ISSN: 3048-9423(Online)
Naveen International Journal of Multidisciplinary Sciences (NIJMS)
© NIJMS | Vol. 2 | Issue 6 | Jun-Jul 2026
Available online at: https://nijms.com/
DOI: https://doi.org/10.71126/nijms.v2i6.139

Deepfake Technology, Misinformation and Media Regulation: Emerging Trends and Challenges

Dr. Manindra Singh Hanspal

Assistant Professor, Presidency University, Bengaluru.

Corresponding Author Email: *mhanspal21@gmail.com*

ABSTRACT:

Deepfake technology has moved from a niche computer vision experiment to a mainstream concern for journalism, electoral processes, and public discourse. Synthetic audio, video, and images now circulate alongside genuine reporting, and the boundary between the two grows thinner each year. This article studies the rise of deepfakes, their entanglement with misinformation, and the regulatory responses unfolding across India, the European Union, and the United States. It draws on a dataset of 2,223 peer-reviewed and preprint publications in Communication and Media Studies retrieved from Dimensions for the period 2017 to 2026, supported by VOSviewer co-authorship maps and citation analysis. The piece blends qualitative discussion with a light bibliometric reading of the field. The analysis shows publication output rising sharply after 2019, climbing from 46 records in 2019 to 768 in 2025, with India contributing 111 papers and ranking second globally after the United States. Scholarly attention concentrates on misinformation, social media platforms, journalism, and democracy, while regulation, governance, and law receive less coverage. The article argues for four planks of a workable response: platform accountability, mandatory labelling, source verification, and media literacy. The findings point to a research gap on enforcement, cross-border coordination, and the position of small language communities in detection systems. The piece closes with directions for further work.

Keywords: *Deepfake, Synthetic media, Misinformation, Fake news, Platform accountability.*

1. INTRODUCTION:

The term deepfake first appeared on a Reddit forum in 2017, when a user posted face-swapped videos generated through a deep learning model.[1] Eight years later, the same underlying technique now powers

fraud calls in Mumbai, fake political robocalls in New Hampshire, and fabricated battlefield footage in Ukraine.[2] The shift from hobbyist code to a tool of routine deception has been swift, and the regulatory machinery is still catching up. Generative adversarial networks, diffusion models, and voice cloning tools have lowered the cost of producing convincing synthetic content to near zero.[3] A six-second voice sample is enough to clone a speaker. A handful of public photographs is enough to build a face model. The barrier to entry is no longer technical skill but access to consumer software. Several free apps offered through mobile stores produce passable face swaps in under a minute. Misinformation on digital platforms has grown in parallel. An MIT Media Lab study published in Science in 2018 documented false news stories on Twitter spreading six times faster than true ones, a finding from before the rise of modern generative tools.[4] Add synthetic video to the mix, and the asymmetry between speed of spread and speed of correction becomes severe. Fact checkers work on a scale of hours. A viral clip reaches millions in minutes.[5]

The damage falls hardest on media credibility. Audiences who once treated video as authoritative now hesitate. The same hesitation extends to genuine footage, a phenomenon researchers call the liar's dividend.[6] When everything looks fake, even real evidence loses force. News organizations report rising skepticism among readers, and surveys conducted by Reuters Institute show declining trust in news across most democracies between 2020 and 2025.[7] The geographic and linguistic spread of the problem also matters. Early concern about deepfakes focused on English-language political content from the United States and Western Europe. By 2024, however, documented cases involved Hindi, Tamil, Telugu, Kannada, Malayalam, Bengali, Mandarin, Spanish, Slovak, Turkish, and Indonesian.[8] The technical pipelines used to generate the content scaled across languages with little adjustment, while the detection pipelines did not. Communities with smaller digital footprints face higher per capita exposure to synthetic harm. A short note on terminology helps frame the rest of the article. Synthetic media is the broader category, covering all algorithmically generated or altered audio, image, and video content.[9] Deepfake is a subset of synthetic media in which the output mimics a real or apparent person in a way liable to mislead.[10] Misinformation is false content shared without intent to deceive. Disinformation is false content shared with the intent to deceive. Fake news, the most common term in public discourse, sits across both. The four terms are used distinctly throughout the article, in line with the working definitions of UNESCO and the European Union AI Act. The problem statement of this article is direct. Deepfake technology and the misinformation it enables outpace existing media regulation in scope, speed, and geographic reach. Courts, ministries, and platform safety teams respond after the fact. Publication output in the field has grown sharply, but the academic conversation tilts toward detection methods and platform behaviour. Regulation, law, and governance receive a smaller share of attention. This study has three objectives. First, to map the growth and shape of academic work on deepfakes, misinformation, and media regulation between 2017 and 2026. Second, to trace the regulatory response in India, the European Union, and the United States through documented cases and statutory instruments. Third, to set out the central challenges and propose a framework linking platform duties, legal remedies, and audience literacy.

1. METHODOLOGY AND RESEARCH TRENDS

The dataset for this article was drawn from Dimensions, a research information platform maintained by Digital Science. The export was conducted on 28 April 2026 using the search string deepfake OR synthetic media AND misinformation OR fake news AND regulation OR law OR policy OR governance. The query was restricted to the field of research code 4701 Communication and Media Studies under the ANZSRC 2020 classification, and to publication years 2017 through 2026. The result was 2,223 publications, including journal articles, conference papers, preprints, book chapters, and edited volumes. Inclusion criteria were straightforward. Records were retained when at least one of the search terms appeared in the title, abstract, or full text indexed by Dimensions, and when the publication carried a Communication and Media Studies tag. Records flagged as duplicates by the platform were removed at the source. Records with hyper authorship beyond 100 authors were retained but lacked author-level metadata, a known constraint of the Dimensions export pipeline. Three analytical layers were applied. The first is a publication trend analysis covering the full period. The second is a keyword frequency analysis on titles and abstracts. The third is a co-authorship network visualization generated through VOSviewer using the same dataset. Country-level affiliation data were extracted from the author affiliation column. Citation counts were taken from the Dimensions citation tracker as of the export date.

2.1 Publication Trend

The publication trend shows a clear inflection after 2019. Output stayed below 50 records per year through 2018, then climbed steadily and accelerated sharply in 2024 and 2025. Figure 1 shows the year-wise count exported from Dimensions.

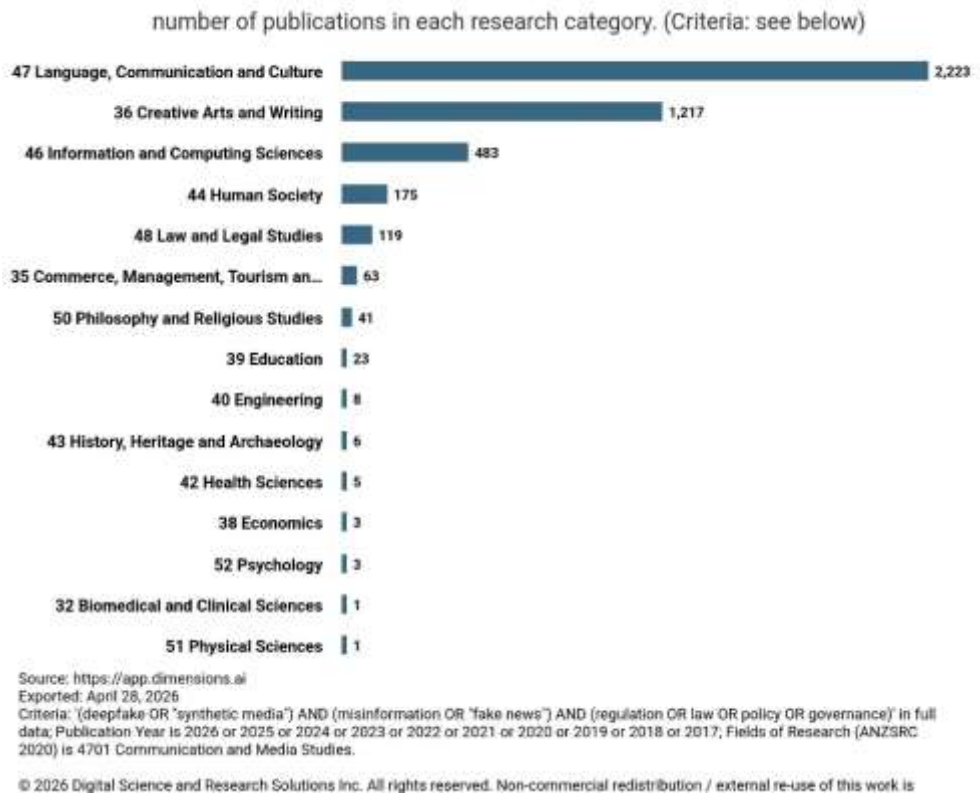

Figure 1. Annual publications on deepfakes, misinformation, and media regulation in Communication and Media Studies, 2017 to 2026. Source: Dimensions, exported 28 April 2026.

The 2025 peak of 768 publications is more than three times the 2022 figure of 196. The 2026 partial count of 256 already exceeds the 2021 total of 165, and the year is only four months in at the time of export. Table 1 sets out the year-wise totals.

Table 1. Annual publication output, 2017 to 2026

Year	Publications	Cumulative
2017	2	2
2018	8	10
2019	46	56
2020	91	147
2021	165	312
2022	196	508
2023	234	742
2024	457	1,199
2025	768	1,967
2026 (partial)	256	2,223

The acceleration aligns with three external developments. The first is the public release of generative AI tools in late 2022 and early 2023, which expanded the supply of synthetic content. The second is the 2024 global election supercycle, which drew researchers toward case studies of political deepfakes. The third is the rollout of regulatory drafts in the European Union, India, and the United States, which generated legal commentary and policy analysis.

2.2 Keyword and Thematic Trends

Keyword frequency in titles and abstracts shows where the field places its weight. Table 2 sets out the top terms.

Table 2. Top thematic keywords in the dataset

Keyword cluster	Mentions	Theme
Artificial intelligence / AI	3,895	Technology
Journalism	2,156	Profession
Social media	2,071	Platform
Disinformation	1,380	Content
Platform	1,376	Platform
Fake news	1,367	Content
Deepfake	1,364	Technology
Misinformation	1,152	Content
Democracy	748	Politics
Literacy	726	Audience

Trust	601	Audience
Verification / fact-check	546	Response
Detection	443	Response
Policy	438	Regulation
Election	368	Politics
Generative AI	350	Technology
Regulation	324	Regulation
Governance	284	Regulation
Law	283	Regulation
Synthetic media	138	Technology

Two patterns stand out. First, content terms such as misinformation, fake news, disinformation, and deepfake together account for over 5,000 mentions, while regulation, governance, law, and policy together yield around 1,300 mentions. The field talks four times as often about the problem as about the formal response. Second, audience-side terms like literacy, trust, and verification have a strong presence, which suggests researchers see civic education as part of the answer. The co-authorship structure confirms this distribution. Figure 2 shows the broad author network with seven coloured clusters, dominated by Hameleers Michael, Tandoc Edson C, Esser Frank, and Ahmed Saifuddin. Figure 3 shows the densely connected core, where a small group of authors with citation counts above twenty serves as the connective tissue of the field.

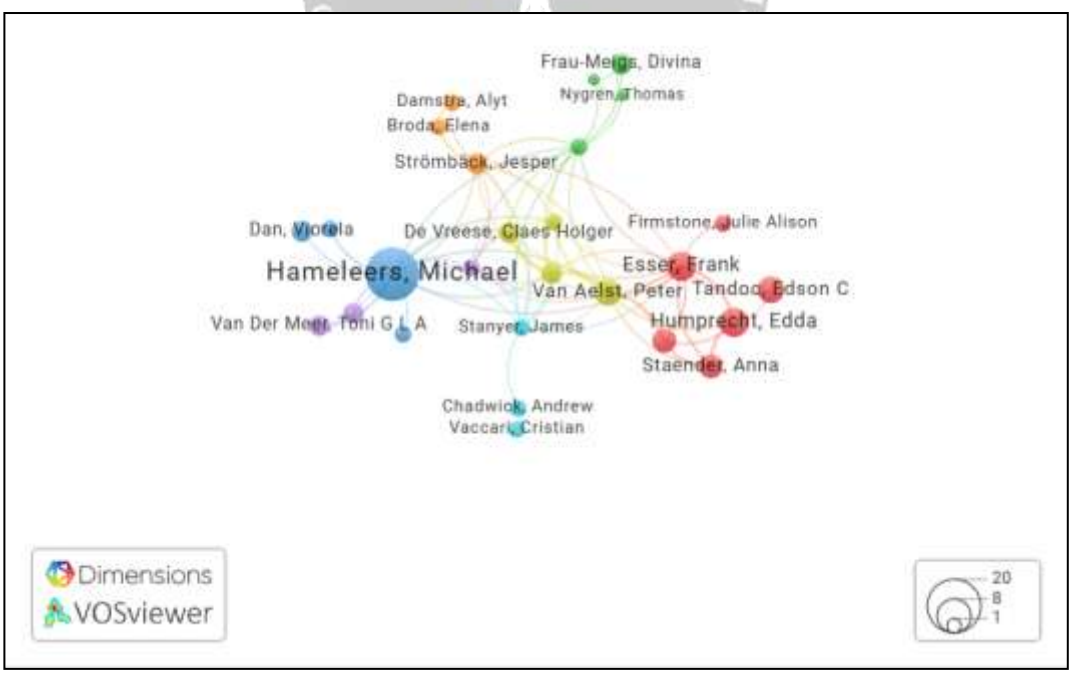

Figure 2. Co-authorship network in the dataset, full view. Node size scaled by document count. Source: VOSviewer rendering of Dimensions data.

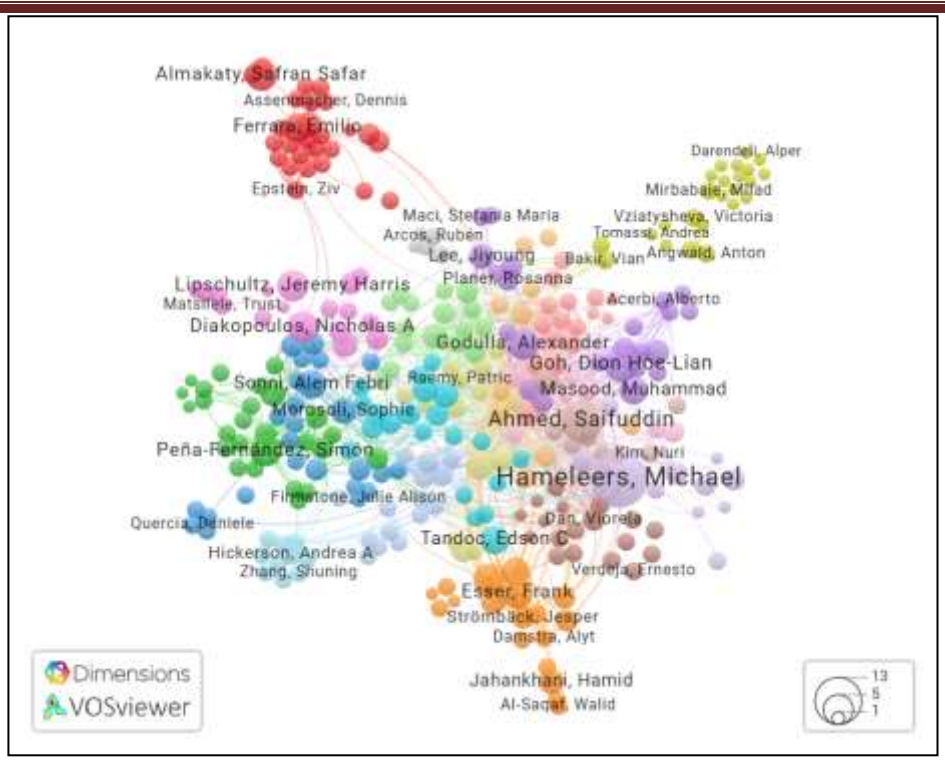

Figure 3. Co-authorship network, dense core view showing the most cited authors in the field.

On the country level production, the United States leads with 276 publications, followed by India at 111, the United Kingdom at 102, Spain at 95, and Germany at 71. India's strong second position is notable given the country's late entry into formal deepfake regulation. Indian output rose from one paper in 2020 to 40 papers in 2025. Citation behaviour shows a long-tailed distribution. The total citation count across the dataset is 17,326. The mean is 7.79, but the median is 1, which indicates that a small number of papers carry most of the scholarly weight.

2.3 Top Sources, Top Countries, and Citation Patterns

The dataset is dominated by preprint repositories and journal outlets in communication and information science. The top five sources are arXiv with 109 records, SSRN Electronic Journal with 78, Journalism and Media with 38, Journalism Practice with 26, and Media and Communication with 24. The presence of arXiv and SSRN at the top is a sign of the field moving toward rapid open dissemination, especially on policy adjacent topics where the publication cycle of traditional journals is too slow for the news agenda. On country level production, the United States leads with 276 publications across the dataset, followed by India at 111, the United Kingdom at 102, Spain at 95, and Germany at 71. China contributes 64 records, Australia 57, the Netherlands 55, Italy 48, and Indonesia 43. The strong showing of Indonesia, Brazil at 27, and Poland at 23 indicates a widening geographic base for the field. Indian output rose from 1 paper in 2020 to 7 in 2021, 5 in 2022, 16 in 2023, 26 in 2024, 40 in 2025, and 16 in the first quarter of 2026, with most of the growth concentrated in the past three years. The most cited paper in the dataset is Vaccari and Chadwick on deepfakes and disinformation in Social Media plus Society from 2020, with 812 citations as of the export date. Other heavily cited works include the Diakopoulos and Johnson paper on artificial intelligence and journalism with 307 citations, the Vraga and Bode paper on misinformation with 211 citations, and the Hwang and Smith paper on social spam detection with 199 citations. The 90 most

cited papers together account for over half of all citations in the dataset, a typical Pareto distribution for an emerging field.

3. CONCEPTUAL FRAMEWORK AND MEDIA REGULATION

3.1 What Are Deepfakes

A deepfake is a piece of synthetic media in which a person's face, voice, or body is generated or altered by a machine learning model in a way an average viewer would treat as authentic.[11] The European Union AI Act, which entered into force on 1 August 2024, defines a deepfake in Article 3(60) as image, audio, or video content generated or modified by AI to imitate real people, objects, places, or events in a way liable to mislead.[12] India's amended Information Technology Rules of 2026 use a similar phrasing under the term synthetically generated information.[13] Three families of techniques produce the most current deepfakes. Face swap models, descended from the original 2017 deepfake code, replace one person's face with another in a video. Lip sync and reenactment models change what a speaker appears to say while keeping the rest of the frame intact.[14] Voice cloning models build a synthetic voice from a short audio sample and then read any input text in the cloned voice.[15] The fourth and newest family covers fully generated avatars, where the speaker does not exist at all.

3.2 The Link with Misinformation

Misinformation is false or inaccurate information shared without intent to deceive. Disinformation is the same content shared with the intent to deceive.[16] Deepfakes feed both streams. A satirical deepfake forwarded by an unaware user becomes misinformation. A campaign deepfake produced to swing voters is disinformation. The form is identical, but the intent and context shift.[17] Three pathways connect deepfakes to misinformation. The first is direct deception, where audiences accept fabricated content as real. The second is the liar's dividend, where genuine content is dismissed as fake. The third is erosion of baseline trust, where audiences disengage from media altogether. The third pathway is the most damaging because it is the hardest to undo.

3.3 The Role of Media Platforms

Social media platforms occupy the central position in the deepfake economy.[18] They host the content, they recommend it through ranking algorithms, and they monetize it through advertising. A 2024 Surfshark study found social media was the channel of distribution in 92 percent of analyzed election deepfake incidents, with messaging apps in 32 percent and online advertisements in 32 percent.[19] Platforms, therefore, sit at the choke point of the distribution graph. Platform responsibility has shifted in the regulatory conversation from a liability shield model to a duty of care model. Under the older approach, platforms were treated as neutral conduits and shielded from liability for user content. Under the newer approach, platforms are treated as active editors of the information environment and held to specific obligations on labelling, removal, and traceability.

3.4 Regulatory Context: India

India's regulatory response has come in three layers. The first is the Information Technology Act, 2000. Section 66C deals with identity theft, Section 66D with impersonation, Section 66E with privacy violations through image capture, and Section 67 with obscene content.[20] The second layer is the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, which set due diligence and grievance redressal duties on platforms. The third and newest layer is the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026. The Ministry of Electronics and Information Technology released the draft on 22 October 2025, notified the final rules on 10 February 2026, and brought them into force on 20 February 2026.[21] The amendment introduces five principal changes.

- A statutory definition of synthetically generated information covering all algorithmically created or altered audio, visual, and audiovisual content presented as authentic.
- A mandatory labelling regime under which visual deepfakes must carry a permanent disclosure occupying at least ten percent of screen area, and audio deepfakes a disclosure in the first ten percent of duration.
- A three-hour takedown window for harmful synthetic content, replacing the earlier 36-hour window for general unlawful content.
- A user declaration duty for significant social media intermediaries, under which posters of AI-generated content must declare the synthetic origin before publication, with platforms required to verify the declaration through automated checks.
- Alignment with the Bharatiya Nyaya Sanhita, 2023, which under Section 353 penalizes misinformation causing public mischief, and the Digital Personal Data Protection Act, 2023, which penalizes the processing of biometric data without consent up to 250 crore rupees.

The judicial branch has been active in parallel. The Delhi High Court granted personality rights protection to actor Anil Kapoor in 2023 in response to deepfake misuse of his image. Similar interim orders followed in cases involving other public figures. The Supreme Court has also taken cognizance of deepfake harms in petitions related to the spread of synthetic intimate images.

3.5 Regulatory Context: European Union

The European Union approach combines the AI Act, the Digital Services Act, and the General Data Protection Regulation. Article 50 of the AI Act, which applies from 2 August 2026, requires deployers of AI systems generating or manipulating images, audio, or video constituting a deepfake to disclose the artificial origin of the content.[22] The exception is content forming part of an evidently artistic, creative, satirical, or fictional work, where disclosure is required in a less intrusive manner. A draft Code of Practice on Transparency of AI-Generated Content, published in December 2025, sets out operational measures including watermarking, machine-readable metadata, and visible icons.[23]

3.6 Regulatory Context: United States

The United States lacks a federal deepfake statute. Regulation operates through state laws and sector-specific federal action. By the end of 2024, twenty states had enacted election deepfake laws.[24] California passed three laws in 2024 in response to a deepfake audio targeting Vice President Kamala Harris.[25] The Federal Communications Commission ruled in early 2024 to bring AI-generated voices in robocalls under the Telephone Consumer Protection Act, after a fake Joe Biden robocall in New Hampshire. First Amendment challenges have followed, and a federal judge blocked one California law in October 2024.[26]

3.7 Comparative Snapshot of Three Regulatory Models

A comparison across the three jurisdictions in the article shows the spread of regulatory choices. India has moved to a centralized intermediary obligations model, with the Ministry of Electronics and Information Technology setting binding rules for platforms operating in the country. The European Union uses a horizontal product regulation model under the AI Act, complemented by the Digital Services Act on platform behaviour and the General Data Protection Regulation on personal data.[27] The United States operates a decentralized model with state-level statutes, sector-specific federal action, and First Amendment-based judicial review. The three models reveal different theories of harm. India treats deepfakes mainly as a public order and personality rights issue, with strong state involvement and rapid takedown timelines.[28] The European Union treats deepfakes mainly as an information ecosystem issue, with structured transparency duties on producers and deployers. The United States treats deepfakes mainly as a free speech issue, with regulation pushed to specific contexts such as elections, financial fraud, and non-consensual intimate imagery, where the speech interest is weakest.[29] Comparison with Chinese rules adds context. China issued the Provisions on the Administration of Deep Synthesis Internet Information Services in 2022, ahead of the European Union and India.[30] The Chinese rules require labelling, traceability, and identity verification for synthesis service providers, and use a registration system for service providers.[31] The Chinese model demonstrates the feasibility of a labelling regime at scale, although enforcement priorities and free speech protections differ sharply from the European and Indian frameworks.

4. KEY CHALLENGES AND ISSUES

4.1 The Authenticity Crisis

The first challenge is the collapse of the default assumption around recorded media as a reflection of reality. For most of the twentieth century, audiences treated photographs and video as evidence. The forensic literature on deepfakes shows that current detection accuracy hovers between 65 and 90 percent, depending on the model and the dataset, and falls further on real-world out-of-distribution samples.[32] Audiences cannot rely on detection tools, and they cannot rely on their own perception either. The PLOS ONE study by Twomey and colleagues on Russo-Ukrainian war tweets found users frequently labelling real footage as deepfakes, a sign of skepticism now defaulted on, sometimes wrongly.[33]

4.2 The Spread of Misinformation

The second challenge is the speed and reach of synthetic content. The 2024 election cycle offered the largest natural experiment yet. In India's general election, an estimated 50 million United States dollars went into AI-generated political content, according to industry tracking by Surfshark.[34] AI-cloned voices of deceased political leaders surfaced in regional campaigns. In the United States, the New Hampshire robocall imitating President Biden reached many voters before takedown.[35] In Slovakia, a fabricated audio clip of a candidate discussing election manipulation went viral two days before the 2023 vote, beyond any practical correction window.[36] Deepfakes were widespread but did not move outcomes in the United States presidential vote in measurable ways.

4.3 Platform Accountability

The third challenge is the gap between platform policy and platform practice. Major platforms have published rules against synthetic political content, but enforcement is uneven. The Internet Freedom Foundation has criticized the compressed Indian takedown windows as creating risks of over-removal and rapid-fire censorship. Platforms face the practical problem of distinguishing parody from deception at scale, and automated systems struggle with cultural and linguistic context, especially in Indian regional languages.[37] The 2024 case of Christopher Kohls in California illustrates the difficulty. A clearly labelled parody video of Vice President Harris was reposted by Elon Musk without the parody disclosure and reached over 129 million views.[38] Was the original speech protected satire? Was the unlabelled repost an election interference? The legal answer remains contested, and the platform answer was inconsistent across services.

4.4 Regulatory Gaps

The fourth challenge is the structural gap between national regulation and a global content economy. A deepfake produced in one jurisdiction reaches users in another within minutes. India's three-hour takedown window applies only to platforms operating in India, but the synthetic content might originate on a server in a third country and be uploaded by users in a fourth.[39] The European Union has the Digital Services Act and the AI Act working in tandem, but coordination with non-European platforms is contractual rather than statutory. Definitional problems compound the structural gap. The Indian rules use the phrase reasonably appears to be authentic, which is not yet tested in litigation. The European Union AI Act draws a line between deepfakes and ordinary AI-assisted edits, but the line shifts with each new model release. United States case law oscillates between First Amendment protection of political speech and consumer protection against deceptive content. The legal terrain is fragmented. Detection itself faces a moving target. Models considered state-of-the-art in 2022 produced visible artifacts at the eyes and hairline. Models released in 2024 and 2025 have closed those gaps.[40] Detection systems must be retrained on each new generator, and small language communities receive less detection investment than English language ones. Indian regional languages, African languages, and many South Asian dialects remain underserved by automated systems.

4.5 The Detection Arms Race

Detection methods follow generation methods at a one to two-year lag. Early detectors picked up on irregular blinking, asymmetric facial features, and inconsistent lighting. Each visible artifact has been closed in newer generation models. Current detectors rely on signal level forensics, including frequency domain analysis, sensor pattern noise, and physiological signal extraction. Provenance-based approaches offer a complementary path. The European Union draft Code of Practice on Transparency of AI-generated Content references provenance metadata as one of the layers of compliance with Article 50.[41] The Indian 2026 rules reference permanent metadata as a labelling requirement. Provenance does not solve the detection problem, but it offers a positive credential for genuine content rather than a negative search for fakes.[42]

4.6 The Impact on Journalism and Public Trust

Journalism sits at the receiving end of the deepfake problem in two ways. Newsrooms face the routine task of verifying user-generated content. Audience trust in the news product depends on the perceived ability of newsrooms to filter synthetic from genuine material.[43] Newsrooms have responded with three sets of practices. First, dedicated verification desks now process incoming user material with reverse image search, geolocation analysis, and forensic tools. Second, editorial guidelines on the use of AI-generated visuals in published stories have been adopted by Reuters, the BBC, the Associated Press, and several Indian outlets, with disclosure required when AI is used in visual reproduction.[44] Third, audience-facing explainers on how synthetic content works form a new genre of public service journalism, with The Hindu, The Indian Express, and The Wire among Indian outlets running such explainers in 2024 and 2025.[45] Public broadcasters carry a particular responsibility in this environment. Doordarshan in India, the BBC in the United Kingdom, and ARD in Germany operate under public interest mandates. The European Broadcasting Union has issued guidance on the use of AI in news production for member organizations.[46] The shared theme is editorial responsibility for any content the organization publishes, regardless of whether the content was generated by humans or machines.

4.8 Costs and Asymmetries

The economics of deepfake production and detection favour the producer side. A consumer-grade voice clone costs less than ten dollars per month in subscription services.[47] A bespoke face swap video produced by a specialist runs to a few hundred dollars. By contrast, a forensic verification of a single contested clip by a qualified analyst takes hours of work and produces a probabilistic answer. Platform-side detection systems run at scale but with the accuracy limitations noted above. The cost asymmetry favours the volume producer over the quality verifier. Distribution economics push in the same direction. The first viewer of a viral clip imposes near-zero marginal cost on the platform. The fact-checking and takedown response involves human reviewers, legal counsel, and engineering work. Each takedown is small in cost, but at the scale of platforms with billions of users, the aggregate cost is substantial. The Indian three-hour

takedown rule shifts a portion of these costs onto platforms, which in turn pass them through to advertisers and to users in the form of stricter content moderation.

5. CONCLUSION

This study shows a rapid expansion of research on deepfakes and misinformation, with output rising from 46 papers in 2019 to 768 in 2025 based on the Dimensions dataset used in this article. India ranks second in publication volume within this dataset. The literature focuses more on misinformation and fake news than on regulation and enforcement, which remain less developed in both scope and empirical evidence. The analysis identifies three patterns. First, the cost of producing synthetic media has declined, which supports continued growth in volume. Second, misinformation systems now treat synthetic content as a routine input. Third, regulation is shifting from advisories to statutory frameworks, including India's 2026 amendment and the European Union AI Act, along with state-level measures in the United States. I cannot confirm consistent enforcement outcomes across these regimes due to limited empirical data. Two research gaps require attention. Empirical work on enforcement remains limited, especially on takedown timelines, error rates, appeals, and platform compliance. Research on regional language ecosystems and cross-border coordination is also underdeveloped. India's multilingual environment requires focused study due to its scale and policy relevance. A four-part response provides a structured approach. Mandatory labelling offers baseline signals to users. Platform duties on traceability and rapid removal reduce exposure time. Clear statutory definitions support legal certainty. Media literacy enables informed interpretation of content. These elements must operate together. Labels depend on user attention. Takedown depends on speed. Legal definitions depend on careful judicial use. Literacy depends on the availability of reliable information. Synthetic media is already embedded in the information environment. Future regulatory effectiveness will depend on measurable enforcement, institutional capacity, and continued adaptation to emerging forms of content.

REFERENCES

[1] Waseem, S., Bakar, S. A. R. S. A., Ahmed, B. A., Omar, Z., Eisa, T. A. E., & Dalam, M. E. E. (2023). DeepFake on face and expression swap: A review. *IEEE Access*, *11*, 117865-117906.

[2] Mohan, S., & Wadhwa, S. (2023). Deepfakes and Shallow Laws: Regulating Distorted Narratives in the Political Cyberspace. *Indian JL & Tech.*, *19*, 85.

[3] Bengesi, S., El-Sayed, H., Sarker, M. K., Houkpati, Y., Irungu, J., & Oladunni, T. (2024). Advancements in generative AI: A comprehensive review of GANs, GPT, autoencoders, diffusion model, and transformers. *IEEe Access*, *12*, 69812-69837.

[4] Raman, R., Nair, V. K., Nedungadi, P., Sahu, A. K., Kowalski, R., Ramanathan, S., & Achuthan, K. (2024). Fake news research trends, linkages to generative artificial intelligence and sustainable development goals. *Heliyon*, *10*(3).

[5] Ünver, A. (2023). Emerging technologies and automated fact-checking: tools, techniques and algorithms. *Techniques and Algorithms (August 29, 2023)*.

[6] Matić, A. (2025). *Video activism: lens of public change that can go both ways* (Doctoral dissertation, Sveučilište u Rijeci, Sveučilište u Rijeci, Filozofski fakultet, Odsjek za kulturalne studije).

[7] Newman, N., Ross Arguedas, A., Robertson, C. T., Nielsen, R. K., & Fletcher, R. (2025). *Digital news report 2025*. Reuters Institute for the study of Journalism.

[8] Babayeva, M., Dunlap, J., Snětinová, M., & Widenhorn, R. (2025). Translating the Force Concept Inventory in the age of AI. *arXiv preprint arXiv:2508.13908*.

[9] Millière, R. (2022). Deep learning and synthetic media. *Synthese*, *200*(3), 231.

[10] Khoo, B., Phan, R. C. W., & Lim, C. H. (2022). Deepfake attribution: On the source identification of artificially generated images. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *12*(3), e1438.

[11] Heidari, A., Jafari Navimipour, N., Dag, H., & Unal, M. (2024). Deepfake detection using deep learning methods: A systematic and comprehensive review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *14*(2), e1520.

[12] Meding, K., & Sorge, C. (2025). What constitutes a deep fake? The blurry line between legitimate processing and manipulation under the EU AI Act (Version 2). arXiv. https://arxiv.org/html/2412.09961v2

[13] Kashyap, S. (2025). The Digital Mirage: India's Evolving Legal Battle Against Deepfake Technology. *SCRIPTed: A Journal of Law, Technology & Society*, *22*(2), 162-226.

[14] Patel, Y., Tanwar, S., Gupta, R., Bhattacharya, P., Davidson, I. E., Nyameko, R., ... & Vimal, V. (2023). Deepfake generation and detection: Case study and challenges. *IEEE Access*, *11*, 143296-143323.

[15] Pallewatta, P., Arachchi, S. M., Cheok, A. D., Rizvan, S., Karunanayaka, K., Kayuni, T., & Aleksandr, T. (2025, November). Human Voice Synthesis and Cloning Using Generative AI Models: A Comprehensive Review of Recent Advances and Applications. In 2025 7th International Conference on Electrical, Control and Instrumentation Engineering (ICECIE) (pp. 378-385). IEEE.

[16] Chadwick, A., & Stanyer, J. (2022). Deception as a bridging concept in the study of disinformation, misinformation, and misperceptions: Toward a holistic framework. Communication Theory, 32(1), 1-24.

17 Payton, T. (2024). Manipulated: Inside the Cyberwar to hijack elections and distort the truth. Bloomsbury Publishing PLC.

[18] Karnouskos, S. (2020). Artificial intelligence in digital media: The era of deepfakes. IEEE transactions on technology and society, 1(3), 138-147.

[19] Surfshark. (2025, July 29). *Global map: Election-related deepfakes reached 3.8B people*. https://surfshark.com/research/chart/election-related-deepfakes

[20] Bharati, D. R. K. (2025). Identity Theft and Impersonation in Cyberspace (Sections 66C and 66D of the IT Act, 2000). *Available at SSRN 5379783*.

[21] Ministry of Electronics and Information Technology. (2026). *FAQs on synthetically generated information (SGI)*. https://www.meity.gov.in/static/uploads/2025/10/065b6deb585441b5ccdf8be42502a49c.pdf

[22] Artificial Intelligence Act. (n.d.). *Article 50: Transparency obligations for providers and deployers of certain AI systems*. https://artificialintelligenceact.eu/article/50/

[23] European Commission. (2025, December 17). *First draft code of practice on transparency of AI-generated content*. https://digital-strategy.ec.europa.eu/en/library/first-draft-code-practice-transparency-ai-generated-content

[24] Panditharatne, M., Norden, L., Zdanys, J., Weiner, D. I., & Abusaif, Y. (2025). *Agenda to Strengthen US Democracy in the Age of AI*. Washington, DC: Brennan Center for Justice at New York University School of Law.

[25] Nguyễn, T. (2024, September 19). *2 of California's 3 new deepfake laws are being challenged in court by creator of Kamala Harris parody videos*. Fortune. https://fortune.com/2024/09/19/california-deepfake-laws-challenged-court-creator-kamala-harris-parody-videos/

[26] Federal Communications Commission. (2024, May 23). *Notice of apparent liability for forfeiture: Illegal AI-generated robocalls (Steve Kramer)*.

https://www.doj.nh.gov/sites/g/files/ehbemt721/files/inline-documents/sonh/item-2-kramer-robocall-nal.pdf

[27] Müller, M., & Kettemann, M. C. (2023). European approaches to the regulation of digital technologies. In *Introduction to digital humanism: A textbook* (pp. 623-637). Cham: Springer Nature Switzerland.

[28] Hanspal, M. S., & Tomar, D. S. (2026). *Digital identity rights: A comparative analysis of Denmark's and India's approaches to deepfakes*. *Maharaja Surajmal Institute Law Journal, 2*(1), 20–30. https://doi.org/10.48165/msilj.2025.2.1.3

29 Kumar, D. (2025). Deepfakes, free speech, and the right to truth: A comparative legal study on regulating synthetic media in the USA, UK, and India. Advanced International Journal for Research, 6(4), 1117.

[30] Kishwar, S. D., Tripathi, A., Khatoon, S., Poddar, D., & Khurana, B. (2025). Regulating deep fakes and synthetic media: Privacy, policy and global regulatory challenges. Journal of Data Protection & Privacy, 8(1), 78-96.

[31] Feng, Z., Li, Z., Cui, H., & Whitty, M. T. (2025). Identity Management Systems: A Comprehensive Review. Information, 16(9), 778.

[32] Francis, N. (2025). Deepfake Detection and Defense: An Analysis of Techniques and Robustness.

[33] Giacomello, A. (2025). Technological deceptions: An analysis of how deepfakes are framed in the press.

[34] Gupta, N., & Mathews, N. (2024, September 25). *India's experiments with AI in the 2024 elections: The good, the bad & the in-between*. Tech Policy Press. https://techpolicy.press/indias-experiments-with-ai-in-the-2024-elections-the-good-the-bad-the-inbetween/

[35] Hailey, A. J. (2025). *Digital deception: how artificial intelligence poses a threat to the United States' domestic elections* (Doctoral dissertation).

[36] Downes, R. T. (2025, September). Generative AI's Implications for Democratic Elections in a World at Risk: A Dahlian Perspective. In *Electoral Integrity Project, Technology and the Future of Democracy, Pre-APSA Workshop*.

[37] Biju, P. R., & Gayathri, O. (2025). Indic approach to ethical AI in automated decision making system: implications for social, cultural, and linguistic diversity in native population. *AI & SOCIETY*, *40*(8), 6645-6670.

[38] Nguyen, T. (2024, July 29). *Elon Musk defends posting a deepfake video of Kamala Harris that may violate X's policies*. Business Insider. https://www.businessinsider.com/elon-musk-reposted-deepfake-kamala-harris-violate-x-policy-2024

[39] Economic Times. (2026, February 18). *Meta flags operational challenges in India's 3-hour takedown rule, backs consultative AI regulations*. https://economictimes.indiatimes.com/tech/artificial-intelligence/meta-flags-operational-challenges-in-indias-3-hour-takedown-rule-backs-consultative-ai-regulations/articleshow/128471460.cms

[40] Apila, A. (2025). *Impacts of facial image distortion on verification accuracy in the context of the Entry/Exit System* (Master's thesis, A. Apila).

[41] Feltes, N. (2025). Article 50 AI Act: Do the Transparency Provisions Improve Upon the Commission's Draft?. *J. Intell. Prop. Info. Tech. & Elec. Com. L.*, *16*, 222.

[42] WITNESS. (2026, February 18). *India's synthetic media rules build enforcement on the wrong foundation*. https://www.witness.org/indias-synthetic-media-rules-build-enforcement-on-the-wrong-foundation/

[43] Bendahan Bitton, D., Ötting, H., Godulla, A., & Hoffmann, C. P. (2026). Implications of deepfake technology for journalistic practice: A technology at the crossroads of challenges and opportunities. Journalism, 14648849261418328.

[44] Mahajan, J. S. (2025). AI-driven media landscape in India: Transforming journalism, media studies and consumption trends. GAP Bodhi Taru: A Global Journal of Humanities, 8, 28-36.

[45] The Hindu. (2025, October 24). *Synthetic media: On labelling of AI-generated content*. https://www.thehindu.com/opinion/editorial/synthetic-media-on-labelling-of-ai-generated-content/article70194342.ece

[46] Lefèvre, B., Errando, A., Afilipoaie, A., Ranaivoson, H., & Wiart, L. (2025). Exploring ethical and regulatory challenges of AI integration in European Union newsrooms. *Media Studies*, *16*(31), 31-55.

[47] DANIEL, S. (2025). *PROSPECTS AND CHALLENGES OF GOOGLE FLOW AND* (Doctoral dissertation, DEPARTMENT OF THEATRE ARTS, FACULTY OF ARTS, UNIVERSITY OF IBADAN, IBADAN).

Cite this Article:

Dr. Manindra Singh Hanspal," **Deepfake Technology, Misinformation and Media Regulation: Emerging Trends and Challenges***", Naveen International Journal of Multidisciplinary Sciences (NIJMS), ISSN: 3048-9423 (Online), Volume 2, Issue 6, pp. 17-31, June-July 2026.*

Journal URL: https://nijms.com/

***DOI:** https://doi.org/10.71126/nijms.v2i5.139*