



© 2026 Balla and Zeko. This article follows the  Open Access policy of CC BY NC under CC v 4.0.



Submitted: 29/05/2026 - Accepted: 30/07/2026 - Published: 28/09/2026

Balancing Act: Constitutional Safeguards for Free Expression Against AI-Driven Disinformation in Emerging Democracies

Rezana Balla¹, Fjoralba Zeko^{1*}

Department of Legal and Political-Administrative Sciences
Faculty of Social Sciences, Albanian University
*f.zeko@albanianuniversity.edu.al

DOI: 10.26417/1htucyt1

Abstract

Artificial intelligence-driven disinformation has emerged as one of the most significant threats to democratic governance, social trust, and freedom of expression in the digital age. While digital technologies facilitate participation and access to information, algorithmic amplification, deepfakes, and targeted misinformation increasingly distort democratic communication and weaken public confidence in institutions. This article examines how constitutional democracies are responding to these challenges through legal regulations, digital governance, and educational resilience strategies. Using Albania as an empirical case study situated within the broader European constitutional framework, the article analyzes the interaction between freedom of expression, platform governance, and systemic disinformation risks. The study applies a qualitative doctrinal and comparative methodology, combining European Court of Human Rights jurisprudence, European Union regulatory instruments, and interdisciplinary literature from sociology, media studies, and digital governance. The article develops the concept of “epistemic harm” to explain how AI-driven disinformation undermines the collective conditions necessary for democratic deliberation. It further argues that digital literacy, media education, and institutional transparency are essential complements to legal regulation. The findings demonstrate that emerging democracies face particular vulnerabilities due to weaker informational infrastructures and lower resilience against coordinated manipulation. The article concludes that effective responses require an integrated model combining constitutional safeguards, platform accountability, media literacy, and democratic education.

Keywords: Artificial Intelligence; digital governance; epistemic harm; media literacy; democratic resilience; platform regulation; social trust.

1. Introduction

The rise of artificial intelligence (AI) and algorithmically mediated communication has transformed the structure of public discourse worldwide. Social media platforms, recommender systems, and generative AI technologies increasingly shape how individuals consume information, form political opinions, and participate in democratic life (Barendt, 2015). Although digital communication technologies have expanded (Balkin, 2018) access to information and participation, they have also enabled the rapid spread of disinformation, conspiracy narratives, and manipulative content capable of undermining democratic institutions and social cohesion.

The global spread of AI-driven disinformation has raised significant concerns regarding electoral integrity, public health, institutional legitimacy, and the protection of fundamental rights (De Gregorio, 2022). The COVID-19 pandemic, electoral misinformation campaigns, and the emergence of deepfake technologies have demonstrated that disinformation is no longer an isolated communication problem but a structural democratic challenge. In particular, algorithmic amplification systems prioritize emotionally engaging content, often privileging sensational or misleading information over verified factual discourse.

This transformation presents a fundamental constitutional dilemma: How can democratic societies protect freedom of expression while simultaneously safeguarding the informational conditions necessary for democracy itself? Classical liberal free speech theory traditionally assumes that truth emerges through open debate in a relatively neutral communicative environment. However, digital infrastructures are no longer neutral spaces. Instead, they are governed by private platforms whose algorithms shape visibility, attention, and information flows.

This article argues that AI-driven disinformation should be understood not merely as false content but as a systemic phenomenon capable of generating “epistemic harm,” namely the degradation of the shared conditions necessary for democratic reasoning and knowledge formation. In this context, constitutional democracies increasingly face the challenge of balancing freedom of expression with the protection of democratic resilience and informational integrity.

Rather than focusing exclusively on a single national constitutional framework, this article situates Albania within a broader international and European discussion concerning digital constitutionalism, platform governance (Organisation for Economic Co-operation and Development [OECD], 2024), and democratic resilience in emerging democracies. Albania serves as an empirical case study through which broader structural and comparative insights can be developed. Its post-communist transition, evolving media environment, and European integration process provide a

valuable context for examining how emerging democracies respond to AI-driven disinformation.

The article develops three central claims:

1. AI-driven disinformation creates systemic forms of epistemic harm that affect democratic legitimacy and social trust.
2. Emerging democracies face heightened vulnerability due to weaker institutional and informational infrastructures.
3. Sustainable democratic responses require an integrated framework combining constitutional safeguards, platform regulation, media literacy and digital education.

The article contributes to contemporary scholarship by integrating legal analysis with social science and educational perspectives, emphasizing that disinformation cannot be addressed solely through legal restriction but requires broader societal resilience strategies.

2. Research Methodology

This article employs a qualitative doctrinal, comparative, and interdisciplinary case-study methodology. Albania is used as an empirical case study to examine broader constitutional and societal challenges associated with AI-driven disinformation in emerging democracies.

The paper combines three methodological dimensions:

2.1 Doctrinal Legal Analysis

The research examines primary legal materials, including:

- European Court of Human Rights (ECtHR) jurisprudence;
- Court of Justice of the European Union (CJEU) case law;
- European Union regulatory instruments such as the Digital Services Act (DSA) and the Artificial Intelligence Act (AI Act);
- Albanian constitutional provisions and Constitutional Court jurisprudence.

The doctrinal analysis evaluates how freedom of expression is interpreted in the context of digital communication, platform governance and systemic disinformation risks.

2.2 Comparative and International Perspective

Although Albania constitutes the principal empirical setting, the article adopts a broader comparative perspective by situating Albanian developments within European and international debates concerning digital constitutionalism. The study compares:

- European risk-based regulatory approaches;
- Emerging democratic systems characterized by transitional institutions and digital vulnerabilities.

This comparative approach allows the article to identify transferable lessons relevant beyond the Albanian context.

2.3 Interdisciplinary Social Science and Educational Perspective

To align legal analysis with broader social science concerns, the study integrates scholarship from sociology, media and communication studies, political science, digital governance, and educational theory and media literacy research. The article specifically examines the social consequences of AI-driven disinformation, including declining institutional trust, political polarization, democratic fragmentation, and the heightened vulnerability of young audiences to manipulative digital content. Special attention is given to the role of digital literacy and education in strengthening democratic resilience.

2.4 Justification for Selecting Albania as a Case Study

Albania provides a unique, representative empirical landscape for four foundational reasons:

1. **Post-Authoritarian Transition:** The nation is continually navigating a complex post-authoritarian constitutional transition that deeply shapes its current baseline rights framework.
2. **European Union Integration:** Albania is undergoing active European Union integration processes, which introduce external legislative pressures and prompt legal alignment.
3. **Platform Dependency:** The society has experienced a rapid digital transformation marked by a deep, unmediated dependency on global private social platforms.
4. **Structural Infrastructural Weaknesses:** There are persistent domestic structural vulnerabilities regarding media pluralism, independent journalism, and institutional public trust.

These intersecting factors make Albania an ideal representative of the broader challenges facing emerging democracies that must navigate swift digital transformations while preserving constitutional legitimacy and free speech.

3. Ai-Driven Disinformation As A Social And Democratic Challenge

3.1 From False Information to Systemic Manipulation

Disinformation in contemporary digital environments functions not simply as isolated false statements but as a coordinated and algorithmically amplified

phenomenon embedded within online communication systems. Social media platforms optimize engagement through recommender systems that reward emotional intensity, outrage, and polarization (Sunstein, 2018). Consequently, misleading information frequently circulates faster and more effectively than verified factual content.

Disinformation is no longer merely false content but a systemic phenomenon driven by algorithmic amplification (Sunstein, 2021). Platforms optimize engagement, often privileging emotionally charged or misleading content (Napoli, 2019). This results in epistemic asymmetries that undermine democratic deliberation (Benkler et al., 2018).

In the contemporary information age, science has an essential role in countering fake news and disinformation—phenomena that threaten public trust, democratic processes, and social stability. Fake news refers to intentionally false or misleading information presented as factual, while disinformation involves the deliberate creation and dissemination of such content to deceive, manipulate, or influence opinions. Their rapid diffusion through social media, where every user becomes a potential “publisher,” increases both their reach and their danger. Understanding these concepts is a first step toward building a resilient and educated society.

Science fights fake news by promoting structured thinking, evidence-based reasoning and ethical responsibility. Ethics as a scientific discipline strengthens awareness by helping individuals evaluate the moral implications of sharing unverified information. Before forwarding a message or posting online, ethical reasoning asks: Is this true? Could this harm someone? Does sharing benefit or mislead the public? By integrating ethical principles with scientific methodology, individuals learn not only how to verify information but also why verification matters.

The importance of community awareness cannot be understated. Fake news often spreads faster than fact because it is emotional, sensational, and aligns with existing biases. Communities with lower educational levels or limited digital skills are particularly vulnerable, while teenagers—highly active online—are exposed to misinformation that can affect mental health, political opinions, health decisions, and social behaviors. Examples include viral health myths, manipulated political content, or misleading advertisements targeting youth. These groups are often unaware of algorithmic influences, echo chambers, and the commercial motives behind online engagement.

The dangers of fake news lie not only in false beliefs but in their capacity to shape public opinion, influence elections, undermine scientific consensus, propagate hate speech, and disrupt social harmony. With widespread access to smartphones, AI-generated content and high-speed information exchange, the speed at which false narratives spread increases the responsibility of each individual. Thus, not spreading fake news is a civic and ethical duty.

A central challenge in understanding digital information disorders is distinguishing between different categories of false or harmful content and evaluating the intent behind them.

Research in media studies and political communication demonstrates that algorithmic amplification contributes to the formation of echo chambers and fragmented public spheres. Users increasingly consume information aligned with pre-existing beliefs, reducing exposure to diverse perspectives and weakening democratic deliberation.

In this context, AI technologies intensify existing democratic vulnerabilities through automated content generation, deepfake technologies, micro-targeted political messaging, synthetic audio and visual manipulation, and personalized algorithmic persuasion.

The societal implications extend beyond misinformation itself. AI-driven disinformation contributes to declining trust in institutions, increased political polarization, and the erosion of shared factual realities necessary for democratic governance.

3.2 Social Trust, Polarization, and Democratic Vulnerability

Social trust represents one of the essential foundations of democratic societies. Democratic systems depend upon confidence in institutions, electoral processes, journalism, and public expertise. However, AI-driven disinformation systematically undermines these foundations.

In emerging democracies, where institutional trust may already be fragile, digital disinformation creates heightened vulnerabilities. False narratives can intensify political divisions, weaken confidence in democratic institutions, fuel conspiracy theories, undermine public health campaigns, and delegitimize elections and judicial systems.

The spread of misinformation during global crises has illustrated the dangers of digitally amplified falsehoods. Health misinformation, manipulated political content and synthetic media increasingly influence public behavior and collective decision-making.

These developments demonstrate that freedom of expression debates can no longer exclusively focus on state censorship. The architecture of digital communication itself has become a constitutional issue.

3.3 The Role of Personal Data Protection

AI-driven disinformation is fueled by systemic privacy failures. Malicious actors exploit unregulated cross-border personal datasets to micro-target vulnerable populations. Behavioral profiles allow algorithms to deliver tailored disinformation directly to users most susceptible to specific emotional triggers.

Strong data protection frameworks (such as the General Data Protection Regulation [GDPR]) function as structural shields against disinformation. Limiting behavioral tracking directly reduces the efficiency of personalized manipulation networks.

4. Education, Media Literacy, And Democratic Resilience

4.1 Media Literacy as Democratic Protection

Legal regulation alone cannot fully address the structural dynamics of AI-driven disinformation. Democratic resilience also requires educational responses capable of strengthening critical thinking and digital awareness.

Media literacy plays a central role in enabling individuals to:

- evaluate information critically;
- identify manipulative content;
- recognize algorithmic bias;
- distinguish verified sources from misinformation;
- understand the societal consequences of online behavior.

Young people are particularly vulnerable to algorithmic influence due to intensive social media exposure (UNESCO, 2021). Educational systems therefore increasingly face the responsibility of integrating digital literacy into civic education.

4.2 Educational Institutions and Digital Citizenship

A central challenge in understanding digital information disorders is distinguishing between different categories of false or harmful content and evaluating the intent behind them. Contemporary literature categorizes them into:

- **Misinformation** – false information shared in *bona fide*, without the intention to harm.
- **Disinformation** – false information shared *in mala fide*, deliberately intended to deceive, manipulate, or cause harm.
- **Malinformation** – genuine information used *maliciously* (e.g., doxxing, private data leaks).

The distinction between *mala fides* (bad faith) and *bona fides* (good faith) is not merely semantic—it grounds responsibility. A key question arises: Can disinformation ever be “innocent”? What if a person unknowingly forwards false news? Scientifically, such actions still produce harmful outcomes (panic, polarization, reputational harm), but ethically and legally, intention affects culpability. Science addresses this through cognitive bias research, epistemic vigilance and digital literacy helping individuals recognize emotional triggers, algorithmic amplification and the human tendency to trust familiar sources.

In line with Antonio Guterres' statement "Disinformation is not just misleading, it is dangerous and potentially deadly."—the ECtHR increasingly recognizes the real-world consequences of digital disinformation: electoral interference, violence, public health risks and reputational destruction.

Science combats these harms not only through technology (AI-based fact-checking, forensic linguistics, data verification tools) but also through behavioral science, ethics, and education, fostering critical thinking and responsible expression (Zuboff, 2019).

As a result, disinformation operates not as isolated speech acts but as coordinated informational strategies embedded in digital infrastructures.

Educational institutions contribute to democratic resilience by promoting critical reasoning, ethical digital participation, evidence-based thinking, and responsible information sharing.

Contemporary democratic education must move beyond traditional civic instruction toward broader forms of digital citizenship education. Students must understand not only constitutional rights but also the technological systems shaping contemporary communication.

Universities and schools can strengthen resilience against disinformation through media verification training, interdisciplinary digital ethics education, fact-checking initiatives, AI literacy programs, and public awareness campaigns.

These educational dimensions are particularly important in emerging democracies where digital transformation often progresses more rapidly than institutional adaptation.

Science and education play a critical role in countering digital manipulation by instilling structured, evidence-based thinking and ethical responsibility in users. Educational frameworks must move beyond basic digital access to teach active epistemic vigilance: training citizens to identify emotional triggers, detect algorithmic confirmation biases, and understand the commercial motivations driving online visibility.

5. Digital Constitutionalism And European Regulatory Responses

5.1 European Risk-Based Governance

The European Union has emerged as a leading actor in the development of digital constitutionalism (Celeste, 2019). Rather than adopting purely libertarian or purely restrictive approaches, the EU increasingly applies a risk-based model focused on platform accountability and systemic safeguards. In a digital environment where personal data move across borders instantly, malicious actors can use that data to create false narratives, manipulate identities, or generate targeted disinformation. These practices affect fundamental rights protected under modern data protection

laws (GDPR, Convention 108+, national legislation) and echo historical principles of privacy protection found in earlier international agreements.

Legal Instruments and Boundaries

The fight against disinformation (Urbinati, 2019) is waged through existing and evolving legislation, jurisdiction, and legal doctrine.

Foundational legal requirements come from various sources:

- Defamation Law: This is our most direct tool. Criminal defamation laws treat false statements that harm a person's reputation as a punishable offense. The legal principle is that expression ceases to be protected when it causes specific, measurable and unlawful harm.
- Intergovernmental treaties: historically, even treaties that appear unrelated can provide contextual guidance. For instance, the **Treaty of Bern (1874)**, which created the **Universal Postal Union (UPU)**, established initial standards for communication across borders, regulating the content carried through international postal services.

Modern Legal Responses: Soft Law

The EU provides a prime example of a non-legislative approach to govern the digital space:

- EU Code of Practice on Disinformation: This is a key example of "soft law." It's a set of voluntary commitments by platforms (Google, Meta, etc.) to combat disinformation through mechanisms like demonetization, content labelling, and enhanced scrutiny of political advertising, rather than just direct government censorship.

Interpreting the Joint Declaration

A critical document guiding legal analysis is the Joint Declaration on Freedom of Expression and "Fake News," Disinformation and Propaganda.

This Declaration interprets how existing principles apply to the digital age, focusing on two key areas:

A. General Principles & Restrictions

- It reaffirms that restrictions on freedom of expression must meet the three-part test (provided by law, pursue a legitimate aim and be necessary in a democratic society).
- Crucially, it emphasizes that general prohibitions on "fake news" are illegitimate, as they are often overly broad and could lead to censorship. The focus must be on the *intent* and the *harm* caused, not merely on falsity.

B. Defamation Standards

- The Declaration cautions against the abuse of Criminal defamation laws. It states that such laws should be abolished or, at a minimum, should never be applied to non-public figures and must require proof of malice or reckless disregard for the truth when applied to public interest matters. The goal is to punish the malicious, not to stifle legitimate public debate.

In late 2024, a Very Large Online Platform (VLOP), operating under the European Union's Digital Services Act (DSA), was found to be hosting a systemic stream of health disinformation. Such fake news promoted an unproven, dangerous "natural cure" for diabetes, leading several users across three member states to stop their prescribed insulin treatment.

The platform was technically compliant with the *voluntary* EU Code of Practice on Disinformation (i.e., it claimed to have policies), but it failed to adequately moderate the content, allowing the "Silent Spread" to bypass AI filters, especially in smaller European languages.

The European Commission (EC, 2021) initiated proceedings under the DSA, citing the platform's failure to mitigate "systemic risks" to public health, a key legal requirement under the new law.

The EU found that the platform violated the systemic risk provisions of the DSA. It imposed a substantial fine and, critically, compelled the platform to audit its content moderation systems specifically for health disinformation in all official EU languages.

The Regulation (EU) 2022/2065 on Digital Services Act (DSA) requires Very Large Online Platforms to address systemic risks associated with disinformation, electoral manipulation and public health harms. Similarly, European Union AI Act December 2023 introduces transparency obligations and regulatory safeguards for high-risk AI systems.

- These developments reflect an important constitutional transformation: platforms are increasingly treated as constitutional actors; democratic integrity becomes a regulatory objective; and procedural transparency gains constitutional significance.

The European approach reflects a broader transformation toward digital constitutionalism, characterized by regulation of private power and integration of fundamental rights into digital governance (UNESCO, 2023) and emphasis on procedural safeguards.

Unlike the U.S. model, the EU recognizes that unregulated infrastructures can undermine freedom itself.

5.2.1 ECtHR Jurisprudence

The European Court of Human Rights has progressively recognized the structural dimensions of speech. In *Handyside v. United Kingdom* (1976), the Court emphasized

pluralism and tolerance. In *Delfi AS v. Estonia* (2015), it acknowledged the role of intermediaries in shaping online discourse.

A further complexity emerges in the tension between defamation and freedom of expression. Where does free speech end and unlawful reputational harm begin? This boundary has been explored extensively by the European Court of Human Rights which interprets the freedom of expression foreseen by Article 10 of European Convention of Human Rights (ECHR, 1950) in harmony with Article 8 (right to reputation). The Court uses a balancing test (Çalı, 2018): weighing public interest, intention, societal value, verification steps, and the severity of harm. Key cases illustrate this balance:

- *VgT v. Switzerland* (2001): ECtHR stressed the importance of pluralism and public debate, overturning a broadcasting ban.
- *Appleby and Others v. United Kingdom* (2003): freedom of expression does not impose unlimited positive obligations on the State but access to communication channels matters.
- *Animal Defenders International v. United Kingdom* (2013): upheld restrictions on political advertising to prevent distortion of public debate.
- *Brzezinski v. Poland* (2019): reaffirmed that expressions harming reputations without a factual basis may lose Article 10 protection.
- *Salov v. Ukraine* (2005): involved intentional false election-related information; the Court emphasized the severe danger of politically motivated disinformation.

This jurisprudence shows that intent, context, and public interest are decisive. Misstatements made in *bona fide* may enjoy protection, while deliberate deception (*mala fide*)—especially when endangering democratic integrity—falls outside Article 10's shield.

The ECtHR has progressively framed freedom of expression as essential to democratic society.

- *Handyside v. United Kingdom* (1976): pluralism and tolerance as democratic foundations.
- *Lingens v. Austria* (1986): enhanced protection for political speech.

More recent jurisprudence reflects a structural awareness:

- *Delfi AS v. Estonia* (2015): platform liability for user-generated content.
- *Magyar Tartalomszolgáltatók Egyesülete v. Hungary* (2016): proportionality in intermediary liability.

These cases signal a shift toward recognizing platforms as key actors in the regulation of speech.

5.2.2 CJEU Jurisprudence

The Court of Justice of the European Union has similarly contributed to digital constitutionalism. In *Google Spain* (2014), the Court recognized the right to be forgotten as an aspect of informational self-determination.

The CJEU has played a crucial role in extending fundamental rights into digital contexts:

- *Google Spain* (2014): recognition of informational self-determination
- *Glawischnig-Piesczek* (2019): global content removal obligations
- *Digital Rights Ireland* (2014): limits on data retention

These decisions demonstrate:

- Expansion of rights into private digital environments
- Increased judicial engagement with technological governance.

5.3 Epistemic Harm and Democratic Integrity

Traditional legal frameworks are fundamentally reactive, designed to address individual, measurable harms like defamation or direct incitement (Pasquale, 2015). However, AI-driven disinformation structures generate a separate category of macro-societal damage, or, as mentioned above epistemic harm, which represents the systemic degradation of the shared socio-technical conditions necessary for authentic knowledge formation, public verification, and democratic legitimacy.

Unlike individual defamation, epistemic harm targets the collective infrastructure of democracy (Castells, 2010). It erodes foundational trust in public institutions, hyper-polarizes public discourse, and disrupts collective decision-making processes.

This article introduces the concept of “epistemic harm” to describe how AI-driven disinformation undermines the collective conditions necessary for democratic reasoning (Floridi et al., 2018).

Unlike traditional legal harms focused on individuals, epistemic harm affects collective trust, democratic legitimacy, public reasoning processes, institutional credibility, and social cohesion.

This conceptual framework helps explain why democratic societies increasingly regulate digital infrastructures rather than simply restricting individual speech acts.

5.4 Freedom of Expression and the Glorification of Terrorism: A Hard Case for Epistemic Infrastructure Theory

The Problem of Terrorism-Related Expression

Few areas test the limits of freedom of expression more severely than speech associated with terrorism. Among these, the glorification or justification of terrorist acts presents a particularly difficult constitutional dilemma.

From an EIT perspective, the key issue is not only whether speech praises terrorism, but whether it contributes to a distortion of the epistemic environment that normalizes or legitimizes violence. This dimension is currently absent from Albanian constitutional reasoning.

Unlike direct incitement, glorification does not necessarily call for immediate violence. Instead, it operates symbolically—legitimizing, normalizing, or valorizing acts of terrorism within public discourse. This ambiguity complicates legal classification and raises concerns about overreach and censorship.

European legal systems have increasingly addressed such speech through criminal and administrative measures, particularly following waves of terrorist attacks across Europe. At the EU level, Directive (EU) 2017/541 on combating terrorism criminalizes public provocation to commit terrorist offences, including indirect forms such as glorification when they create a risk of terrorist acts.

The French Eurodeputy Case and the Expansion of Risk-Based Logic

A paradigmatic illustration is the controversy surrounding a French Member of the European Parliament who publicly engaged in discourse interpreted as legitimizing or contextualizing acts of terrorism. Although the legal outcomes depend on specific factual determinations, the case highlights a broader trend: The shift from punishing incitement to addressing symbolic endorsement and narrative framing. This evolution reflects a preventive logic, where speech is regulated not solely for what it directly causes, but for the risk environment it contributes to.

From a doctrinal perspective, this aligns with the ECtHR's recognition that freedom of expression may be restricted where speech undermines democratic values or contributes to violence. In cases such as: *Zana v. Turkey* (1997), *Garaudy v. France* (2003), *Perinçek v. Switzerland* (2015, Grand Chamber) the Court has grappled with the boundaries between protected expression and dangerous speech. In *Zana v. Turkey* (1997), statements supporting a terrorist organization were deemed capable of exacerbating violence in a volatile context, justifying restriction.

5.5 The Insufficiency of Self-Regulation

The limits of purely voluntary platform policies were starkly illustrated in late 2024. A major platform, despite its technical compliance with the voluntary EU Code of Practice on Disinformation, allowed a dangerous stream of health disinformation to

bypass its automated content moderation systems. The content specifically exploited gaps in smaller, non-English European languages, leading directly to real-world medical emergencies across multiple Member States.

The European Commission's subsequent intervention and fine under the DSA established a critical global precedent: self-regulation is structurally insufficient. Platforms bear an active legal duty of care to safeguard the epistemic environments they profit from.

6. Albania As A Case Study Of An Emerging Democracy

6.1 Constitutional Protections and Structural Challenges

Freedom of expression in Albania is constitutionally protected under Article 22 of the Constitution of the Republic of Albania (1998), which protects freedom of speech, press, and information, and is interpreted consistently with Article 10 of the European Convention on Human Rights. This provision is reinforced by Article 17, which establishes the general limitation clause based on proportionality and necessity in a democratic society.

Albanian criminal legislation addresses terrorism-related speech, including incitement to terrorism, public endorsement or apology of terrorist acts. However, there is no consolidated constitutional doctrine clarifying how such restrictions interact with Article 22 of the Constitution.

Unlike the European Court of Human Rights, which has developed nuanced standards distinguishing abstract opinion, contextual incitement, and the legitimization of violence, the Albanian framework remains fragmented, reactive, and insufficiently theorized.

Although the Constitutional Court of Albania has repeatedly emphasized the importance of pluralism, democratic tolerance, the proportionality of restrictions, and the protection of political speech, Albanian State still illustrates broader structural vulnerabilities common to emerging democracies. Specifically, these advancements are frequently undermined by severe media concentration, political influence on journalism, limited institutional trust, insufficient digital literacy, and weak platform accountability mechanisms.

These structural conditions create an environment in which AI-driven disinformation may exert disproportionate democratic influence.

In its jurisprudence, particularly in Decision no. 16/2004, the Constitutional Court, echoing the doctrine of the European Court of Human Rights, emphasized that freedom of expression constitutes one of the foundations of a democratic state. Without the protection of this right, pluralism, tolerance, and the formation of free political will, which are essential elements of a democratic society, cannot be guaranteed. Furthermore, in Decision no. 20/2006, the Court held that limitations on fundamental rights and freedoms must comply with the requirements prescribed by

Article 17 of the Constitution, including the requirement that such limitations be imposed only by law and that their necessity and extent be assessed by the highest legislative authority, thereby serving as an essential safeguard against arbitrary interference with constitutional rights.

Based on Decision no. 11/2021, the Court examined the restrictions imposed on the freedom of assembly and collective freedom of expression, emphasizing that, even when justified by the protection of public health, such restrictions must be prescribed by law, proportionate to the circumstances that justify them, and must not impair the essence of these fundamental rights in a democratic society.

These cases demonstrate a strong rights-protective tradition, but one still rooted in a classical liberal paradigm, focusing on state interference rather than systemic distortions.

The inclusion of terrorism glorification within the scope of free speech limitations demonstrates a broader transformation. This shift reinforces the article's central claim: Freedom of expression is increasingly governed by the need to preserve the epistemic conditions of democracy, not merely to protect individual speech acts.

6.2 Albania within the Europeanization Process

Albania's European Union accession process increasingly influences domestic approaches to media regulation, platform governance, and democratic standards (EC, 2024). European integration pressures encourage stronger media transparency, harmonization with EU digital regulation, improved democratic safeguards, and institutional modernization.

Consequently, Albania represents not merely a national case but a broader example of how emerging democracies adapt constitutional systems to digital transformation and international regulatory expectations.

7. Findings

The findings demonstrate that AI-driven disinformation represents both a legal and societal challenge that cannot be addressed through traditional free speech frameworks alone. Emerging democracies face a particularly difficult balancing exercise because they must simultaneously protect democratic freedoms and strengthen institutional resilience.

The European model increasingly recognizes that unrestricted digital infrastructures may undermine democratic participation itself. This contrasts with more absolutist approaches to freedom of expression that focus primarily on protection from state interference. At the same time, excessive regulation carries serious risks regarding political abuse, censorship, chilling effects on journalism, and suppression of dissent.

Therefore, democratic responses must remain grounded in proportionality, transparency, and judicial oversight.

The Albanian case demonstrates that constitutional guarantees alone are insufficient when informational infrastructures remain vulnerable to manipulation. Democratic resilience depends equally on institutional trust, educational capacity, media pluralism, digital literacy, and transparent governance.

This analysis also highlights the importance of interdisciplinary approaches. Legal regulation, educational policy, social trust, and technological governance increasingly intersect within contemporary democratic systems.

8. Conclusion

AI-driven disinformation has fundamentally transformed the relationship between freedom of expression, democratic participation and digital governance. Contemporary democracies face a structural challenge in which algorithmic amplification, synthetic media and platform-driven communication increasingly shape public understanding and political behavior.

This article argues that freedom of expression can no longer be understood solely as an individual liberty protected against state interference. Instead, democratic systems must also protect the informational conditions necessary for meaningful democratic deliberation.

Using Albania as a case study situated within broader European developments, the article demonstrates that emerging democracies face heightened vulnerabilities due to weaker institutional resilience and less developed informational infrastructures.

The concept of epistemic harm offers a useful framework for understanding how AI-driven disinformation affects democratic legitimacy beyond traditional individual harm models.

The findings further demonstrate that sustainable democratic responses require an integrated framework combining constitutional safeguards, platform accountability, transparent digital governance, media literacy and democratic education.

Ultimately, the future of democratic freedom depends not only on protecting speech itself but also on preserving the integrity, trust and resilience of the digital public sphere.

References

- [1] Animal Defenders International v. United Kingdom, 2013-II Eur. Ct. H.R. 35 (Application No. 48876/08).
- [2] Appleby and Others v. United Kingdom, 2003-VI Eur. Ct. H.R. 65 (Application No. 44306/98).
- [3] Balkin, J. M. (2018). Free speech in the algorithmic society: Big data, private governance, and new school speech regulation. *UC Davis Law Review*, 51(3), 1149–1210.

- [4] Barendt, E. (2005). *Freedom of speech* (2nd ed.). Oxford University Press.
- [5] Benkler, Y., Faris, R., & Roberts, H. (2018). *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. Oxford University Press.
- [6] Brzezinski v. Poland, European Court of Human Rights (2019) (Application. No. 38654/14).
- [7] Çalı, B. (2018). Explaining variation in the intrusiveness of regional human rights remedies in domestic orders. *International Journal of Constitutional Law*, 16(1), 214–234. <https://dx.doi.org/10.1093/icon/moy009>.
- [8] Castells, M. (2010). *The rise of the network society* (2nd ed.). Wiley-Blackwell.
- [9] Celeste, E. (2019). Digital constitutionalism: a new systematic theorisation. *International Review of Law, Computers & Technology*, 33(1), 76–99. <https://doi.org/10.1080/13600869.2019.1562604>.
- [10] Constitution of the Republic of Albania. (1998), as amended.
- [11] Council of Europe. (1950, November 4). *European Convention for the Protection of Human Rights and Fundamental Freedoms, as amended*. E.T.S. No. 5. https://www.echr.coe.int/documents/d/echr/convention_eng.
- [12] Council of Europe. (2018, October 10). *Protocol amending the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data*. C.E.T.S. No. 223 [Convention 108+].
- [13] De Gregorio, G. (2022). *Digital constitutionalism in Europe: Reframing rights and powers in the algorithmic society*. Cambridge: Cambridge University Press.
- [14] Delfi AS v. Estonia, 2015-II Eur. Ct. H.R. 319 (Application No. 64569/09).
- [15] Digital Rights Ireland Ltd v. Minister for Communications, Energy and Natural Resources and Others, Joined Cases C-293/12 and C-594/12, ECLI:EU:C:2014:238 (2014).
- [16] Directive (EU) 2017/541 of the European Parliament and of the Council of 15 March 2017 on combating terrorism and replacing Council Framework Decision 2002/475/JHA and amending Council Decision 2005/671/JHA, OJ L 88, 31.3.2017, p. 6–21.
- [17] European Commission. (2021). *Proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts* (COM(2021) 206 final). <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206>.

- [18] European Commission. (2024). *Albania 2024 report: Accompanying the document communication from the commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions* (SWD(2024) 590 final). https://neighbourhood-enlargement.ec.europa.eu/albania-report-2024_en.
- [19] Eva Glawischnig-Piesczek v. Facebook Ireland Limited, Case C-18/18, ECLI:EU:C:2019:821 (2019).
- [20] Floridi, L., Cowls, J., Beltrametti, M. *et al.* AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds & Machines* 28, 689–707 (2018). <https://doi.org/10.1007/s11023-018-9482-5>.
- [21] Garaudy v. France, 2003-IX Eur. Ct. H.R. 21 (Application No. 65831/01).
- [22] Gjykata Kushtetuese e Republikës së Shqipërisë [Constitutional Court of the Republic of Albania]. (2004, November 11). *Vendim Nr. 16/2004*. <https://www.gjykatakushtetuese.gov.al/>.
- [23] Gjykata Kushtetuese e Republikës së Shqipërisë [Constitutional Court of the Republic of Albania]. (2006, July 11). *Vendim Nr. 20/2006*. <https://www.gjykatakushtetuese.gov.al/>.
- [24] Gjykata Kushtetuese e Republikës së Shqipërisë [Constitutional Court of the Republic of Albania]. (2021, March 9). *Vendim Nr. 11/2021*. <https://www.gjykatakushtetuese.gov.al/>.
- [25] Google Spain SL and Google Inc. v. Agencia Española de Protección de Datos (AEPD) and Mario Costeja González, Case C-131/12, ECLI:EU:C:2014:317 (2014).
- [26] Handyside v. United Kingdom, 1 E.H.R.R. 737 (1976).
- [27] Lingens v. Austria, 8 E.H.R.R. 407 (1986).
- [28] Magyar Tartalomszolgáltatók Egyesülete and Index.hu Zrt v. Hungary, 64 E.H.R.R. 4 (2016).
- [29] Napoli, P. M. (2019). *Social media and the public interest: Media regulation in the disinformation age*. Columbia University Press.
- [30] Organisation for Economic Co-operation and Development. (2021). *OECD digital economy outlook 2020*. OECD Publishing. <https://doi.org/10.1787/bb167041-en>.
- [31] Organisation for Economic Co-operation and Development. (2024). *Building digital trust: The role of data governance and transparency in democratic societies*. OECD Publishing.

- [32] Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press.
- [33] Perinçek v. Switzerland, 63 E.H.R.R. 6 (2015).
- [34] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). OJ L 119, 4.5.2016, p. 1–88. <http://data.europa.eu/eli/reg/2016/679/oj>.
- [35] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act). OJ L 277, 27.10.2022, p. 1–102. <http://data.europa.eu/eli/reg/2022/2065/oj>.
- [36] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2016/1629 (Artificial Intelligence Act). OJ L 2024/1689, 12.7.2024. <http://data.europa.eu/eli/reg/2024/1689/oj>.
- [37] Salov v. Ukraine, 2005-VIII Eur. Ct. H.R. 209 (Application No. 65518/01).
- [38] Sunstein C. R. (2018). *#Republic: Divided democracy in the age of social media* (NED-New edition). Princeton University Press. <https://doi.org/10.2307/j.ctv8xnhtd>.
- [39] Sunstein, C. R. (2021). *Liars: Falsehoods and free speech in an age of deception*. Oxford University Press. <https://doi.org/10.1093/oso/9780197545119.001.0001>.
- [40] Treaty Concerning the Formation of a General Postal Union (Treaty of Bern), October 9, 1874, 19 Stat. 577. <https://www.govinfo.gov/content/pkg/STATUTE-19/pdf/STATUTE-19-Pg577.pdf>.
- [41] UNESCO. (2021). *Media and information literate citizens: Think critically, click wisely!* <https://unesdoc.unesco.org/ark:/48223/pf0000377068>.
- [42] UNESCO. (2023). *Guidelines for the governance of digital platforms. Safeguarding freedom of expression and access to information through a multi-stakeholder approach*. <https://unesdoc.unesco.org/ark:/48223/pf0000387339>.
- [43] Urbinati, N. (2019). *Me the people: How populism transforms democracy*. Harvard University Press. <https://doi.org/10.2307/j.ctvk12sz4>.

- [44] VgT Verein gegen Tierfabriken v. Switzerland, 2001-VI Eur. Ct. H.R. 243 (Application No. 24699/94).
- [45] Zana v. Turkey, 27 E.H.R.R. 667 (1997).
- [46] Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.