematic but do not necessarily qualify as imminently dangerous.
4. Since social media provides the platform on which disinformation is spread, it needs to take responsibility for the content they distribute. The form of responsibility will be discussed hereafter at paragraph 5. It should be noted that the threat of disinformation must be mitigated by means of pro-active measures. Social media will always struggle to verify the truthfulness of content or assess the potential harm of each of the hundreds of thousands of posts their users generate every minute of the day. Social media platforms and their users must therefore share the responsibility of combating harmful or false content (Lapping, 2021).
5. There are many social media disinformation posts daily. It is not possible for law enforcement to respond to all social media disinformation postings. Determining the circumstances in which law enforcement must respond to social media disinformation as a cybersecurity threat may be problematic but the following lessons have been learnt from the 2021 Washington, DC, South African, and 2023 Brazil riots:

- Social media can be used as a tool to incite violence in the "real world". There are people who are looking to exploit and manipulate a situation by spreading disinformation and inciting violence to forward their own agendas (Malinga, 2021).
- The police should take a warning of impending violence seriously. For example, local officials, FBI informants, social media companies, former national security officials, researchers, lawmakers, and tipsters alerted law enforcement of possible violence on 6 January 2021, but law enforcement officials in Washington, DC did not respond with urgency to the cascade of warnings about possible violence (Almany *et al,* 2022). Washington, DC law enforcement may not have envisaged that social media communication could result in violence.

The 2021 US Capitol Hill riot following social media disinformation has now set a precedent and governments globally should be aware that elections on a national level may coincide with disinformation and possible protests and riots. Brazil law enforcement disregarded the warnings of possible violence.

6. Proving the origin of specific pieces of disinformation may be difficult (Hook and Verdeja, 2022).
7. It is difficult to track or completely prevent the spreading of disinformation. Content sharing has moved from open to encrypted platforms, such as WhatsApp.
8. There are many tools available to spread disinformation and it is not easy to keep up with all the technological tools used to impart disinformation. For example, the use of deep-fake technology can be harmful since people cannot tell whether content is genuine or false. The Cyberspace Administration

of China has started implementing China's new rules that prohibit the use of AI-generated content for spreading fake news or information deemed disruptive to the economy or national security (Hutton, 2023). In 2019 the US passed legislation to regulate AI-generative information concerning deepfakes and the EU is in the process of considering additional regulations for deepfakes and AI-generated imagery (Hutton, 2023).

5. Response to the cybersecurity threat of disinformation

5.1 Introduction

The challenges discussed at paragraph 4 shows that disinformation is a threat and requires pro-active measures.

Colomina, Margalef and Youngs (2021) identify three groups of actors in responding to disinformation, namely

- legislative and regulatory bodies;
- private sector such as social media platforms; and
- civil society.

Certain responses target the actors deemed responsible for disinformation; some target the disruptive techniques used whereas others aim at improving citizens resilience to disinformation.

The response must be human rights centred (Colomina, Margalef and Youngs, 2021). Although disinformation threatens human rights, the inverse challenge is that counter-disinformation policies can also restrict freedoms and rights. There may have to be a trade-off between security and the violation of the rights to freedom of expression and access to information but it will depend on the seriousness of the harm that the disinformation constitutes.

5.2 Threat actors held responsible for disinformation per se by means of anti-disinformation law

A complex issue is whether a government should implement anti-disinformation laws to hold the actors of disinformation criminally responsible. Those who support enacting anti-disinformation law criminalising disinformation argue that the right to freedom of speech does not include the freedom to spread disinformation. On the other hand, anti-disinformation law may violate freedom of speech and amount to censorship.

Different global contexts have diverse norms, ethical standards, and values around the blend of civil liberties, national security, and civilian protection as reflected by some countries that have criminalised disinformation *per se* (Hook and Verdeja, 2022).

In 2016, China criminalised the spread of information that may "undermine economic and social order". To further censor any possible misinformation, a law enacted in 2017 also demands that social media platforms must only circulate news from registered news outlets. Additionally, an app has been launched in China which allows its users to report those who are potentially spreading fake news (Elton, 2020).

Singapore enacted the Protection from Online Falsehoods and Manipulation Act (POFMA) in 2019. The POFMA strives to combat false statements of fact through online communication and provide support for the consequences of such communications. The POFMA also aims to censor false information that is against public interest. Contravening of the act may lead to a possible conviction of 10 year imprisonment and/or a $100,000 fine (Elton, 2020).

In 2022, Turkey adopted anti-disinformation law. Article 29 of the law provides that those who spread false information online about Turkey's security to "create fear and disturb public order" will face a prison sentence of one to three years. Sentences may be increased if the disinformation is spread through anonymous accounts (Shan, 2022).

In 2022, Russia lawmakers implemented a law imposing imprisonment of up to 15 years for spreading intentionally "fake" news about the military. They also impose fines for public calls for sanctions against Russia (Bloomberg News, 2022).

I opine that only the most serious disinformation content, such as the calls for incitement or the threat of violence to persons or damage to property based on the disinformation, should result in criminal liability. Criminalisation is a reactive measure and the emphasis should be on pro-active measures.

5.3 Social media platforms

5.3.1 Introduction

Social media platforms must take responsibility for disinformation as their platform is used for the dissemination of it, but the manner in which the responsibility is determined, is evolving (see par. 5.3.2)

There is a risk that the activities of the platforms in combating disinformation may restrict freedom of expression (Colomina, Margalef and Youngs, 2021). In 2021, Twitter banned former US president Trump from its platform amid concerns about Trump and his followers using social media to spread disinformation (Addler and Frier 2022; Jurecic, 2022; Lapping, 2021; Romano, 2021). Deplatforming former US president, Trump raised questions about limitations to freedom of expression and the role of gatekeepers of disinformation (Colomina, Margalef and Youngs, 2021). Romano (2021) indicates that deplatforming is effective at removing extremists from mainstream Internet spaces and he opines that it is not a violation of free speech, since free speech does not mean that there are no consequences or boundaries.

During the 2021 South African riots, a proposal was made to declare a state of emergency and to prevent access to certain communication towers in areas where the community was being mobilised to riot and loot (Nxumalo, 2022). However, social media can be used to access authentic information and serve as an efficient communication tool for urgent, ongoing and rapid up-to-date communication. By denying access to social media, the right to free speech and access to information would be violated unless the trade-off is security. There is a fine balance between security and protection of rights.

5.3.2 Evolution of social media self-regulation to legislative platform regulation

Platforms have terms and conditions that regulate the use of their platforms. Over the years attention has increasingly focussed on social media moderation and the responsibilities that platforms can and should have to control what people do on their sites.

Social media platforms have been attempting to address disinformation (Adler and Frier, 2022; Jurecic, 2022 Hetler, 2022). For example, Facebook runs two initiatives, News Integrity Initiative and Facebook Journalism Project, to highlight problems with disinformation and spread awareness. The platform also takes action against pages and individuals that share fake news and remove them from the site. Instagram and Facebook have a new "false information" label to combat disinformation. Third-party fact checkers review and identify potential false claims and posts. If this team determines this information is untrue, they flag it with a label to notify social media users it contains disinformation. When readers want to view a post with this label, they must click an acknowledgement that says the information is not true. If they try to share this information, they get a warning they are about to share disinformation (Hetler, 2022).

Many feel that social media companies are not doing enough in addressing disinformation. There is now a move away from platform self-regulation to platform legislative regulation to establish platform legal accountability pertaining to specifically illegal content. This move means that a social media platform can be held legally accountable if they do not comply with the rules.

Some governments have moved quicker to legal regulation than other governments. India put Twitter, Facebook, and the like under direct government oversight, enacting regulations requiring Internet platforms to help law enforcement identify those who post “mischievous information” (Adler and Frier, 2022). In China, Russia, and other countries subjected to authoritarian rule, governments actively censor the Internet, including blocking or greatly restricting access to American-owned social media sites (Adler and Frier, 2022).

The following legislation serves as examples of platform regulation:

- EU platform regulation in terms of the Digital Services Act (DSA)

On 5 July 2022, the European Union Parliament approved the DSA (https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment_en). The DSA gives member states new power to take down illegal content such as hate speech and terrorist propaganda. The platforms now have to comply with rules for content moderation by applying the so-called “notice and action mechanism”. If users contest a decision by the platforms regarding the illegality of user-generated content, they will have the possibility to seek judicial redress. In addition to illegal online content being removed, where it is criminal, it should be followed up by law enforcement and the judiciary. Online platforms will be obliged to report serious crimes to the competent authority. The DSA would help to address

harmful content (which might not be illegal) and the spread of disinformation by including provisions on mandatory risk assessments, risk mitigation measures, independent audits, and the transparency of so-called "recommender systems" (algorithms that determine what users see).

- United Kingdom (UK) platform regulation in terms of the Online Safety Bill (OSB)

Platforms will be required to (see https://www.gov.uk/government/publications/online-safety-bill-supporting-documents/online-safety-bill-factsheet):

- remove illegal content,
- remove material that violates their terms and conditions, and
- give users controls to help them avoid seeing certain types of content to be specified by the bill.

The OSB defines which conduct constitutes as illegal conduct and provides a general monitoring requirement for social media platforms. The main aim of the OSB is to remove illegal content and to protect children online. The social media platforms will continue mitigating against disinformation which does not fall under the concept of "illegal".

- United States reform of section 230 of the Communications Decency Act of 1996

Section 230 provides that an "interactive computer service" cannot be treated as the publisher or speaker of third-party content. This protects websites from lawsuits if a user posts illegal information, although there are exceptions for copyright violations, sex work-related material, and violations of federal criminal law. The spread of misinformation and disinformation through Internet services and denials of access to some individuals and materials have raised questions about the scope of Section 230 protections and many want reform of section 230 (Kozinets and Pheiffer, 2023; Hetler, 2022; Jurecic, 2022).

On October 3, 2022, the Supreme Court granted the petition for *Gonzalez v. Google*, an appeal from a decision of the United States Court of Appeals for the Ninth Circuit. The justices will weigh cases that could test legal protections for social media companies concerning posts on their platforms (Adler and Frier, 2022). The first case involves Google's alleged responsibility for terrorist propaganda on its subsidiary, YouTube. The plaintiffs argue that YouTube allegedly recommended ISIS content to users and therefore should not be protected by section 230 for its own algorithmic recommendation.

5.4 Private companies or organisations

Companies and organisations should anticipate that they may be victims of disinformation campaigns and show people how to verify content about their business or organisation (Lapping 2021). For example, a disinformation campaign specifically targeting the Black Lives Matter (BLM) organisation resulted in BLM calling on the public to assist with disinformation against the movement (see https://blacklivesmatter.com/help-us-fight-disinformation/). The disinformation may be a consequence of the 2020 BLM protests that erupted globally after the death of George Floyd on May 25, 2020 under the knee of a Minneapolis police officer, Derek Chauvin. The protests were also characterised by a great deal of mis- and disinformation which caused confusion, stoked tension, and distracted from factually correct information (**Georgacopoulos and Poche, 2020).**

5.5 Cybersecurity resilience

Cybersecurity is the ability to withstand, respond to, and recover from all types of information and communication technology (ICT) related disruptions and threats which include the threat of disinformation. A government has to identify the threat and risk that disinformation presents and determine how it will respond. For example, the manner in which the South African government and law enforcement responded to the 2021 riots was inadequate. There should have been up-to-date communication on a consistent basis during the riots.

6. Conclusion

There is a fine line between balancing social media's potential use for good and bad. Unfortunately, the characteristics of social media assist with the creation and spread of disinformation.

Social media disinformation forms part of the cybersecurity threat and risk landscape on national level. The take-away from the discussion is that disinformation will not abate and is exacerbated by certain events. Disinformation requires pro-active measures to mitigate against it.

Governments, social media platforms, organisations, and social media users are in the early stages of responding to disinformation. They should consider the possible outcomes of their response carefully. The response to disinformation can easily result in the violation of free speech contrary to the essence of social media, such as freedom of expression and access to information.

References

Adler, M. and Frier, S. (2022), "Twitter, Musk and Why Online Speech Gets Moderated", [online], https://www.washingtonpost.com/business/twitter-musk-and-why-online-speech-gets-moderated/2022/10/03/0cb0ae68-434f-11ed-be17-89cbe6b8c0a5_story.html.

Alemany, J. *et al*; (2022) "Jan. 6 insurrection: Washington Post investigation of the cost, causes and aftermath" [online], https://www.washingtonpost.com/politics/interactive/2021/jan-6-insurrection-capitol/.

Baptista, J.P. and Gradim, A. (2022) "A working definition for fake news", [online], https://www.mdpi.com/2673-8392/2/1/43.

Bentzen, N. (2022) "Understanding propaganda and disinformation", [online], https://epthinktank.eu/2015/11/17/understanding-propaganda-and-disinformation/.

Bloomberg News, (2022) " Russia criminalises sanctions calls and fake news on military", [online], https://www.bloomberg.com/news/articles/2022-03-04/russia-to-punish-sanctions-appeals-and-fake-news-on-military?leadSource=uverify%20wall.

Colomina, CL, Margalef, HS. and Youngs, R. (2021) "The impact of disinformation on democratic processes and human rights in the world", [online], https://www.europarl.europa.eu/RegData/etudes/STUD/2021/653635/EXPO_STU(2021)653635_EN.pdf.

Dwoskin, E. (2023) "Come to the 'war cry party': How social media helped drive mayhem in Brazil" **[online],** https://www.washingtonpost.com/technology/2023/01/08/brazil-bolsanaro-twitter-facebook/.

Elton, W. (2020) "Is it illegal to post fake news?", [online], https://zegal.com/en-nz/blog/post/is-it-illegal-to-post-false-information/.

Georgacopoulos, C. and Poche, T. (2020) "Fake news, disinformation and the George Floyd protests", [online], https://faculty.lsu.edu/fakenews/about/protestfakenews.php.

Hetler, A. (2022) "10 ways to spot disinformation on social media", [online], https://www.techtarget.com/whatis/feature/10-ways-to-spot-disinformation-on-social-media.

Hook, K. and Verdeja, E. (2022) "Social Media Misinformation and the Prevention of Political Instability and Mass Atrocities", [online], https://www.stimson.org/2022/social-media-misinformation-and-the-prevention-of-political-instability-and-mass-atrocities/.

Hutton, C. (2023) "China implements new rules to regulate 'deepfakes' and AI art", [online], https://www.washingtonexaminer.com/policy/technology/china-implements-new-rules-over-deepfakes-chatgpt-ai-art.

Jurecic, Q. (2022) "The politics of Section 230 reform: Learning from FOSTA's mistakes", [online], https://www.brookings.edu/research/the-politics-of-section-230-reform-learning-from-fostas-mistakes/.

Kozinets, R. and Pheiffer, J.(2023) "Beyond Section 230: A Pair of Social Media Experts Describes How to Bring Transparency and Accountability to the Industry", [online] https://www.nextgov.com/ideas/2023/01/beyond-section-230-pair-social-media-experts-describes-how-bring-transparency-and-accountability-industry/381532/.

Lapping, G. (2021) "The July riots: an inflection point for digital media", [online], https://www.businesslive.co.za/redzone/news-insights/2021-09-01-the-july-riots-an-inflection-point-for-digital-media/.

Lasiello, E. (2021) "What is the Role of Cyber Operations in Information Warfare?", [online], https://digitalcommons.usf.edu/cgi/viewcontent.cgi?article=1931&context=jss.

Malinga, S. (2021) "Social media used as a dangerous tool to mobilise ongoing attack", [online], https://www.itweb.co.za/content/DZQ58MVPpK3vzXy2.

Nxumalo, L. (2022) "July unrest: social media fuelled unrest" [online], https://www.iol.co.za/sunday-tribune/news/july-unrest-social-media-fuelled-unrest-481bbafe-1ce1-4ca7-87c9-7da6cfc652eb.

Pijoos, I. (2022) "Harmful disinformation sowed anger, violence, and anarchy during July unrest - media expert ", [online], https://www.news24.com/news24/southafrica/news/harmful-disinformation-sowed-anger-violence-and-anarchy-during-july-unrest-media-expert-20220304.

Romano, A. (2021) "Kicking people off social media isn't about free speech", [online], https://www.vox.com/culture/22230847/deplatforming-free-speech-controversy-trump.

Rubin, O, Mallin, A. and Steakin, W. (2022) "By the numbers: How the Jan. 6 investigation is shaping up 1 year later", [online], https://abcnews.go.com/US/numbers-jan-investigation-shaping-year/story?id=82057743.

Shan, L.Y. (2022) "Turkey has long been hell for journalists': Reporters slam country's new 'fake news' law", [online], https://www.cnbc.com/2022/10/21/turkish-reporters-slam-countrys-new-fake-news-law.html.

Watney, M.M. (2018) "The Legal Position of Social Media Intermediaries in Addressing Fake News" **European Conference on Cyber Warfare and Security; Reading**, [online], https://books.google.co.za/books?id=-kFmDwAAQBAJ&pg=PA497&lpg=PA497&dq=murdoch+watney+fake+news&source=bl&ots=ZERXYgrf_n&sig=ACfU3U3XdbTGA27HmOT6PLnpmg_Om7W4HA&hl=en&sa=X&ved=2ahUKEwiExJWR_q38AhXOesAKHSERA68Q6AF6BAgkEAM#v=onepage&q=murdoch%20watney%20fake%20news&f=false.