

ZU-LAT
Equality and Human Rights

Regulating Social media platforms
in the Age of False Information

Between Freedom of Expression & Democratic Resilience

Adv. Adi Shai | Adv. Tal Hillel | Advisor Prof. Amit Schejter

July 2026

ZULAT
Equality and Human Rights

Regulating Social media platforms in the Age of False Information

Between Freedom of Expression and Democratic Resilience

Author

Adv. Adi Shai

Adv. Tal Hillel

Academic Advisor

Prof. Amit Schejter

Translation

Shoshana Michkin

Graphic Design

Einhar Design

July 2026

© **Zulat for Equality and Human Rights**, All Rights Reserved

Table of Contents

Introduction 04

False Information as a Structural Threat to Democracy 07

The Regulatory Challenge 09

The Normative Foundation of Regulation: Freedom of Expression 10

Key Considerations for Regulation 12

Principles of Regulation 13

1. Equality of Access and Participation 13

2. Transparency 13

3. Pluralism and Exposure to Diversity 14

4. Rationality and Information Quality 14

5. Accountability and Fairness 15

6. Meaningful Civic Oversight 15

Regulatory Principles in Practice: Comparative Models 16

Taiwan 16

The European Union 19

The United States 20

A Democratic Framework for Regulating Israel's Online Environment 22

1. Equality of Access and Participation 23

2. Transparency 24

3. Pluralism and Exposure to Diversity 25

4. Rationality and Information Quality 26

5. Accountability and Due Process 27

6. Meaningful Civic Influence 28

Conclusion 30

Introduction

Over the past two decades, social media platforms have undergone a profound transformation. Initially, they were perceived as a democratic promise, as an engine for opening up the public discourse, decentralizing media power, and expanding opportunities for public participation. Today, however, it is increasingly recognized that these platforms have themselves become key actors, shaping the boundaries of the public discourse, patterns of civic participation, and the quality of information. **The World Economic Forum's Global Risks Report 2024 ranked false information, both deliberate (disinformation) and non-deliberate (misinformation), as the most significant short-term global risk, warning that it threatens to disrupt elections, deepen social polarization, and erode trust in public institutions and authoritative sources of information.**[1] Similarly, according to the OECD, the spread of false and misleading information fuels polarization and undermines both the public discourse and social cohesion.[2]

The key development that turned these platforms into major forces shaping the public sphere was the transition to the interactive web. The World Wide Web laid the groundwork for networked communication, making it possible, for the first time, to build online communities and share information on a large scale without the mediation of traditional intermediaries. With the emergence of Web 2.0 in the early 2000s, the Internet evolved from a collection of websites and information pages into a dynamic space where users themselves create, share, and distribute content. Over time, these platforms developed into arenas where people consume news, form political views, participate in protests, scrutinize institutions, and shape their identities in relation to public issues.[3]

To wit, a global survey conducted in 2024 found that more than 40% of users in selected countries worldwide rely on social media as a source of news.[4] Similarly, in Israel too, the use of social networks as a source for news consumption is relatively widespread. Data from a survey conducted that same year by the Israel Internet Association and the Technion indicates that approximately 58% of Israelis frequently consume news through social media. The survey further found that 78% of secular young adults (both Arabs and Jews) under the age of 39 prefer social media as their primary source of news. As age increases, however, reliance on social media declines, with news consumption shifting primarily to television (71% among those aged 50 and over) and to dedicated news websites and applications (65%).[5]

---

* English-language reference. All other references in this document are in Hebrew.

** Special thanks are due to the Taiwan Mission in Israel for facilitating meetings with Taiwanese experts working on the regulation of social media platforms, the response to disinformation, and the strengthening of democratic resilience. These discussions formed an integral part of the preparation of this document. Zulat is grateful to the Taiwan Mission in Israel for its willingness to cooperate and for making Taiwan's unique experience and expertise in these fields available to us.

1. *The Global Risks Report 2024, 19th Edition: Insight Report, *World Economic Forum* (January 2024).
2. *Facts Not Fakes: Tackling Disinformation, Strengthening Information Integrity, *OECD* (4 March 2024).
3. Jose van Dijck, *The Culture of Connectivity: A Critical History of Social Media, *Oxford University Press* (24 January 2013).
4. *Share of Adults Who Use Social Media as a Source of News in Selected Countries Worldwide as of 2024, *Statista* (June 2026).
5. *Measuring Disinformation in Israel: Exposure, News Consumption Preferences, and Media Trust, *Israel Internet Association* (14 December 2025).

The migration of the public discourse into diverse Internet platforms fundamentally changed the nature of the democratic debate. In the past, much of this debate took place through regulated intermediaries, such as professional journalism, broadcasting organizations, political parties, academic institutions, and civil society organizations. Although imperfect, these institutions employed mechanisms of accountability, including editorial oversight, fact-checking, institutional accountability, and reputational considerations. Today, by contrast, a significant portion of the public debate, news distribution, and agenda-setting occurs within private commercial spaces controlled by social media platforms.
Thus, alongside the promise of greater pluralism and increased access for groups previously excluded from centers of media power, this new environment has also created serious risks to the basic conditions of a healthy democratic discourse. Erroneous or misleading information can now spread rapidly, on a massive scale and often virally, without effective oversight mechanisms, and in many cases with built-in incentives encouraging its dissemination. Platforms do not merely "host" the discourse, but they rather rank, filter, recommend, remove, shape incentives, and steer public attention.

The power of these platforms stems from their modus operandi. They are not neutral environments, but are built around the competition for human attention, a scarce and highly valuable resource. In this reality, the value of content is measured not necessarily by its accuracy, contribution to the public debate, or democratic worth, but by its ability to evoke an immediate emotional response.[6] **Herein lies the problem posed by false information: untruthful or misleading content, particularly when it is emotional, sensational, polarizing, or conspiratorial, is almost perfectly aligned with the economic logic of social media platforms.** Such content arouses curiosity, suspicion, outrage, or fear; it is easy to share, is brief and easily digestible, and it fits seamlessly into the rapid flow of users' feeds. In this sense, false information is an almost inevitable byproduct of a system where attention is the primary resource, engagement is the currency, and virality is the central mechanism of distribution.

6. Brendan Nyhan, Jaime Settle, Emily Thorson et al, *Like-Minded Sources on Facebook Are Prevalent But Not Polarizing, *Nature* (27 July 2023).

Democracies around the world are today grappling with the question of how to protect the digital public sphere without hurting freedom of expression or entrusting the state with the power to determine what constitutes political truth. This is the central question addressed in this paper. In this context, Taiwan's experience provides an important point of reference for Israel. Having for years faced persistent disinformation and foreign influence campaigns, it developed an approach aimed at strengthening democracy without excessively expanding the state's censorship powers. The Taiwanese model demonstrates that combating disinformation cannot be reduced to prohibitions, content removal, or the state's role as an arbiter of truth. Instead, it requires building broad democratic resilience based on transparency, cooperation with civil society, stronger media literacy, greater access to data for research, and a clear distinction between legitimate political expression versus coordinated foreign manipulations. Against this background, the following chapter examines the spread of false information as a structural threat to democracy, and the principles that should guide the democratic regulation of social media platforms.

False Information as a Structural Threat to Democracy

False information impairs the democratic discourse far beyond misleading individual users. It is not merely a matter of a person coming across false information and making decisions accordingly, but rather a broad phenomenon with systemic consequences for the very conditions that make a healthy democratic discourse possible.

It compromises the conditions necessary for free and informed political decision-making, increasingly impedes meaningful political debate on a shared factual foundation, and erodes confidence in institutions, information intermediaries, and the very ability of separating truth from manipulation. Moreover, in the age of social media platforms - where information is disseminated through algorithms that reward engagement, emotional appeal, and polarization - false information often enjoys a structural advantage in both distribution and amplification. Consequently, the damage extends beyond the false content itself to the architecture of the public discourse, the shaping of public opinion, and the integrity of the democratic process as a whole.

A report published by the State Comptroller in July 2026 adds an important dimension to understanding this structural risk, and shows how the inherent features of the digital environment - anonymity, identity concealment, rapid and low-cost dissemination, microtargeting, use of bots and fake identities, and echo chambers - enable hostile actors to create a false image of public opinion, amplify exposure to particular messages, and distort authentic discourse. Therefore, the problem is not merely the existence of false content, but also the ability of the platform architecture to transform such content into a tangible political force. The challenge is further intensified by the development of generative artificial intelligence tools. These technologies enable hostile actors to produce, rapidly and at minimal cost, texts, images, videos, and audio files tailored to different target audiences, overcome linguistic and cultural barriers, and create large numbers of fictitious identities that impede identification of the source. Thus, artificial intelligence not only enhances the sophistication of deception but also expands its scale, speed, and supposed credibility.[7]

Against this backdrop, **the following section briefly outlines the principal threats that the dissemination of false information poses to the fundamental foundations of democracy.**

First, the shared factual foundation is undermined. Democracy does not require consensus, but it does require the ability to disagree on the basis of at least a minimal common set of facts. Citizens may differ on solutions, values, or policy, but meaningful political disagreement depends on sharing at least some basic facts or being able to verify them via consensual methods. When different groups are exposed to separate, contradictory, and sometimes entirely false information ecosystems, the ability to conduct a rational debate diminishes. The problem is not merely that citizens are "wrong," but that the conditions enabling meaningful public deliberation are compromised.

7. National Response to Foreign Influence in the Digital Sphere, *State Comptroller* (7 July 2026).

Second, institutional trust is impaired. A democratic society depends on a basic level of confidence in sources of knowledge, electoral processes, public institutions, and the ability to distinguish truth from error and deception. The widespread and persistent propagation of false information does not only persuade people to believe false claims, but it also erodes trust in the entire system of information intermediaries. When everything is perceived as subject to manipulation, when every fact is dismissed as propaganda, and when the very possibility of verifying claims is undermined, democracy is weakened from within.

Third, polarization is intensified. Social media users tend to consume content that reinforces their views and to connect with people who share similar opinions. This process creates "echo chambers" that reinforce existing beliefs, reduce exposure to opposing viewpoints, and increase hostility between groups. When false information is injected into this environment, at times with strategic intentions, it is no longer evaluated simply as facts, but it instead becomes intertwined with identity, group affiliation, and political allegiance. In such circumstances, refutation ceases to function as a factual correction and may instead be perceived as an attack on the group itself.

Fourth, political equality is subverted. Visibility within social media platforms is not distributed equally, but is mediated by recommendation engines, popularity metrics, financial resources, favorite formats, patterns of sharing, and unequal access to digital tools. Circumventing the notion of transparency, actors with greater resources, technological expertise, access to influence networks, bots, fake accounts, or advertising budgets can amplify their visibility, including the visibility of false information, thereby undermining the principle of equality of political voice and influence.

Fifth, the legitimacy of the democratic process is emasculated. Elections, referendums, public debate, and parliamentary deliberation all presuppose a basic capacity for persuasion, criticism, and informed choice. When the information environment in which political opinions are formed is structured in ways that favor manipulation, deception, and confusion, the formal procedures of democracy may remain intact, but their substantive legitimacy is compromised. Citizens may continue to vote, yet it becomes increasingly unclear whether their choices reflect the product of free and informed democratic deliberation.

The Regulatory Challenge

As social media platforms have become the primary arena for shaping public opinion, political participation, information dissemination, and public influence, it has become increasingly important to examine their impact on society, its foundational values, and the democratic social order. Accordingly, **the central question that should guide the regulation of social media platforms is whether they create the conditions necessary for freedom of expression and, above all, for the public discourse that sustains the democratic order.**

This question leads us to examine the principles that should guide the regulation of social media platforms – in other words, the normative characteristics of the kind of public discourse that democracy requires, commonly referred to as democratic dialogue.

The regulatory challenge becomes particularly acute when the focus of the debate is disinformation. On the one hand, the absence of regulation leaves platforms with enormous power to shape the public discourse. When private platforms determine how content is ranked, recommended, amplified, or removed based on their own economic or other interests and without meaningful transparency, the result is not neutrality but a regulatory vacuum that effectively places the governance of the public discourse in private hands. On the other hand, excessive regulation raises equally serious concerns. Any attempt by the state to determine what is true or false, or what content may or may not be disseminated, risks sliding into censorship, suppressing legitimate disagreement, and expanding state control over the public discourse. This concern is particularly severe in the political sphere, where factual, normative, and interpretive disputes are deeply intertwined, making it difficult to distinguish between disinformation and genuinely contested claims in a clear, timely, and politically neutral manner.

Herein lies one of the central lessons of the Taiwanese model, which emphasizes the need to build democratic resilience around the public discourse itself through clearly delineated definitions, rapid and transparent clarification mechanisms, cooperation with civil society organizations and independent fact-checkers, stronger media literacy, and access to data that enables research and public oversight. In this sense, the challenge is not to choose between the absence of regulation and state control over the truth, but rather to design a regulatory framework that protects the conditions necessary for a democratic discourse without turning the state into the ultimate arbiter of political disagreement.
Accordingly, sound regulation should primarily center on the structure of the information environment. It should focus less on direct control of the content, and more on transparency, accountability, disclosure of content-distribution mechanisms, labeling requirements, due process, access to data for research, and the management of systemic risks.

The Normative Foundation of Regulation: Freedom of Expression

Any discussion about regulation must begin with freedom of expression, not merely as an individual right but as a fundamental constitutional principle that serves the core purposes of a democratic society. Three principal rationales for freedom of expression are commonly recognized: the pursuit of truth, individual autonomy and self-fulfillment, and the democratic rationale.

The pursuit of truth: Freedom of expression is a prerequisite for the competition of ideas, exposure of error, and the advancement of public understanding. Under this view, a free society requires openness to discussion, debate, refutation, and correction, because even mistaken opinions may help reveal the truth by showing what is false, prompt critical reflection, or strengthen the truth by confronting it.

The question of whether the conditions of that competition still help discover the truth becomes more acute in the digital era. When content-distribution mechanisms reward not the quality of arguments, credibility, or fact-checking, but rather brevity, dramatics, emotionalism, shock value, and virality, the marketplace of ideas operates under distorted conditions, fundamentally impairing the rationale for the pursuit of truth. In this context, false information challenges not only the content of the public discourse but also the very infrastructure that is supposed to sustain a "free marketplace of ideas."

Individual autonomy: This rationale focuses on both the speaker and the listener. Freedom of expression protects an individual's ability to form opinions, express identity, criticize institutions, encounter diverse viewpoints, and reach independent judgments through free choice in an environment that supports an informed public discourse. When the information environment is shaped by manipulation, biased interface design, concealment of material information, dark patterns, and incentives that encourage rapid, emotional, and uncritical responses, the individual's capacity to exercise independent judgment is diminished. The U.S. Federal Trade Commission (FTC) has indeed described dark patterns as deceptive design practices likely to manipulate consumer decisions and undermine informed choice.[8] Disinformation thus impairs individual autonomy not only by conveying false data, but also by creating an information ecosystem that weakens a user's capacity for critical thinking and independent judgment.

8. Christine Lyon and Emily Parfitt, *'Bringing Dark Patterns to Light': The FTC Warns Companies Against Manipulative User Interfaces, *Freshfields* (1 November 2022).

The democratic rationale. This is widely regarded in academic literature as the central and most influential rationale for freedom of expression. The underlying premise is that without freedom of expression, citizens cannot meaningfully develop and evaluate political views, build consensus, sharpen disagreements, or participate in shaping their shared political life, and that in the absence of such participation, the legitimacy of democratic decision-making itself is compromised.

In the past, safeguarding democratic participation meant limiting the state's power to suppress speech. However, that is no longer the case today. Social media platforms have shifted the focus away from state censorship toward the quality of the new public sphere itself - namely, whether the digital infrastructure where the public discourse now takes place genuinely enables participation, exposure to diverse viewpoints, listening, criticism, and influence. When that environment is saturated with false information, polarization, manipulation, and lack of transparency, freedom of expression can no longer fulfill its primary purpose and its democratic mission.

Key Considerations for Regulation

The regulation of social media platforms in a manner consistent with the public interest, first and foremost the democratic rationale for freedom of expression, should be guided by several key considerations:

First, the paradox: On the one hand, social media platforms have created unprecedented opportunities to deepen democracy by expanding access, enabling individuals and groups to make their voices heard, bypassing traditional gatekeepers, facilitating civic mobilization, allowing rapid public engagement, and seemingly fostering greater pluralism. On the other hand, these very same characteristics – of content, speed of dissemination, absence of major barriers to entry, and interactivity - also create an environment particularly conducive to the spread of misinformation and disinformation.

Second, erosion of traditional gatekeeping mechanisms: Unlike the era of traditional media, where professional bodies were responsible for filtering and verifying information, the digital environment enables the rapid and large-scale dissemination of false or misleading information without effective oversight. This danger is compounded by the fact that such content is especially well suited to the economic logic of social media platforms. Emotional, sensational, and often conspiratorial content tends to receive broader distribution because it generates higher levels of user engagement. Algorithms designed to maximize user attention and time spent on a platform are therefore inclined to amplify such content even when its credibility is highly questionable.

Third, the incentive structure: False content earns wide dissemination, which accords it a semblance of credibility or importance. Users repeatedly encounter it through shares, comments, and algorithmic promotion, while the platform itself - even if it does not endorse the message -- effectively becomes a mechanism for amplifying it and benefiting financially from the resulting volume of dissemination. This is not merely a moral failing but a structural one. Therefore, disinformation is not (only) a problem of "irresponsible users," but is the fuel that feeds the engines of the digital communications marketplace.

Fourth, the attention economy: Disinformation accelerates an attention economy driven by emotional escalation, echo chambers, polarization, and content manipulated for strategic purposes by political, economic, and even state actors. The spread of false information can be used to influence public opinion, hurt political opponents, destabilize society, undermine trust in elections, and erode confidence in public institutions. In an era in which information can be distributed rapidly and cheaply, and where the boundary between authentic media and manipulation has become increasingly blurred, the effectiveness of these tactics has grown significantly.

The central challenge, therefore, is not simply how to remove false content, but how to redesign the operating conditions of social media platforms to align with democratic values. This is where the model of democratic dialogue becomes relevant.

Principles of Regulation

The quality of the digital public sphere should not be assessed solely by the level of participation it enables, but above all by the quality of the discourse it fosters. Drawing on deliberative conceptions of democracy, public discourse can be described as a normative framework premised on a set of cumulative conditions, without which meaningful democratic dialogue cannot exist. These conditions extend beyond the formal protection of freedom of expression and require the codification of the media environment in a manner that promotes an equal, open, critical, and data-based discourse.

1. Equality of Access and Participation

Freedom of expression essentially consists not only of the right to speak but also of the opportunity to be heard. In the digital environment, where distribution and visibility depend on algorithms and commercial considerations, there is a risk of systematic marginalization of certain groups, speakers, or viewpoints. Democratic regulation should therefore examine not only the rules governing access to platforms, but also the actual conditions of participation: Who gets exposure, which voices are pushed to the margins, and whether structural, economic, technological, or algorithmic barriers undermine equality of opportunity in the public discourse.

The connection to false information is clear. It does not spread simply because it exists, but because it often enjoys an edge within the visibility systems. Consequently, equality is not merely a matter of the formal right to publish content, but also a question of how visibility and influence are allocated. If false information enjoys a structural advantage over reliable and verified information, a significant gap emerges between formal freedom of expression and effective equality in the discourse.

2. Transparency

Transparency is a fundamental prerequisite for democratic accountability. In the social media realm, where decision-making processes are driven by complex and often opaque algorithmic systems, transparency requires far more than the publication of general terms of service. Meaningful disclosure is required on how decisions are made in regard to the visibility, distribution, and moderation of content. Platforms should clearly explain their enforcement principles, the operation of their recommendation engines, the role of sponsored content, and the principal criteria that shape the digital agenda.

Transparency is particularly important in the context of false information. As long as users, researchers, and the public do not know why certain information reached a particular audience, how it was promoted, what role sponsored advertising played, how borderline content was classified, or what is the criteria for reducing visibility or removing content, it is impossible to seriously assess the mechanisms that amplify false information.

3. Pluralism and Exposure to Diversity

Democracy is not merely a forum for expressing opinions, but also a space where different viewpoints encounter one another. Without exposure to diverse perspectives, the public discourse may collapse into "echo chambers" that intensify polarization and weaken the capacity for dialogue. Therefore, the objective of regulation in this context is not to impose an artificial balance among viewpoints, but rather to create structural conditions that encourage exposure to diverse ideas.

With respect to disinformation, the problem is not only that false information is disseminated, but also that it is often circulated within closed information environments where corrective information rarely penetrates. Exposure to a diversity of sources, decrease of mechanisms that reinforce ideological isolation, and promotion of access to content that crosses political and social boundaries are therefore important tools not only for pluralism, but also for reducing the vulnerability of democratic discourse to false information.

4. Rationality and Information Quality

Democratic dialogue presupposes the existence of a public sphere where claims can be evaluated, information can be assessed, and arguments can be challenged and refuted. This does not mean imposing an "official truth", but rather requires strengthening the conditions that enable users to distinguish between different types of information: editorial versus sponsored content, verified information versus rumors, and substantiated claims versus misleading assertions.

In this context, regulation may include source labeling, contextual information, disclosure of distributors, digital forensics tools, and measures to counter practices that impair critical judgment. Disinformation directly undermines this principle by exploiting an environment that encourages rapid and emotional reactions rather than verification, reflection, and deliberation.

5. Accountability and Fairness

When platforms function as gatekeepers of the public discourse, their decisions to remove content, reduce visibility, or block users carry significant public consequences. Therefore, internal discretion alone is insufficient and procedural safeguards are required to ensure fairness. Users should be notified of enforcement actions, understand the reasons behind them, and have access to meaningful avenues of appeal. In addition, external oversight mechanisms are necessary.

This is especially important in the context of disinformation, because efforts to combat it can easily evolve into arbitrary and opaque control. It is precisely because disinformation poses a genuine challenge that it is essential to ensure that the measures addressing it do not undermine freedom of expression without adequate scrutiny.

6. Meaningful Civic Oversight

If social media platforms constitute part of the infrastructure of public discourse, citizens cannot remain merely passive users. Democratic regulation should facilitate access to data, promote independent research, and ensure meaningful opportunities for public scrutiny of systemic risks. Without access to relevant data, it is impossible to understand how disinformation spreads, who benefits from it, which groups are most vulnerable to it, or how recommendation engines operate.

Taken together, these principles reflect a shift from viewing regulation as a response to isolated harms toward a systemic approach focused on shaping the conditions of the public discourse itself. Accordingly, the democratic dialogue model does not seek to control content, but rather to create a codified environment enabling an open, diverse, critical, and data-based discourse.

Regulatory Principles in Practice: Comparative Models

Taiwan

Taiwan is a prominent case study of how a democratic state can address disinformation without compromising the fundamental principles of freedom of expression, political pluralism, and open public debate. It views the fight against disinformation as an ongoing process of public learning, based on broad cooperation between government and civil society, with the aim of strengthening social cohesion and resilience over time.

The Taiwanese model rests on the premise that restrictions on individual freedom of expression in a democratic system must be approached with particular caution, and that the law alone cannot effectively address the challenge of disinformation. Thus, Taiwan has developed a decentralized, multi-layered system that combines targeted legislation addressing cases of clearly defined public harm, rapid and transparent government responses, monitoring of influence operations, independent fact-checking, cooperation with digital platforms, and media literacy education. This division of responsibilities reflects the view that resilience to disinformation cannot be achieved through a single authority determining what is true, but rather through a system of actors who balance, complement, and, in some cases, scrutinize one another. [9]

The first principle of the model is targeted legal intervention. Rather than creating a universal offense of disseminating false information, the law focuses on specific circumstances in which such dissemination may cause tangible public harm. In 2019, for example, laws pertaining to such areas as disaster prevention, communicable diseases, food safety, and agriculture were amended to allow sanctions to be imposed for the dissemination of rumors or false information, [10] thus binding legal intervention to concrete harm and a specific public interest.

Alongside sector-specific legislation, Taiwan enacted the Anti-Infiltration Act, which was signed into law and entered into force in January 2020. Section 1 states that its purpose is to prevent infiltration and intervention of "foreign hostile forces," to ensure national security and social stability, and to safeguard the sovereignty and liberal democratic constitutional order of the Republic of China.[11]

Taiwan has also pursued legislation aimed at imposing greater responsibility on online platforms and requiring them to put in place stronger mechanisms to curb the spread of false information. A prominent example is the "Deepfake Clause," enacted in May 2023, which allows candidates to ask law enforcement authorities to investigate suspected fake audio or video recordings featuring them. Platforms are required to remove content verified as deepfake and may be imposed fines for noncompliance.[12]

---

9. Kai-Shen Huang and Muyi Chou, *Resilience in Truth: Public-Private Collaboration in Taiwan's Response to Disinformation, *DSET* (14 April 2025).

10. *Taiwan Counters FIMI – Governmental and Parliamentary Responses, *IORG* (31 October 2024).

11. *Anti-Infiltration Act, *Laws & Regulations Database of Taiwan (Ministry of Justice),* (last amended 15 January 2020).

12. See footnote #10.

The Public Officials Election and Recall Act and the Presidential and Vice Presidential Election and Recall Act also underwent significant revision. The legislative amendments focused on two main objectives. "Real Name Authentication for Election Advertisements," which requires all media outlets to disclose the sponsor, funder, and other relevant information in any election or recall advertisements they publish; and "Preventing Foreign Interference in Elections," which mandates newspapers, magazines, broadcast media, and online service providers to verify whether the sponsors of election or recall advertisements are foreign entities. If a sponsor is identified as a foreign influence, media and service providers are legally prohibited from displaying these advertisements.[13] While these requirements are not directly targeted at disinformation campaigns, they help trace funding sources and identify malicious actors, preventing foreign entities from improperly influencing or distorting public opinion through financial intervention.

However, legislation is only one layer of the Taiwanese model, and not necessarily the most important one. The second principle is to rapidly fill the public space with reliable, accessible, and verifiable information. Beginning in 2018, the government developed coordination mechanisms enabling government departments and public agencies to respond swiftly to rumors and issue clarifications within their respective areas of responsibility. Since 2019, the Taiwanese government has operated an "immediate clarification mechanism," under which government departments actively monitor information across online platforms, broadcast media, and print sources to quickly identify and address disinformation. A press release must be issued within four hours and posted in a dedicated real-time clarification section on official websites. In addition, executive agencies have developed a standardized process to guide administrative departments in effectively responding to and debunking disinformation.[14] During the COVID-19 pandemic, for example, the Central Epidemic Command Center held frequent press conferences, presented data, answered questions, and addressed rumors.[15] This practice reflects a broader approach based on the understanding that disinformation thrives in information vacuums. Accordingly, the continuous, timely, and transparent supply of reliable information is itself a tool for strengthening democratic resilience.

The third principle is collaboration with civil society, which plays a central and independent role in the Taiwanese model. Fact-checking organizations and civic initiatives, including Taiwan FactCheck Center, MyGoPen, and Cofacts, verify information, debunk rumors, and make fact-checking accessible to the public, thereby complementing the efforts of the government and digital platforms.[16]

---

13. Ibid

14. Ibid

15. Irving Yi-Feng Huang, *Fighting COVID-19 Through Government Initiatives and Collaborative Governance: The Taiwan Experience, *Public Administration Review*, Vol. 80, No. 4, 2020, pp. 665–670.

16. Chiaoning Su and Wei-Ping Li, *Three Musketeers Against Mis/Disinformation: Assessing Citizen-Led Fact-Checking Practices in Taiwan, *Taiwan Insight* (31 March 2023).

The fourth principle is to hold digital platforms accountable while ensuring that they do not assume governmental functions without adequate oversight. The NCC's 2024 Communications Market Report notes that as early as June 2019, the Taipei Computer Association (a Taiwanese non-governmental organization dealing with the technology sector) released Self-Discipline Guidelines for False Information, which were subsequently endorsed by such digital platforms as Yahoo, Meta, LINE, and Google.[17]

The fifth and most long-term pillar is strengthening the public's own ability to recognize and address misleading information. Media and digital literacy programs focus on identifying manipulations in images and videos, understanding AI-generated content, and providing people with the tools to assess sources and information critically. These initiatives are based on cooperation among public institutions, researchers, schools, and civil society organizations. The underlying rationale is that not all misleading information is disseminated as part of an organized campaign, and not everyone who shares it does so maliciously. Much of its spread is driven by citizens who believe in the content and pass it on. Consequently, a system that does not strengthen citizens' ability to assess information independently will remain perpetually dependent on enforcement, content removal, and post-factum debunking.[18]

From this perspective, the Taiwanese model can be understood as an approach to democratic resilience whose aim is not to prevent all instances of false information - a goal that may well be unattainable in a free society - but rather to reduce the conditions that allow such information to take root, spread, and exert significant influence on the public discourse. The Taiwanese experience thus offers a key lesson: addressing a problem as complex as disinformation requires a multifaceted approach rather than reliance on any single, one-dimensional solution.

17.*2024 Communications Market Report, Chapter 4, discussing platform self-regulation against disinformation, *National Communications Commission* (December 2024).

18.You-Hao Lai, *Disinformation and Civil Defence: How Did Taiwan's Civil Society Counter Foreign Information Manipulation, *DSET* (5 February 2024).

The European Union

The European Union currently represents the most comprehensive effort to regulate social media platforms through a broad institutional approach. The Digital Services Act (DSA) applies to online services, including social networks, online marketplaces, app stores, and other platforms, and seeks to protect fundamental rights, establish a framework of transparency and accountability, and reduce systemic risks. The European Commission asserts that, under the DSA, very large online platforms are subject to additional obligations, including systemic risk assessments, risk mitigation measures, advertising transparency requirements, and access to data for accredited researchers.[19]

With regard to false information, the significance of the European model lies in the fact that it does not focus solely on unlawful content in the narrow sense, but also on systemic risks. In this respect, it reflects the realization that disinformation is not only an isolated false post, but rather a product of platform structures and technological and economic incentives.

In terms of the principles of democratic dialogue, the European model shows several significant points of alignment. First, it explicitly adopts the principle of transparency. Second, it comes close to the principle of accountability via oversight mechanisms, risk assessments, and review procedures. Third, it incorporates the principle of research and public oversight by creating mechanisms for data sharing with accredited researchers.

However, the alignment is not complete. The European model still focuses largely on risk mitigation and user protection, and less on reinforcing positive conditions required for a democratic discourse, such as genuine equality of influence, effective exposure to diversity, or the correction of visibility hierarchies. In addition, the expansion of supervisory powers raises questions regarding the balance between protecting democracy and institutional intervention in the freedom of expression.

In addition to the DSA, the Digital Markets Act (DMA) too is indirectly relevant to the discussion about false information, as it seeks to address the structural power of digital gatekeepers. The DMA does not directly regulate content, but rather addresses the concentration of digital power and, in this sense, the concentration of intermediary power over the flow of information.[20]

19. *Digital Services Act, *European Commission* (19 October 2022).

20. *Digital Markets Act, *European Commission* (14 September 2022).

The United States

Unlike the European Union, the United States does not currently have a comprehensive federal framework for regulating social media platforms in terms of transparency, systemic risks, or structural obligations relating to the public discourse. The American approach largely rests on the principle of minimal intervention, due in part to Section 230 of the Communications Decency Act, which set out to promote the development of the Internet and offer broad protection to interactive service providers regarding third-party content and certain content moderation practices.[21]

In the context of false information, this means that the American model is particularly cautious about any regulation that could be perceived as government interference in the public discourse. The advantage of this model lies in its strong protection against state censorship, while its disadvantage is that it leaves open a significant vacuum in those areas that the democratic dialogue model identifies as critical: transparency, accountability, exposure to diversity of views, access to research, and recognition of the systemic risks created by social media platforms.

It is true that certain sector-specific and targeted regulatory frameworks exist, such as Children's Online Privacy Protection Act (COPPA), California Consumer Privacy Act (CCPA), and the Federal Trade Commission's (FTC) enforcement operations against deceptive practices and dark patterns. However, these measures do not amount to a coherent and comprehensive framework for regulating the platform-based public sphere. [22]

In other words, the United States provides robust protection for the negative liberty of expression, but offers only a partial response to the question of whether the conditions in the public sphere actually enable meaningful democratic expression.

An illustration of this can be found in the U.S. Supreme Court's 2024 ruling in Moody v. NetChoice, which demonstrates the constitutional and regulatory complexity entailed in attempts to regulate social media platforms and their content management discretion.

---

21. *Communications Decency Act, 47 U.S.C. §230, *Columbia University* (1996).

22. *Children's Online Privacy Protection Act (COPPA), *National Archives* (updated 9 July 2026); *California Consumer Privacy Act (CCPA), *California Department of Justice* (updated 13 March 2024); 'Bringing Dark Patterns to Light': The FTC Warns Companies Against Manipulative User Interfaces, *Freshfields* (1 November 2022).

The case concerned two state laws enacted in Florida and Texas in 2021, in the wake of the surging spread of online disinformation during the 2020 presidential election and the COVID-19 pandemic. These state laws sought to restrict platforms' authority to remove, label, or reduce the visibility of content based on its viewpoint or subject matter. The legislative intervention raised a fundamental question: Can a state compel private platforms to disseminate or remove content contrary to their own editorial judgment? Opponents of the laws argued that decisions regarding content moderation, ranking, and removal are integral aspects of freedom of expression and the editorial autonomy of platforms. The U.S. Supreme Court did not resolve the underlying constitutional conflict on the merits. Instead, it held that a more thorough and comprehensive constitutional review of the regulatory schemes was required and remanded the cases to the lower courts. This decision illustrates one of the central dilemmas in regulating the digital public sphere: The need to address disinformation, incitement, and harmful content without disproportionately impairing freedom of expression or the autonomy of private actors that mediate the public discourse. [23]

23. *Moody v. NetChoice, *Supreme Court of the United States* (1 July 2024).

A Democratic Framework for Regulating Israel's Online Environment

Israel is currently characterized by a significant regulatory gap in the area of social media platforms.[24] Citing a mapping project conducted by the National Cyber Directorate in 2024, a State Comptroller's report notes that Israel lacks regulation requiring social media platforms to manage risks, report transparently, or follow specific guidelines on content removal, except where the content itself constitutes a violation of the law. This means that the primary arena where a substantial portion of the Israeli public discourse occurs nowadays operates without a binding framework of transparency, accountability, and systemic risk management.[25]

As a result, the legal response to their activities relies primarily on existing general legal frameworks, such as consumer protection laws, tort laws, defamation laws, and privacy protection laws. These bodies of law are important, but they were not originally designed to address the unique characteristics of social media platforms, including algorithmic recommendation engines, viral dissemination, targeted advertising, large-scale data collection and processing, private content moderation, and the ability of platforms to effectively shape the conditions of the public discourse.

Israel's Supreme Court has sought to partially address this legislative vacuum. Recent rulings reveal a trend toward recognizing Israel as the proper venue for resolving disputes between global social media platforms and Israeli users, while preventing foreign jurisdiction and choice-of-law clauses from emasculating mandatory rights afforded under Israeli law.[26] Nevertheless, however significant this development may be, it addresses individual cases rather than establish a comprehensive regulatory framework.

Consequently, the existing regulatory vacuum reduces platforms' incentives to invest adequately in creating a safe, reliable, and fair online environment. When their liability for harms resulting from user-generated content remains highly limited, their economic model continues to be driven primarily by maximizing attention, engagement, and rapid content distribution, even when that content is false, inciting, harmful, or polarizing. This creates a clear regulatory failure in which platforms reap the economic benefits generated by user activity and engagement, while Israeli society as a whole bears the social and democratic costs of the spread of false information.

The discussion below follows the six regulatory principles presented in this paper. We will first examine how the structure of online platforms undermines each principle in practice, followed by a concise list of the regulatory measures required to protect it, adapted to Israel's legal, social, and geopolitical context. Taken together, these measures outline a principles-based regulatory framework suited to the Israeli context and informed by the comparative experience described above.

---

24. Two governmental committees were established in 2021 to look into digital regulation: the Ben-Horin Committee, appointed by the Ministry of Communications, and the Davidi Committee, appointed by the Ministry of Justice. The Davidi Committee did not progress to the publication of a draft bill, while the Ben-Horin Committee submitted its recommendations in December 2022 in: Report of the Advisory Team to the Minister of Communications on the Regulation of Digital Content Platforms, *Ministry of Communications* (14 December 2022).

25. National Response to Foreign Influence in the Digital Sphere, *State Comptroller* (7 July 2026).

26. Ruling of Supreme Court, sitting as Court of Civil Appeals, in Case 6992/22 Agoda Company v. Shai Tzvia, *Supreme Decisions* (27 May 2024).

1. Equality of Access and Participation

The principle of equality of access and participation is undermined in the digital environment in two cumulative ways. The first concerns inequality in visibility. While the formal right to publish content is open to all users, actual visibility is determined by algorithmic recommendation engines that prioritize certain characteristics of content (length, visuals, engagement potential, and compatibility with the profile of the end user) and of the content creator (account history, size of audience, and investment in sponsored promotion). The result is a structural asymmetry: actors with greater resources, technological expertise, access to influence networks, advertising budgets, or the ability to operate bot networks and fake accounts can acquire large-scale public visibility, while private citizens, minority groups, and non-commercial organizations are pushed to the margins. In the Israeli context, this harm is particularly significant for Arabic speakers and peripheral communities, whose needs are often inadequately served by content moderation systems.

The second form of harm concerns the way false information exploits this very structure, when emotional, polarizing, or sensational content enjoys a structural advantage in visibility over verified and evidence-based information. This creates a gap between formal equality of access and substantive inequality of participation, where voices that serve important democratic purposes struggle to be heard versus content that has gained a competitive advantage not because of its quality, but because of its compatibility with the logic of the attention economy.

Here is where a key lesson from the Taiwanese experience becomes relevant. Disinformation and foreign influence spread not only through single pieces of content, but also via distribution mechanisms, coordinated networks, and technological advantages. Therefore, protecting political equality requires more than formal equality regarding the right to publish: It calls for identifying and mitigating coordinated inauthentic behavior and foreign manipulation within the information environment.

Proposed Rules for Upholding the Principle of Equality

- Prohibit platforms from giving a structural advantage to sponsored content over organic content unless it is clearly and prominently labeled as such.
- Require identification and mitigation of fake networks and accounts, coordinated inauthentic behavior, and foreign information manipulation and interference (FIMI) based on periodic risk assessments.
- Require the establishment of moderation systems, reporting channels, and appeal mechanisms that provide an equivalent level of service in both Hebrew and Arabic, including moderators proficient in each language and automated detection systems trained on adequate language-specific datasets.
- Prohibit systematic over-enforcement against specific linguistic or identity-based communities.Transparency reports should provide enforcement data disaggregated by language and be subject to external review.

2. Transparency

Transparency, which is a prerequisite for any form of public or regulatory oversight, is violated when the decision-making processes that shape the digital public sphere take place within internal algorithmic systems that remain hidden from view. Although platform community guidelines are generally public, they represent only the tip of the iceberg. Actual decisions are made through a combination of algorithms, internal ranking systems, human moderation decisions, and business considerations that are not available for external review. In the context of false information, this lack of transparency is not neutral. As long as external observers are unable to examine how a particular piece of information was distributed, who promoted it, what role sponsored promotion played, and which types of content were subject to automatic reductions in visibility, it is impossible to conduct a meaningful discussion about the mechanisms that amplify false information.

The transparency that exists today, where it exists at all, is often little more than a semblance, relying on aggregated reports, vague statements of principle, and claims that cannot be independently verified. Therefore, what is required is substantive and structured transparency, based on uniform standards that facilitate processing, external review, and effective oversight.

In this context, the European Digital Services Act (DSA) offers an important model of structural transparency by requiring platforms to disclose the key parameters underlying their recommendation engines and to enable users to understand and influence how content is presented to them. With regard to sponsored content, the DSA imposes enhanced transparency obligations, including clear labeling that content is advertisement, disclosure of its sponsor, and information regarding the parameters used to target it to the user.

Proposed Rules for Upholding the Principle of Transparency

- Require platforms to publish the operating principles of recommendation and ranking engines, including the main weighting factors and how borderline content is handled.
- Require personal explanation: Every user is entitled to know why specific content was presented to them, whether it was sponsored, and which profile characteristics influenced its presentation.
- Require periodic transparency reports based on a uniform, machine-readable standard, including language-disaggregated data on content removal, reduction of visibility, account blocking, categories of violations, and error rates.

3. Pluralism and Exposure to Diversity

Algorithmic recommendation engines are inherently designed to prioritize content that aligns with the user's preferences. The more successful these systems are in their narrow objective of maximizing time spent on the platform and user engagement, the more they reduce exposure to intellectual, political, and social diversity. This is a structural byproduct of attention optimization, with "echo chambers" emerging over time where users are primarily exposed to content that reinforces their existing views and encounter fewer opposing perspectives or ideas in which they have not previously expressed interest.

This phenomenon changes the way political views develop. Instead of being shaped via encounters between competing perspectives, political identities increasingly form within relatively closed environments. In the context of false information, the harm is twofold: not only is it propagated, but it is spread within information environments that are nearly impenetrable to refutation. A factual disclaimer is not merely rejected, but at times may be perceived as an attack on an in-group. Israel has an additional distinctive feature in this context: intense political polarization, social stratification, and fragmentation into identity-groups (ultra-Orthodox, seculars, Arabs, settlers, LGBTQ+, and immigrants) intensify this phenomenon and create a situation in which different communities effectively inhabit different information environments.

Proposed Rules for Upholding the Principle of Pluralism

- Require platforms to provide the option of a chronological or a non-personalized feed, free of charge and without reducing service accessibility or functionality.
- Require clear labeling of borderline content that has undergone shadow banning (the practice of silently restricting the reach or visibility of a user's account or content on a social media platform without notifying them), thereby allowing users to decide whether they nevertheless wish to access such content.

4. Rationality and Information Quality

The principle of rationality assumes that the public sphere enables citizens to distinguish between different types of information, evaluate claims, and exercise critical judgment. In the online environment, this principle is eroded when the same feed sequentially presents professional news reports, personal posts, sponsored advertisements, viral videos, satirical rumors and content, often without sufficient visual distinction. In such circumstances, the context required for critical evaluation (the identity of the source, its authority, and whether the content is sponsored) can easily be lost.
It is exactly in this environment that false information thrives, exploiting the jumble of formats, rapid pace, and absence of clear indicators of reliability. Thus, rather than posing as a well-established newspaper, false information can simply appear in users' feeds in a format that fits the surrounding context.

In this regard, the Taiwanese experience demonstrates that strengthening information quality requires building a public infrastructure that helps citizens evaluate information independently. Taiwan's response to disinformation relies, among other things, on independent fact-checking organizations, institutional clarification mechanisms, cooperation with civil society, and investment in media literacy. The lesson for Israel is that democratic regulation should focus on providing context, clear labeling, and strengthening people's ability to assess information, rather than on the state directly determining what constitutes political truth.

Proposed Rules for Upholding the Principle of Rationality

- Establish regulatory incentives to encourage cooperation between social media platforms and time-honored professional journalistic sources that provide public-value content.
- Require clear and distinct visual labeling of sponsored content, political advertisements, AI-edited content, and content identified as misleading by fact-checking organizations.
- Require contextual labeling of content particularly susceptible to disinformation (such as wartime news, public health information, and election-related content) with links to verified sources.

5. Accountability and Due Process

Social media platforms function as gatekeepers of the public discourse, but they do so without the accountability mechanisms that characterize traditional public gatekeepers. When a public authority makes a decision that infringes upon an individual's rights, it is required to provide reasons, allow the affected person to respond, and offer an opportunity for appeal. When a platform removes content, reduces its visibility, or suspends an account, there is often no reasoned explanation, no factual review, and no effective appeal mechanism. Decisions may at times be made by automated systems, by human moderators operating under tight time constraints, or on the basis of reports submitted by users or authorities that are themselves not necessarily bound by transparent procedures.

In the context of false information, the problem takes on an additional dimension. Precisely because disinformation poses a genuine challenge, platforms face public, political, and regulatory pressure to act swiftly. This pressure creates an incentive to over-remove borderline content without sufficient scrutiny. Such over-removal not only infringes on individual freedom of expression but also creates a chilling effect on public discourse, as speakers preemptively self-censor for fear of sanctions. Even where the removal of content is substantively justified, the absence of due process undermines the legitimacy of the decision itself.

The European Digital Services Act (DSA) provides an instructive example of how platform accountability requires the codification of procedural safeguards. Under this framework, platforms are required to provide reasoned explanations for decisions concerning the removal, blocking, or reduced visibility of content or accounts; establish internal complaint mechanisms; ensure that decisions on appeals are not made solely through automated processes; and provide access to an out-of-court dispute settlement body.

Proposed Rules for Upholding the Principle of Accountability and Due Process

- Require a reasoned notification that includes the relevant facts, the rule allegedly violated, and available appeal options in every case of content removal, reduction of visibility, or account restriction.
- Establish an accessible multilingual appeals mechanism with predefined response times and the right to approach an independent out-of-court dispute settlement body.
- Apply Israeli law and jurisdiction as the default framework for disputes between Israeli users and platforms (consistent with the emerging approach in Israel's Supreme Court jurisprudence).
- Require documentation and publication of cases where content removal or restrictions followed a request by a state authority, including the identity of the authority and the grounds for the request.

6. Meaningful Civic Influence

The principle of civic influence is based on the premise that citizens, researchers, journalists, and civil society organizations are not merely end users, but partners in shaping the public sphere. In the traditional media environment, this partnership could be exercised through journalistic scrutiny, academic research, the work of human rights organizations, and public advocacy. In the social media environment, however, this capacity is significantly weakened, primarily due to lack of access to information: a substantial portion of the data concerning how platforms operate, patterns of content distribution, algorithmic effects, and the structure of public risks remains controlled by the companies themselves. In recent years, this situation has deteriorated further: API interfaces that were previously relatively accessible to researchers have been closed or restricted, the cost of accessing data has increased, and platforms often prefer to share information only with internal researchers, or not to share it at all.

In the context of false information, without access to data it is impossible to meaningfully research how it spreads, which groups are more vulnerable to it, the interplay between organic dissemination and coordinated distribution, and how platforms' own mitigation mechanisms operate. This situation creates significant knowledge gaps that impair the ability of a democratic society to conduct a data-based discussion about its own public sphere.

In this regard, it is imperative to incorporate the Taiwanese lesson that a strong civil society is a prerequisite for a democratic response to disinformation. Accordingly, alongside platform regulation, stable and independent public funding should be ensured for research, fact-checking, and information literacy education.

Proposed Rules for Upholding the Principle of Civic Influence

- Establish an accreditation mechanism for qualified researchers (academics, civil society organizations, and fact-checking bodies), granting access to both public and non-public data for the purpose of researching systemic risks.
- Require platforms to maintain structured API interfaces for researchers, without additional fees and without arbitrary restrictions on the volume of requests.
- Require publication of the findings of public research conducted on this basis, free of charge and in an accessible format.
- Provide stable public funding for independent research and fact-checking infrastructure in Israel, while maintaining institutional safeguards ensuring independence from private capital, social media platforms, and government authorities.

The six principles and the rules derived from them do not purport to constitute an exhaustive list of all the measures required for regulating social media platforms, but are rather intended to serve as a normative foundation of core principles that any appropriate regulatory framework should satisfy. Taken together, these principles establish a coherent framework based on equal opportunity for visibility, transparency in distribution and governance mechanisms, pluralism in exposure to content, rationality and reliability in the information environment, accountability in the exercise of power and authority, and public access to research, oversight, and public scrutiny. Embedding these principles in legislation constitutes an important condition for creating an effective, democratic regulatory framework adapted to the challenges that the digital sphere presents to Israeli society and its democratic system.

Conclusion

The phenomenon of misinformation and disinformation lies at the heart of the democratic challenge in the current digital era. It creates a structural problem arising from the way in which the digital public sphere is organized today, via algorithmic recommendation engines, the attention economy, commercial incentives to maximize engagement, targeted advertising, coordinated influence networks, and foreign interference efforts. Within this environment, false, emotional, sensationalist, and polarizing information often enjoys a structural advantage over verified, complex, and evidence-based information. Accordingly, the democratic harm caused by false information impacts on the very foundational conditions of the public discourse itself.

From the perspective of freedom of expression, this is a particularly complex challenge. Freedom of expression is intended not only to protect an individual's right to speak, but also to ensure the conditions through which citizens can access information, evaluate claims, criticize government, participate in shaping political life, and engage in democratic disagreement on the basis of a shared factual foundation. When the space in which the public discourse takes place is governed by opaque mechanisms that amplify polarization, undermine public trust, reduce exposure to diverse viewpoints, and make it more difficult to distinguish reliable information from manipulation, the democratic promise of freedom of expression is also compromised. The central question, therefore, is not merely how to remove false content, but how to shape the operating conditions of social media platforms so that they serve an open, equal, transparent, and information-based democratic discourse.

A comparison between the European Union, the United States, and Taiwan demonstrates that there is no single model that can be adopted in its entirety. The European Union offers the most comprehensive framework currently in existence, in the form of the Digital Services Act, which is based on transparency, accountability, systemic risk assessment, and a certain degree of access for researchers. The United States, by contrast, continues to emphasize strong protection of freedom of expression against state intervention, but has a hard time addressing the structural power of online platforms and their impact on the conditions of the democratic discourse.

Taiwan offers a multi-layered, decentralized model that combines sector-specific legislation, rapid government responses, monitoring of coordinated activity, cooperation with platforms, independent fact-checking, and information literacy education. The 2024 election experience demonstrates that this combination can curb certain rumors and limit identifiable effects, but does not establish an overall and sustained reduction in disinformation.

The central lesson from Taiwan's experience is that democracy cannot confront disinformation through the state alone, and certainly not by turning the state into the ultimate arbiter of political truth. A multi-layered framework is needed, one that distinguishes legitimate political expression from coordinated, covert, foreign, or manipulative activity; responds swiftly but on the basis of evidence; and strengthens independent fact-checking, transparency, information literacy, and access to research. At the same time, success in addressing a particular rumor should not be equated with a systemic solution. Rather, the impact of regulatory measures on public beliefs, trust in institutions, and freedom of expression must be assessed over time. In the Israeli context, given the country's polarization, social divisions, security threats, and vulnerability to foreign influence, what is needed is a democratic, gradual, and carefully calibrated regulatory framework that addresses the conditions shaping public discourse, incorporates independent monitoring and oversight mechanisms, and does not rely on direct state control over content.

Zulat institute works to promote a policy of equality and human rights and to restore the legitimacy of the human rights discourse in Israel.

President | Zehava Galon
Executive Director | Einat Ovadia

All rights reserved to **Zulat for Equality and Human Rights ©**

zulat.org.il

ZU-LAT Equality and Human Rights