OGMUN - SHSID

Model United Nations Conference

牛津大学-上海中学国际部国际模拟联合国大会

2026

Disarmament and International Security Committee (DISEC)

#BACKGROUND GUIDE

Contents

Letter to Delegates

Dear Delegates,

We are very much looking forward to meeting you all at the SHSID OGMUN conference. DISEC is one of the six main committees at the UN General Assembly, and plays a crucial role in maintaining global peace and security. We are thrilled to have you join us in this committee.

This year, you will have the opportunity to research, debate, and forge international agreements concerning the evolving dangers of information warfare in an increasingly interconnected world. This issue is an ever-growing threat to international security, and impacts all countries, and simultaneously, it is linked to technological advancements that aid development. Finding a way to address a multifaceted and fast-paced security concern is the primary focus of this committee. You will be engaging with one of the most pressing and complex challenges in the international arena, and success will require both attention to detail and creativity.

We are excited to see you rise to the challenge. Whether this is your first Model UN conference or one of many, we are sure you will all contribute meaningfully to the debate and have a memorable experience.

Sincerely,

SHSID OGMUN team

Introduction to the Committee

DISEC, the Disarmament and International Security Committee, is the UN General Assembly's first committee. It was established in 1945 under the ratification of the UN charter and serves as a platform for multilateral dialogue and policy making with the objective of maintaining peace, de-escalating tensions between nations and pursuing the non- proliferation of conventional and non-conventional weapons.

Rooted in Article 11 of the UN Charter, the committee's mandate encompasses the regulation, reduction, and eventual elimination of all forms of armaments. This broad remit empowers it to consider treaties and conventions, mechanisms for verification and compliance, and confidence-building measures among member states.

Unlike the United Nations Security Council, DISEC's resolutions are non-binding, and it cannot enforce punitive measures (e.g. sanctions or military actions). However, it serves as an important medium for debate and policy making among the nations. The history of international arms control as reflected in this committee dates back to the aftermath of World War II, when the horrors of Hiroshima and Nagasaki catalysed global efforts to prevent future nuclear catastrophes. Throughout the Cold War, disarmament negotiations took place primarily through bilateral channels between superpowers, yet the General Assembly's First Committee remained a critical venue for smaller states to voice concerns and propose measures.

Famously, its first resolution in 1946 established a commission to address the challenges posed by the use of atomic energy. It attempted to ensure that nuclear power would be used only for energy purposes, and not for weapons. Landmark agreements, such as the Nuclear Non-Proliferation Treaty of 1968, the Chemical Weapons Convention of 1993, and the Comprehensive Nuclear-Test-Ban Treaty of 1996 bear the imprint of multilateral debates that sharpened the legal language and verification standards essential to their efficacy.

In the modern day, DISEC continues to tackle challenges related to evolutions in warfare.

In our case, the committee will address either the role of non-state actors in conflict zones, or automated disinformation campaigns as a form of information warfare. Emerging security threats compel the committee to address newer domains such as cybersecurity, artificial intelligence, and outer space. Engaging with these cutting-edge topics requires creative diplomacy to bridge divergent views on state responsibility and verification.

Engagement in this committee offers a unique chance to influence the framework of international peace and security. By grappling with arms control's ethical dilemmas and technical complexities, delegates will acquire a deeper appreciation for the delicate architecture of global governance.

Automated Disinformation Campaigns in the Digital Age

Background of the Problem

Since the dawn of the digital age, the nations of the world have fought not just in the physical world, but in the digital world as well. 'Information warfare' has become a tool of nations and private entities to influence citizens and politicians to gain a favourable outcome. 'Disinformation' has become a potent weapon in the arsenal of entities fighting the information war, and with the rise of Artificial Intelligence and Large Language Models, it has become easier than ever to create and disseminate disinformation.

Disinformation refers to deliberately manipulated or falsified information, used as a tool of psychological manipulation in times of peace, crisis, and war. It is created and spread to further an entity's political, economic or military goals.[10] It may be emotionally manipulative, contain cognitive traps or biased information, or be outright false in an attempt to drown out the truth. Disinformation is distinct from 'misinformation', in that it is planned and deliberate, whereas misinformation may arise from rumours or misunderstandings and is without the intention of manipulation.

Disinformation has been prevalent long before the invention of the internet. Although TV and radio were generally subject to greater regulation, intranational and international disinformation existed. Intranational was particularly common within nations, with examples like political slander or the spread of false research (e.g. disinformation of the effects or tobacco products sponsored by cigarette manufacturers, or climate change denial supported by fossil fuel industry corporations). There were successful international disinformation campaigns however. A notable example is "Operation Denver', or 'Operation Infektion', a Disinformation campaign by the USSR's KBG in the 1980s, alleging that HIV/AIDS was a bioweapon developed by the United States. The story received media coverage in 80 countries.[13]

With the rise of information sharing technologies, it has become easier than ever for

extranational entities to pedal disinformation. One such online source of disinformation is ‘fake news’ websites. In 2016, just months before US election day, The Guardian22 and Buzzfeed[18] discovered a network of 150 domains linked to Veles, Macedonia, hosting news sites that spread baseless pro-Donald Trump and anti-Hillary Clinton stories. Many of those hosting the sites were teenagers, and it appears they were motivated by revenue earned from their sites being shared by unwitting Facebook users. The extent to which these stories affected the 2016 American Election is unclear, but the incident highlights how non-state actors can be motivated to engage in perpetuating disinformation.

Just as Facebook has been used to spread fake news articles, X (formerly known as twitter) has been a consistent target of disinformation. Even before the rise of the sophisticated Large Language Models we know today, bots have been used alongside human troll accounts. Mass bot-perpetuated disinformation can drown out truthful information and create Astroturfed (i.e. artificial and planned) controversy where there would be none organically.[17] In 2017 The Washington Post detailed Pro-Kremlin bots on Russian-Language twitter to influence the Russian-speaking diaspora and Russians who did not use state-owned social media sites, which could be easily censored. A specific case is the 2014 shooting down of a Malaysian Airlines flight over southeast Ukraine, which killed 298. The plane was shot down with a Russian-made missile in a region held by Russian-backed separatist forces. Bots originating in Russia falsely spread that the plane was shot down by Ukrainian or American Forces. In this way, a government may use bots to effectively censor information on platforms they do not own and which are extranational.

In recent years, we have also seen the trade of ‘spreader accounts’ emerge. It has become common for tech-savvy individuals, typically in low-income countries where opportunities are sparse, to spend time building up an X account by reposting memes and sexually explicit content. The owners of the ‘spreader accounts’ will sell their account on to a disinformation entity once they reach 100,000 followers, at which point the account will begin posting disinformation to its follower-base.[19]

Bots may also be used to promote already existing content.29 Bots can be used to like and comment on a post en masse, creating engagement which the feed algorithm recognises, and spreads the post to genuine users of the platform. This process is known

as ‘co-ordinated inauthentic behaviour’ (CIB).
Now with the popularisation of sophisticated Large Language Models, bot accounts can more persuasively interact with users on social media. The largest language models like GPT3.5, and Llama have safeguarding in place to prevent harmful responses, but this can be bypassed through the process of ‘finetuning’, where specific biased data is fed to an LLM to increase the likelihood of a desired response.[11]

In Tony Ma’s (2024) study, he found that GPT3.5 produced harmful content when prompted 81.2% of the time after fine tuning, where it only compiled 1.8% of the time normally. It has also become easier to bypass safeguarding with the more recent rise of open-source LLMs, such as DeepSeek, Gemma, and Llama 2.

Disinformation needs to be countered, as it poses a threat to stability and safety. It can be perpetuated by foreign powers to encourage companies and citizens to act against their own interest, or the interest of the state. This can be done with medical disinformation, disinformation that skews a democratic election, or disinformation that leads to hate crime or exclusion against a particular community.

Historical Solutions

In recent years, there has been a surge of interest in combating disinformation, particularly following the disastrous effects of disinformation and misinformation. In particular, efforts have been focused on social media. Tech companies running social media platforms like Alphabet, Meta, and X have made attempts to stem the problem at a low cost, while governments have considered and implemented laws to pass liability for disinformation onto the platforms that host them.

Social Media Innovations

The first major social media platform to introduce third-party fact checking was Facebook in 2016.[14] Facebook partnered with a number of independent fact-checkers including Snopes, Politifact, ABC News, and Factcheck.org. Human fact-checkers from these organisations would flag false stories, and a banner would be added to them linking to a fact-checker's verified article where relevant. The system effectively flagged the most shared false stories, but its reliance on human labour meant it couldn't reach all false posts. The system fell under scrutiny from right-wing groups in the United States, who felt the fact-checking was biased against them.[12] As a result of this (and likely to save costs, offset liability, and cosy up to a newly elected Donald Trump), Facebook (now Meta) discontinued third-party fact checking in 2025, replacing it with the 'Community Notes' system as on X (see below).

The next platform to take anti-disinformation measures was YouTube, which in 2018 introduced algorithmic detection of key words, and bannered videos discussing controversial topics with links to Wikipedia or National/WHO medical sources. The effectiveness of this measure was unclear but conspiracy theorists and content containing disinformation remain prevalent on the platform.

On Twitter (now X), the 'Community Notes' section was introduced in 2021. The Community Notes system involves users leaving suggested notes on a post they believe to contain disinformation, which can then be rated, edited, or have citations added to it by other users. An algorithm then determines which submission is the most appropriate. The selection algorithm takes into account previous actions on the website in an attempt

to determine a user's political persuasion, and would take this into consideration, promoting answers which have consensus along the political spectrum.6 The system is flawed in that its need for cross- political-spectrum consensus means that only issues without political contention can be effectively fact-checked. With a faction of the American-right subverting medical consensus during covid19 for example, this algorithm became an obstacle to medical fact checking. It is also unclear to what extent the black-box algorithm can accurately identify and represent the depth and nuance of political thought. The system also puts power in the hands of amateur fact-checkers, rather than experts, though its appeal is evident in that it shifts liability to the user-base of a platform and does not require paying for professional fact-checkers.

Attempts at Regulation

The issue with attempting to regulate disinformation is infringement of free speech. Governments must draw the line at some point and decide when disinformation becomes so bad that it should be illegal. Below are a few national approaches.

Germany

On the first of January 2018, Germany passed the 'NetzDG' Online hate speech act (Network Enforcement Act), obliging social media platforms with over 2 million users to remove 'clearly illegal' content within 24 hours, and 'illegal' content within 7 days or face a fine of up to 50 million euros. According to a joint Centre for European Policy Studies, and Counter Extremism Project study, three-quarters of reports are not upheld by social media companies, and no company has yet been fined.[8] The policy has not been effective as it cannot realistically levy its fines, and because the majority of disinformation is protected under free speech and so not 'illegal'.

United Kingdom

The 2023 'Online Safety Act' designates more online acts of speech as illegal; namely disinformation that may be harmful to children, and disinformation perpetuated by foreign governments. The monitoring body 'Ofcom' monitors social media platforms for

illegal content and may levy a fine of £18 million or 10% of global revenue if a platform does not have mechanisms in place to prevent disinformation.[7] So far the act has been slow acting, and it will take years to be able to evaluate its effectiveness.[23]

Brazil

Brazilian Congressional Bill No. 2630, also known as the 'Fake News Bill' was a bill intended specifically to fight disinformation. As well as a number of other points, the bill would require that:

- Service providers identify and flag all automated content
- Social media, messaging and search services provide half-yearly transparency reports
- Messaging services would have to define measures to limit forwarding messages to multiple recipients. Artificially boosted content would need to be identified as well
- The dissemination of disinformation which may 'compromise the electoral process of cause physical harm'

These measures go a long way to address disinformation, but concerningly, would also extend the so-called 'parliamentary immunity' to those in public office, potentially enabling disinformation from authorities.[20]

The bill was passed by the Federal Senate (the lower chamber) on the 30th of June 2020, but has not been voted on by the Chamber of Deputies (the upper chamber). As of December 2024, discussion on the bill has been inactive. One reason for this appears to be free speech concerns, with the bill being called the 'censorship bill' by its opponents.[16] The further a government extends the definition of 'illegal' speech, the more power it has to combat disinformation.

Current Situation

Case Study: Anti-Vaccination and Medical Disinformation in the United States

The modern anti-vaccination and scientific skepticism movements in the United States have become prevalent due to a number of domestic and foreign influences. This has had disastrous human consequences, with the USA falling below other developed nations in terms of Covid 19 Vaccination3, and the return of measles to the country by 2019, after the disease had been eliminated from the country for 20 years.

In 2014, protests began in California as the state government moved to tighten the criteria for non-medical vaccination exemptions. The protests were largely motivated by a desire to protect the right to avoid vaccination because of personal belief, although vaccine sceptics were also rallying behind the movement. Protests for vaccine choice followed in Texas in 2015. At the same time, Russian bot and troll networks (with alleged connections to the Russian government) ramped up activities, amplifying the voices of protestors through Coordinated Inauthentic Behaviour.[9]

Russian actors then perpetuated the disproven claim that vaccines cause autism in children. In 2015, accounts linked to Russian troll networks promoted a video of a black minister in Los Angeles at an anti-vaccination rally, who claimed, 'They're not just shooting us with guns. They're killing us with needles". Overlaid on the video was text falsely claiming that immunisations caused autism in 200,000 black children.5 In promoting this video, Russian actors amplified racial tensions and unjust vaccination scepticism by playing into the distrust of the US medical system held by some members of racial minorities following historical abuses and conspiracies.

In 2016, disgraced ex-physician Andrew Wakefield released his film, 'Vaxxed: From Cover-Up to Catastrophe', in which he reiterated the points of his peer-disproven study that the MMR vaccine may cause autism in children, and is linked to intestinal issues. The film also alleged a conspiracy in which the medical community moved to suppress his findings, lending itself credibility by mimicking the style of a legitimate documentary. The film was repeatedly linked to and promoted on social media by Russian trolls.

By 2019, anti-vaccination groups were prevalent online, occupying over 500 websites continually spread by Russian actors, which led to a surge of Anti-vaccination protests in the Southern US in 202,0 contributing to a surge of Covid 19 deaths. Russian actors further attempted to undermine confidence in Western vaccines for Covid-19 in an attempt to promote the Russian-made 'Sputnik V' Vaccine.[3]

At the same time, elements within the US political system perpetuated narratives of disinformation relating to Covid 19. Donald Trump claimed that the malaria drug 'hydroxychloroquine' could be used to treat covid 19, when in reality it was linked to an increase in deaths.[4] Likewise, the White House Coronavirus Task Force was established in January 2020 misled Americans into prematurely dismantling social distancing guidelines with exaggerated claims about herd immunity.

In addition to foreign and internal actors, blame might also be placed on the social media companies which hosted disinformation. YouTube and Twitter (now X) partnered with the WHO in 2020 to delete covid19 misinformation and promote articles from legitimate healthcare organisations, but reports showed they failed to delete misinformation in a timely manner.

Overall, the spread of disinformation online likely contributed to massive deaths during the Covid 19 pandemic. The case of the United States in particular, is interesting as blame can be assigned to foreign and internal actors, as well as social media companies. Delegates must take into consideration the wide range of sources of disinformation when resolving to prevent such disinformation in the future and quelling the medical scepticism movement, which persists in the various countries, including the United States.

Relevant UN Actions

In 2021, Secretary General António Guterres released his report on 'Countering Disinformation' with advice for states and tech companies. He advised that states should:

- Avoid regulations with vague definitions and criminalising legitimate content
- Refrain from internet shutdowns
- Ensure public officials are held accountable for false information
- Involve civil society (e.g. NGOs) in the design of counter-disinformation policies

He further advised that tech companies should:

- Avoid contributing to adverse human rights impacts
- Disclose policies for countering disinformation
- Review their business models
- Ensure transparency and relevant data access
- Ensure content moderation is consistent and sufficiently resourced26

The statement is sensible but reveals an inability on the part of the UN to enforce policy or effect change on its own. More helpful are the direct attempts at countering disinformation.

In 2020, the UN launched its 'Verified initiative', calling on people around the world to become 'information volunteers', sharing UN-Verified science based content in response to Covid 19. This system is commendable in that it tackles the issue of translating content into local languages by recruiting locals themselves.24 However, its feasibility is hampered by the lack of funding behind the programme. Volunteers do not receive compensation and so will always be limited in number.

During the pandemic, the WHO partnered with a number of tech companies, as mentioned in the case study. Despite limited compliance from the tech companies in addressing flagged content, algorithmic detection of COVID-19-related content allowed for the automatic linking to WHO articles.

Finally, the UN has taken measures when dealing with disinformation relating to its peacekeeping operations. Local corporations with peacekeeping operations were slowed down in the Central African Republic, Mali, and the Democratic Republic of the Congo when it was falsely alleged on social media that UN Peacekeepers were trafficking weapons to local land and intending to exploit local resources.[21] In response, the Department of Peace Operations (DPO) implemented a workstream, in which they would monitor disinformation using an app called 'Talkwalker', engage with local communities through WhatsApp, and conduct analysis of disinformation to help understand its source.

These solutions were limited in their success. Talkwalker was often unhelpful as it could not identify a reason or source for the disinformation; it could only flag up frequently occurring stories. The peacekeepers also could not stop these stories. Monitoring and engaging on WhatsApp was also difficult, as it was dependent on the local community agreeing to add peacekeeping agents to their WhatsApp groups, and the analysis portion of the plan was hampered by limited access to local journalists and an over-reliance on social media. The DPO also lacked the resources to monitor local language radio stations.

Proposed Solutions

The use of AI to detect AI and Disinformation

Artificial Intelligence is already more than capable of detecting content generated by Large Language Models. LLM-detecting AI could be deployed by tech companies or third-party monitors to flag content generated by AI. This approach may be particularly potent in combating the AI fake news websites when shared via social media. CIB attacks can also be identified and tracked this way, before being flagged and stopped.

Further, in a study conducted by computational experts Alsmadi, Rice, and O'Brien.3 The study explored the possibility of identifying dis/misinformation with AI based on the content of posts rather than LLM linguistic signifiers like described above. That means it could
potentially be used to identify disinformation written by humans as well. It appears as though human involvement is still necessary however, as is the need for disinformation data sets relating to each individual topic before AI can somewhat reliably detect disinformation. This approach comes with a risk of misidentifying true information.

Both of these approaches also raise the question of what to do when disinformation is found:

- Should it be flagged? Would a human consider the flagging reliable if they knew it was done by AI?
- Should humans be involved, and how can we attract enough people to monitor the entire website? Is that even possible?
- Should content be automatically deleted if it is flagged as containing disinformation? How to ensure this is done accurately?
- How can we determine what is illegal speech when every country has their own definition?

Regulation

As discussed in the 'historical solutions' segment, it may be possible to curb the spread of

disinformation through regulation, though not without its drawbacks. The most common legal precedent in democracies appears to be that only content designated as 'illegal' can be removed by authorities, and in designating content as 'illegal', free speech is threatened.

Disinformation perpetuated by a foreign power against one's own country seems to be the most clear-cut case of uncontroversial illegal content. However, large amounts of disinformation are not aimed at a country in particular, and powers which utilise disinformation do so under 'deniable entities' not officially associated with the government. Illegalising disinformation harmful to human life, or human rights, is not practical eithe,r as it infringes on free speech, and requires evidence of intent to cause harm, which is impractical on large scales.

Medical disinformation is difficult to regulate, as one has to distinguish between falsehood and controversial opinion, the latter of which would be protected under free speech. Some disinformation may be regulated under hate speech, as it has in some countries, but it has also proven a controversial free speech point. Rather than criminalising disinformation speech acts, it may be more feasible to force tech companies to implement regulatory procedures. However, overregulation of foreign companies may cause them to pull out of a region, hurting jobs and markets, while domestic regulation may hinder economic innovation.

Regulation is not impossible as a solution, but a delegate must carefully consider and balance a number of aspects. It must also be noted that DISEC cannot enforce punitive measures against a country or company, and so can only make policy suggestions to nations, or recommendations to the United Nations Security Council.

Resilience

In 'International Disinformation: A Handbook for Analysis and Response, authors Kupiecki, Bryjka, and Chlon suggest that the most effective way to counter disinformation may be to build resilience to disinformation in the global population.10 A delegate may consider setting up an educational body, or a set of resources for global citizens to learn to identify

misinformation, or advise governments to include such programmes in their education systems.

That being said, a delegate must consider how they intend to fund such systems or persuade countries of their relevance. They must also address the issue of access and distribution. Citizens of wealthy liberal democracies with computer and internet access may easily access an online resource, but it's not something most people want to do in their free time, and so it might be ineffective. How can information be distributed to impoverished and rural communities? How can resources be translated into local languages, and who will fund them? Education may be a useful tool, but overreliance on it may distract the committee from other viable solutions.

Questions a Resolution Must Answer

- How much responsibility should private social media companies take for disinformation on their platforms?
- How can we encourage private social media companies to take measures to prevent disinformation?
- How can disinformation be classified and which kinds of disinformation require action?
- Where should the line be drawn between policing disinformation and protecting free speech?
- How can we provide reliable information in low-income areas, and is it even necessary?
- How can disinformation be combated in countries with tightly regulated information, and where the government may be a source of disinformation?
- How can disinformation be monitored in local languages?
- How can we prevent verified information being drowned out in the age of generative AI?
- How can new technologies like AI be used to combat disinformation?

Bloc Positions

Liberal Democracies

Democracies are uniquely vulnerable to disinformation, as voters need access to verified information in order to vote in their own best interest and their country's best interest. For this reason, liberal democracies may be keen to strengthen regulation against disinformation. These countries also have the resources to improve their citizens' resilience to disinformation.

On the other hand, many of the largest social media platforms used around the world were created in these countries, and so there may also be an incentive to protect these companies from regulation for economic benefit.

Medium and High-Income State-centric Countries

For many non-democracies, the state often promotes stronger initiatives in media. Therefore, combating disinformation may mean defending against information warfare from foreign powers in the interest of their own stability. These countries may therefore be particularly interested in measures that protect their domestic platforms, and address the influence of foreign interests.

Low-income countries

Strengthening the state and its stability is often a concern for low-income countries. Disinformation from outside, or from one group within the state**,** may leave a country at risk of political instability. Low-income countries should resist sweeping reforms imposed by more powerful countries, where they might cause instability. At the same time, they may wish for international regulation on foreign tec**h** companies, whom they often have difficulty resisting. Alternatively, they may wish to work with international schemes or companies to make their populations more resistant to disinformation, particularly in areas with inconsistent levels of education.

Suggestions for Further Research

It's recommended to read widely on any topic of debate. For disinformation, it's particularly important to keep with the news and recent publications as this is a rapidly developing topic.

- United Nations. 2025. Countering Disinformation https://www.un.org/en/countering-disinformation
- United Nations. 2024. 'Press Release: UN launches recommendations for urgent action to curb harm from spread of mis and disinformation and hate speech. Global Principles for Information Integrity address risks posed by advances in AI' United Nations Information Service Nairobi https://www.un.org/en/unis-nairobi/press- release-un-launches- recommendations-urgent-action-curb-harm-spread-mis-and
- Kupiecki, R. Bryjka, F. Chlon, T. 2025. International Disinformation: A Handbook for Analysis and Response. Leiden: Brill
- O'Brien, M. Alsmadi, I. 2021. 'Misinformation, disinformation and hoaxes: What's the difference?' Salon https://www.salon.com/2021/05/02/misinformation-disinformation-and-hoaxes-whats-the- difference_partner/

Bibliography

1. Allcott, H. Gentzkow, M. 2017. 'Social Media and Fake News in the 2016 Election' The Journal of economic perspectives Vol.31 (2), pp. 211-235

2. Alsmadi, I. Rice, N. O'Brien, M. 2024. 'Fake or not? Automated detection of COVID- 19 misinformation and disinformation in social networks and digital media' Computation and mathematical organization theory Vol.30 (3), pp.187-205

3. Bosely, S. 2020. 'Hydroxychloroquine: Trump's Covid-19 'cure' increases deaths, global study finds' The Guardian, Accessed 21/12/2025 https://www.theguardian.com/science/2020/may/22/hydroxychloroquine-trumps- covid-19-cure-increases-deaths-global-study-finds

4. Broad, W. 2021. 'Putin's Long War Against American Science' The New York Times, Accessed 21/12/2025 https://www.nytimes.com/2020/04/13/science/putin-russia-disinformation-health-coronavirus.html

5. Czopek, M. 2023. 'Why Community Notes mostly fails to combat misinformation' Poynter, Accessed 21/12/2025 https://www.poynter.org/fact-checking/2023/why- twitters-community-notes-feature-mostly-fails-to-combat-misinformation/

6. Department for Science, Innovation & Technology. 2025. 'What does the Online Safety Act Do?' Gov.uk Accessed 21/12/2025 https://www.gov.uk/government/publications/online-safety-act-explainer/online- safety-act-explainer#how-the-act-will-be-enforced

7. Echikson W. Ibsen, D. 2018. 'Germany's New Anti-Hate Speech Law Needs Teeth If It Has Any Hope Of Stamping It Out Online' Euronews, Accessed 21/12/2025 https://www.counterextremism.com/press/icymi-new-report-germany%E2%80%99s-netzdg-online-hate-speech-law-shows-no-threat-over-blocking

8. Hotez, P 'Anti-science kills: From Soviet embrace of pseudoscience to accelerated attacks on US biomedicine' PloS biology Vol.19 (1)

9. Kupiecki, R. Bryjka, F. Chlon, T. 2025. International Disinformation: A Handbook for Analysis and Response. Leiden: Brill

10. Ma, T. 2024. ‘LLM Echo Chamber: personalized and automated disinformation’ arxiv, Accessed 21/12/2025 https://arxiv.org/abs/2409.16241

11. McMahon, L. Kleinman, Z. Subramanian, C. 2025. ‘Facebook and Instagram get rid of fact checkers’ BBC News, Accessed 21/12/2025 https://www.bbc.co.uk/news/articles/cly74mpy8klo

12. Nato. 2005. ‘Media - (Dis)information - Security’ Accessed 21/12/2025 https://www.nato.int/nato_static_fl2014/assets/pdf/2020/5/pdf/2005-deepportal4-information-warfare.pdf

13. Newton, C. 2016. ‘Facebook partners with fact-checking organizations to begin flagging fake news’ The Verge, Accessed 21/12/2025 https://www.theverge.com/2016/12/15/13960062/facebook-fact-check-partnerships- fake-news

14. O’Brien, M. Alsmadi, I. 2021. ‘Misinformation, disinformation and hoaxes: What’s the difference?’ Salon Accessed 21/12/2025 https://www.salon.com/2021/05/02/misinformation-disinformation-and-hoaxes-whats-the-difference_partner/

15. Paul, K. 2023. ‘Brazil receives pushback from tech companies on ‘fake news’ bill’ The Guardian, Accessed 21/12/2025 https://www.theguardian.com/world/2023/may/03/alphabet-google-fake-news-law

16. Roth, A. 2017. ‘Pro-Putin bots are dominating Russian political talk on Twitter’ The Washington Post, Accessed 21/12/2025 https://www.washingtonpost.com/world/europe/pro-putin-politics-bots-are-flooding-russian-twitter-oxford-based-studysays/2017/06/20/19c35d6e-5474-11e7-840b-512026319da7_story.html

17. Silverman, C. Lawrence, A. 2016. ‘How Teens In The Balkans Are Duping Trump Supporters With Fake News’ Buzzfeed, Accessed 21/12/2025 https://www.buzzfeednews.com/article/craigsilverman/how-macedonia-became-a-global- hub-for-pro-trump-misinfo

18. The Economist Intelligence Unit. 2024. ‘Disinformation is on the rise. How does it work?’ The Economist. Accessed 21/12/2025 https://www.economist.com/science- and-

technology/2024/05/01/disinformation-is-on-the-rise-how-does-it-work

19. Tomaz, T. 2023. 'Brazilian Fake News Bill: Strong Content Moderation Accountability but Limited Hold on Platform Market Power' Journal of the European Institute for Communication and Culture Vol.30

20. Trithard, A. 2022. 'Disinformation against UN Peacekeeping Operations' International Peace Institute, Accessed 21/12/2025 https://www.ipinst.org/wp-content/uploads/2022/11/2212_Disinformation-against-UN-Peacekeeping-Ops.pdf

21. Tynan, D. 2016. 'How Facebook powers money machines for obscure political 'news' sites' The Guardian, Accessed 21/12/2025 https://www.theguardian.com/technology/2016/aug/24/facebook-clickbait-political-news- sites-us-election-trump

22. UK Parliament. 2024. 'Online Safety Act may take years to have noticeable impact despite public's high expectations' Uk Parliament News, Accessed 21/12/2025 https://committees.parliament.uk/committee/127/public-accounts-committee/news/199900/online-safety-act-may-take-years-to-have-noticeable-impact-despite-publics-high-expectations/

23. United Nations Department of Global Communications. 2022. ''Verified' initiative aims to flood digital space with facts amid COVID-19 crisis' United Nations COVID- 19 Response, Accessed 21/12/2025 https://www.un.org/en/coronavirus/%E2%80%98verified%E2%80%99-initiative- aims-flood-digital-space-facts-amid-covid-19-crisis

24. United Nations. 2024. 'Press Release: UN launches recommendations for urgent action to curb harm from spread of mis and disinformation and hate speech. Global Principles for Information Integrity address risks posed by advances in AI' United Nations Information Service Nairobi Accessed 21/12/2025 https://www.un.org/en/unis-nairobi/press-release-un-launches-recommendations- urgent-action-curb-harm-spread-mis-and

25. United Nations. 2025. Countering Disinformation, Accessed 21/12/2025 https://www.un.org/en/countering-disinformation

26. United States Department of State. 1987. 'Soviet Influence Activities: A Report on

Active Measures and Propaganda, 1986-87' Accessed 21/12/2025 https://jmw.typepad.com/files/state-department---a-report-on-active-measures-and-propaganda.pdf

27. Zhdanova, M. Orlova, D. 2017. 'Computational Propaganda in Ukraine: Caught Between External Threats and Internal Challenges' Computational Propaganda Research Project Accessed 21/12/2025 https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2017/06/Comprop-Ukraine.pdf


INTERNATIONAL
DIVISION