Curbing Information Disorder and Protecting Global Democracy

Opportunities for Development Cooperation

Published by:

giz Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

Curbing Information Disorder and Protecting Global Democracy: Opportunities for Development Cooperation

March 2025

Author:
Marilia Maciel
Director of Digital Trade and Economic Security at DiploFoundation. She holds a PhD in Information and Communication Sciences from the University of Bordeaux Montaigne, France.

The publication "Curbing Information Disorder and Protecting Global Democracy: Opportunities for Development Cooperation" was commissioned by the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, Sector Programme Governance, on behalf of the German Federal Ministry for Economic Cooperation and Development (BMZ).

Acknowledgements: We would like to express our sincere gratitude to the Competence Center 4C30 and the Sector Programme Human Rights for their valuable contributions to this Policy Paper. Their insightful comments, critical feedback, and expertise have greatly enriched the analysis and helped shape the final outcome.

Table of Contents

Executive Summary 4

1. Information disorder: Defining the problem 5
   - 1.1 Unpacking information disorder 6
   - 1.2 The role of social media platforms in today's information ecosystem 8
   - 1.3 Drivers of disinformation 9
2. The interplay between information disorder and democratic principles 11
   - 2.1. Electoral processes 11
   - 2.2. Civil liberties and political rights 13
   - 2.3. The functioning of the government 14
   - 2.4. Civic and political participation 15
   - 2.5. Democratic political culture 15
3. Combating disinformation: Strategic approaches 18
   - 3.1. Promoting individual and societal resilience by strengthening agency 18
   - 3.2. Protecting society by reducing disinformation exposure 20
   - 3.3. Boosting resilience and protection through a multidimensional approach to regulation 23
4. Political and operational recommendations: Opportunities for development cooperation 25

Annex I: Approaches to protect components of democracy 30

References 33

Imprint 35

Executive Summary

With the rise of the internet, any actor with a connected device is able to produce and distribute content. This has empowered individuals, yet it has also led to a surge in information overload and pollution. Disinformation is part of a broader phenomenon of 'information disorder'. It poses a significant threat to global democracy by eroding trust in institutions, diminishing political engagement, exacerbating societal polarization, and disturbing electoral processes. Information disorder also widens some pre-existing societal gaps by disproportionately affecting marginalized groups, women, and activists, for example.

The mechanisms driving disinformation are multifaceted. Social media algorithms amplify sensationalist content, while economic motivations, political interests, and social identity dynamics provide fertile ground for its proliferation. This global problem is particularly fueled by the dominance of social media platforms and their attention-driven business models, and by the profitability of actors operating in an emerging 'disinformation industry'.

Addressing disinformation requires not only reactive measures but also proactive efforts to reinforce democratic principles and institutions, tackling its root causes at their core. Efforts to combat disinformation focus on two key approaches: fostering individual and societal resilience, and regulating the digital information environment in order to reduce social exposure to disinformation. Media and information literacy (MIL) initiatives, strengthen critical thinking and equip individuals to recognize and resist false or misleading narratives. On the regulatory front, approaches aimed at reducing societal exposure to disinformation often rest on three main pillars: a) asserting greater liability of intermediaries for content posted by third parties on their platforms; b) introducing positive obligations for intermediaries, which may consist of content removal, transparency, and due process, and c) empowering fact-checkers. Technological solutions, including AI-based fact-checking tools and watermarking AI-generated content, are also emerging as countermeasures.

This policy paper identifies opportunities for development cooperation, and makes recommendations spanning both political and operational dimensions. Among the strategic approaches suggested, it is worth highlighting: the importance of tailoring initiatives to the level of political freedom and institutional strength in different countries; the need to work with partners to align measures to counter disinformation with human rights standards; the urgency of supporting journalists and fact-checkers through funding, technological tools, and safety measures; and the need for a multidimensional strategy – integrating legal, social, economic, and technological solutions. Collaborative, multistakeholder, and long-term efforts will be key to countering this hybrid threat and fostering information integrity.

1 Information disorder: Defining the problem

The deceptive use of information is a longstanding phenomenon, with examples stretching from Ancient Egypt through the widespread use of information operations during the Cold War, and beyond. While historical lessons offer valuable insights, they fall short in addressing the complexities of today's information disorder, driven by profound shifts in the communication ecosystem.

The rise of the internet has disrupted traditional, centralized information sources and enabled any actor with a connected device to produce, distribute, and consume content. Today, digital communication forms the backbone of identity construction, information exchange, and socio-political discourse. On the one hand, these developments have empowered individuals, fostering political participation and civic mobilization, as seen during the Arab Spring. On the other hand, they have given rise to information overload and pollution, exacerbating the spread of false, misleading, or manipulated content capable of undermining social cohesion and democratic governance.

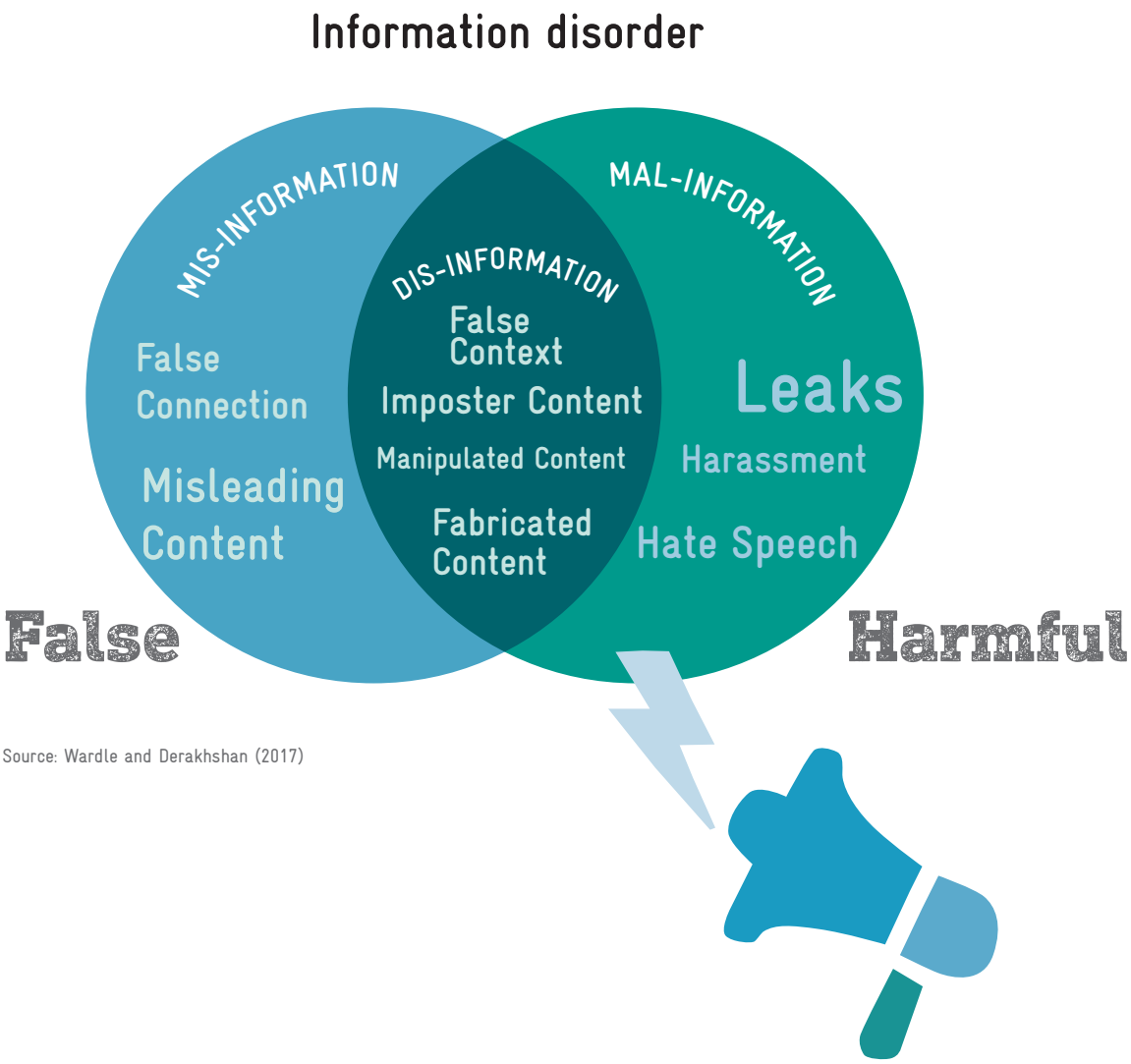
In the context of development cooperation, addressing this 'information disorder' is essential to safeguarding democratic processes, ensuring access to reliable information, and promoting resilient information ecosystems in partner countries. The promotion of 'information integrity' aims to address not only the accuracy and reliability of information, but also the systemic and structural dynamics that shape the flow and accessibility of content in the digital space, including through the use of algorithms and AI. This understanding is supported by the UN Global Principles on Information Integrity (United Nations, 2024), geared at promoting a "pluralistic information space that champions human rights, peaceful societies, and a sustainable future". This focus not only supports democratic governance but also underpins progress across a wide array of sustainable development goals.

1.1 Unpacking information disorder

Information disorder is an overarching concept that can be used to make reference to multiple challenges. For the United Nations (2023), for example, it encompasses disinformation, misinformation, and hate speech. According to Wardle and Derakhshan (2017), it encompasses disinformation, misinformation, and mal-information, and the latter includes some types of harassment and hate speech.

While hate speech has an internationally established definition, consisting of any kind of communication that attacks or uses pejorative or discriminatory language with reference to a person or a group on the basis of an identity factor, such as religion, ethnicity, nationality, or gender, for example (United Nations, 2019), there is no consensus on how to define disinformation and misinformation. One commonly adopted distinction between the two concepts lies in assessing intent. While disinformation is intended to deceive and create harm, misinformation refers to the unintentional spread of inaccurate information shared in good faith. Wardle and Derakhshan also refer to mal-information, which would take place when genuine information is shared to cause harm, often by taking it out of context, or moving information designed to stay private into the public sphere.

Figure 1: Unpacking the concept of information disorder



Source: Wardle and Derakhshan (2017)

In focus: Defining and tackling disinformation

Disinformation can often be ill-defined (UNGA, 2021). As an example, a study conducted on over 62 disinformation bills proposed in **Brazil** during the administration of right-wing politician Jair Bolsonaro (2019–2022) showed that most legislators proposed disinformation bills without specifying exactly what disinformation-related problem they were trying to tackle (Anastácio, 2024).

When definitions are present, they are often structured around the notions of truth/falsehood and intent, which are difficult to assess in practice. The former brings about philosophical as well as practical challenges in terms of defining what 'truth' is, as well as naming authoritative actors to be 'arbiters of truth', a role that can lead to abuse. In **Singapore**, for example, the government has discretion over defining false statements, raising concerns related to freedom of expression (Amnesty International, 2023).

Assessing intent is also challenging, and requires knowledge about the speaker's awareness of the falsehood, and about the intention to deceive or mislead others and create harm. This proof is difficult to establish in relation to the original speaker, and almost impossible in relation to subsequent speakers who disseminate the message (Appleman et al., 2022).

Because of these challenges, an alternative way of tackling mis- and disinformation from policy and legal perspectives is to question whether the information produces actual or potential harm, and endangers or violates legally protected rights or interests. The **European Union Digital Services Act (DSA)** provides an example of a risk-based approach to disinformation that also addresses the potential harms that disinformation can cause, including to democratic processes and to public health.

Some countries focus on combating influence operations coming from abroad, while maintaining a high level of freedom of speech for their own citizens. The **Swedish Psychological Defence Agency (PDA)**, for example, does not focus on determining whether information is false, but places emphasis on mitigating undue control over the information environment by foreign actors, which would enhance their capacity to conduct information campaigns.[1] This separation between foreign and domestic can often be observed in societies with high social cohesion and high levels of democratic quality.

1 Information campaigns are defined by the PDA as "a set of activities coordinated by a foreign power that involves the promotion of misleading or inaccurate information or other specially-adapted actions aimed at influencing the decisions of politicians or other public decision-makers, the opinions of all or a part of the population, and opinions or decisions taken in other countries" (Sörensen and Pamment, 2023).

1.2 The role of social media platforms in today's information ecosystem

False information diffuses significantly farther, faster, and more widely than the truth on social media (Vosoughi et al., 2018). Fabricated stories that evoke shock, surprise, or strong emotions tend to capture greater attention and are more likely to be shared widely, in contrast to factual information, which is often perceived as less dramatic or engaging. This has both psychological and behavioural, as well as technical (e.g., prioritizing through algorithms) reasons, and causes concern in a context in which platforms have become a significant part of the infrastructure of the information ecosystem. For example, over 70% of individuals in some developing countries use social media as a source of news (Statista, 2022). This high percentage could be, in part, related to the pervasiveness of zero-rating practices, especially in Latin American and African countries.[2]

In **Latin America**, a whopping 86% of the population use social media, and they especially resort to Facebook, Instagram, and WhatsApp for news. **Argentina** occupies the first place in terms of average social media use, followed by **Brazil and Mexico**. TikTok is particularly popular among the young population (18-24 years old) – more than 40% use TikTok as a news source (Ferreyra and Segarra, 2024). In **African countries** such as **Nigeria, South Africa and Kenya**, WhatsApp, Telegram and Facebook are important for finding, sharing, and discussing news (Nanfuka et al., 2024). In some European countries, such as **Greece, Bulgaria and Hungary**, the use of social media for information is above 60% (Statista, 2022), and the figure is above 50% in the **United States** (Khalid, 2019).

There is an interplay between these findings and the loss of space of established journalism in the media ecosystem. The 2022 edition of the Reuters Institute Digital News Report shows an overall decline in the news trust index. On a global average, 42% of respondents trust the news. In Latin America, **Brazil** is the only country with an above-average score of 48%, while, in **Argentina**, only 35% trust in the media (Ferreyra and Segarra, 2024). In settings where trust in established journalism is declining, people often turn to less regulated sources of information, which do not follow journalistic standards, with social media often serving as a popular substitute. At the same time, the spread of disinformation itself undermines trust in the media, creating a vicious cycle.

This shows that discussions on disinformation are intimately associated with ensuring press freedom and media pluralism. Traditional media is weakened in contexts in which authoritarian governments exert control over the media, or when media is concentrated in the hands of a few political players. Traditional media also faces financial pressure because of the shift of advertisement revenue to social media. In this context, journalistic channels often resort to state advertisement revenue for survival, potentially compromising their independence and credibility. In **Africa**, lack of resources has led some private media channels to sacrifice their editorial independence, and start to disseminate partisan disinformation content online and offline (Nanfuka et al., 2024).

2 Through zero-rating arrangements, telecommunications companies offer unlimited free use of major social media services as part of limited mobile data plans. Often, platforms pay telecommunication companies to include their services in zero-rating contracts, giving an advantage to large companies with the resources to pursue such strategies (Berger, 2024). In economic and cultural contexts where access to information is already scarce, such as among the low-income population, zero-rating services end up enabling disinformation campaigns, because people have access to no other sources to check or verify what they are reading on social media. In 15 Latin American countries, zero rating arrangements are currently in place, providing free access to social media platforms, notably WhatsApp or Facebook (Ferreyra and Segarra, 2024).

While individuals increasingly resort to social media for information, they still use platforms with a recreational mindset, which lowers critical thinking and makes them more vulnerable to content that provokes an emotional response, that has a powerful visual component or a strong narrative, or that is shown repeatedly. A diverse range of actors exploit this model to spread false or misleading information for economic, political, ideological, or social gains.

1.3 Drivers of disinformation

The motivating factors to the spread of disinformation may vary, but they can usually be placed under three potential types of incentives: economic, political and/or ideological, and social incentives.

Economic incentives:

The spread of disinformation can be a profitable activity for individuals, organizations, and platforms. For example:

- The underlying **business models of online platforms** are frequently driven by engagement metrics that depend on retaining users within the platform. Many platform algorithms are designed to prioritise content that generates higher user engagement, which may include sensationalist posts, clickbait, polarising content, and disinformation. The massive collection of personal data by platforms allows for a detailed profiling of users, leading to a high degree of personalisation of social media feeds and microtargeting. Different people may have completely different experiences online, creating filter bubbles and echo chambers – a space in which a person's own beliefs or opinions are, or appear to be, echoed back at them by other users or content that they consume.
- **Financially motivated 'disinformation entrepreneurs'** have transformed the spread of disinformation into a business model. In 2022, a group named 'Team Jorge' (named after its leader) was accused of spreading disinformation to manipulate elections in over 20 African countries, and a further 10 countries outside Africa for financial gain (Andrzejewski, 2023). In **Latin America**, the 2023 'Digital Mercenaries' research, coordinated by the Latin American Center for Investigative Journalism (CLIP, 2023), identified the main consultants and agencies that have carried out disinformation strategies designed to support the emergence of particular political parties in the region, mostly from the far-right. Disinformation entrepreneurs usually act across countries and regions. For example, Fernando Cerimedo, an Argentine consultant, provides services for Latin American far-right politics in **Argentina, Brazil**, and **Chile** (Ferreyra and Segarra, 2024). Disenso Foundation, created by Spain's far-right Vox party, has also been used to try to delegitimize the electoral results in several Latin American countries. In **Africa**, disinformation campaigns in **Congo** have been carried out from neighboring states Uganda and Rwanda, which border the conflict regions in the East of the country. Moreover, London-based Bell Pottinger stirred up racial tensions and narratives of an 'economic apartheid' in **South Africa** (Nanfuka et al., 2024). Paid influencers – individuals with a large social media following – may also play an important role in spreading disinformation by leveraging their reach, credibility, and trust with their followers to amplify false or misleading narratives.

Political and/or ideological incentives:

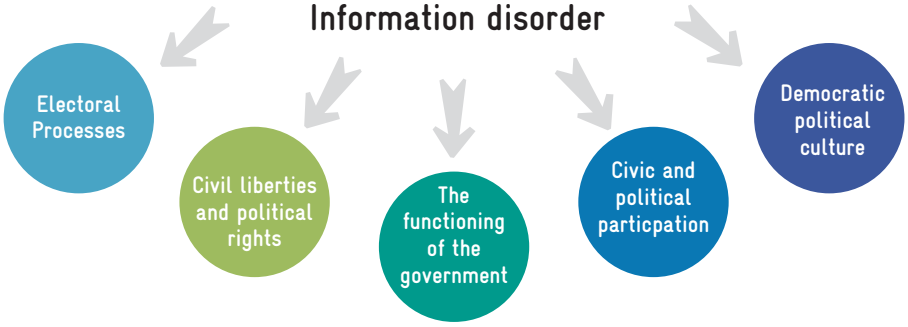
Disinformation may be shared in an attempt to influence public opinion and political decisions, or to de-stabilize a country's political system by undermining trust in public institutions or processes. Discrediting a political candidate in an election, or conducting foreign information operations to exploit rifts within societies are examples of these practices. Targeted disinformation campaigns seek to polarize, agitate, and weaken liberal democracies. For example, according to the Africa Center for Strategic Studies, in Africa, almost 60% of disinformation is sponsored by foreign states – mostly Russia, China, the United Arab Emirates, Saudi Arabia and Qatar. Additionally, the Russian government is actively funding disinformation campaigns in several countries in Latin America, including **Argentina, Bolivia, Chile, Colombia, Cuba, Mexico, Venezuela, Brazil, Ecuador, Panama, Paraguay, Peru**, and **Uruguay**, according to a 2023 U.S. Department of State statement.

Social incentives:

Disinformation may be spread as a way of reinforcing identity by expressing affiliation with a certain group, or of seeking social prestige. This 'affectual dimension' (Monsees, 2023) helps us to understand why some people share knowingly false information. Narratives stick not because the information itself is true or false, but because it ties into people's view of the world.

2 The interplay between information disorder and democratic principles

Disinformation erodes trust in democratic institutions and processes, undermining the credibility of political decision makers, the judiciary, and the media. It compromises the integrity of public discourse and decision-making, while fuelling attacks on politicians, journalists, and human rights activists – disproportionately affecting women and other vulnerable groups. As a driver of polarization and societal division, disinformation can escalate tensions and lead to unrest, posing a significant challenge even to stable democracies as part of hybrid threats.[3] Democracy, understood as a political continuum, can be assessed by the stronger or weaker presence of certain components, such as: electoral process and pluralism, civil liberties, the functioning of government, political participation, and political culture.[4] Information disorder has an across-the-board impact on all these dimensions.



2.1 Electoral processes

The most basic cornerstone of democratic governance is the right to choose political representatives in free and fair elections. Mis- and disinformation can distort public perception, threaten the integrity of electoral outcomes – as can be exemplified by the annulation of the first round of Presidential elections in Romania, in December 2024 – and may weaken democratic participation and government legitimacy. The importance of digital media for communication during elections creates the potential for abuse, both by foreign and domestic actors. This potential is being leveraged by the disinformation industry, which is becoming professionalised and monetized, undermining equal campaigning opportunities, and, in the most serious cases, the peaceful transfer of power from one government to another.[5] Platforms are often used to circumvent national electoral laws against disinformation, as they operate outside the jurisdiction of most countries. For example, in 2023, Facebook and X continued to run political advertising in **Kenya**, contradicting electoral laws that ban campaigning on polling day (Kitili et al., 2023).

3 A hybrid threat refers to the use of a mix of interconnected tactics – such as cyberattacks, information campaigns, and economic pressure – to achieve strategic objectives. Disinformation campaigns are often part of a broader strategy, seeking to weaken a target nation's resilience.

4 These five categories structure the assessment provided by the Democracy Index, produced by The Economist Intelligence Unit (EUI). According to the methodology that guides the formulation of the Index, each of the categories can be unpacked into specific items of assessment.

5 Examples can be found in the events of January 6, 2021, in the United States, when supporters of then-President Donald Trump sought to disrupt the certification of the 2020 presidential election results and attacked the U.S. Capitol in Washington, D.C., as well as in the events of January 8, 2023, in Brazil, when supporters of former President Jair Bolsonaro stormed key government buildings in the capital, Brasília, in an effort to overturn the results of the 2022 presidential election.

In focus: AI-powered disinformation and elections

The first use of deepfake technology – manipulated images, videos, or audio files that appear strikingly realistic – in an electoral setting was a short video by Indian politician Manoj Tiwari in 2020, translating a campaign message into a Hindi dialect to target new voters. In 2022, Yoon Suk-yeol, the president of South Korea, used a deepfake avatar of himself to appeal to younger voters during his presidential election campaign. However, not all uses of deepfakes are benign. They have also been deployed to craft false narratives, including fabricated speeches or actions attributed to political leaders, and to damage the reputations of individuals, posing significant ethical, legal, and societal challenges. An additional emerging concern is related to the capability of Artificial Intelligence (AI) to operate a large number of fake profiles on social media, and to produce misleading deepfakes or other AI-generated disinformation.

In spite of a few examples of AI misuse (Swenson and Chan, 2024), the impact of the technology was deemed modest in the 2024 'super election year'. This could be due to a combination of governmental norms and guidelines, and action taken by AI Labs including research centers of large companies such as OpenAI and Anthropic. It is also difficult to accurately gauge the impact of AI-based disinformation without access to data that will clarify the interaction between users' preferences (information pull) and algorithms (information push). The impact of platforms' algorithms on polarization and social cohesion remains contested, with studies pointing to other deeper and non-technological factors as being more influential, such as pre-existing social divisions, the absence of robust social security systems, gender inequality, and limited digital literacy. This highlights the need to address foundational inequities to mitigate polarization effectively. However, isolating the specific impact of technology on democratic processes remains challenging, as it operates within a complex web of social, cultural, political, and economic factors. These dynamics often have long-term effects, including the polarization of public discourse and the fragmentation of social cohesion, rather than affecting short-term behavioural changes.

While automated systems may not be the only cause of polarization, they shape news consumption and media diversity through algorithmic gatekeeping. Moreover, the prevalence of hate speech can lead vulnerable groups to withdraw from online discussions, decreasing social cohesion. Because of this, it is difficult to predict the relevance of AI-based disinformation on upcoming elections. For example, Meta has recently announced changes in their content policies that may lead to a less stringent approach to countering disinformation across their platforms, and a more permissive approach towards hate speech. This may reduce their willingness to adjust their AI models and to cooperate with public authorities.

In addition, while most attention has been paid to large language models, comparatively simpler AI technologies can also be used to influence elections. A report from the Center for Countering Digital Hate (CCDH, 2024) investigated the vulnerabilities and risks associated with six popular AI voice-cloning tools. In 80% of 240 tests, these tools successfully generated convincing audio statements of high-profile politicians. The study found that none of the tools had adequate safeguards against the misuse of their technology.

2.2 Civil liberties and political rights.

Disinformation impinges on freedom of opinion and freedom of expression. It leads to the blurring of the lines between facts and falsehoods, leading to a non-consensual manipulation of the thinking process necessary to develop one's opinion. The right to hold opinions and the right to freedom of expression are intertwined, since the former entails an environment which preserves the capacity to freely access reliable information, necessary to form one's opinion and to change one's mind. Measures put in place against information disorder may impact freedom of opinion if they restrict access to certain media channels, or promote internet shutdowns to counter the problem, for example. In Africa, both autocratic and democratic governments have increasingly resorted to shutdowns in response to concerns about disinformation around elections, or when confronted with the potential for online hate speech to encourage violence (Gagliardone and Stremlau, 2022).

Measures against disinformation may also curtail the right to freedom of expression, either unintentionally or because they can be easily abused by public authorities to impose arbitrary or politically motivated limits to freedom of expression, or to undermine the oppositions' prospects for victory.

In focus: Disinformation and freedom of expression

According to the UN Human Rights Committee, the right to freedom of expression applies to all kinds of information and ideas, including those that may shock, offend, or disturb, irrespective of the truth or falsehood of the content. This means that, under international human rights law, combating false information is not, in itself, a legitimate aim that justifies restricting freedom of expression. Freedom of expression may be restricted only in accordance with Article 19 (3) of the ICCPR, which requires all restrictions to be exceptional, provided by law, narrowly construed, and necessary for a legitimate aim (UNGA, 2021).

The **directness of the causal relationship** between the speech and the harm to a legitimate aim, and the **severity and immediacy** of the harm, are key considerations in assessing whether the restriction of speech is necessary. If restrictions are introduced, they must be appropriate and proportionate, using the **least restrictive means** possible. If measures to combat mis- and disinformation unduly restrict freedom of expression, they also undermine democracy by weakening civil and political rights which are some of its most fundamental pillars. Vague laws that confer excessive discretion to fight disinformation are particularly concerning, as they can lead to arbitrary decision-making (UNGA, 2021).

2.3 The functioning of the government.

The ability of a state to make independent decisions, manage its internal affairs, and protect its own interests, while also ensuring that needs and rights of its citizens are met, is a crucial aspect of its sovereignty and stability. In a democratic context, this function is particularly important as it directly influences the quality of governance, citizen trust, and overall effectiveness of public institutions. It also evaluates the effectiveness of domestic checks and balances on the exercise of government authority. External threats, such as foreign information campaigns, may significantly affect the functioning of the government. These campaigns increasingly exploit domestic vulnerabilities to sway domestic audiences and politicians. Additionally, threats can emerge from within the state, particularly in situations where checks and balances are weak or absent, enabling harmful actions by domestic authorities. In some partner countries in Asia, Africa, and Latin America, established political parties are involved in spreading disinformation. In Argentina, for example, narratives of electoral fraud were circulated by members of the libertarian party La Libertad Avanza, led by then-candidate Javier Milei (now president of Argentina). Claims were spread on platforms such as TikTok and X with the aim of casting doubt on the electoral process (Ferreyra and Segarra, 2024).

2.4 Civic and political participation.

Mis- and disinformation may also affect the right to political participation, especially of marginalised and vulnerable groups. Disinformation campaigns often target ethnic minorities, immigrants, and other marginalised communities, exacerbating social tensions and creating a chilling effect that undermines their predisposition to take part in public discussions.[6] In **Africa**, social media platforms are used to mobilise support along ethnic lines, and to promote violence against minority groups or political opponents (Nanfuka, 2024).

In focus: Examples of the impact of disinformation campaigns on women, journalists, and activists

Disinformation campaigns have disproportionately targeted women, contributing to widening the gender gap in politics. For example, Freedom House (2024) reported that Janet Love, a member of the electoral commission in **South Africa**, was accused of vote rigging by supporters of Jacob Zuma in the May 2024 elections. During elections in **Kenya**, female candidates faced gender-based attacks questioning their qualifications (Nanfuka et al., 2024).

Disinformation also disproportionately affects human rights activists and journalists. In the **Philippines**, journalist and Nobel Peace Prize laureate Maria Ressa, CEO of Rappler.com, an investigative journalism organization dedicated to fact-checking and exposing state-sponsored disinformation, suffered online harassment, particularly online-gender based violence, in retaliation. Overall, the precarious conditions for practicing journalism do not only compromise professional integrity, but also create an environment of self-censorship and reluctance to conduct investigative journalism.

2.5 Democratic political culture.

Attitudes, beliefs, and behaviours of citizens and institutions support the functioning of a democratic system. They include tolerance and pluralism, civic responsibility, transparency, accountability, and trust. Mis- and disinformation have a significant impact on social cohesion, which encompasses both the horizontal relationship between different individuals and groups within a society, as well as the vertical relationships between individuals or groups and the state. They provoke a general state of uncertainty within society that some authors have called the 'post-truth' predicament (Chambers, 2019), which amplifies distrust and damages the capacity to engage in communicative rationality, leading to political disengagement and apathy.

6 Article 20 (2) of the International Covenant on Civil and Political Rights (ICCCP) provides that any advocacy of national, racial, or religious hatred that constitutes incitement to discrimination, hostility, or violence is to be prohibited by law, regardless of any assessment of its truthfulness.

Information disorder is particularly concerning in a context in which many countries are experiencing a democratic backslide. Latin America and Africa provide examples of the phenomenon, but democracy is also being challenged in developed regions. This takes place in parallel to attempts by authoritarian governments to disseminate government-sponsored disinformation and further restrict civic space, undermining the capacity of societies in these countries to counter disinformation through bottom-up mobilization.

In focus: Information disorder in democratic, semi-democratic, and autocratic contexts

Disinformation operates differently in countries that display a stronger or weaker presence of certain components of democratic governance. This is not only shaped by the political system itself, but also by media landscapes and societal dynamics in each country.

Democratic countries face disinformation campaigns from multiple actors, such as political parties, private entities, such as businesses or influencers, and foreign actors conducting interference operations. False claims of election fraud, or narratives that exploit particularly controversial and divisive points within societies seem to be the most common choice of narratives. A particular characteristic of disinformation campaigns conducted in democratic societies is that they exploit and, to some extent, weaponise freedom of expression and the diverse media landscape. They create fragmentation of the public sphere, as well as social divisiveness, especially through astroturfing – which can be understood as the practice of hiding the sponsors of a message to make it appear as if it originates from, and is supported by, grassroots participants. This multiplies echo chambers where disinformation can thrive among like-minded groups. In democratic countries, disinformation may erode social cohesion and political stability and potentially weaken democratic governance.

In semi-democratic countries – i.e. countries in which a few of the components of democratic governance are weak or inexistent – institutional weakness allows disinformation to become a pervasive practice. Disinformation campaigns become a go-to strategy promoted by governments seeking to legitimize themselves and undermine opposition groups, and by opposition parties trying to challenge the government. In this scenario, public media is often under-resourced, undermining its editorial independence and diminishing its credibility. Other media outlets often become polarized, with state-controlled media pushing official narratives and independent outlets countering with alternative claims. If countries display low levels of public education or awareness, citizens may be even more susceptible to manipulation. Regional or global powers find fertile ground to meddle in domestic affairs, seeking to influence polarized actors and align the country with their strategic interests.

In autocratic countries – i.e countries in which most of the components of democratic governance are missing – governments use disinformation as a top-down tool of control. The dissemination of disinformation is often carefully managed by state authorities, relying on a controlled media landscape and a sophisticated surveillance system. The main disinformation narratives seek to legitimise the regime, often presented as the sole path towards prosperity and stability, and the protector of national, cultural, or religious identity – while portraying any opposition as illegitimate or traitorous. Such narratives are supported by the suppression of alternative viewpoints, with censorship and punitive measures against dissenting voices. As a flipside, these narratives seek to vilify external enemies, often blaming them for domestic problems, and suppressing domestic dissent, such as opposition leaders, activists, and journalists. In many cases, autocratic regimes actively cooperate with aligned foreign governments or organizations to amplify their disinformation campaigns and neutralize external criticism.

Table 1. An assessment of elements of disinformation across democratic, semi-democratic and autocratic countries.

Elements of disinformation	Democratic	Semi-democratic	Autocratic
Main actors	Political parties, foreign states, private entities, disinformation entrepreneurs	Governments, opposition, polarized media, disinformation entrepreneurs, foreign states	The state, state-sponsored actors
Narratives	Electoral fraud, public health disinformation, exploiting dissent	Election fraud, regime legitimacy, discrediting dissent, ethnic or religious tensions	Regime propaganda, vilifying external enemies
Facilitators of disinformation spread	Electoral competition, fragmented media, exploitation of social polarization, foreign interference	Weak institutions, authoritarian tendencies, polarized media, weak support to journalists	Censorship, social control, suppression of dissent, pervasiveness of propaganda
Main tools	Social media, polarized outlets	State media, polarized outlets, and social media	State-controlled media, propaganda machinery across media channels, including social media

Annex I shows a correlation between the democratic components mentioned in this section, with strategic approaches and concrete actions that could assist in curbing information disorder.

3 Combating disinformation: Strategic approaches

Measures introduced to combat mis- and disinformation are related to two main sets of strategic goals:

- **Promoting individual and societal resilience by strengthening agency**
- **Protecting society by reducing societal disinformation exposure.**

This two-pronged strategy has led to more public engagement in shaping media and information literacy initiatives on the one hand, and more government regulation on the other. The success of these strategies requires a multidimensional regulatory approach, which takes into account the influence of laws, norms, economic incentives and technical architecture on behaviour.

3.1 Promoting individual and societal resilience by strengthening agency

Policies that foster **media and information literacy (MIL)** as a way of combating disinformation are underpinned by a participatory approach, in which the solution to the problem is necessarily co-constructed with the interpreters of the message. Empowering individuals is seen as a necessity, particularly in a scenario in which the speed of disinformation surpasses the capabilities of traditional regulatory and oversight mechanisms.

MIL aims at strengthening analytical skills and critical thinking across stakeholder groups (i.e. government, civil society organisations, journalists, and social media users). It places individuals as front-liners capable of 'prebunking' and 'debunking' practices. While **debunking** involves exposing an already disseminated false claim – by, for instance, fact checking it – **prebunking** tackles disinformation before it has been spread. Through MIL training, audiences are exposed to weakened versions of disinformation in order to strengthen their capacity to identify and resist this type of information and to prevent its amplification. This can be achieved, for example, by revealing the main mechanisms and techniques employed in disinformation strategies, and by using pedagogical tools such as simulations and games.[7]

7 An example is Bad News, a free online game in which players take the perspective of a fake news tycoon. https://www.getbadnews.com/books/english/

In focus: Examples of MIL initiatives

In Africa, **Dubawa** – a fact-checking and media literacy initiative operating in West Africa – promotes workshops and trainings to empower journalists and the public in general with the skills to critically evaluate information (Nanfuka et al., 2024). A similar role is played by **RedLATAM**, a Latin American fact-checking network, focused on educating journalists and the public about disinformation (Ferreyra and Segarra, 2024). In Argentina, **Chequeado** is an independent fact-checking organization that runs MIL workshops and helped more than 60 universities to include fact-checking in their curricula. The organization also developed tools such as Chequeabot, which has empowered other organizations, reducing the time and effort required to fact-check information. In Brazil, **Aos Fatos** conducts fact-checking and provides workshops for teachers and students (Ferreyra and Segarra, 2024). In Europe, **Finland** has been particularly recognised for policies and educational strategies aimed at equipping citizens with the critical thinking skills necessary to navigate a complex media landscape. In 2019, the country was ranked first in a study that compared social levels of media literacy and resilience in 35 countries (Lessenski, 2018).

MIL may also play a strategic role in countering foreign information influence operations, a situation that reveals the interplay between disinformation and national security concerns. In these cases, MIL initiatives may provide a strategy to strengthen resilience at a societal level, which can be understood as the symbiosis of a group of individual capacities coming together. Sweden provides an example of a whole-of-society approach to combating disinformation, which can be associated with the concept of total defence (Sörensen and Pamment, 2023). This approach necessarily entails the collaboration between government and non-governmental actors, including organisations in the educational system, civil society organisations, journalists, and the private sector.

Among the responsibilities of Sweden's Psychological Defence Agency (PDA), created in 2022, is strengthening the population's resilience and to provide overall coordination to Sweden's strategy to counter disinformation. The PDA constantly monitors the information environment in order to identify and counter external threats at an early stage, as well as internal social vulnerabilities that could be unduly exploited (i.e. a growing polarisation over a social issue, such as immigration, for example), taking into account social or affectual motivations to spread disinformation (Monsees, 2024). The long-term approach adopted by the PDA is important, since observing the information environment through an amount of time is important to be able to draw a distinction between external sources and domestic sources of information disorder, especially in the context of ever more sophisticated examples of astroturfing.

3.2 Protecting society by reducing disinformation exposure

Individuals depend on online intermediaries[8] to access information, maintain social interactions, and express their views. This creates an asymmetry of power between individuals and intermediaries, which are *de facto* gatekeepers of the information flowing through their infrastructure. In the context of mounting information disorder, governments have adopted a protectionist stance, seeking to compel intermediaries to reduce societal exposure to disinformation.

Approaches aimed at reducing societal exposure to disinformation often rest on three main pillars: asserting greater **liability of intermediaries** for content posted by third parties on their platforms, introducing positive **obligations for intermediaries**, which may consist of content removal, transparency, and due process, and **empowering fact-checkers**. Some examples of measures being adopted in relation to these three pillars can help us to understand present and future challenges of curbing information disorder on online platforms.

Recent developments in the United States

In the USA, Section 230 of the US Communications Decency Act – which limits the liability of intermediaries for content posted by third parties on their platforms as a way to protect free speech and innovation – has been under mounting criticism in recent years. Scandals, such as the Cambridge Analytica affair, as well as news that social media has contributed to child exploitation and suicide, have rendered views on the impact of social media more negative.

In spite of that, after taking control of Twitter (today X), in 2022, Elon Musk committed to turn the platform into a "digital town square" where diverse viewpoints could be expressed with minimal restrictions. He reinstated formerly banned accounts, including Donald Trump's, and significantly reduced the staff employed in Twitter's Trust and Safety team, which oversaw content moderation. He also put emphasis on crowd-sourced fact-checking (referred to as Community Notes) instead of fostering collaboration with fact-checking organizations. After these changes, X became the social network with the largest proportion of disinformation, according to a European study (Vallance, 2023).

8 In the digital context, intermediaries are mostly private agents, and they exist in the three layers of the internet: telecommunications infrastructure, technical standards, and applications. This means that the notion of intermediary may comprise various services, including internet service providers (ISPs), domain name system (DNS) service providers, cloud providers, search engines, e-commerce platforms, and social media platforms.

Social media platforms and disinformation: What will the next US presidency hold?

In January 2025, Meta's CEO, Mark Zuckerberg, announced a major shift in the company's content policy in the US across its platforms – Facebook, Instagram, and Threads. The company decided to discontinue its third-party fact-checking program in the United States, relying on Community Notes instead, a model that transfers content moderation responsibilities to users. While Meta committed to observe strict control of illegal and highly severe content violations, fewer restrictions will be imposed on other types of content, including anti-immigrant and anti-LGBTIQ+ content, according to the CEO (Hendrix, 2025). Zuckerberg claimed that censorship pressures are mounting in multiple regions, through "secret courts" in Latin America and in Europe's "ever-increasing number of laws, institutionalizing censorship". He promised to work with President Donald Trump "to oppose global 'censorship' pressures" (Hendrix, 2025).

Decisions were centered in the USA and in an anglophone political scenario, without giving due consideration to their impact in countries where institutions and democracy are less strong. The role played by Facebook in social crises, such as the dissemination of hate speech inciting ethnic violence in Ethiopia in 2023, or in the genocide that took place in in Myanmar, in 2017, were ignored.

In the meantime, more than 150 civil society organizations in Latin America signed a letter in opposition to the changes, noting that the proposals jeopardize vulnerable groups, undermine global efforts to promote a safer, more inclusive, and democratic digital space, and "openly attack the sovereign and democratic efforts of nations to protect their populations from the harms caused by Big Tech" (Coalizão Direitos na Rede, 2025).

Increasingly different views on content policy and disinformation are being adopted by the United States and the European Union. This is not only related to the more stringent legal approach to disinformation adopted by the EU (if compared to the traditional reliance of the USA on platform self-regulation), but also to the prominent role given to fact-checkers in EU regulation.

Regulation in the European Union

In the European Union, the Digital Services Act (DSA) of 2023 introduced legal obligations to combat disinformation, as well as liability in the event of nonfulfillment of these obligations. Under the DSA, there are two main possibilities to combat disinformation: a) if the content is illegal – which refers to information that is incompatible with EU law or with the law of any member state, as per Article 3(h) – then all types of online intermediaries are obliged to act upon it;[9] and b) if the content is 'socially harmful', the DSA imposes specific obligations only on a selected category of online intermediaries – very large online platforms (or VLOPs) and very large online search engines (or VLOSEs). These companies should assess the systemic risks arising from the design, operation, and use of their services, as well as from the potential misuse of services (Article 34). Among the systemic risks are the ones related to actual or foreseeable negative impact on democratic processes, civic discourse, and electoral processes, as well as on public security (Recital 82).[10]

The **importance of fact-checkers** in strategies to counter disinformation has grown in recent years. In Europe, the DSA has institutionalized the role of a category called 'trusted flaggers' across the EU. The DSA mandates that online platforms prioritise and respond promptly to reports from trusted flaggers. This prioritisation is designed to expedite the removal of harmful content, enhancing user safety and platform accountability. The DSA also encourages greater transparency by requiring platforms to publish regular reports on their content moderation activities, including the role and impact of trusted flaggers.

Fact-checking initiatives

Fact-checking initiatives are widespread across different regions, and are based on close connections and mutual learning, which keep fact-checking networks dynamic and agile (Berger et al, 2024). For example, the **Latam Chequea Network** connects more than 35 organizations, including fact-checkers, journalists, and technologists (Ferreyra and Segarra, 2024). The **African Fact Checking Alliance** (AFCA) is Africa's largest fact-checking network, bringing together fact-checking organizations across the continent to share resources, tools, and strategies for addressing disinformation (Nanfuka et al., 2024).

Over the years, many fact-checking organizations evolved into hubs for research, training, and development of cutting-edge tools to counter disinformation. The Brazilian **Aos Fatos** created a technological solution called Fatima, a fact-checking chatbot which helps users by sending them verified information (Ferreyra and Segarra, 2024). Chequeabot is an AI tool developed by **Chequeado**, capable of assessing checkable phrases, detecting information that has already been fact-checked, and flagging suspected cases of disinformation in social networks, media, or podcasts. It can also provide real-time transcripts of videos to text, facilitating fact-checking of live speeches, debates, and other relevant events.

9 In the case of illegal content, platforms have several obligations, such as providing mechanisms enabling any person or entity to report content deemed illegal (Article 16), taking action against illegal content upon judicial or administrative order (Article 9(1)), and ensuring priority handling of notices submitted by entities referred to as trusted flaggers (Article 22).

10 Other systemic risks include b) risks related to an actual or foreseeable negative effect on the protection of public health, minors and serious negative consequences to a person's physical and mental well-being, or on gender-based violence (Recital 83). In particular, these latter risks may stem from coordinated disinformation campaigns related to public health, or from online interface design that may stimulate behavioural addictions among service recipients (Recital 83).

The growing importance of fact-checking organisations has not been matched by commensurate resources. Fact-checking initiatives remain underfunded, are faced with budgetary cuts in critical moments, and often struggle to keep up with speed and volume information on social media platforms. Moreover, the U-turn of some US platforms on fact-checking could become a blow to the future development of this work worldwide, given the centrality of US platforms.

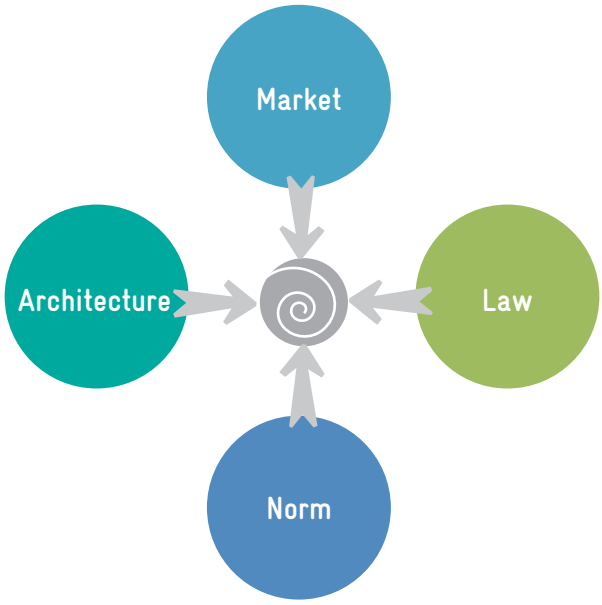
3.3 Boosting resilience and protection through a multidimensional approach to regulation

In digitalised societies, regulation may be carried out not only by top-down laws enacted by governments, but also by the use of four main mechanisms that can influence behaviour: laws, norms, market, and technological architecture (Lessig, 2006).

Taken together, these mechanisms could strengthen the effectiveness of strategies to curb information disorder.

- **Laws** on disinformation may have many different goals, such as clarifying responsibilities, creating obligations for intermediaries, or seeking to protect electoral integrity from information disorder (i.e. by establishing certain conditions for content removal, or to restrict algorithmic mainstreaming of certain content).
- **Norms** are social conventions that one often feels compelled to follow, which include non-binding agreements, frameworks, and principles. The UN principles on Information Integrity and the UNESCO Guidelines for Governance of Digital Platforms are two examples of documents that aim to foster the consolidation of social norms on disinformation. MIL also helps to make disinformation campaigns more likely to fail by shaping social norms and fostering a culture of critical thinking and responsible information consumption.
- **Market forces** regulate by acting upon supply, demand and the pricing system. Financial incentives play an important role in the creation and dissemination of disinformation. There is growing awareness about the importance of intervening upon the 'platform economy'– the UN Guiding Principles, for example, calls for "a fundamental shift in incentive structures and for the adoption of human rights-driven business models that do not rely on algorithm targeted advertising". Nevertheless, more could be done to tackle the incentives of actors operating in the 'disinformation industry', who take advantage of platforms to spread disinformation.
- **The 'code' or architecture** refers to technical elements, such as software instructions and protocols that underpin the functioning of online systems. The architecture constrains behaviour by limiting what can be done online, therefore it could be used to make the spread of disinformation more difficult or less opaque. The use of watermarks to flag AI generated content, and the automated identification of mis- and disinformation with the use of AI by fact-checkers, are some ways in which technology could be used to limit the spread of disinformation.

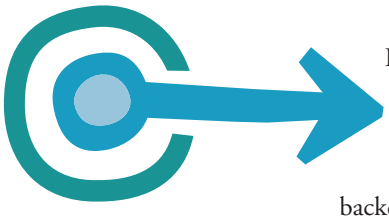
Figure 2. Lessig's socioeconomic theory of regulation



Source: adapted from Lessig, 2006

This multidimensional approach to regulation may also provide a compass to activities in the field of development cooperation. A holistic strategy should assess if and how these four different mechanisms could be leveraged in order to reduce exposure to disinformation and to strengthen social resilience.

4 Political and operational recommendations: Opportunities for development cooperation



Information disorder is undermining important principles of democracy, and contributing to democratic backsliding. Addressing this issue requires more than just reactive measures to curb disinformation; it demands proactive strategies that simultaneously reinforce democratic institutions and societal resilience. Against the backdrop of previous sections, it is possible to distil: a) political recommendations, understood as high-level policy and strategic guidelines, as well as b) operational recommendations, consisting of focused and short-term practical actions aimed at fulfilling the broader political goals. These recommendations, along with their relevance to enhancing the five democratic components outlined in Section 2 – electoral processes, civil liberties and political rights, governmental functioning, civic and political participation, and democratic political culture – are presented in the matrix in the annex.

4.1 Addressing disinformation in different country-contexts

Tackling disinformation in a specific country context is a prerequisite for all subsequent recommendations. This requires taking into account the main elements that underpin disinformation – actors, narratives, facilitators and tools – across democratic, semi-democratic, and autocratic countries (Table 2). From a general perspective, strategies should adapt to the level of political freedom and institutional strength in different countries, and seek to establish long-term partnerships that empower local actors, rather than promoting one-off interventions. Trust-building with local stakeholders is especially important in semi-democratic and autocratic settings where governments may resist external involvement.

Table 2. Examples of specific initiatives of cooperation to combat information disorder that could be undertaken in democratic, semi-democratic, and autocratic countries

Strategy of engagement	Democratic	Semi-democratic	Autocratic
Main focus	Protecting free speech while combating disinformation	Strengthening institutions, safeguarding media independence	Expose disinformation campaigns, put pressure through multilateral channels, support activists
Examples of cooperation	Support to a broader understanding of information integrity; foster the development of laws that address disinformation and uphold free speech; strengthen social resilience; share good practices on dealing with electoral disinformation and foreign interference; push for platform and algorithm transparency; tackle the problem of disinformation entrepreneurs.	Support the development of non-partisan and balanced laws; strengthen institutions responsible for election monitoring, media regulation, and judicial oversight; support journalists and media diversity; support fact-checkers in their range of activities including tech-based innovation against disinformation.	Collaborate with international fact-checking alliances to debunk state-sponsored disinformation, support independent media and activists in exile, use international pressure to push for incremental reforms, specifically for media and journalists, help actors to leverage technology for circumvention of censorship and to protect journalists and activists.

4.2 Strategies to curb disinformation should be multidimensional, taking into account the role of law, norms, market, and architecture.

- Conduct an assessment of existing disinformation mitigation strategies adopted by partner countries in order to evaluate if they take all four dimensions into account. Encourage countries that do not yet have a strategy to adopt this multidimensional approach.
- Support dialogue among multidisciplinary and multistakeholder teams capable of identifying opportunities to leverage the four mechanisms.

4.3 If specific laws on disinformation are put in place, they should observe certain criteria, such as:

a) criminal sanctions should be avoided (especially if vague definitions of disinformation are used), as they constitute serious interference with the freedom of expression and are disproportionate responses in most cases;

b) any restrictions to freedom of expression should be necessary and proportionate to achieve a well-defined legitimate aim, using the least restrictive means;

c) political speech should undergo a high threshold of legality, legitimacy, necessity, and proportionality;

d) consider a legal approach based on risk and harm, rather than truth/falsehood.

- Develop a check-list to assist partner countries in developing laws that are effective and in alignment with human rights and civic freedoms.
- Advocate for the development of laws which are context-specific. The 'Brussels effect' has made the DSA a regulatory inspiration in other regions, but transposing the DSA to different legal and cultural realities may lead to suboptimal solutions, especially in countries that display a relatively weak economic and political position vis-à-vis large global corporations.

4.4 Connecting media and information literacy initiatives with the promotion of democratic principles.

This would ensure that MIL assists in promoting vertical trust in public institutions and the strengthening of democratic values.

- Include elements of a democratic political culture in MIL initiatives.
- Develop a comprehensive MIL strategy, implemented not only in formal education and school curricula, but also among the adult population.
- Use simulations and games that reproduce the difficult choices and high stakes situations that take place during elections.
- Promote AI literacy, demystifying the technology.
- Discuss the responsibility of social media in mainstreaming MIL content through their platforms.

4.5 Adopting a whole-of-government approach to promoting information integrity.

- Promote a whole-of-government approach by supporting coordination mechanisms, strategic frameworks, and capacity-building programmes tailored to the needs of the Public Administration.
- Strengthen institutions responsible for election monitoring, media regulation, and judicial oversight, while raising awareness about good practices in tackling disinformation.

4.6 Strengthening the role of journalists and fact-checkers.

Both types of actors are essential, and play different roles when it comes to curbing information disorder.

- Actively promote the importance of press freedom and media pluralism, including by holding discussions on the ways to financially support journalism.
- Promote the use of technology tools to enhance the safety of journalists and activists.
- Provide additional support to fact-checkers to counter the actions of some platforms (i.e. Meta and X) to undermine this role.
- Make independently fact-checked sources available at scale, and labelled more prominently on platforms.
- Support projects that aim to develop technologies (i.e. automated fact-checking solutions and user-friendly interfaces) to support fact-checking organizations.

4.7 Support social segments which are being particularly targeted by disinformation campaigns.

- Introduce clear regulation and policy against hate speech.
- Oppose the erosion of safeguards against hate speech on social media (i.e. recent decisions by Meta and X).
- Include gender and LGBTIQ+ lenses in programs aimed at combating information disorder. This is particularly important in relation to partner countries in which an erosion of LGBTIQ+ rights can be observed, such as the United States and Argentina.

4.8 Creating the conditions to better distinguish genuine domestic social discontent from astroturfing.

It may be difficult to differentiate external information influence campaigns from legitimate domestic opinion, but the distinction facilitates the choice of the best course of action.

- Mobilise the financial, human, and technical resources to promote long-term monitoring of expressions of dissatisfaction and social cleavages in the information environment that can be easily exploited by disinformation actors. Tackle them through appropriate policy channels. The monitoring adopted by the Swedish PDA provides an example of this approach.

4.9 Maintaining government accountability and a whole-of-society approach in the midst of growing concern over foreign information campaigns.

Problems related to information operations originating from abroad are receiving increasing attention, especially during elections. As a consequence, combating disinformation is beginning to be considered a matter of upholding sovereignty and national security. This triggers the involvement of security-oriented apparatuses within government structures (i.e. ministries of defence and intelligence organisations), which tend to side-step other areas of the government. As a collateral effect, this could lead to a decrease in accountability and public control over decisions.

- Raise awareness about the drawbacks of designing policies against information disorder that are oriented by a 'crisis mindset'. Decisions should be made according to transparent and inclusive decision-making processes, incorporating feedback from all relevant stakeholders, and in line with international human rights.

4.10 Fostering tailored long-term international cooperation.

A closer look at the problem of disinformation reveals that measures taken at the domestic level are important, but may not be sufficient, as the information ecosystem has gained transnational and global dimensions. While countermeasures need to be embedded in local contexts and consider local cultures in order to be effective, they must also be coordinated and integrated at the international level.

- Design follow-up mechanisms for the implementation of principles and guidelines such as the UN Global Principles for Information Integrity and UNESCO's Guidelines for the Governance of Digital Platforms. This could help in ensuring that all stakeholders are held accountable for their role in combating disinformation and that their actions align with agreed-upon standards and principles.

4.11 Enhancing platforms' transparency obligations, notably by requesting privacy-preserving data access for a diverse range of researchers or civil society organisations.

The interaction between users' preferences (information pull) and algorithms (information push) is not well understood. More data and research are particularly needed in non-Western and non-English speaking countries, where the lack of sufficient content moderation tailored to linguistic and cultural needs makes users particularly exposed to disinformation.

- Support the full implementation of transparency obligations created by the DSA in the context of the EU.
- Support data-sharing arrangements and multinational research groups, in particular those that embody a collaboration between Global North and Global South.
- Leverage ongoing cooperation initiatives in the Global South, such as the United Nations Development Programme (UNDP) Action Coalition on Information Integrity, as a way to promote more inclusive data analysis.

4.12 Curbing financial incentives to disinformation.

This could provide an important leverage against disinformation by making it less profitable. Proposed solutions to the problem usually fall short of tackling the business models of online platforms, and of taking into consideration the growing disinformation industry.

- Echo calls for "a fundamental shift in incentive structures" of online platforms made by the UN Guiding Principles for Information Integrity.
- Support the creation of a multistakeholder working group which includes businesses, the technical community, civil society and academia, to discuss alternative incentive structures.
- Raise awareness among key actors who could play a role in mapping and dismantling transnational networks of disinformation entrepreneurs (i.e. investigative journalists, civil society, governments, law enforcement, parliamentarians), and supporting initiatives of collaboration.

Annex I: Approaches to protect components of democracy

Democracy component	Elements within component in closest relation to disinformation	Facilitating conditions to curb disinformation	Strategy	Mechanisms	Actions
Electoral process and pluralism	Free elections (freedom to vote and a range of choice); equal campaigning opportunities; transparent financing of political parties; clear, established, and accepted constitutional mechanisms for the orderly transfer of power; opposition's realistic prospect of winning elections	Vertical trust in public institutions; separation of powers; checks and balances; democratic culture	Media and information literacy	Social norms, architecture	■ Include elements of a democratic political culture in MIL initiatives ■ Use simulations and games that reproduce the difficult choices and high stakes situations that take place during elections ■ Promote AI literacy, demystifying the technology
			Content policy	Laws, market, architecture	■ If laws against disinformation are put in place, political speech should undergo a high threshold of legality, legitimacy, necessity, and proportionality ■ Platforms should ensure that election content is treated with the highest standards of care ■ Label political advertising more clearly with details of the sponsor, advertising spending, and display period ■ ackle the professionalization of 'disinformation services' ■ Support the work of independent fact-checkers, especially in situations of media capture

Democracy component	Elements within component in closest relation to disinformation	Facilitating conditions to curb disinformation	Strategy	Mechanisms	Actions
Civil liberties	Free electronic media and free print media; freedom of expression and protest; robust media coverage; open and free discussion of public issues, with a reasonable diversity of opinions	Deep-rooted culture of respect for human rights and freedoms across society	Media and information literacy	Social norms, architecture	■ Understand the role of information, digital technology, and media in upholding democracy and human rights ■ Promote the use of technology tools to enhance the safety of journalists and activists
			Content policy	Laws, market	■ Consider a legal approach based on risk and harm, rather than truth/falsehood ■ Any restrictions to FoE should be necessary and proportionate to achieve a well-defined legitimate aim, using the least restrictive means ■ Support for an open and free press, including by promoting new sources of revenue for traditional media ■ Uphold high standards of data protection against practices of profiling ■ Monitor the information ecosystem and create safeguards against capture
The functioning of the government	An effective system of checks and balances on the exercise of government authority; important government functions or policies are not determined by foreign powers; public confidence in government; public confidence in political parties	Vertical trust, checks and balances, due process	Media and information literacy	Social norms	■ A whole-of-government approach to strengthening information integrity within the administration: coordination mechanisms, strategic frameworks, and capacity-building programmes ■ Promote peer-learning and international co-operation between democracies (and semi-democracies) facing similar disinformation threats
			Content policy	Laws, market	■ Strengthen institutions responsible for election monitoring, media regulation, and judicial oversight ■ Adopt active measures to ensure press freedom and media pluralism, including financial/market incentives ■ Establish respect for fundamental rights and freedoms as a pillar of strategies to counter disinformation

Democracy component	Elements within component in closest relation to disinformation	Facilitating conditions to curb disinformation	Strategy	Mechanisms	Actions
Political participation	Voter participation/ turn-out; autonomy and voice of ethnic, religious and other minorities; women in parliament; citizens' engagement with politics; adult literacy	Promoting empowerment and agency of individuals; equality as a social value	Media and information literacy	Laws, social norms	■ Develop a comprehensive MIL strategy, implemented not only in formal education and school curricula, but also among the adult population ■ Discuss the responsibility of social media in promoting MIL through their platforms
			Content policy	Laws, market, architecture	■ Introduce clear regulation and policy against hate speech, regardless of whether the information is true ■ Oppose the erosion of safeguards against hate speech on social media (i.e. recent decisions by Meta and X) ■ Adopt active measures to ensure press freedom and media pluralism, including financial/market incentives ■ Make independently fact-checked sources available at scale, and labelled more prominently on platforms
Political culture	Sufficient degree of societal consensus and cohesion to underpin a stable, functioning democracy; popular support for democracy	Horizontal and vertical trust, individual and societal resilience	Media and information literacy	Norms	■ Include elements of democratic political culture in MIL initiatives ■ Use MIL exercises and games to expose the mechanisms that accentuate polarisation on social media
			Content policy	Market, laws, architecture	■ Promote a fundamental shift in incentive structures in the business models of social media platforms ■ Monitor the expression of dissatisfaction and social cleavages in the information environment, which can be easily exploited by disinformation actors. Tackle them through appropriate policy channels

References

Amnesty International. (2023, May 16). Singapore: Suppression of activists, critics continue ahead of elections. Public Statement. https://www.amnesty.org/ar/wp-content/uploads/2023/05/ASA3667882023ENGLISH.pdf

Andrzejewski, C. (2023, February 15). "Team Jorge": In the heart of a global disinformation machine. Forbidden Stories. https://forbiddenstories.org/team-jorge-disinformation/

Anastácio, K. (2024). Framing disinformation through legislation: Evidence from policy proposals in Brazil. Harvard Kennedy School Misinformation Review, Vol 5 (3). https://misinforeview.hks.harvard.edu/wp-content/uploads/2024/06/anastacio_framing-misinformation-brazil_20240627-2.pdf

Appleman, N. et al. (2022). Truth, intention and harm: Conceptual challenges for disinformation-targeted governance. Internet Policy Review.

Berger, C. et al. (2024). Effectively countering disinformation: Perspectives from every continent. Bertelsmann Stiftung.

Budak, C. et al. (2024). Misunderstanding the harms of online misinformation. Nature, Vol 630 no. 6.

Center for Countering Digital Hate [CCDH]. (2024). Attack of the Voice Clones.

Centro Latinoamericano de Investigación Periodística [CLIP]. (2023). Mercenarios Digitales. July 2023. www.elclip.org/mercenarios-digitales.

Coalizão Direitos na Rede. (2025, January 8). Against the Regression in Meta's Content Moderation and the Attacks on Democratic Regulation of the Digital Space. Statement. https://direitosnarede.org.br/2025/01/08/against-the-regression-in-metas-content-moderation-and-the-attacks-on-democratic-regulation-of-the-digital-space/

Collaboration on International ICT Policy in East and Southern Africa [CIPESA]. (2019). Desposts and Disruptions: Five Dimensions of Internet Shutdowns in Africa. https://cipesa.org/wp-content/files/briefs/report/Despots-And-Disruptions_March-20.pdf

Ferreyra, E. and Segarra, A. E. (2024). Truth in Turmoil: Countering Disinformation in Latin America. Bertelsmann Stiftung.

Freedom House. (2024). Freedom of the Net 2024. https://freedomhouse.org/country/south-africa/freedom-net/2024

Khalid, A. (2019, May 15). Americans can't stop relying on social media for their news. Quartz. https://qz.com/1720695/pew-study-shows-more-americans-rely-on-social-media-for-news

Kitili, J. et al. (2023). Contextualising political advertising policy to political micro-targeting in Kenyan elections. Center for Intellectual property and Information technology Law, University of Strathmore. https://cipit.org/wp-content/uploads/2023/03/Political-Advertising_compressed.pdf.

Lessenski, M. (2018). Common Sense Wanted: resilience to 'post-truth' and its predictors In: The New Media Literacy Index 2018. Open Society Institute, Sofia. https://osis.bg/wp-content/uploads/2018/04/MediaLiteracyIndex2018_publishENG.pdf

Lessig, L. (2006). Code: version 2.0. New York: Basic Books.

Monsees, L. (2023). Information disorder, fake news and the future of democracy. Globalizations, Vol 20. https://www.tandfonline.com/doi/abs/10.1080/14747731.2021.1927470

Nanfuka, J. et al. (2024). A tapestry of actors, attitudes, and impact: Countering disinformation in Africa. Bertelsmann Stiftung.

Statista. (2024). Share of adults who use social media as a source of news in selected countries worldwide as of February 2024. https://www.statista.com/statistics/718019/social-media-news-source/

Sörensen, S. and Pam ment, J. (2023). Operationalising the Framework for Evaluating Capability Against Information Influence Operations A Case Study of the Psychological Defence Agency's Courses. NATO Strategic Communications Centre Of Excellence. https://stratcomcoe.org/publications/operationalising-the-framework-for-evaluating-capability-against-information-influence-operations-a-case-study-of-the-psychological-defence-agencys-courses/295

Swenson, A. and Chan, K. (2024, March 14). Election disinformation takes a big leap with AI being used to deceive worldwide. Associated Press. https://apnews.com/article/artificial-intelligence-elections-disinformation-chatgpt-bc283e7426402f0b4baa7df280a4c3fd

The Economist Intelligence [EUI]. (2023). Democracy Index 2023: The Age of Conflict. https://www.eiu.com/n/campaigns/democracy-index-2023/

United Nations. (2019). United Nations Strategy and Plan of Action on Hate Speech. f

United Nations. (2023). Our Common Agenda Policy Brief 8 Information Integrity on Digital Platforms. https://www.un.org/sites/un2.un.org/files/our-common-agenda-policy-brief-information-integrity-en.pdf

United Nations. (2024). United Nations Global Principles For Information Integrity: Recommendations for Multi-stakeholder Action. https://www.un.org/sites/un2.un.org/files/un-global-principles-for-information-integrity-en.pdf

United Nations Educational, Scientific and Cultural Organization [UNESCO]. (2023). Guidelines for the governance of digital platforms: safeguarding freedom of expression and access to information through a multi-stakeholder approach. https://unesdoc.unesco.org/ark:/48223/pf0000387339

United Nations Educational, Scientific and Cultural Organization [UNESCO]. (2024). User empowerment through media and information literacy responses to the evolution of generative artificial intelligence (GAI). UNESCO CI/FMD

United Nations General Assembly [UNGA]. (2021). Disinformation and freedom of opinion and expression. Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, Irene Khan. A/HRC/47/25.

US Department of State. (2023, November 7). The Kremlin's Efforts to Covertly Spread Disinformation in Latin America.

Vallance, C. (2023, September 26). Disinformation most active on X, formerly known as Twitter, EU says. BBC. https://www.bbc.com/news/technology-66926080

Vosoughi, S. et al. (2018). The spread of true and false news online. Science Vol 359, Issue 6380. https://www.science.org/doi/full/10.1126/science.aap9559

Wardle, C. and Derakhshan, H. (2017). Information disorder: Toward an interdisciplinary framework for research and policymaking. Council of Europe Report DGI(2017)09.

As a federally owned enterprise, GIZ supports the German Government in achieving its objectives in the field of international cooperation for sustainable development.

Published by:
Deutsche Gesellschaft für
Internationale Zusammenarbeit (GIZ) GmbH

Registered offices
Bonn and Eschborn

Division Global Policy, Governance
Section Governance, Human Rights
Sector Programme Governance

Friedrich-Ebert-Allee 36
53113 Bonn, Germany
T +49 228 44 60-0
F +49 228 44 60-17 66

Dag-Hammarskjöld-Weg 1-5
65760 Eschborn, Germany
T +49 61 96 79-0
F +49 61 96 79-11 15

E info@giz.de
I www.giz.de

Responsible:
Sector Programme Governance

Author:
Marilia Maciel | DiploFoundation | Strasbourg | France

Contact:
sv-governance@giz.de

Design und Layout:
Barbara Reuter | Oberursel | Germany | barbarareuter-grafik@web.de

Photo credits:
Photo on the cover and backpage © AdobeStock, pkproject
Photo on page 5 © AdobeStock, Victor Bertrand | on page 10 © AdobeStock, Thipphaphone | on page 13 © AdobeStock, Peopleimages - AI | on page 18 © AdobeStock, Suriyo | on page 21 © AdobeStock, BAYC | on page 24 © AdobeStock, Noschey | on page 29 © AdobeStock, Stock Source Studio

On behalf of
German Federal Ministry for Economic Cooperation and Development (BMZ)
Division G12: Governance
Bonn

GIZ is responsible for the content of this publication.

Eschborn/Bonn March 2025

Deutsche Gesellschaft für
Internationale Zusammenarbeit (GIZ) GmbH

Registered offices
Bonn and Eschborn

Friedrich-Ebert-Allee 36 + 40
53113 Bonn, Germany
T +49 228 44 60-0
F +49 228 44 60-17 66

Dag-Hammarskjöld-Weg 1 - 5
65760 Eschborn, Germany
T +49 61 96 79-0
F +49 61 96 79-11 15

E info@giz.de
I www.giz.de

On behalf of

Federal Ministry
for Economic Cooperation
and Development