



**UNITED NATIONS HUMAN RIGHTS
COUNCIL**

**AGENDA- Addressing human rights
challenges arising from misinformation,
disinformation and deepfake
technologies.**

DISCLAIMER: The authors of this document make no claim to copyright ownership. All information contained herein is utilized solely for educational purposes in accordance with fair use principles. The content does not represent the individual ideologies of the Executive Board. This background guide has been developed to offer delegates structured guidance and support their research within the committee framework.

LETTER FROM THE EXECUTIVE BOARD:

Delegates, it is with utmost pleasure that we welcome you all to the Conference. We are eager to see you all for three days of intense debate, deliberation, and discussions in this simulation of the UNHRC at SAI INTERNATIONAL MODEL UNITED NATIONS. We hope your preparations are going smoothly and are sure that the efforts put in by the management, secretariat, organising committee, and all the Executive Board members will culminate in an unforgettable conference for all of you. This background has been written to provide insight into the committee, its agenda, and its relevance to the modern scenario. It also consists of all the information necessary for it to act as a foundation for all your further research. Please note that this background guide is neither the beginning nor the end of your research. It should simply act as a catalyst for you to go above and beyond. It is the first step towards understanding the agenda. Remember, understanding a problem is the first step towards solving it. From a first-timer's perspective, we understand that an MUN conference can be overwhelming. That being said, please do not hesitate to contact any of your Executive Board members during the duration of your research for any sort of clarification. New ideas, by their very nature, are disruptive. Finally, we cannot wait to see all the new ideas and solutions all of you are going to come up with and all the very best.

Best regards
Executive Board Members
Niharika Jaiswal and Shlok Samant

Note: We strongly encourage you to conduct a PESTAL analysis of your assigned portfolio, covering the following areas:

1. Political Landscape: Understanding the domestic and international policies shaping education.
2. Economic Factors: Evaluating the financial infrastructure, investments, and funding models supporting development
3. Social Demographics: Examining cultural, ethnic, and social factors that influence public participation.
4. Technological Development: Investigating advancements in science, facilities, and training techniques.
5. Analysis of Current Events: Staying up to date with the latest trends and data relevant to the agenda.
6. Legal Framework: Understanding the legal environment governing regulations and policies at both national and international levels.

Basic Preparation and Research:

1. The most basic thing to do in a Model UN is to research, speak and rebut.
2. Do a basic PESTAL analysis of your country, which is connected with the agenda.
3. Prepare around 2 to 3 speeches for the General Speakers' List (GSL). Such speeches are generally 60 - 90 seconds long. Their length should never exceed more than 200 words.
4. Always give your remaining time to questions while ensuring that you try to get more speaking time from others through something called "yields".
5. Prepare a list of at least 3 subtopics. Have a 60-second speech ready for each mentally.
6. You must listen to the speeches of other delegates. This helps you find questions and rebuttal points for future speeches. More than research, questions involve logic.
7. Have some solutions ready beforehand in an "Actionable clause" format. Solutions in a MUN follow a very legal and diplomatic language, which can be noticed in the past UN documents and resolutions. Refer to the following site - bestdelegate.com

Techniques to Write Good Speeches:

1. HPA Method - Hook, Point, Action

How to Write an GSL Speech -Visit Best Delegate Model United Nations site.

2. LEET Method - Label, Explain, Example, Tieback

3. FLL Approach - Fact, Legality and Logic

This is a very technical and substantive approach to certain subtopics. You need to analyse the facts of the situation along with the legal realities surrounding it. Then one needs to connect both the facts and legalities through logic to weave a story and frame a narrative that they wish to portray to the public/audience.

4. PEN Method - Punchline, Example, Nail in the Coffin

This is a more aggressive style of debate, which is relevant for defensive arguments or rebuttals to allegations. One can start with a “punchline”, which is generally the main observation or summary of the speech delivered in a humorous, satirical, aggressive or fairly dramatic fashion. One needs to back such an observation with substantive arguments, examples or legal references. Then, similar to a Tieback, you need to put the nail in the coffin with a concluding statement to remind the audience why you believe your particular interpretation of the subject matter is the right one.

5. IRAC Method - Issue, Rule, Application and Conclusion

This method is usually used in a lot of MUNs. The methodology of IRAC is something which covers the aspects of the agenda.

Issue - Issue here means that you are bringing up the problem of the agenda to the purview.

Rule - Now that you’ve brought the issue to the purview. Defend this issue with relevant UN international laws, treaties, declarations and conventions.

Application - After the Rule, talk about your country’s stance on the agenda, and if required, use the laws established by your country’s government in relevance to the agenda and committee.

Conclusion - As the name suggests, the conclusion is to conclude what has been stated so far

“A good speaker is not one who only knows how to start it well, but also to end it well.”

Important Points to Remember:

A few aspects that delegates should keep in mind while preparing: 1. Procedure: The purpose of putting in procedural rules in any committee is to ensure a more organised and efficient debate. The committee will follow the UNA-USA Rules of Procedure. Although the Executive Board shall be fairly strict with the Rules of Procedure, the discussion of the agenda will be the main priority. So, delegates are advised not to restrict their statements due to hesitation regarding procedure.

2. Foreign Policy: Following the foreign policy of one’s country is the most important aspect of a Model UN Conference. This is what essentially differentiates a Model UN from other debating formats. To violate one’s foreign policy without adequate reason is one of the worst mistakes a delegate can make.

3. Role of the Executive Board: The Executive Board is appointed to facilitate debate. The committee shall decide the direction and flow of debate. The delegates are the ones who constitute the committee and hence must be uninhibited while presenting their opinions/stances on any issue. However, the Executive Board may put forward questions and/or ask for clarifications at all points in time to further debate and test participants.

4. Nature of Source/Evidence: This Background Guide is meant solely for research purposes and must not be cited as evidence to substantiate statements made during the conference. Evidence or proof for substantiating statements made during formal debate is acceptable from the following sources:

a. United Nations: Documents and findings by the United Nations or any related UN body are held as credible proof to support a claim or argument. Multilateral Organisations: Documents from international organisations like UNHRC , OIC, NATO, SAARC, BRICS,

EU, ASEAN, the International Court of Justice, etc., may also be presented as credible sources of information.

b. Government Reports: These reports can be used in a similar way as the State Operated News Agencies reports and can, in all circumstances, be denied by another country.

SOURCES - Legitimate Sources of Information In order to ensure the authenticity and credibility of the research, the delegates should depend on the following sources. Evidence or proof is acceptable from sources:

Primary Sources:

- Reuters
- State-operated News Agencies – These reports can be used in the support of or against the state that owns the agency.
- United Nations Reports and Resolutions (UNHRC, ECOSOC, SDG Progress Reports)
- National Government Reports and Ministry Publications
- World Bank, UNDP, and World Economic Forums.

Secondary Sources:

- Academic Journals, Research Papers, and University Publications
- Peer-reviewed Think Tank Publications (Brookings, Chatham House, etc.) Other websites and news agencies can be a source of initial information but will be excluded as evidence for the debate and your working papers/resolutions.

***Please Note: Reports from NGOs working with UNESCO, UNICEF and other UN bodies will be accepted. Under no circumstances will sources like Wikipedia or newspapers like the Guardian, Times of India, etc. be accepted. However, notwithstanding the criteria for acceptance of sources and evidence, delegates are still free to quote/cite from any source as they deem fit as a part of their statements.**

Hierarchy of evidence. Evidence can be presented from a wide variety of sources, but not all sources are treated as equal. Here is the hierarchy in which evidence is categorised:

Tier 1: Includes any publication, statement, resolution, or document released by any of the Nations' official organs or committees; any publication, statement, or document released by a UN member state in its own capacity. The evidence falling in this tier is considered most reliable during the simulation.

Tier 2: Includes: any news article published by any official media source that is owned and controlled by a UN member state. E.g.: Xinhua News (China), Prasar Bharti (India), BBC (United Kingdom), etcetera. The evidence falling in this tier is considered sufficiently reliable in case no other evidence from any Tier 1 source is available on that particular fact, event, or situation.

Tier 3: Includes: any publication from news sources of international repute such as Reuters, The New York Times, Agence-France Presse, etc. The evidence falling under this tier is

considered the least reliable for this simulation. Yet, if no better source is available in a certain scenario, it may be considered. Foreign Policy and Foreign Relations Foreign policy, in simple terms, is what your country aims to achieve regarding the issue at hand or in general with its relations with other countries.

1. What role must foreign policy play in your research?

Understanding the foreign policy of your country must be a checkbox that you tick off at the very beginning of your research. Your foreign policy should dictate everything from the arguments you make, the reasoning you give for making those arguments, and the actions you take in the Council.

2. Where do I look to find foreign policy?

Most of the time, foreign policy is not explicitly stated. It must be inferred from the actions and statements issued by the country. Reading the meeting records from previous meetings of the ICW(or any other UN body where your country might have spoken on the issue) is a great place to start. If such records are unavailable, look for statements from your country’s Foreign Ministry (or equivalent like Ministry of External Affairs, Ministry for Foreign Affairs, etc.) and top leadership (PM, Pres., Secretary of State, Defence Minister). Foreign Relations, on the other hand, refers to the diplomatic ties that one country has with another and considers elements such as the mutual presence of embassies, consulates, ambassadors & diplomatic dialogue. More often than not foreign policy is what will be your primary concern during your MUN, but it is also important to consider any extremities in your allotted country's foreign relations.

DOCUMENTATION:

NOTE- Before DR, delegates have to prepare a working paper as well which is a rough work for DR. It doesn’t have any ideal format.

UNHRC has documentation called draft resolution. Here comes the format:

DRAFT RESOLUTION FORMAT		
ELEMENT	FORMAT	EXAMPLE
TITLE	Creative name or acronym	SAIMUN/RES/01
COMMITTEE	Full committee name	United Nations Human Rights Council
SPONSORS	Main authors (max 5% of the committee)	Germany, France, Japan
SIGNATORIES	Want it debated (min 20%)	Canada, Australia, Kenya
PREAMBULATORY	<i>Italicized</i> phrase, ends	<i>Recognizing.....</i>

DRAFT RESOLUTION FORMAT

ELEMENT	FORMAT	EXAMPLE
CLAUSES	with comma, preexisted frameworks	
OPERATIVE CLAUSES	Solutions, Bolded phrase, numbered, ends with semicolon	1. Calls Upon.....

For further information about draft resolution, kindly go through these sites-

<https://www.un.org/en/model-United-nations/drafting-resolutions>

<https://bestdelegate.com/model-un-made-easy-how-to-write-a-resolution/>

<https://www.un.org/unispal/document/right-of-palestinian-people-to-self-determination-unhrc-34th-session-draft-resolution/>

TIPS:

Learn about your Country from the official *CIA World Factbook* website of your State's permanent mission to the United Nations, and from other government websites (for example the website and social media profiles of the Royal Family or President, Prime Minister, Ministry of Foreign Affairs...). Make sure to know as much as possible about the foreign relations of your country! Consider that at the United Nations, Countries generally lean towards their own geopolitical blocks: Western, Former Soviet, African, Middle Eastern, Asian, Latin American... Consider also organisations such as the European Union (EU), African Union, Arab League, Association of Southeast Asian Nations (ASEAN), The Union of South American Nations (USAN).

ABOUT UNHRC

Introduction to the Committee

The United Nations Human Rights Council is an intergovernmental body within the United Nations established in 2006 by the General Assembly, that is responsible for the promotion and protection of human rights across the globe and for addressing situations of human rights violations and making recommendations on them. The council is composed of 47 member states. History of the UNHRC The UNHRC was established in 1946 by ECOSOC and was one of the first two Fundamental roles and responsibilities of the council. Adopting Resolutions: The HRC adopts resolutions during regular sessions that express the will of the international community on given human rights issues or situations. Adopting a resolution sends a strong political signal which can prompt governments to take action to remedy those situations.

Special Sessions: The HRC holds crisis meetings known as special sessions to respond to urgent human rights situations. Universal Periodic Review (UPR): The Universal Periodic Review is a unique mechanism of the HRC that calls for each member of the United Nations to undergo a peer review of its human rights records every 4.5 years. The UPR provides each state the opportunity to regularly report the actions it has taken to improve human rights situations in their countries and to overcome challenges.

Special Procedures: The special procedures of the HRC are independent human rights experts who are mandated to report and advise on human rights from a thematic or country-specific perspective. They are non-paid and elected for 3-year mandates that are reconducted repeatedly.

Inquiries, debates, and discussions: Authorises commissions of inquiry and fact-finding missions that produce hard-hitting evidence on war crimes against humanity. The HRC also conducts debates where civil, social, economic, and cultural rights issues are being addressed.

Mandate of UNHRC

The UNHRC is mandated by the United Nations General Assembly to promote and protect the enjoyment and full realization, by all people, of all human rights. These rights are established by bodies such as The Charter of the United Nations, The Universal Declaration of Human Rights, and International Human Rights Laws and Treaties. The council also manages the Universal Periodic Review (UPR), a process through which each UN member state's overall human rights record is reviewed. It was established as a peer review mechanism for the assessment and advancement of human rights in all the UN member states. UPR reviews are formally conducted by the UPR Working Group of the council which comprises all 47 Council Member States. The UPR process is composed of:

1. Preparation and submission of the written reports by the State, civil society, and the Office of the High Commissioner for Human Rights which provides a foundation for the peer review.
2. During the UPR Session, the reports are presented orally, followed by an interactive dialogue where UN Member States make comments and recommendations to the State under review.
3. The Troika (group of three states) prepares a summary report of the comments and recommendations; the State may indicate its acceptance or rejection of the Recommendations.

Working of the UNHRC

The council serves as a forum for dialogue among states, with input from other stakeholders. As a result of its discussions, the council may issue resolutions calling on states to take specific actions or uphold certain principles, or it may create mechanisms to investigate or monitor questions of concern. The Human Rights Council has created or renewed the mandates of various “Special Procedures”. The special procedures are experts appointed to monitor human rights around priority themes or in specific countries with serious countries with serious human rights problems. The special procedures may be individual experts or working groups. Special Procedures have either thematic or country-specific mandates. As of October 2022, The HRC oversees about 45 thematic special procedure mandates and 13 country-specific special procedures mandates. Special procedures mandate holders to serve in their personal capacities, meaning they do not represent their countries of citizenship. Each mandate holder may serve for a maximum of six years. This independent status is intended to allow these experts to carry out their functions unbiased.

INTRODUCTION TO THE AGENDA:

The proliferation of Artificial Intelligence (AI) and generative synthetic media represents one of the most critical challenges facing international human rights governance. While digital technologies hold immense promise for human progress, the current trajectory of data-intensive AI applications and different media platforms poses severe, multifaceted risks to fundamental rights, including the bedrock principles of human dignity, autonomy, and the protection of truth in the public sphere (OHCHR, 2023). This report analyzes the mandate of the United Nations Human Rights Council (UNHRC) in addressing the specific threats posed by deepfakes and digital misinformation, focusing on the resulting human rights abuses and the urgent need for a comprehensive global response. There are in some cases correlation between the presence of widespread disinformation, misinformation and hate speech and resulting protection risks such as stigmatisation and dehumanisation, polarization, heightened insecurity, normalisation of violence, targeting of minority or marginalized groups, and silencing of dissenting voices (e.g. human rights activists, journalists and local media, and community leaders).

For example, a campaign of hate speech and misinformation on Facebook against the Rohingya Muslim minority in Myanmar was linked to the commission of grave human rights violations that prompted the flight of some 750,000 Rohingya into Bangladesh.

Risks related to the information space are far from new phenomena - lies and deceptions have always circulated. However, the advent and expansion of digital platforms have played a significant part in enabling the scale of their reach. Furthermore, a recent survey – the largest ever conducted capturing sentiments on attitudes towards refugees – has shown that social media is playing a role in polarizing societies.

Addressing risks to information integrity increasingly need to be considered as part of holistic protection approaches, including situation analysis, strategy development and responses. Analysis of such risks can inform protection analysis (e.g. where sources of the risks are identifiable) and increase the impact of UNHCR's Digital Transformation (e.g. the provision of online services and digital APP).

Beyond risks to the forcibly displaced and stateless, such risks can constrain the operating space for UNHCR through damaging the organisation's reputation and possibly leading to communication crisis, weaken trust and support amongst communities, and threaten staff safety and security through targeting of staff members due to their role or personal identity characteristics. In recognition, the United Nations Global Principles for Information Integrity were developed to promote the integrity of the information ecosystem – “where freedom of expression is fully enjoyed and where accurate, reliable information, free from discrimination and hate, is available to all in an open, inclusive, safe and secure information environment”.

Ensuring information integrity is crucial within a pluralistic information landscape that advocates for peaceful societies, human rights, and the 2030 Agenda for Sustainable Development. It embodies the potential of the digital age to cultivate trust, knowledge, and individual autonomy. Promoting information integrity includes the duty of states to ensure an enabling space for individuals to exercise their right to seek, receive, and share diverse information and ideas, while safeguarding their right to hold opinions without interference. In our intricate digital information environment, this includes enabling people to navigate information spaces securely, preserving the right privacy and the right to freedom of expression.

Digital platforms play a pivotal role in reshaping social, cultural, and political interactions worldwide. They serve as conduits connecting engaged global citizens around critical issues. For UNHCR, these platforms facilitate direct communication with communities and the provision of protection information and services, and support our pursuit of protecting refugees, displaced and stateless people through advocacy and resource mobilisation. In times of crisis and adversity, they offer hope, amplify marginalized voices, and invigorate global movements.

Yet these same platforms have also exposed and contributed to a darker side of the digital ecosystem. While they connect people globally, they've also facilitated the rapid dissemination of falsehoods and hate, resulting in both online and offline consequences for forcibly displaced and stateless people.

The danger cannot be overstated. Social media enabled hate speech and disinformation can lead to offline acts of violence and can in some cases result in injuries or deaths. The ability to disseminate large-scale disinformation to undermine scientifically established facts and stoke xenophobia and racism not only pose a risk to forcibly displaced and stateless people but threaten fundamental human rights. However, for the forcibly displaced and stateless these risks are even more acute as they are particularly exposed to the risks of racism. These dangers have further intensified in part due to rapid technological advancements, such as generative artificial intelligence which can be used to quickly and cheaply generate realistic imagery and audio. It has become clear that business as usual is not an option.

KEY DEFINITIONS OF THE TERMS-

Information integrity refers to the accuracy, consistency and reliability of information. It is threatened by disinformation, misinformation and hate speech. While there are no universally accepted definitions of these terms, United Nations entities have developed working definitions.

Disinformation is information that is not only inaccurate, but is also intended to deceive and is spread in order to inflict harm. Disinformation can be spread by State or non-State actors in multiple contexts, including during armed conflict, and can affect all areas of development, from peace and security to human rights, public health, humanitarian aid and climate action.

Misinformation refers to the unintentional spread of inaccurate information shared in good faith by those unaware that they are passing on falsehoods. Misinformation can be rooted in disinformation as deliberate lies and misleading narratives are weaponized over time, fed into the public discourse and passed on unwittingly. In practice, the distinction between mis- and disinformation can be difficult to determine.

Hate speech, according to the working definition in the United Nations Strategy and Plan of Action on Hate Speech, is “any kind of communication in speech, writing or behaviour, that attacks or uses pejorative or discriminatory language with reference to a person or a group on the basis of who they are, in other words, based on their religion, ethnicity, nationality, race, colour, descent, gender or other identity factor”.

Case Studies-

Myanmar: Anti-Rohingya Disinformation Campaign and Ethnic Cleansing (2016–2018)

The situation in Myanmar remains one of the most extensively documented examples of disinformation contributing to large-scale human rights violations. Between 2016 and 2018, military-linked actors, ultranationalist groups, and coordinated online networks used social media platforms particularly Facebook to disseminate false narratives portraying the Rohingya Muslim minority as terrorists, illegal immigrants, and existential threats to Myanmar's national identity.

The Rohingya, a predominantly Muslim ethnic minority residing in Rakhine State, had already faced decades of discrimination and statelessness. However, the rapid expansion of social media dramatically amplified anti-Rohingya propaganda. False stories alleging Rohingya attacks on Buddhist communities circulated widely online. Fabricated reports, manipulated images, and inflammatory content encouraged hostility against the minority population and contributed to an environment where violence became normalized.

The UN Independent International Fact-Finding Mission on Myanmar concluded that social media had become a significant tool for spreading hate speech and incitement. UN investigators specifically observed that Facebook had played a “determining role” in the dissemination of anti-Rohingya narratives and the escalation of ethnic tensions. Investigations later revealed that military-linked networks operated coordinated propaganda campaigns through dozens of accounts and pages reaching millions of users.

The consequences were catastrophic. In August 2017, Myanmar's military launched so-called “clearance operations” against Rohingya communities. Reports documented mass killings, sexual violence, village burnings, and forced displacement. More than 700,000 Rohingya fled to neighboring Bangladesh. The UN Fact-Finding Mission subsequently concluded that there were reasonable grounds to investigate senior military officials for genocide, crimes against humanity, and war crimes.

From a human rights perspective, the campaign implicated multiple internationally protected rights. Article 3 of the Universal Declaration of Human Rights (UDHR), guaranteeing the right to life and security of person, was gravely violated through widespread killings and attacks. Article 2, prohibiting discrimination, was undermined through targeted hate propaganda. The persecution of Rohingya communities also engaged protections contained within the International Covenant on Civil and Political Rights (ICCPR), particularly concerning equality and minority rights.

Resource:

1. <https://www.ohchr.org/en/hr-bodies/hrc/myanmar-ffm>
2. UN Independent International Fact-Finding Mission on Myanmar (A/HRC/39/64)

Ethiopia: Online Hate Speech and the Tigray Conflict (2020-2022)

The Tigray conflict in Ethiopia provides another major example of how misinformation and disinformation can intensify armed conflict and contribute to human rights abuses. Beginning in November 2020, conflict erupted between the Ethiopian federal government, allied forces, and the Tigray People's Liberation Front (TPLF). As military operations expanded, digital platforms became central arenas for information warfare.

Throughout the conflict, social media platforms hosted large volumes of inflammatory rhetoric, manipulated content, false accusations, and ethnically targeted propaganda. Various actors circulated narratives portraying ethnic communities as enemies, traitors, or security threats. Fabricated allegations of atrocities were frequently disseminated without verification, while inflammatory posts encouraged hostility toward specific ethnic groups.

The conflict occurred within Ethiopia's highly polarized political environment, where ethnicity often serves as a primary marker of political identity. Consequently, misinformation frequently reinforced pre-existing tensions. Human rights observers documented instances where online content dehumanized entire populations and encouraged retaliatory violence.

The Joint Investigation conducted by the Office of the United Nations High Commissioner for Human Rights (OHCHR) and the Ethiopian Human Rights Commission identified widespread violations committed by multiple parties to the conflict. Simultaneously, UN experts repeatedly warned that hate speech and digital incitement were increasing risks to civilian populations.

Resource:

1. <https://www.ohchr.org/en/press-releases/2024/06/ethiopia-un-human-rights-chief-calls-sustained-efforts-halt-violations-and>
2. <https://www.ohchr.org/en/statements-and-speeches/2024/03/turks-global-update-human-rights-council>
3. <https://www.ohchr.org/en/meeting-summaries/2026/01/experts-committee-rights-child-commend-ethiopia-law-providing-free>
4. <https://www.ohchr.org/en/meeting-summaries/2024/06/independent-expert-notes-concerning-trend-ever-expanding-restrictions>

Russia–Ukraine War: Disinformation, Information Warfare, and Human Rights Violations (2022–Present)

The Russia–Ukraine War is one of the most significant contemporary examples of how misinformation, disinformation, and information warfare can contribute to serious human rights violations during armed conflict. Since the Russian Federation launched its full-scale invasion of Ukraine on 24 February 2022, information operations have accompanied military actions, shaping public perception, justifying military objectives, and influencing access to accurate information.

Prior to and during the invasion, Russian state media and affiliated networks circulated narratives alleging that Ukraine was controlled by “Nazis,” that Russian-speaking populations were facing persecution, and that military intervention was necessary for their

protection. These narratives were widely disseminated through television, social media platforms, and digital communication networks. The spread of such claims contributed to the creation of an information environment in which military actions could be portrayed as legitimate or defensive despite widespread international condemnation. The conflict also witnessed extensive use of propaganda, manipulated content, and coordinated disinformation campaigns across online platforms.

The human rights consequences of the conflict have been extensively documented by the Independent International Commission of Inquiry on Ukraine, established by the United Nations Human Rights Council. The Commission concluded that Russian authorities committed numerous violations of international human rights law and international humanitarian law, including war crimes. Investigations documented indiscriminate attacks against civilians, unlawful killings, torture, rape, sexual violence, arbitrary detention, deportation of children, and attacks on civilian infrastructure. The Commission found that Russian armed forces repeatedly used explosive weapons in populated civilian areas, demonstrating disregard for civilian protection obligations.

Particularly concerning has been the situation in occupied territories. United Nations reports documented extensive restrictions on freedom of expression, suppression of independent media, intimidation of journalists, and efforts to impose Russian-controlled information systems. Residents in occupied regions faced pressure to accept Russian citizenship, while expressions of Ukrainian identity were restricted. Human Rights Monitoring Mission reports also documented enforced disappearances, arbitrary detention, torture, and coercion directed against individuals perceived as opposing occupation authorities.

The Commission of Inquiry further concluded that torture was widespread and systematic in areas under Russian control. Investigators documented severe beatings, electric shocks, sexual violence, prolonged detention, and incommunicado confinement of civilians and prisoners of war. In 2024, the Commission stated that torture committed by Russian authorities amounted to crimes against humanity due to its widespread and systematic nature.

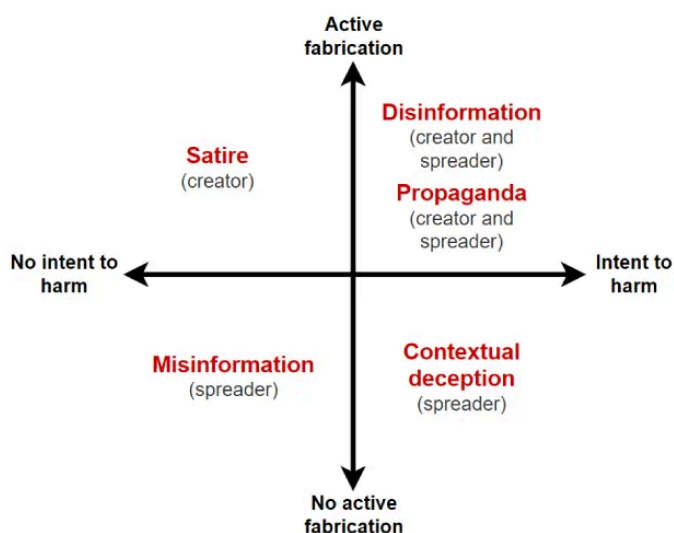
The conflict implicates numerous provisions of international human rights law, including Article 3 of the Universal Declaration of Human Rights (right to life), Article 5 (freedom from torture), Article 9 (freedom from arbitrary detention), and Article 19 (freedom of expression). It also raises concerns under the Geneva Conventions regarding civilian protection during armed conflict.

Resources:

1. UN Human Rights Council Resolution 49/1 (2022)
2. Report of the Independent International Commission of Inquiry on Ukraine (A/HRC/52/62)
3. Report of the Independent International Commission of Inquiry on Ukraine to the UN General Assembly (A/78/540)
4. <https://ukraine.un.org/en/203733-un-commission-has-found-array-war-crimes-violations-human-rights-and-international>

5. <https://ukraine.ohchr.org/en/Situation-of-human-rights-in-the-temporarily-occupied-territories-of-Ukraine-including-the-Autonomous-Republic-of-Crimea-and-the-city-of-Sevastopol-1-July-2024-to-30-June-2025>

Taxonomy of false and misleading content online



OECD Going Digital Toolkit Note, No 23, 2022

WHAT IS “DISINFODEMIC?”

Understanding the dynamics of mis- and disinformation is vital for the enablement of effective policy responses. Due to the complexity of the subject matter, the field is fraught with diverging definitions and conceptualizations. There are no internationally agreed-upon definitions of misinformation or disinformation. In wading through this potential confusion, it is important to consider the key dynamics of the spread of false and potentially harmful content online.

WHO defines an infodemic as “an overabundance of information – some accurate and some not – that occurs during an epidemic... [and] makes it hard for people to find trustworthy sources and reliable guidance when they need it” (WHO [2020](#)). Within the larger category of the infodemic, what, though, is a disinfodemic?

“Disinfodemic” is a term used by UNESCO to specify the potential harms caused not only by too much information in general, but by false and misleading information specifically (Posetti and Bontcheva [2020](#)). It focuses on the potential harmful consequences of mis- and disinformation, as well as the specific challenges associated with an information landscape polluted with false and misleading content.

Mis- and disinformation related to the COVID-19 pandemic was spread virally. According to data presented in UNESCO’s 2021/2022 edition of the *World Trends in Freedom of*

Expression and Media Development Report, over 1 million posts circulated on Twitter in September 2021 with inaccurate, unreliable, or misleading information related to the pandemic (UNESCO [2022](#)). From the start of the pandemic up until August 2021, Facebook reported that it had removed over 20 million posts on Facebook and Instagram on the grounds of promoting COVID-19-related misinformation (Rosen [2021](#)). Encrypted messaging apps also provided a platform for the rapid spread of false information that proved difficult to monitor and trace. The scale of the problem and its social consequences should not be underestimated.

The issue of false and misleading information is not only in its scale as measured by total amount of content, but also in the number of users it reaches online. In addition, the way that biased and selective sharing (fuelled by algorithms designed to maximize user engagement) of both true and false news skews overall constructions of reality for users (Roozenbeek and van der Linden [2019](#)). Combined with a lack of trust in quality information, and a lack of overall media and information literacy, this creates an information environment in which citizens are vulnerable to false and misleading content with potentially negative social consequences (Kim et al. [2020](#); Pennycook et al. [2021](#)).

Within the sea of content encountered online, and particularly during moments of great uncertainty, people are most likely to share false content that is novel, emotionally evocative, and confirms existing ideologies, biases, and attitudes (Guess et al. [2019](#); Rathje [2021](#)) – the type of content that social media algorithms have been designed to deliver to targeted individuals. As an often-cited study by researchers at the Massachusetts Institute of Technology, false news travels faster than true stories (Soroush et al. [2018](#)). When citizens do not trust, cannot access, or do not have the necessary media and information literacy skills to analyze verified information provided by media outlets, they are especially susceptible to mis- and disinformation.

COVID-19 showed us the potential varied social harms these dynamics can cause. While there is limited evidence to suggest widespread public belief in COVID-related misinformation, those susceptible to mis- and disinformation (e.g., due to low levels of trust in traditional media outlets or science institutions) report lower levels of compliance with public health guidance, including accepting vaccination (Roozenbeek et al. [2020](#)).

The way we understand the drivers behind the disinfodemic matters. If we assume that the main problem behind mis- and disinformation is simply that there is too much of it circulating online, then the solution may seem to lie in simply reducing the amount of false and misleading content in circulation. However, this solution is far from straightforward as content often reappears through shares, screenshots, and other forms of dissemination that can be difficult to track. Attempts to reduce the amount of false information online have, as noted previously, often been implemented through content moderation by social media companies, penal sanctions by governments, and even large-scale legal bans. These have had complicated implications.

Shifting our attention to the supply side, trustworthy information, however, reveals multiple benefits of this approach. The role of the media is particularly important in crisis situations, when accurate information can help alleviate human suffering and save lives. The media can also reduce risks by contributing to preventing, mitigating, and preparing people for disasters. Among the many key roles in this context, the media can help rectify and prevent the spread of mis- and disinformation at a sensitive time, provide access to public information adapted to the needs of different groups across societies, monitor the respect for human rights during crises, and act as a watchdog of public entities in charge of crisis response and disaster risk

reduction. In short, the media empowers individuals, humanitarian workers, and decision-makers to take informed decisions, and during the COVID-19 pandemic, there was an increased demand for reliable information as readership and viewership of trusted news sources surged around the world (Pollack [2020](#); UNESCO [2022](#)).

At the same time, reporting on sensitive topics requires skills and knowledge. During the pandemic, journalists needed greater skills in areas such as crisis communications and fact-checking. Medical knowledge and skills in science journalism also proved essential. With this in mind, it is important to consider how efforts to stifle the flow of false and misleading content may create a backlash, preventing the very antidotes needed to counter them: true and reliable information (Bontcheva and Posetti [2020](#)). It is vital that any measures aimed at tackling mis- and disinformation do not hinder the production and accessibility of quality information that citizens can trust in order to navigate a crisis. This means protecting and respecting freedom of expression. The following section underlines the international human rights standards that outline when freedom of expression can and cannot be legitimately restricted.

THE ROLE OF TECHNOLOGY IN SPREADING MISINFORMATION, DISINFORMATION AND DEEPPAKES:

While governments around the world have failed to address the challenges of disinformation and misinformation in a way that respects and protects the right to freedom of expression, social media companies have also failed to uphold their responsibility to respect human rights while countering misinformation. There is no doubt that the exponential spread of misinformation and disinformation has been facilitated by new digital technologies and that all major search engines and social media platforms have played a role in the spread of erroneous and misleading information. Social media platforms are purposefully designed to captivate users' attention and maximise their engagement to a degree that can become highly addictive. This engagement, coupled with the very detailed information these companies are collecting on people from across the digital world, enables them to infer detailed pictures of peoples' lives and behaviours which can be monetized – primarily for highly targeted advertising. This algorithmically driven targeting, profiling, and personalised content play an enormous role in shaping the online experience of each user and in determining the information they see, which may often include high levels of misinformation. Since people are more likely to be drawn to novel information that confirms their biases, or sensationalist and incendiary information, algorithms that curate what users see on social media may end up promoting or amplifying false or misleading content.

By virtue of repetition of similar content, some of these platforms can end up persuading users that falsehoods are true. For example, YouTube's recommendation algorithm automatically promotes similar content in a way that can reinforce false information, conspiracy theories and rumours by joining together different videos that repeat similar false narratives. This creates the illusion that there are multiple sources for the same idea, persuading the user of the veracity of the information. In some cases, users can develop a sort of "tunnel vision" and lose their capacity for discerning between scientific facts, misleading information and outright lies. The combination of algorithmically-driven ad targeting and personalized content has allowed Google and Facebook's platforms to play an enormous role in shaping people's online experience and determining the information they see. This can influence and modify opinions and thoughts, which risks affecting peoples' ability to make autonomous choices and form opinions free from manipulation.

Moreover, the algorithms are designed to find the best ways to nudge people towards particular outcomes based on an individual's predicted personal characteristics and behaviors. These capabilities mean there is a high risk that the companies could be directly harming the rights to freedom of thought, conscience and religion and to freedom of opinion and expression through their use of algorithmic systems. As human rights lawyer Susie Alegre argues, "the scale and range of interferences with our ability to think and feel freely that technology can and might facilitate is in many ways beyond our imagination. But it is happening now. We have forgotten that rights need protections to be real and effective". Furthermore, the way in which these algorithms work risk providing a platform for other actors to access, utilize or weaponize their platforms to abuse the rights of other users.

UNESCO has also warned that intentionally misleading content spread on social media platforms can affect peoples' understanding of reality and undermine trust, informed dialogue, a shared sense of reality, mutual consent, and participation. In other words, in a world in which erroneous or misleading information is so prevalent, the right to be informed and form an opinion based on facts, as well as the ability to have a debate around those facts, becomes fundamentally undermined. Similarly, the Council of Europe's Committee of Ministers has also warned that "fine grained, sub-conscious and personalised levels of algorithmic persuasion may have significant effects on the cognitive autonomy of individuals and their right to form opinions and take independent decisions". As individuals are deprived of the ability to discern truth from fiction and to hold an informed opinion based on the best available information and a diversity of sources, and as they are nudged towards certain types of information based on harvesting of their personal data, their right to freedom of thought is deeply affected. As UNICEF has pointed out, children and young people are at heightened risk of being affected by online misinformation.

Social media companies bear a direct responsibility for how their platforms are making the public vulnerable to misinformation and thus posing a direct threat to their human rights, including the right to seek, receive and impart information. The UN Guiding Principles on Business and Human Rights state that the responsibility to respect rights extends to all companies, wherever they operate, regardless of their size, sector, operational context, ownership and structure. This means that social media companies must consider the human rights risks that arise from their operations and services and take concrete action when their activities might have a negative effect on the enjoyment of human rights. After repeated calls and increased pressure to address the problem of false and misleading information driven by social media and search engines, some tech companies started to take some measures to reduce the impact of misinformation and disinformation on their platforms.

HUMAN RIGHTS APPROACH ON TACKLING MISINFORMATION, DISINFORMATION AND DEEPPAKES:

False and misleading information cannot be easily censored or simply expunged, particularly in the age of social media and messaging apps. Restricting information and the free expression of opinions and ideas through censorship, punitive laws, internet shutdowns, closing down of media outlets, and persecution of journalists, human rights defenders and others expressing their views, are not only in contravention of international human rights law but are also ineffective measures that do not tackle the root causes of why the public remains vulnerable to misinformation. On the contrary, censorship and a blunt overregulation of the exercise of the right to freedom of expression increases mistrust in the authorities and drives more people to seek out "alternative", "anti-establishment" or covert sources of information

that may not have been put through scrutiny and cannot be debunked in the public arena. Restrictions on the right to freedom of expression that impose blanket prohibitions on the dissemination of information, including those based on vague and ambiguous concepts such as “false news” or “spreading misinformation”, are incompatible with international human rights law. As stated by the UN Human Rights Committee, international law does not permit general prohibitions of expressions of an erroneous opinion or an incorrect interpretation of events.

Legislation prohibiting and criminalizing “false news” also risks having a chilling effect on the general population and the media, leading to self-censorship out of fear of reprisals. As noted by the UN Special Rapporteur on the right to freedom of expression, such limitations often appear not to be imposed for the legitimate purpose of promoting accurate information but in order to suppress relevant information uncomfortable for the government or to use the situation as a pretext to crack down on opposition politicians, critical media outlets or human rights defenders.

Instead of punishment and censorship, human rights offer a different path which States should take in order to minimize the impact of false and misleading information. By upholding the right to freedom of expression, States would ensure that individuals and groups, including journalists and other civil society actors, can exchange information, air and debate diverse opinions, discuss effective ways of tackling social problems, hold governments accountable, defend human rights, debunk false information and challenge propaganda. Public officials play a particular role in stemming the impact of misinformation, and it is crucial that they do not make, sponsor, encourage or further disseminate statements which they know or reasonably should know to be false or misleading. As emphasised by regional and international experts on the right to freedom of expression, public officials should take care to ensure that they disseminate reliable and trustworthy information, including about matters of public interest.

International human rights mechanisms have further recommended that state authorities build up a reliable and prompt system of accurate information that leads to increased trust by the general public. In order for the public to have that trust, the public needs to have access to all relevant and available information. States are therefore required to step up their efforts to ensure that they disseminate reliable, accessible, evidence-based and trustworthy information, which is crucial to counter false and misleading information. States also have an obligation to ensure an enabling environment for freedom of expression, including by promoting a free, independent and diverse communications environment which is a key means of addressing misinformation and propaganda. In 2017 international and regional experts on freedom of expression laid down a series of obligations and general principles for States to follow in order to combat misinformation.

As a key step, the independent experts identified the obligation of States to create an enabling environment for freedom of expression and highlighted the importance of ensuring access to a wide variety of sources of information and ideas, including through “promoting, protecting and supporting diverse media”, which can facilitate public debates and the open confrontation of ideas, as well as acting as a watchdog of government and others in power. In addition, this would require States to ensure the presence of strong, independent and adequately resourced public service media operating under a clear mandate to serve the overall public interest and to set and maintain high standards of journalism. UN human rights experts further called on States to take measures to address misinformation in a way that complies with their human rights obligations, including by reaffirming their commitment to the freedom, diversity and independence of the media, and guaranteeing the safety of journalists. Moreover, experts

have urged States to invest in media, information and digital literacy to equip individuals with the critical thinking tools to distinguish between verifiable and unverifiable information, which they recommend should become part of the national school curriculum, and to ensure full, honest and evolving communication with the public.

KEY PILLARS OF GLOBAL GOVERNANCE

European Union: The EU leads with two major frameworks: The Digital Services Act (DSA) mandates strict content moderation and risk assessments for large platforms, while the EU Artificial Intelligence Act requires explicit watermarking and transparency labeling for all AI-generated or manipulated media.

United Nations & UNESCO: Focuses on information integrity and human rights. UNESCO has developed guidelines for digital platforms to combat algorithmic amplification of disinformation and protect freedom of expression while tackling the societal impacts of deepfakes.

Voluntary Tech Standards: Initiatives like the Coalition for Content Provenance and Authenticity (C2PA) establish technical standards using cryptographic metadata (provenance) so users and systems can verify the origin and authenticity of digital media.

Election & Media Integrity Laws: Driven by the proliferation of AI-generated political deepfakes, global frameworks increasingly target election interference, demanding that platforms enforce stricter takedown timelines and block non-consensual altered images (especially non-consensual intimate imagery)

RECOMMENDATIONS :

TO STATES:

RESTRICTIONS ON THE RIGHT TO FREEDOM OF EXPRESSION

- Ensure that restrictions on the right to freedom of expression are provided for by a law formulated with sufficient precision and are necessary and proportionate to the purpose of protecting one of the legitimate purposes under international human rights law.
- Refrain from imposing blanket bans on the dissemination of information, including those based on vague and ambiguous concepts such as “false news” or “spreading misinformation”.
- Repeal or amend legislation that imposes criminal penalties simply for sharing or distributing information.

ACCESS TO RELIABLE INFORMATION

- Step up efforts to disseminate reliable, accessible, evidence-based and trustworthy information, which is crucial to counter false and misleading information and to build trust among the population in the integrity and reliability of the journalistic environment.
- Ensure people can have access to a wide variety of sources of information and ideas, including from a diverse media and through reliable and unfettered access to the internet, to facilitate public debates and the open confrontation of ideas.

- Promote media, information and digital literacy to equip individuals with the critical thinking tools to distinguish between verifiable and unverifiable information.
- Take appropriate measures to ensure public officials refrain from making, sponsoring, encouraging or further disseminating statements which they know or reasonably should know to be false or misleading. Public officials should take care to ensure that they provide reliable and trustworthy information, including about matters of public interest.

PROTECTION OF JOURNALISTS

- Guarantee a safe and enabling environment in which journalists and other media workers can work independently in a diverse communications environment and without fear of reprisals.
- Enable and support journalists, media outlets, human rights defenders and civil society organizations to contribute in efforts to counter misinformation and disinformation.
- Provide effective protection to journalists, media outlets, human rights defenders and civil society organizations from threats and attacks to allow them to carry on their work without fear of punishment, reprisal or intimidation.

REGULATION OF THE ONLINE SPACE

- Ensure that companies do not abuse the right to freedom of expression while countering misinformation and disinformation, including online.
- Uphold the principle that intermediaries should not be required to substantively evaluate the legality of third-party content, in line with the Manila Principles on Intermediary Liability.
- Refrain from imposing duties on social media companies to proactively monitor online content or intermediary liability regimes that incentivize overbroad censorship.
- Legally require technology companies to carry out human rights due diligence to identify and address human rights impacts related to their global operations, including risks and abuses linked to their algorithmic systems or arising from their business model as a whole.
- Enact and enforce strong digital regulation, including a ban on surveillance advertising that relies on invasive tracking and the processing of personal data; independent oversight over the algorithmic recommendation systems used by online platforms and require these to be profiling-free by default; and measures to ensure people can practically choose rights-respecting alternatives to online platforms.

TO SOCIAL MEDIA COMPANIES:

- Uphold the corporation's human rights responsibilities by engaging in human rights due diligence, including to address risks arising from their business model, and taking concrete action to respond to the dissemination of false or misleading information.
- Ensure greater transparency regarding content moderation practices and policies, as well as the algorithmic systems underpinning their platforms. This must include enabling third parties to scrutinise and assess the functioning of the platforms and their underlying algorithmic systems.

QUESTIONS A RESOLUTION MUST ANSWER:

1. How will the international community distinctively define "misinformation," "disinformation," and "malicious deepfakes" without creating loopholes?
2. How can member states align domestic digital regulations with Article 19 of the ICCPR to ensure protection of freedom of expression?
3. Where should governments draw the line between harmful disinformation and protected speech, satire, or political dissent?
4. What legal or financial liabilities should social media platforms and tech conglomerates face for the algorithmic promotion of disinformation?
5. How can the UN enforce transparency from tech companies regarding their content moderation policies and data-gathering algorithms?
6. What frameworks can regulate the creators and distributors of generative AI tools without stifling technological innovation?
7. How can member states collaboratively fund and develop open-source, global deepfake detection software?
8. What international standards should be set for mandatory digital watermarking and provenance tracking of AI-generated content?
9. How can independent, multi-stakeholder fact-checking networks be protected and funded globally without government interference?
10. How will the resolution protect journalists, human rights defenders, and political dissidents who are disproportionately targeted by targeted deepfakes?
11. What mechanisms will mitigate the democratic harms of disinformation during national election cycles?
12. How can the international community prevent state-sponsored disinformation campaigns used to incite ethnic violence or discrimination?
13. What specific educational frameworks can the UN provide to member states to implement digital and media literacy programs in schools?
14. How can public awareness campaigns be structured to rebuild societal trust in credible journalism and democratic institutions?

FURTHER READINGS:

1. <https://www.unesco.org/en/articles/deepfakes-and-crisis-knowing>
2. https://pubs.aip.org/aip/acp/article-pdf/doi/10.1063/5.0234995/20199986/050016_1_5_0234995.pdf
3. <https://www.un.org/en/un-chronicle/safeguarding-human-rights-and-information-integrity-age-generative-ai>
4. [https://www.europarl.europa.eu/RegData/etudes/BRIE/2026/785000/IUST_BRI\(2026\)785000_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2026/785000/IUST_BRI(2026)785000_EN.pdf)
5. <https://pmc.ncbi.nlm.nih.gov/articles/PMC9910783/>
6. <https://www.lawjournals.org/assets/archives/2026/vol12issue2/12117.pdf>
7. <https://www.cambridge.org/core/elements/addressing-misinformation-and-disinformation/66BE72E9F1FC74DE2CD6286B8383C146>
8. <https://www.un.org/en/un-chronicle/safeguarding-human-rights-and-information-integrity-age-generative-ai>
9. <https://www.gov.uk/government/publications/deepfake-detection-technology/deepfake-detection-technology>
10. <https://www.lawjournals.org/assets/archives/2026/vol12issue2/12117.pdf>