

Misinformation and Disinformation: Threats to Information Ecosystem in the Modern Era

Sumiti Ahuja*

ABSTRACT

In an age where information is instantly accessible, the spread of false and misleading content—whether intentional (disinformation) or unintentional (misinformation)—poses a significant threat to the global information ecosystem. During the COVID-19 pandemic, WHO Director-General Tedros Adhanom Ghebreyesus warned that the world was confronting not only an epidemic but an “infodemic,” as misleading content circulated with the same speed and danger as the virus itself. The internet and social media have amplified this challenge, enabling false information to reach vast audiences at minimal cost and unprecedented scale. Such distortions of truth undermine human security, obstruct justice, and create barriers to accessing reliable information. The rapid expansion of Artificial Intelligence has further intensified vulnerabilities within digital information spaces worldwide. Sustainable Development Goal 16, promoting “Peace, Justice and Strong Institutions,” particularly Target 16.10, emphasizes public access to accurate information and the protection of fundamental freedoms. In India, Article 19(1)(a) guarantees free speech while simultaneously placing a responsibility on institutions and individuals to ensure the dissemination of truthful information. Several legal and policy initiatives—such as the IT Rules 2021, the Disaster Management Act 2005, and the Digital Personal Data Protection Act 2023—seek to address these challenges. Globally, the UN’s “Global Principles for Information Integrity” aim to build healthier information environments grounded in human rights and sustainable development. Ensuring information integrity and advancing the SDGs ultimately require the coordinated efforts of governments, tech companies, media, and the public.

Keywords: Fake News, Misinformation, Disinformation, Co-regulation, Algorithmic Accountability.

* Assistant Professor, Campus Law Centre, University of Delhi.

1. Introduction

Disingenuous information can now be produced easily and realistically, thanks to the rapid advancements in information and communication technologies (ICT) and their growing pervasiveness.¹ It can also be disseminated to a targeted audience at a never-before-seen speed and scale, including through artificial intelligence (AI) techniques. Misinformation and disinformation pose immense risks to the information ecosystem in the contemporary era. There are severe repercussions with wide-ranging effects. Effective action is essential to strengthen the integrity of information spaces to maintain freedom of expression and democratic participation.

Misinformation is defined as incorrect or misleading information that may or may not have a specific harmful intent. It encompasses inaccurate, incomplete, misleading, or deceptive information, as well as selective or half-truths.² Disinformation, on the other hand, is misinformation coupled with the intent to deceive. Both misinformation and disinformation may be false. The essential distinction is the intent behind them. Misinformation is shared without malevolent purpose, whereas disinformation is false information shared intentionally to deceive.³ The key causes of the rise in disinformation and misinformation are diminishing faith in the mass media, the impact of blogging, the social media, and the proliferation of “clickbait media”.⁴

The key challenges in dealing with misinformation are identifying and separating it from genuine content.⁵ This complexity derives from the rapid expansion and dissemination of information *via* digital platforms. Furthermore, balancing the fundamental right to free speech and expression with content regulation by state actors presents a significant

¹ Jon Bateman and Dean Jackson, *Countering Disinformation Effectively: An Evidence-Based Policy Guide* 1-4 (Carnegie Endowment for International Peace, 2024), available at: https://carnegie-production-assets.s3.amazonaws.com/static/files/Carnegie_Countering_Disinformation_Effectively.pdf (last visited on Dec. 20, 2024).

² Yanmengqian Zhou and Lijiang Shen, “Processing of Misinformation as Motivational and Cognitive Biases” 15 *Frontiers in Psychology* 1 (2024).

³ Insights, “Misinformation vs. Disinformation”, available at: <https://insights.taylorandfrancis.com/social-justice/misinformation-vs-disinformation/> (last visited on Dec. 20, 2024).

⁴ Clickbait media is content that uses dramatic headlines or visuals to entice readers to click a link. The idea is to drive traffic to a website or blog to earn money through advertisements. [See Michael Pauly, “What Is the Meaning of Clickbait and Is It Dangerous?” Apr. 20, 2024, available at: <https://www.avg.com/en/signal/what-is-clickbait-is-it-dangerous> (last visited on Dec. 23, 2024)].

⁵ Esma Aïmeur, Sabrina Amri, *et.al.*, “Fake news, Disinformation and Misinformation in Social Media: A Review” 13 *Social Network Analysis and Mining* 8 (2023).



challenge. It involves considerable planning to avoid censorship while effectively addressing harmful misinformation.⁶

Misinformation and disinformation can negatively impact the Sustainable Development Goals (SDGs) by undermining the efforts to promote sustainable development, peace, and security.⁷ Misinformation and disinformation can threaten: (i) 'democracy' by undermining trust in institutions and elections, and fueling societal divisions; (ii) 'public health' by jeopardizing efforts to address pandemics and other public health issues; (iii) 'environment' by weakening the efforts to conserve oceans, seas, and marine resources, and to address climate change; (iv) 'peace and security' by subverting the efforts to promote peace and security, and to resolve conflicts; and (v) 'vulnerable communities' by the targeting of marginalized and vulnerable communities, and further their social, economic, and political exclusion.⁸

The author has tried to elaborate upon the importance of information and access to information for citizens as part of SDGs, risks posed by mis- and disinformation due to continuous growth and advancement in information technology, measures taken at both national and global level to counter this ever-growing menace.

2. Importance of Information

Information is a fact, thought, or data given or described *via* many modes of communication, such as written, oral, visual, and auditory communications.⁹ It is knowledge acquired or communicated by study, instruction, investigation, or news, and is shared through the act of communicating, whether audibly, non-verbally, visually, or in writing.¹⁰

⁶ Prof. Rasmus Kleis Nielsen, "How to Respond to Disinformation While Protecting Free Speech", *Reuters Institute for the Study of Journalism*, Feb. 19, 2021, *available at*: <https://reutersinstitute.politics.ox.ac.uk/news/how-respond-disinformation-while-protecting-free-speech> (last visited on Dec. 29, 2024).

⁷ United Nations, "Information Integrity and the Sustainable Development Goals", *available at*: <https://www.un.org/sites/un2.un.org/files/information-integrity-and-sdgs-en.pdf> (last visited on Dec. 29, 2024).

⁸ OECD, "Mis- and disinformation", *available at*: <https://www.oecd.org/en/topics/disinformation-and-misinformation.html> (last visited on Dec. 29, 2024).

⁹ Colin James, "The 4 Types of Communication – Definitions and Examples", *available at*: <https://colinjamesmethod.com/the-4-types-of-communication/> (last visited on Dec. 26, 2024).

¹⁰ Indeed Editorial Team, "6 Types of Information (With Examples)", *available at*: <https://www.indeed.com/career-advice/career-development/types-of-information> (last visited on Dec. 26, 2024).

*Misinformation and Disinformation: Threats to Information
Ecosystem in the Modern Era*

Information is of vital importance for the public at large and is equated to a public good.¹¹ It is so because, information is both, non-rivalrous and non-excludable. In other words, there is no dearth of information, it is something which if used by one, does not affect the ability of others to make use of it (non-rivalrous); it is also of such a nature that no individual should be prevented from accessing information (non-excludable).¹²

The importance of information for the public at large is mainly five-fold.¹³ First, it promotes democracy, as accessibility to information is an essential component of democracy and is inextricably tied to freedom of speech. It enables citizens to make informed decisions about problems that are pertinent to them and to determine if the governments are keeping their commitments. Second, it enhances the accountability of the government, since information helps to guarantee that the government is transparent and accountable to its citizens. Third, access to information contributes to trust building by strengthening the bond of trust between any government and its citizens. Fourth, it promotes the rule of law, as access to information is both a right as well as a multiplier of other rights¹⁴, forming an integral part of the freedom of speech and expression and working towards the promotion of the rule of law. Finally, it furthers the empowerment of citizens, as the basic object of the right to information is empowerment of citizens and the promotion of vigilance amongst them, so that the government is transparent and accountable.

India, in the year 2005, passed the Right to Information Act (RTI Act)¹⁵, which grants citizens the right to access information. Long Title of the RTI Act, *inter alia*, states that

“.....to secure access to information under the control of public authorities, in order to promote transparency and accountability in the working of every public authority.....”¹⁶

¹¹ For a good to qualify as 'public good', the former must be non-rivalrous and non-excludable. It means, something which if consumed by one, does not affect the ability of the other to consume it (non-rivalrous); it should not be of such a nature that any individual can be prevented from its consumption (non-excludable). [See Stanford Encyclopedia of Philosophy, “Public Goods”, Jul. 21, 2021, *available at*: <https://plato.stanford.edu/entries/public-goods/> (last visited on Dec. 27, 2024)].

¹² John M. Abowd, “Private and Public Data Provision in the U.S.” (Jan. 2007), *available at*: <https://ecommons.cornell.edu/server/api/core/bitstreams/082711d7-ecfa-42c1-bbb3-99719afeeab8/content> (last visited on Dec. 27, 2024).

¹³ Laura Neuman (ed.), *Access to Information: A Key to Democracy* (The Carter Center, United States, 2002), *available at*: <https://www.cartercenter.org/documents/1272.pdf> (last visited on Dec. 28, 2024).

¹⁴ UNESCO, “Right to Information”, *available at*: <https://www.unesco.org/en/right-information> (last visited on Dec. 30, 2024).

¹⁵ The Right to Information Act, 2005 (Act 22 of 2005).

¹⁶ See The Right to Information Act, 2005 (Act 22 of 2005), *available at*: <https://www.indiacode.nic.in/bitstream/123456789/2065/5/a2005-22.pdf> (last visited on Dec. 29, 2024).



Apart from this, information also plays crucial role in personal growth of an individual, as it helps the latter in making informed decisions towards achievement of his/her goals. The flow of information facilitates social integration and development, as well as the coordination and control of people's actions.¹⁷

3. Misinformation and Disinformation: Meaning

False or erroneous information, a misrepresentation of the facts, is known as misinformation. False information purposefully meant to mislead or, intentionally misrepresent the facts is known as disinformation.¹⁸ In other words, disinformation is false, inaccurate, or misleading information that is shared with the intent to deceive the recipient, as opposed to misinformation which refers to false, inaccurate, or misleading information that is shared without any intent to deceive.¹⁹

Misinformation and disinformation are phenomena that are being examined more and more; in other words, they are now identified as a significant social issue that is being researched a lot around the world and seems to be becoming a bigger worry. Given the abundance of erroneous number of hoax claims on COVID-19 that were made and disseminated during the pandemic, such a measure does not come as a surprise. The Director General of the World Health Organization (WHO), Tedros Adhanom Ghebreyesus,²⁰ while addressing the issue of misinformation said, “We're not just fighting an epidemic; we're fighting an infodemic²¹”, and pointed out that misleading information that was doing rounds during such sensitive times was spreading as fast as the virus, and was as dangerous²² as the

¹⁷ LIS Education Network, “The Needs and Importance of Information”, Jan. 06, 2014, *available at*: <https://www.lisedunetwork.com/the-needs-and-importance-of-information/> (last visited on Dec. 29, 2024).

¹⁸ American Psychological Association, “Misinformation and Disinformation”, *available at*: <https://www.apa.org/topics/journalism-facts/misinformation-disinformation> (last visited on Dec. 29, 2024).

¹⁹ Canadian Centre for Cyber Security, “How to Identify Misinformation, Disinformation, and Malinformation (ITSAP.00.300)” (May 2024), *available at*: <https://www.cyber.gc.ca/en/guidance/how-identify-misinformation-disinformation-and-malinformation-itsap00300> (last visited on Dec. 30, 2024).

²⁰ Ethiopian diplomat, scholar, biologist, public health officer, who assumed the office of Director-General of WHO in 2017, becoming the First African to hold that position. [See Kara Rogers, “Tedros Adhanom” *Encyclopedia Britannica*, *available at*: <https://www.britannica.com/biography/Tedros-Adhanom> (last visited on Dec. 30, 2024)].

²¹ The novel coronavirus (nCoV) has brought back this phrase, which was first used during the SARS pandemic. It is defined as “an abundance of information about a problem that is considered detrimental to its solution” [See Dictionary.com, “Infodemic”, *available at*: <https://www.dictionary.com/browse/infodemic> (last visited on Dec. 31, 2024)].

²² United Nations Department of Global Communications, “UN Tackles 'Infodemic' of Misinformation and Cybercrime in COVID-19 Crisis”, Mar. 31, 2020, *available at*: <https://www.un.org/en/un-coronavirus-communications-team/un-tackling-%E2%80%98infodemic%E2%80%99-misinformation-and-cybercrime-covid-19> (last visited on Dec. 31, 2024).

latter. Several fake videos and articles were being circulated on social media: how to cure COVID-19 by drinking alcohol, or, how vegetable vendors were licking vegetables to spread the virus, on how cow urine kills the virus, on complete lockdown lasting for more than three months, xenophobia²³, etc.

The virtue of knowledge is threatened by mis/disinformation. In the digital age, where consumers of media platforms are exposed to a variety of convincing assertions with dubious provenance and a lack of authenticity, this scenario is especially pertinent.²⁴ At its core, disinformation is an assault on the veracity of knowledge. The information ecosystem becomes contaminated when people believe incorrect information to be true and it spreads throughout society.²⁵ This is especially the case in the field of politics, which is where the concept of post-truth²⁶ originated. According to a study,²⁷ the purpose of disinformation could be scientific purpose, political propaganda, cultural purpose, financial purpose, advertising or clickbait, humorous purpose, religious propaganda, and at times, there may not be any specific purpose at all.

4. Misinformation and Disinformation: Threats to the Information Ecosystem & Measures To Combat

4.1 Threats to the Information Ecosystem

An information ecosystem is a complex system that describes the flow and transformation of information through a network of people, tools, and institutions.²⁸ The intricate web of information sources, audiences, and technology that influence the production, dissemination, and consumption of information is referred to as an

²³ Attitudes, biases, and actions that despise, exclude, and even demonize individuals because they are considered to be outsiders or foreigners to the community, culture, or sense of national identity. [See EMN Asylum and Migration Glossary, “Xenophobia”, available at: https://home-affairs.ec.europa.eu/networks/european-migration-network-emn/emn-asylum-and-migration-glossary/glossary/xenophobia_en (last visited on Jan. 01, 2025)].

²⁴ Marta Pérez-Escolar and Darren Lilleker, “A Systematic Literature Review of the Phenomenon of Disinformation and Misinformation” 11(2) *Media and Communication* 76 (2023).

²⁵ Gleb Tsipursky, “Towards a Post-lies Future: Fighting “Alternative Facts” and “Post-truth” Politics” 77(2) *The Humanist*, 12 (2017).

²⁶ According to Cambridge Dictionary, post-truth is defined as “...a situation in which people are more likely to accept an argument based on their emotions and beliefs, rather than one based on facts” [See Cambridge Advanced Learner's Dictionary & Thesaurus, “Post-truth”, available at: <https://dictionary.cambridge.org/dictionary/english/post-truth> (last visited on Jan. 01, 2025)].

²⁷ *Supra* note 24.

²⁸ Internews: Center for Innovation & Learning, “Mapping Information Ecosystems to Support Resilience” (2015), available at: https://internews.org/wp-content/uploads/legacy/resources/Internews_Mapping_Information_Ecosystems_2015.pdf (last visited on Jan. 01, 2025).



information ecosystem.²⁹ The public, citizen journalism,³⁰ social media, and traditional media all engage dynamically to shape the information flow in this ecosystem.³¹

From a technological perspective, it can be inferred that the major factors to have contributed to the increased generation and circulation of mis/disinformation can be said to be as follows. These include the democratization of content generation, which has enabled widespread production and sharing of information; economic incentives and fast-paced news cycles, which encourage rapid dissemination without adequate verification; widescale and prompt reach facilitated by digital platforms; deliberately created filter-bubbles that limit users' exposure to diverse viewpoints; algorithmic content curation and the absence of transparency in how such systems operate; and the scalability and anonymity in online accounts, both of which allow mis/disinformation to spread quickly and often without accountability.

Democratization of content generation: The democratization of content generation refers to making it easier for people to generate and distribute digital content.³² It allows ordinary users to produce, edit, personalize, and distribute digital content, something which earlier could only have been possible by traditional content suppliers in established media businesses. This has been made feasible by the introduction of new technology such as AI and cloud computing,³³ as well as the increased availability and lower cost of new media tools.³⁴ Some benefits of democratization of content creation are that it is more inclusive, more diverse, it has leveled the playing field for businesses of all sizes, and has blurred the lines of distinction, to some extent, between amateur and professional content creators.

Economic incentives and fast-paced news cycles: The more hits a story receives, the more money is generated through ad revenue. To increase the number of hits/clicks, content

²⁹ See Fiveable: Honors Journalism, "Information Ecosystem", available at: <https://library.fiveable.me/key-terms/hs-journalism/information-ecosystem> (last visited on Jan. 03, 2025).

³⁰ Type of journalism in which citizens or non-professionals collect and disseminate news and information. Citizen journalists disseminate their work via websites, blogs, and social media. [See Sonny Albarado, "Citizen Journalism" *Encyclopedia Britannica*, Jan. 25, 2024, available at: <https://www.britannica.com/topic/citizen-journalism> (last visited on Jan. 05, 2025)].

³¹ Kehinde Segun, "Exploring the Impact of Citizen Journalism on Traditional Media" 1(3) *International Journal of Human Research and Social Science Studies* 63-73 (Sep., 2024).

³² Joshua, "The Democratization of Content Creation: How AIGC Makes it Accessible and Affordable", available at: <https://quickcreator.io/articles2/aigc-democratization-content-creation/> (last visited on Jan. 03, 2025).

³³ Delivery of computing services via the internet, or "the cloud". It enables customers to access computing resources such as storage, servers, and databases on demand. [See Microsoft Azure, "What is Cloud Computing?: A Beginner's Guide", available at: <https://azure.microsoft.com/en-us/resources/cloud-computing-dictionary/what-is-cloud-computing> (last visited on Dec. 30, 2024)].

³⁴ Oxford Reference, "Democratization of Content", available at: <https://www.oxfordreference.com/display/10.1093/acref/9780191800986.001.0001/acref-9780191800986-e-3140> (last visited on Dec. 30, 2024).

creators compete to provide attention-grabbing material that appeals to consumers' emotions or cognitive biases,³⁵ given the fast-paced news cycle and multiple information sources.

Widescale and prompt reach: Content created in any part of the world, can be viewed in any other part of the world, nearly instantaneously.³⁶ Furthermore, content creators receive rapid feedback on the performance of their content *via* likes, share, comments, reactions, etc. The latter emphasizes upon the interactivity part of the digital content creation and dissemination.³⁷

Deliberately created filter-bubbles: When algorithms provide us with information that aligns with our beliefs based on our past behaviour and search history, it can lead to a state of intellectual or ideological isolation known as a “filter bubble”.³⁸ Organisations that want to disseminate mis/disinformation might purposefully produce filter bubbles by precisely and directly targeting potentially vulnerable people, for example, through sponsored posts or advertisements.³⁹

Algorithmic content curation and absence of transparency: The practice of selecting, organizing, and presenting content to consumers based on their preferences and behaviours is known as algorithmic curation.⁴⁰ Online platforms, like social media and search engines, employ it. However, these algorithms are not transparent to end users. Due to this intricacy, viewers may find it difficult to comprehend why they are experiencing particular online watching experiences, where the information was actually coming from, or even that their experiences differ from those of other users.⁴¹

Scalability and anonymity in online accounts: The weak identity and account management systems of many online platforms can be used by individuals who want to

³⁵ Buster Benson, “Cognitive Bias Cheat Sheet” *Medium*, Sep. 01, 2016, available at: <https://buster.medium.com/cognitive-bias-cheat-sheet-55a472476b18> (last visited on Dec. 30, 2024).

³⁶ Centre for Media Transition, *The Impact of Digital Platforms on News and Journalistic Content* 42 (University of Technology Sydney, NSW, 2018).

³⁷ Hamidreza Shahbaznezhad, Rebecca Dolan, *et.al.*, “The Role of Social Media Content Format and Platform in Users' Engagement Behavior” 53(1) *Journal of Interactive Marketing* 47 (2021).

³⁸ Richard Fletcher, “The Truth Behind Filter Bubbles: Bursting Some Myths” *Reuters Institute*, Jan. 24, 2020, available at: <https://reutersinstitute.politics.ox.ac.uk/news/truth-behind-filter-bubbles-bursting-some-myths> (last visited on Jan. 02, 2025).

³⁹ Janna Anderson and Lee Rainie, “The Future of Truth and Misinformation Online”, Oct. 19, 2017, available at: <https://www.pewresearch.org/internet/2017/10/19/the-future-of-truth-and-misinformation-online/> (last visited on Jan. 01, 2025).

⁴⁰ Ambika Sharma, “Algorithmic Content Curation: Crafting the Ultimate User-First Experience” *Pulp Strategy*, Aug. 22, 2023, available at: <https://www.pulpstrategy.com/algorithmic-content-curation-crafting-the-ultimate-user-first-experience> (last visited on Jan. 02, 2025).

⁴¹ Arunesh Mathur, Arvind Narayanan, *et.al.*, “Endorsements on Social Media: An Empirical Study of Affiliate Marketing Disclosures on YouTube and Pinterest” 2(CSCW) *Proceedings of the ACM on Human-Computer Interaction* 119:11 (Nov., 2018).



disseminate misinformation to create many accounts (also referred to as 'Sybils' or 'bots')⁴² in order to, for instance, pretend to be authentic followers of a political or socio-welfare movement and/or fabricate the appearance that a certain piece of content is well-liked.⁴³

It is believed that psychological factors make people susceptible to think and act on misinformation. Misinformation that presents opponents in an unfavorable light is more likely to be believed by people than misinformation about their group. Lastly, even when information contradicts prior knowledge, people are more inclined to trust it if repeated.⁴⁴ Exposure to incorrect information, according to behavioural models, increases the likelihood that people would believe it, and thus the likelihood that they will spread it. Spreading false information, on the other hand, does not require individuals to accept it; they can do so to express their political allegiance, disparage opponents they regard as rivals, or obtain social benefits.⁴⁵ While older people are more likely to follow and disseminate false information on social media, they may also be more adept at identifying disinformation than younger people.⁴⁶

4.2 Measures to Combat

Two main types of interventions have been suggested to avert the growth and dissemination of false information, *viz.*, system-level interventions, such as laws and technological standards which concentrate on significant systemic changes; and individual-level interventions, which aim to modify specific behaviours. Individual-level initiatives often have fewer implications for freedom of expression and rely less on collaboration from tech businesses, but system-level interventions may be more effective in limiting the spread of erroneous information.⁴⁷

⁴² Sybils are illegal entities that pose as legitimate individuals, whereas bots are automated programs that can participate in misleading or dangerous behaviour. Their existence can result in erroneous estimates of wallet reputation and power, potentially jeopardizing the dependability of decentralized ecosystems. [See Emmmanuella Dan, "Solving the Sybil and Bot Problem: How 0xScore Enhances Decentralized Systems" *Medium*, Jan. 23, 2024, available at: <https://medium.com/@danemmanuella1999/solving-the-sybil-and-bot-problem-how-0xscore-enhances-decentralized-systems-0328e4bdeed6> (last visited on Jan. 03, 2025)].

⁴³ Zoe Adams, Magda Osman, *et.al.*, "(Why) Is Misinformation a Problem?" 18(6) *Perspectives on Psychological Science* 1445 (2023).

⁴⁴ American Psychological Association, "What Psychological Factors Make People Susceptible to Believe and Act on Misinformation?", Nov. 29, 2023, available at: <https://www.apa.org/topics/journalism-facts/misinformation-belief-action> (last visited on Dec. 26, 2024).

⁴⁵ American Psychological Association, "How and Why Does Misinformation Spread?", Nov. 29, 2023, available at: <https://www.apa.org/topics/journalism-facts/how-why-misinformation-spreads> (last visited on Dec. 27, 2024).

⁴⁶ *Supra* note 44.

⁴⁷ American Psychological Association, "What Interventions Can be Used to Counter Misinformation Effectively?", Nov. 29, 2023, available at: <https://www.apa.org/topics/journalism-facts/misinformation-interventions> (last visited on Dec. 28, 2024).

Debunking, pre-bunking, literacy training, and nudging approaches focus on changing individual behaviours.⁴⁸ Correcting false information is known as debunking or fact-checking.⁴⁹ It is applied after people have been exposed to false information, and it works best when it has a thorough justification that disproves false information and substitutes it with accurate information. Fact-checking takes a lot of time, deteriorates over time, and may need to be corrected repeatedly.⁵⁰ The goal of pre-bunking, also known as pre-emptive debunking, is to stop individuals from initially believing false information.⁵¹ Psychological inoculation, the most popular pre-bunking technique, involves exposing people to a mild form of a lie to increase their resistance to future persuasion.⁵² Literacy training assists people in improving their ability to judge the quality and accuracy of information and complete tasks. Nudges are tiny environmental modifications designed to influence behaviour in predictable and desirable ways.⁵³ Researchers have employed a variety of nudges to try to discourage people from distributing misinformation: i) Accuracy nudges, to encourage people to think about the truthfulness of information before sharing it,⁵⁴ whereas ii) social norms nudges highlight community standards of behaviour when reporting information,⁵⁵ and iii) motivational nudges reward people for being as accurate as possible.⁵⁶

5. Role of Artificial Intelligence in the Expansion of Threats

The technological miracle of electronic intelligence has catapulted us into an era in which we are building a global society whose reliance on computers is growing

⁴⁸ *Ibid.*

⁴⁹ Clara Juarez Miro and Jonathan Anderson, "Correcting False Information: Journalistic Coverage During the 2016 and 2020 US Elections" 25(2) *Journalism Studies* 220 (2024).

⁵⁰ Ethan Porter and Thomas J. Wood, "The Global Effectiveness of Fact-checking: Evidence From Simultaneous Experiments in Argentina, Nigeria, South Africa, and the United Kingdom", Sep. 10, 2021, available at: <https://www.pnas.org/doi/10.1073/pnas.2104235118> (last visited on Jan. 02, 2025).

⁵¹ Mauro Bertolotti and Patrizia Catellani, "Counterfactual Thinking As A Prebunking Strategy To Contrast Misinformation on COVID-19" 104 *Journal of Experimental Social Psychology* 2 (2023).

⁵² Toby D. Pilditch, Jon Roozenbeek, *et al.*, "Psychological Inoculation Can Reduce Susceptibility to Misinformation in Large Rational Agent Networks" 9 *Royal Society Open Science* 2 (2022), available at: <https://pmc.ncbi.nlm.nih.gov/articles/PMC9363981/pdf/rsos.211953.pdf> (last visited on Jan. 03, 2025).

⁵³ The Decision Lab, "Nudge Theory", available at: <https://thedeclaration.com/reference-guide/psychology/nudge-theory> (last visited on Jan. 04, 2025).

⁵⁴ Ethan Porter, Thomas J. Wood, *et al.*, "Shoves, Nudges and Combating Misinformation: Evidence On A New Approach" *Behavioural Public Policy* 3 (2024), available at: <https://www.cambridge.org/core/services/aop-cambridge-core/content/view/C2C2837C576F69A0F37F02FD5F9B1CF8/S2398063X24000423a.pdf> (last visited on Jan. 04, 2025).

⁵⁵ Yifat Nahmias, "Privacy Preserving Social Norm Nudges" 26 *Michigan Technology Law Review* 46 (2019).

⁵⁶ Altaf Mazhar Soomro, Gnana Bharathy, *et al.*, "A Review On Motivational Nudges For Enhancing Building energy Conservation Behavior" 1 *Journal of Smart Environments and Green Computing* 3-20 (2021).



exponentially.⁵⁷ AI technologies play a significant role in the current global disinformation epidemic. Such systems compound the problem not just by increasing the likelihood of developing realistic AI-generated fake content, but also, more crucially, by enabling hostile stakeholders to spread disinformation to a targeted audience on a massive scale. In response, more AI systems are being developed to detect and control internet misinformation. Such technologies also do not avoid ethical and human rights considerations, particularly those related to freedom of expression and information.⁵⁸

When there is too much information available, people experience cognitive overload and only engage in matter that confirms their preconceived notions. This effect is exacerbated by AI technology, which makes it possible to create false information and synthetic content. Increased societal division and difficulties in public discourse have resulted from this.

5.1 AI as a Facilitator in the Creation & Dissemination of Fake Content

AI has transformed many aspects of human existence and impacted countless businesses. This usage of technology has also generated new obstacles and risks, the most noteworthy of which is deepfakes. “Deepfakes”⁵⁹ or “synthetic media”⁶⁰ refer to the use of modified digital content, such as hyper-realistic synthetic video, audio, photos, or texts created with powerful AI capabilities, to undermine specific decision-making processes.⁶¹ The primary epistemic hazard is that deepfakes can readily induce individuals to adopt erroneous beliefs. This technique has the ability to create information that is substantially identical to the original material, and impacts target public opinion, social groups, political discourse, and both personal and national security.⁶² The use of AI to produce Deepfakes has

⁵⁷ Tom Stonier, *Information and Meaning: An Evolutionary Perspective 2* (Springer, New York, 2012).

⁵⁸ Noémi Bontridder and Yves Poulet, “The Role of Artificial Intelligence in Disinformation” 3 *Data & Policy* e32 (2021).

⁵⁹ Realistic images, videos, or sounds made with AI to make individuals or events appear real that did not occur. They are a form of synthetic media that can be used to disseminate misleading information and propaganda. [See Ian Sample, “What are Deepfakes – And How Can You Spot Them?” *The Guardian*, Jan. 13, 2020, available at: <https://www.theguardian.com/technology/2020/jan/13/what-are-deepfakes-and-how-can-you-spot-them> (last visited on Jan. 06, 2025)].

⁶⁰ Media that is created or modified using AI, and may include images, videos, text, voice, etc. Deepfake is a popular type of synthetic media. Synthetic media can be used to spread misinformation and can erode public trust. [See Giulio Corsi, Bill Marino, *et.al.*, “The Spread of Synthetic Media on X” 5(3) *Harvard Kennedy School Misinformation Review 2* (June, 2024).

⁶¹ Mekhail Mustak, Joni Salminen, *et.al.*, “Deepfakes: Deceptions, Mitigations, and Opportunities” 154 *Journal of Business Research 1* (2023).

⁶² Rakash L. Kharvi, “Understanding the Impact of AI-Generated Deepfakes on Public Opinion, Political Discourse, and Personal Security in Social Media” 22(4) *IEEE Security & Privacy* 115 (July-Aug. 2024).

become concerning for varied reasons, including the fact that results are becoming more realistic, that they can be produced quickly, and that they can be produced at a low cost using publicly available software and the ability to rent processing power through cloud computing.⁶³ Even non-professional users can obtain the necessary software tools and, by using publicly available information, can produce convincing fake content.

The Global Risks Report 2024⁶⁴ by the World Economic Forum⁶⁵ has identified misinformation and disinformation as serious dangers in the coming times, noting the possibility of domestic propaganda and censorship.⁶⁶ Political exploitation of AI offers serious hazards, with the rapid spread of deepfakes and AI-generated content making it increasingly impossible for voters to distinguish between truth and lie, thereby influencing voter behaviour and compromising the democratic process. Elections can be manipulated, public trust in institutions eroded, social discontent sparked, and violence erupted.⁶⁷

It is believed that 'false media has existed for as long as there has been media to falsify': forgers have forged documents or works of art, youngsters have falsified driving licenses, and so on.⁶⁸ Although propaganda and similar issues are not new, it is becoming increasingly obvious that the development of technology has acted as a catalyst for the mass production, distribution, and consumption of mis/disinformation.⁶⁹

5.2 AI as a Tool to Detect & Regulate False Information Online

The rise of AI in the digital age has brought both opportunities and concerns, particularly in terms of information integrity. AI technology capable of producing 'deepfakes'

⁶³ Drew Robb, "Deepfake Awareness Riding on Top Gun's Coattails" *CIOInsight*, June 02, 2022, *available at*: <https://www.cioinsight.com/news-trends/deepfake-awareness-top-gun/> (last visited on Jan. 04, 2025).

⁶⁴ Presents an in-depth study of the most serious risks confronted by the world today. It is intended to help in the identification of the major risks for 2024 and thereafter. It serves as a standard for global risk management and has an impact on policies and plans at both the organisation and the government levels. [See MarshMcLennan, "Global Risks Report 2024", *available at*: <https://www.marshmclennan.com/insights/publications/2024/january/global-risks-report-2024.html> (last visited on Jan. 06, 2025)].

⁶⁵ Non-profit organization that convenes leaders from business, politics, academia, and civil society to address global concerns. [See World Economic Forum, "Stakeholders", *available at*: <https://www.weforum.org/stakeholders/> (last visited on Jan. 04, 2025)].

⁶⁶ World Economic Forum, "The Global Risks Report 2024: 19th Edition" 8 (Jan. 2024), *available at*: https://www3.weforum.org/docs/WEF_The_Global_Risks_Report_2024.pdf (last visited on Jan. 04, 2025).

⁶⁷ Gabriel R. Sanchez and Keesha Middlemass, "Misinformation is Eroding The Public's Confidence in Democracy" July 26, 2022, *available at*: <https://www.brookings.edu/articles/misinformation-is-eroding-the-publics-confidence-in-democracy/> (last visited on Jan. 08, 2025).

⁶⁸ John Akers, Gagan Bansal, *et. al.*, "Technology-Enabled Disinformation: Summary, Lessons, and Recommendations" 4 (Dec., 2018), *available at*: <https://dada.cs.washington.edu/research/tr/2018/12/UW-CSE-18-12-02.pdf> (last visited on Dec. 28, 2024).

⁶⁹ *Id.* at 2.



can be used to spread both misinformation and disinformation. However, at the same time, AI can help battle erroneous information by studying patterns, language, and context to aid in content moderation, fact-checking, and detection of false information.⁷⁰ The same AI and data-driven approaches that are used to spread mis/disinformation can also be used to detect and regulate it.⁷¹

Machine learning (ML)⁷² algorithms enable automated and scalable identification of false information online, which is critical given the massive amount of information exchanged on social media sites. Overall, ML is an effective method for detecting and mitigating the spread of bogus content on social media.⁷³

6. Public Access to Information and SDGs

Since the Rio Declaration of 1992⁷⁴, access to information has been acknowledged as essential to sustainable development.⁷⁵ It was underlined in the 2020 Resolution of the Human Rights Council⁷⁶ as a duty of public institutions, and found place under the SDGs of 2015, to encourage participatory governance.

⁷⁰ Cathy Li and Agustina Callegari, “Stopping AI Disinformation: Protecting Truth in the Digital World” *World Economic Forum*, June 14, 2024, available at: <https://www.weforum.org/stories/2024/06/ai-combat-online-misinformation-disinformation/> (last visited on Jan. 02, 2025).

⁷¹ The University of Queensland, “How AI is Being Used to Fight Fake News”, available at: <https://sponsored.chronicle.com/how-ai-is-being-used-to-fight-fake-news/index.html> (last visited on Dec. 30, 2024).

⁷² AI is a broad field that includes ML as a subset. ML allows machines to learn from data and improve on their own. ML uses algorithms and statistical models to process large amounts of data and identify patterns. [See AWS, “What's the Difference Between AI and Machine Learning?”, available at: <https://aws.amazon.com/compare/the-difference-between-artificial-intelligence-and-machine-learning/> (last visited on Jan. 03, 2025)].

⁷³ Maialen Berrondo-Otermin and Antonio Sarasa-Cabezuelo, “Application of Artificial Intelligence Techniques to Detect Fake News: A Review” 12(24) *Electronics* 5041 (2023).

⁷⁴ The 1992 United Nations Conference on Environment and Development (UNCED), also referred to as the 'Earth Summit', adopted the Rio Declaration on Environment and Development. The Declaration identified 27 guiding principles to give countries a starting point for decisions and policies that take impact of socio-economic development on the environment into account. [See Convention on Biological Diversity, “Rio Declaration on Environment and Development”, available at: <https://www.cbd.int/doc/ref/rio-declaration.shtml> (last visited on Jan. 04, 2025).

⁷⁵ Principle 10 of the Rio Declaration states, “...at the national level, each individual shall have appropriate access to information concerning the environment that is held by public authorities” [See UN General Assembly, *Report of the United Nations Conference on Environment and Development*, GAOR, UN DocA/CONF.151/26 (Vol. I) (Aug. 12, 1992) p.no. 2].

⁷⁶ UN General Assembly, *The Role of Good Governance in the Promotion and Protection of Human Rights*, GA Res 45/9, GAOR, UN Doc A/HRC/Res/45/9 (Oct. 09, 2020).

Misinformation and Disinformation: Threats to Information Ecosystem in the Modern Era

The United Nations (UN) adopted the SDGs, sometimes referred to as the 'Global Goals'⁷⁷, in 2015 as a global call to take action towards the eradication of poverty, to safeguard the environment, and to guarantee that everyone lives in peace and prosperity by 2030. The SDGs, or Agenda 2030, are an ambitious attempt to improve the lives of the world's poorest people within 15 years, and all 193 UN Member States⁷⁸ agreed to implement them in 2015. However, the Agenda goes beyond poverty; among its 17 agreed-upon objectives are ambitious targets to address issues like gender inequality, environmental degradation, corrupt governance, and healthcare hurdles.⁷⁹ The 17 SDGs are integrated in nature, i.e., they acknowledge that development must strike a balance between social, economic, and environmental sustainability and that actions taken in one area will impact results in other areas as well.⁸⁰

The primary requirement for sustainable development is the union of three elements: economic development, environmental protection, and social equity. Thus, it is imperative that the public is allowed to access relevant information freely to ensure the successful integration of these elements and the achievement of the set goals.

Public access to information is a key part of the SDGs of the UN and is a driver for achieving the 2030 Agenda for Sustainable Development.⁸¹ Open and inclusive access to information empowers people and sets the groundwork for the creation of inclusive knowledge societies.⁸² Access to information is essential for evaluating public officials in implementing and monitoring the SDGs, as well as for empowering the public to make

⁷⁷ Set of universally accepted goals, adopted by all UN member states, that aim to address global issues such as poverty, inequality, climate change, and environmental degradation. To accomplish these goals by 2030, all nations must strive to work together, effectively making them a “global” call to action to ensure that no one is left behind. [See United Nations: Sustainable Development Goals, “The Sustainable Development Agenda”, *available at*: <https://www.un.org/sustainabledevelopment/development-agenda-retired/> (last visited on Jan. 08, 2025)].

⁷⁸ List of Member States, *available at*: <https://www.un.org/en/about-us/member-states> (last visited on Jan. 08, 2025).

⁷⁹ United Nations, “Sustainable Development Goals: 17 Goals to Transform our World”, *available at*: <https://www.un.org/en/exhibits/page/sdgs-17-goals-transform-world> (last visited on Jan. 09, 2025).

⁸⁰ United Nations Development Programme, “The SDGs in Action”, *available at*: <https://www.undp.org/sustainable-development-goals> (last visited on Jan. 01, 2025).

⁸¹ Tawfik Jelassi, “Advancing Universal Access to Information Within the 2030 Agenda for Sustainable Development” *UN Chronicle*, Sep. 27, 2021, *available at*: <https://www.un.org/en/un-chronicle/advancing-universal-access-information-within-2030-agenda-sustainable-development> (last visited on Jan. 08, 2025).

⁸² UNESCO, “Inclusive Knowledge Societies for Sustainable Development”, Mar. 2012, *available at*: https://www.un.org/en/development/desa/policy/untaskteam_undf/groupb_unesco_knowledge_societies.pdf (last visited on Jan. 09, 2025).



decisions, hold governments accountable, and promote “just, peaceful, and inclusive societies”,⁸³ which is SDG Goal 16 of the UN 2030 Agenda for Sustainable Development.⁸⁴

According to the Instruction Manual of the United Nations Educational, Scientific and Cultural Organization (UNESCO)⁸⁵ Survey on 'Public Access to Information',⁸⁶ the latter is founded on the fundamental freedom of expression and association, which is an established human right and States are responsible for upholding this right.⁸⁷ In order to uphold this right, States are obliged to create a legal framework which gives the public the right to request access to information, and the State must promptly respond to such requests. The relevant authorities must proactively publish information of public interest, in public domain.⁸⁸

The right to information (RTI) is not merely an objective; it is a necessary condition for attaining sustainable development in its entirety.⁸⁹ RTI aims to increase access to water and healthcare by giving people the knowledge they need to demand services, enabling communities to hold businesses and governments responsible for contaminating their water supplies, and holding governments responsible for unethical behaviour.⁹⁰ Thus, RTI encourages accountability in development-related concerns just as it does in more political ones, and encompasses more than just government pledges. It gives people the ability to take part, advocate for, and track significant advancements towards development objectives.⁹¹

⁸⁵ UNESCO has been identified as that UN custodian agency, for SDG Indicator 16.10.2, which is responsible to track the overall progress in the number of those countries that have adopted and effectively implemented legal guarantees with respect to 'Access to Information' [See Andrey Richter, *The need to accelerate worldwide progress: UNESCO 2023 report on public access to information (SDG 16.10.2)* (UNESCO, Paris, 2024)].

⁸⁶ Right to obtain information held by public authorities. It is a fundamental right which is part of freedom of expression. (See The Universal Declaration of Human Rights, 1948, art. 19).

⁸⁷ UNESCO, “Instruction Manual: UNESCO Survey on Public Access to Information (SDG Indicator 16.10.2)” 9-10 (2021), *available at*: <https://eyeonglobaltransparency.net/wp-content/uploads/2021/04/Instruction-Manual-UNESCO-Survey-on-Public-Access-to-Information-English-1.pdf> (last visited on Jan. 02, 2024).

⁸⁸ United Nations, “SDG Indicator Metadata”, *available at*: <https://unstats.un.org/sdgs/metadata/files/Metadata-16-10-02.pdf> (last visited on Jan. 09, 2025).

⁸⁹ Shamsul Bari and Ruhi Naz, “Revisiting The Role of RTI For Sustainable Development” *The Daily Star*, Apr. 16, 2021, *available at*: <https://www.thedailystar.net/opinion/news/revisiting-the-role-rti-sustainable-development-2078061> (last visited on Jan. 10, 2025).

⁹⁰ Sahina Mumtaz Laskar, “Importance of Right to Information for Good Governance in India” *Bharati Law Review* 216-229 (Oct.-Dec., 2016).

⁹¹ Harsh Mander and Abha Joshi, “The Movement for Right to Information in India”, *available at*: <https://www.humanrightsinitiative.org/programs/ai/rti/india/articles/The%20Movement%20for%20RTI%20in%20India.pdf> (last visited on Jan. 10, 2025).

SDG Goal 16, 'Target 16.10' refers to ensuring public access to information and protection of fundamental freedoms, in accordance with national legislation. The proposed indicator for the same is "...existence and implementation of a national law and/or constitutional guarantee on the right to information".⁹² For instance, article 19(1) (a) of our Constitution grants freedom of speech and expression. At the same time, various institutions must release reliable information to avoid both, misinformation, and disinformation. If citizens have a right to access information, then the relevant institutions must have a corresponding duty to provide them with reliable information.

The 2015 SDG agreement places a strong emphasis on transparency, and countries everywhere have pledged to include the right to information in their legal frameworks. Although more 'right to information' laws have been passed (90 percent of the world's population currently resides in a country with an RTI law), implementation has not always kept up with the trend, and progress is undoubtedly not universal.⁹³ To secure development by the year 2030, the SDGs must foster partnership among the commercial sector, governments, civil society, and educational institutions.⁹⁴

7. United Nations Global Principles for Information Integrity

The idea behind the 'UN Global Principles for Information Integrity' was first presented in the UN Secretary-General's 'Our Common Agenda' Report,⁹⁵ and it was expanded upon in the 'Information Integrity on Digital Platforms' policy brief⁹⁶. 'Our Common Agenda' presents the vision of the UN Secretary-General on the future of global cooperation through an "inclusive, networked, and effective multilateralism"⁹⁷. A new social contract based on human rights, improved management of vital global commons, global public goods that provide equitable and sustainable benefits to all, and a revived sense of

⁹² Sustainable Development Solutions Network, "Indicators and a Monitoring Framework", *available at*: <https://indicators.report/targets/16-10/> (last visited on Jan. 01, 2025).

⁹³ Article 19, "Access to information is critical to achieving the SDGs", Sep. 28, 2018, *available at*: <https://www.article19.org/resources/access-to-information-is-critical-to-achieving-sdgs/> (last visited on Jan. 02, 2025).

⁹⁴ Maltez Alberto Mabuie, "Access to Information: A Key Driver To Achieve Sustainable Development Goals" 4(12) *American Research Journal of Humanities & Social Science* 26 (2021).

⁹⁵ The UN Secretary-General's landmark report presented at the 75th session of the General Assembly, *available at*: https://www.un.org/en/content/common-agenda-report/assets/pdf/Common_Agenda_Report_English.pdf (last visited on Jan. 04, 2025).

⁹⁶ United Nations, "SG's Policy Brief on Information Integrity on Digital Platforms", *available at*: <https://www.un.org/en/civil-society/information-integrity-digital-platforms> (last visited on Jan. 09, 2025).

⁹⁷ United Nations, "Secretary-General's Report "Our Common Agenda"" (2021), *available at*: https://www.un.org/en/content/common-agenda-report/assets/pdf/Common_Agenda_Key_Proposals_English.pdf (last visited on Jan. 05, 2025).



solidarity between peoples and future generations are the four main themes covered by the recommendations in 'Our Common Agenda'.⁹⁸

The Secretary-General in the policy brief on 'Information Integrity on Digital Platforms', has called for guardrails for addressing the global threat of online hate speech, misinformation, and disinformation. A need for coordinated international action has been emphasized in order to ensure safety and inclusiveness of digital space, at the same time focusing upon effective protection of human rights.⁹⁹ The policy brief delves deep into the adverse effect of threats to information integrity on advancements in local, national, and international issues, and has been promoted to provide a gold standard to guide action for strengthening information integrity.¹⁰⁰

In June 2024, the UN Global Principles for Information Integrity, came in force. This document's main purpose was to serve as a resource for the UN Member States, before the 'Summit of the Future'¹⁰¹, which was to be held in September 2024. These principles aimed to combat the spread of online misinformation, disinformation, and hate speech, and alleviate the broad harm caused by incorrect information on digital platforms. Its five guiding principles are: “societal trust and resilience; healthy incentives; public empowerment; independent, free, and pluralistic media; and transparency and research”.¹⁰²

Thereafter, in September 2024, 'Summit of the Future' was held with the goal to improve and bolster global governance. The outcome of this summit was “Pact for the Future”, alongwith two annexes, namely “Global Digital Compact” and “Declaration on Future Generations”.¹⁰³ The goal of the Pact is to accelerate SDGs, among other things, and

⁹⁸ *Ibid.*

⁹⁹ United Nations, “Guardrails Urgently Needed to Contain “Clear and Present Global Threat” of Online Mis- and Disinformation and Hate Speech, says UN Secretary-General” June 12, 2023, *available at*: <https://www.un.org/sustainabledevelopment/blog/2023/06/press-release-guardrails-urgently-needed-to-contain-clear-and-present-global-threat-of-online-mis-and-disinformation-and-hate-speech-says-un-secretary-general/> (last visited on Jan. 09, 2025).

¹⁰⁰ *Supra* note 96.

¹⁰¹ United Nations: Regional Information Centre for Western Europe, “The Summit of the Future (22-23 September 2024)” Apr. 15, 2024, *available at*: <https://unric.org/en/the-summit-of-the-future-in-2024/> (last visited on Jan. 10, 2025).

¹⁰² IISD, “UN Launches Global Principles for Information Integrity” July 12, 2024, *available at*: <https://sdg.iisd.org/news/un-launches-global-principles-for-information-integrity/> (last visited on Jan., 10, 2025).

¹⁰³ United Nations, “Pact for the Future, Global Digital Compact and Declaration on Future Generations” (Sep. 2024), *available at*: https://www.un.org/sites/un2.un.org/files/sotf-pact_for_the_future_adopted.pdf (last visited on Jan. 05, 2025).

the Global Digital Compact encourages fair access to technology and ensures that everyone benefits from technology.

8. Indian Initiatives to Curb Mis- and Dis-information

8.1 Legislative Provisions

India lacks specific legislation to regulate the dissemination of mis/disinformation. However, several initiatives have been taken by the government to tackle this menace – from time to time, relevant provisions of the Epidemic Diseases Act, 1897, the Disaster Management Act, 2005, and the Bharatiya Nyaya Sanhita, 2023, have been invoked to prosecute wrongdoers accused of spreading rumors and/or creating panic amongst the populace. Provisions of the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, the Information Technology Amendment Rules, 2023, the Digital Personal Data Protection Act, 2023, and the Telecommunications Act, 2023 also restrict access to the alleged misinformation.

i) *The Epidemic Diseases Act, 1897*: This law was originally designed to control the spread of contagious diseases, and can also be used to address misinformation during public health crises.¹⁰⁴ It grants the government the authority to implement measures that prevent disease transmission, including regulating the dissemination of false information which could worsen the situation, and provides for penalties against individuals spreading harmful misinformation. Additionally, it facilitates legal action against those deliberately spreading falsehoods that threaten public health, thus helping to mitigate the impact of misinformation during epidemics. The cobwebs surrounding this law were dusted off during the Covid19 pandemic. Originally designed to address the spread of infectious diseases during the British colonial period, this law has been criticized for granting broad and often unchecked powers to authorities, which can lead to potential abuse. The absence of safeguards for transparency and accountability further exacerbates its shortcomings, rendering it an inadequate and potentially dangerous instrument for addressing contemporary public health challenges.¹⁰⁵

ii) *The Disaster Management Act, 2005*: It contains provisions to punish persons who raise a false alarm or circulate a false warning related to a disaster or the severity/extent of such

¹⁰⁴ Rakesh P.S., “Implementing the Epidemic Diseases Act to Combat Covid-19 in India: An Ethical Analysis” *Indian Journal of Medical Ethics*, available at: <https://ijme.in/articles/implementing-the-epidemic-diseases-act-to-combat-covid-19-in-india-an-ethical-analysis/?galley=print> (last visited on Feb. 25, 2025).

¹⁰⁵ Parikshit Goyal, “The Epidemic Diseases Act, 1897 Needs an Urgent Overhaul” 55(45) *Economic and Political Weekly: Engage* (Nov. 07, 2020).



disaster, causing panic among the public. Punishment extends to imprisonment for one year or fine.¹⁰⁶ Fifteen years after its enactment, this Act was invoked for the very first time during the Covid19 pandemic, since lots of false or unverified information was being circulated through dubious and unverified channels.¹⁰⁷ Some such information pertained to the cure and precautions against Covid19 virus, whereas, other mis/disinformation had the potential to create fear and panic in the minds of the already scared people. Thus, the law which was enacted as a response to the 2004 Tsunami that struck eastern coast of India, and saw thousands of casualties, was put to use in the country.¹⁰⁸

iii) *Bharatiya Nyaya Sanhita, 2023* (BNS): Recognizing the serious implications of misinformation and the crucial necessity to combat it, the BNS has adopted significant new measures to punish the deliberate manufacture, distribution, or publication of incorrect information. It makes it illegal to make, publish, or disseminate statements, rumours, false information, or reports (including through electronic means) with the intention or potential of causing harmful consequences.¹⁰⁹ Misinformation has traditionally been utilized to incite public fear that could result in crimes against the state or public tranquillity, or to incite one class or community to commit crimes against another, this section consequently includes misinformation within its purview. The Sanhita stipulates imprisonment for up to three years, a fine, or both for commission of such offence. Promoting animosity, hostility, or malice amongst various racial, religious, linguistic, or regional groups is also punishable under this provision.¹¹⁰ Longer period of imprisonment of up to five years and a fine has been imposed to deter such offences in places of worship or during religious ceremonies.

The BNS also aims to regulate misinformation by punishing anyone who makes or publishes false or misleading information, whether it be in the form of spoken or written words, signs, visible representations, or through electronic medium, which results in 'jeopardizing the sovereignty, unity, integrity, or security of India'.¹¹¹ The accused in such a case is liable to face imprisonment for up to three years, a fine, or both. But if it occurs in a

¹⁰⁶ The Disaster Management Act, 2005 (Act 53 of 2005), s. 54.

¹⁰⁷ Chetan Chauhan, "Covid-19: Disaster Act Invoked for the 1st Time in India" *Hindustan Times*, Mar. 25, 2020, available at: <https://www.hindustantimes.com/india-news/covid-19-disaster-act-invoked-for-the-1st-time-in-india/story-EN3YGrEuxhnl6EzqrleWM.html> (last visited on Mar. 10, 2025).

¹⁰⁸ Neeraj Chauhan, "Disaster Act Provisions Lifted After 2 yrs of Covid" *Hindustan Times*, Mar. 24, 2022, available at: <https://www.hindustantimes.com/india-news/disaster-act-provisions-lifted-after-2-yrs-of-covid-101648060066847.html> (last visited on Mar. 10, 2025).

¹⁰⁹ The Bharatiya Nyaya Sanhita, 2023 (Act 45 of 2023), s. 353. The provision is titled 'Statements conducing to public mischief'.

¹¹⁰ *Id.*, s. 353(2).

¹¹¹ *Id.*, s. 197(1)(d). The provision is titled 'Imputations, assertions prejudicial to national integration'.

place of worship or during religious ceremonies, the maximum imprisonment of up to five years can be levied which may include a fine.¹¹²

When a person, who is under legal obligation to disclose information to a public servant, knows or reasonably believes that the information is false and yet provides it, can face six months in prison, a fine of up to five thousand rupees, or both.¹¹³ However, if such false information is with respect to: an offence committed, prevention of an offence, or is crucial to apprehend an offender, the punishment is increased to up to two years in prison, a fine, or both.¹¹⁴

Whoever, with the intent to influence the outcome of an election, makes or publishes any statement purporting to be a statement of fact that is false and which, he either knows or believes to be false or does not believe in the truth of it, concerning the personal character or conduct of any candidate, shall be fined¹¹⁵ under the Sanhita.

iv) *The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021* (IT Rules, 2021):¹¹⁶ The IT Rules, 2021 impose obligations on intermediary platforms to prohibit the hosting, sharing, uploading, transmitting, and so on of any prohibited information, including misinformation and evidently false information on the Internet, as well as impersonating another person.¹¹⁷

Non-compliance by intermediary platforms may result in loss of intermediary status and consequent liability for hosted content, as well as penalization under the Information Technology Act, 2000 (IT Act).¹¹⁸ If any information falls into the categories banned by this Rule, the user may contact the Grievance Officer of the relevant intermediary.¹¹⁹ Upon receipt of the request, the intermediary must act within the timeframe¹²⁰ specified in the IT Rules, 2021.

¹¹² *Id.*, s. 197(2).

¹¹³ *Id.*, s. 212(1). The provision is titled 'Furnishing false information'.

¹¹⁴ *Id.*, s. 212(2).

¹¹⁵ *Id.*, s. 175. The provision is titled 'False statement in connection with an election'.

¹¹⁶ The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 were notified under s.87(2) of the Information Technology Act, 2000, superseding the earlier Information Technology (Intermediary Guidelines) Rules, 2011.

¹¹⁷ Ministry of Electronics & IT, "Government of India Committed Towards Preventing Dissemination of Misinformation Including Deep Fakes on Internet" *PIB Delhi*, July 31, 2024, available at: <https://pib.gov.in/PressReleasePage.aspx?PRID=2039640> (last visited on Jan., 11, 2025).

¹¹⁸ The Information Technology Act, 2000 (Act 21 of 2000), s. 45; The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, rule 7.

¹¹⁹ The IT Rules, 2021, *Id.*, rule 3 sub-rule 2.

¹²⁰ *Ibid.*



v) *The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2023* (IT Amendment Rules, 2023): According to the IT Amendment Rules, 2023, social media platforms could be compelled to take down any content that has been verified as untrue by the Press Information Bureau's fact-check unit (FCU).¹²¹ These rules aim to stop false information and fake news from spreading on social media. The central government can remove whatever information that the fact-checking unit deems untrue. The amendment included a provision that revokes the legal protection afforded to social media platforms and intermediaries,¹²² if they fail to comply with the directives of the government.¹²³

On January 31, 2024, a Division Bench of the Bombay High Court comprising Patel and Gokhale, JJ., delivered a split decision in the case of *Kunal Kamra v. Union of India*.¹²⁴ This case was concerned with the constitutional validity of the 2023 Amendment, since it allegedly took away the immunity or, safe harbour, granted to intermediaries from liability for third-party content uploaded on their platforms, if the intermediary failed to remove such content flagged by the FCU of the central government as fake, false, or misleading. As per Patel, J., the Amendment introduced restrictions to freedom of speech which were not reasonable in nature, and jeopardized the commercial interests of the intermediaries. Gokhale, J. ruled in favour of the Amendment, and upheld the setting up of fact-check units. The matter was referred to third judge, Chandurkar, J., who rejected the interim relief application. Thereafter, the central government notified the coming in force of FCU.¹²⁵

Thereafter, the Supreme Court temporarily stayed the government from utilizing the services of fact-check units over social media content.¹²⁶ This ruling came after comedians,

¹²¹ The Press Information Bureau (PIB) has been at the forefront of proactive attempts to counteract fake news about the Indian government. In November 2019, PIB formed FCU to combat fake news about the Government of India, its different ministries, departments, public sector undertakings, and other central government organizations. See Press Information Bureau, "PIB Fact Check Unit", available at: <https://pib.gov.in/aboutfactchecke.aspx?reg=3&lang=1> (last visited on Jan. 12, 2025).

¹²² *Supra* note 118, s. 79.

¹²³ Hemendra Singh, "The IT Amendment Rules, 2023: Censorship in the Guise of Fact-checking" 58(43) *Economic & Political Weekly* (Oct. 28, 2023), available at: <https://www.epw.in/journal/2023/43/commentary/it-amendment-rules-2023.html> (last visited on Jan. 12, 2025).

¹²⁴ 2024 SCC OnLine Bom 360.

¹²⁵ Ministry of Information & Broadcasting, "Government Notifies PIB's Fact Check Unit Under IT Rules 2021" *PIB Delhi*, Mar. 20, 2024, available at: <https://pib.gov.in/PressReleasePage.aspx?PRID=2015792> (last visited on Jan. 13, 2025).

¹²⁶ See *Editors Guild of India v. Union of India & Ors*, decision dated March 21, 2024, of the Supreme Court of India [SLP(C) No. 6871-6873/2024]. Also see Anmol Kaur Bawa, "Supreme Court Stays Centre's Notification Of 'Fact Check Unit' Under IT Rules Till Final Decision By Bombay HC" *Live Law*, Mar. 21, 2024, available at: <https://www.livelaw.in/top-stories/supreme-court-kunal-kamra-editors-guild-notifying-fact-check-unit-it-rules-2023-252998> (last visited on Jan. 13, 2025).

journalists, and media organizations challenged one of the rules¹²⁷ under the IT Amendment Rules, 2023, which would allow the government to classify information about its activity as 'fake news', and requiring social media platforms to either remove such content or face losing legal protection.

A three-judge Bench, comprising Dr. D.Y. Chandrachud, J.B. Pardiwala, and Manoj Misra, JJ. overturned the Bombay High Court's ruling dated March 11, 2024, which refused to stop the execution of the 2023 Amendment Rules, as well as the subsequent order permitting the central government to notify the formation of FCUs. The Bench further highlighted that the concerned Rules raised some critical constitutional questions pertaining to freedom of speech and expression, and the Bombay High Court was asked to decide upon the challenges raised against the IT Amendment Rules, 2023. Chandurkar, J. upheld the judgment of Patel, J. and struck down the 2023 Amendment¹²⁸ for violating the principles of equality before law¹²⁹, freedom of speech and expression,¹³⁰ and freedom to practice any profession¹³¹ etc., and for exceeding the scope of the Information Technology Act, 2000 (the "IT Act").

vi) *The Digital Personal Data Protection Act, 2023* (DPDP Act): A data principal¹³², *i.e.*, an individual whose data is being processed has been granted certain rights¹³³ and duties¹³⁴ under the Act. Apart from the right to access information about processing, the right to seek correction and erasure of personal data, grievance redressal, etc., the data principals have some duties as well. Duty to refrain from registering any false or frivolous complaint, and duty not to furnish any false particulars or impersonate another person in specified cases. In case of violation of these duties, the Act levies a penalty of upto rupees 10,000¹³⁵.

The Act clearly defines 'personal data' as 'names, photos, email addresses, social media posts, biometric data,¹³⁶ and online activities'; this ensures that deepfake incidents can be effectively addressed within the legal framework offered by the Act, and mandates the data fiduciaries,¹³⁷ *i.e.*, entities responsible for determining the purpose and means of

¹²⁷ Rule 3(1)(b)(v) of the IT Amendment Rules, 2023. The Rule mandates intermediaries to ensure that users do not host or share information identified as fake or misleading by a fact-check unit of the central government.

¹²⁸ *Kunal Kamra v. Union of India*, 2024 SCC OnLine Bom 3025.

¹²⁹ The Constitution of India, art. 14.

¹³⁰ *Id.*, art. 19(1)(a).

¹³¹ *Id.*, art. 19(1)(g).

¹³² The Digital Personal Data Protection Act, 2023 (Act 22 of 2023), s. 2(j).

¹³³ *Id.*, ss. 11-14.

¹³⁴ *Id.*, s. 15.

¹³⁵ *Id.*, s. 33(1) read with the Schedule.

¹³⁶ *Id.*, s. 2(t).

¹³⁷ *Id.*, s. 2(i).



processing personal data, to have strong security measures to avoid any form of breach of personal data.

Suppose a social media website collects and retains users' facial biometric information for verification of their identities. Under the DPDP Act, this website needs to maintain strong security steps¹³⁸ like encrypting biometric data when storing and sending it, limiting access to only authorized staff, and carrying out frequent security checks to find its weaknesses. If the data fiduciary does not have requisite security measures in place which leads to data, it may lead to creation of deepfake photographs and videos using facial biometric data of users. In such a situation, the data fiduciary would be held responsible for such breach. The DPDP Act promotes pro-activity when it comes to data security, and reduces dangers related to deepfakes.¹³⁹

vii) *The Telecommunications Act, 2023*: It allows for temporary suspension of internet services to ensure law and order, public safety, and other related concerns. This action is carried out under the Telecommunications (Temporary Suspension of Services) Rules, 2024, issued under section 20(2) of the Telecommunications Act. Additionally, section 14(4) of the Act regulates the shutdown of telecommunication networks, except in cases of natural disasters or public emergencies. Earlier, the Indian Telegraph Act, 1885 provided the government with the authority to suspend internet services. Till 2017, internet shutdowns were primarily issued under section 144 of the recently repealed Code of Criminal Procedure, 1973. The content of the erstwhile provision remains unchanged, and can now be found under section 163 of the Bharatiya Nagarik Suraksha Sanhita, 2023. The relevant provision granted the District Magistrate, the authority to prohibit unlawful assemblies and instruct individuals to refrain from specific activities.

8.2 Executive Actions: Internet Shutdowns in India

Data from the Internet Shutdown Tracker maintained by Software Freedom Law Centre India,¹⁴⁰ reveals that our country witnessed a total of 60 mobile internet shutdowns in the calendar year 2024, which was the lowest in the last eight years.¹⁴¹ This reduction was

¹³⁸ *Id.*, s. 8(5).

¹³⁹ Arhant Kumar and Akash Kumar Sahu, "Deepfakes and The DPDP Act: Can The DPDP Act Effectively Combat AI-Generated Misinformation" *Law, Humanities and Social Sciences Collective*, Oct. 14, 2024, available at: <https://lhssccollective.in/deepfakes-and-the-dpdp-act-can-the-dpdp-act-effectively-combat-ai-generated-misinformation/> (last visited on Jan. 14, 2025).

¹⁴⁰ The first legal services organization in India focused solely on technology, law, and policy. It is a non-profit organization that works to empower Indian citizens about their digital freedom and rights. See Sfel.in: Internet Shutdowns, "About Us", available at: <https://internetshutdowns.in/about/> (last visited on Mar. 20, 2025).

¹⁴¹ Aroon Deep, "Internet Shut Down 60 Times in 2024, Fewer Than Last Year" *The Hindu*, Dec. 29, 2024, available at: <https://www.thehindu.com/news/national/internet-shut-down-60-times-in-2024-fewer-than-last-year/article69039755.ece> (last visited on Mar. 20, 2025).

occasioned on account of fewer shutdowns in the troubled provinces of Manipur and Jammu & Kashmir in 2024¹⁴².

Why does an internet shutdown take place first? The major arguments given by the government are on the lines of upkeep of national security and counter insurgency, as a response to communal violence and ethnic clashes, to counter the problems of mis/disinformation and hate speech, and maintenance of law and order.¹⁴³ Following the abrogation of article 370 of the Constitution in August 2019, internet shutdowns in Jammu and Kashmir were repeatedly prolonged in order to combat separatist propaganda and militant activity.¹⁴⁴ In 2023, to avert additional violence after ethnic hostilities in Manipur, mobile internet services were terminated.¹⁴⁵ During the anti-Citizenship Amendment Act (CAA) protests in 2019, and farm bill protests of 2020-2021, internet services were shut down to maintain public order in the sites of protest.¹⁴⁶ In September 2021, Rajasthan enforced a state-wide internet shutdown, specifically targeting mobile internet, SMS/MMS, and social media, in several districts to prevent cheating during the Rajasthan Eligibility Exam for Teachers (REET).¹⁴⁷ The Jharkhand state government had also imposed a blanket ban on internet services across the state in 2024 to guard against rumors with respect to prevalence of unfair practices in the Jharkhand General Graduate Level Combined Competitive Examination (JGGLCCE) in 2024.¹⁴⁸

The issue which now arises is that if the Hon'ble Supreme Court of India has affirmed internet freedom under the article 19(1) (a) of the Constitution,¹⁴⁹ then to what extent is it

¹⁴² Kanav Narayan Sahgal, "India's Internet Shutdown Crisis: A Growing Threat to Digital Rights" *Global Voices Advox*, Jan. 06, 2025, available at: <https://advox.globalvoices.org/2025/01/06/indias-internet-shutdown-crisis-a-growing-threat-to-digital-rights/> (last visited on Mar. 23, 2025).

¹⁴³ Ivy Dhar, "Commentary: Navigating Tensions and Stabilising Public Safety with Internet Shutdowns" 5(2) *Indian Public Policy Review* 59 (2024).

¹⁴⁴ Geeta Seshu, "Kashmir Media Policy: Accentuating the Curbs on the Freedom of Press" *Economic and Political Weekly: Engage*, available at: <https://www.epw.in/engage/article/kashmir-media-policy-accentuating-curbs-freedom-press> (last visited on Mar. 22, 2025).

¹⁴⁵ Apar Gupta, "Apar Gupta Writes: In Manipur, Another Internet Shutdown, A Conflict Intensified" *The Indian Express*, Oct. 06, 2023.

¹⁴⁶ Manisha Madapathi, "Digital Barricades and Blackouts: A Case of Internet Shutdowns in India" 12 *Media and Communication* 8511 (Dec. 03, 2024), available at: <https://www.cogitatiopress.com/mediaandcommunication/article/download/8511/4096> (last visited on Mar. 10, 2025)

¹⁴⁷ ANI, "Rajasthan Suspends Mobile Internet In Numerous Districts To Prevent Cheating In Teachers' Exam" *The Print*, Sep. 26, 2021, available at: <https://theprint.in/india/education/rajasthan-suspends-mobile-internet-in-numerous-districts-to-prevent-cheating-in-teachers-exam/740572/> (last visited on Mar. 10, 2025).

¹⁴⁸ TOI News Desk, "Jharkhand Suspends Mobile Internet to Check Malpractice During Exam: All You Need to Know" *The Times of India*, Sep. 21, 2024, available at: <https://timesofindia.indiatimes.com/india/jharkhand-suspends-mobile-internet-to-check-malpractice-during-exam-all-you-need-to-know/articleshow/113539597.cms> (last visited on Mar. 10, 2025).

¹⁴⁹ *Anuradha Bhasin v. Union of India* (2020) 3 SCC 637.



justifiable and how long can such shutdown prevail? The public has right to access information and as highlighted in this Paper earlier, in contemporary times people rely upon internet mainly for such information because of low-cost, speed, and reach. Repetitive internet shutdowns apart from violating fundamental rights, also have certain other serious implications like economic impact,¹⁵⁰ unemployment,¹⁵¹ difficulty in online filing of complaints by abused women, undermining of media and press freedom, disruption of education and health care, etc.

In *Foundation for Media Professionals v. Union Territory of Jammu and Kashmir & Anr.*,¹⁵² the Supreme Court addressed a critical issue concerning the balance between national security and fundamental rights, particularly the right to freedom of speech and expression, and access to education, health, and business during a period of restricted internet access in Jammu and Kashmir. Referring to the earlier case of *Anuradha Bhasin*,¹⁵³ the Court emphasized that internet restriction should meet minimum standards, such as being proportionate, transparent, and temporary. While the Court found that the government had not provided sufficient reasoning for blanket enforcement of internet shutdown across all districts, it acknowledged the compelling circumstances of cross-border terrorism. Consequently, the Court did not find a constitutional violation, but ordered the formation of a Special Committee, led by the Union Home Secretary, to assess the necessity of the internet restrictions.¹⁵⁴

9. Conclusion

Fake and distorted information can escalate social, political, and economic conflicts, erode public confidence in democratic institutions, and even put citizens' lives in jeopardy. This is particularly problematic in the age of social media when information spreads quickly. Maintaining democratic discourse and social cohesiveness requires striking a balance between the advantages of information availability and the threats posed by social media and AI. Ongoing efforts in technology standards and legislation are essential to control the possible risks of false information and synthetic content.

New technologies pose changing opportunities and problems in the information arena. The development of generative AI will amplify changes in the information

¹⁵⁰ *Ibid.*

¹⁵¹ According to the Internet Society's Net Loss Calculator, a single day of shutdown in India can result in the unemployment of up to 379 persons. [See Aditi Agrawal, "India Leads in Global Internet Shutdowns for 6th Year in A Row: Report" *Hindustan Times*, May 15, 2024].

¹⁵² (2020) 5 SCC 746.

¹⁵³ *Supra* note 149.

¹⁵⁴ *Supra* note 152.

environment even further. AI solutions can accurately detect a wide range of misleading information and recognize disinformation strategies used by bots and deepfakes. At the same time, the technical limits of AI and other technologies highlight the necessity for a hybrid strategy that combines human engagement with technological tools. To effectively develop and implement such techniques, governments must facilitate collaboration among researchers, industry professionals, and the scientific community.

The government should actively recognize its crucial role in ensuring the integrity of the information space. This can be achieved through regulations that enhance transparency, accountability, and diversity of information sources, covering both traditional media and digital platforms. Regulatory measures should focus on promoting media pluralism, ensuring media independence, and holding online platforms accountable for their role in the information ecosystem. There is a need to push for greater transparency within digital platforms. This includes initiatives that require platforms to disclose more about their risk management processes, algorithms, data flows, and content moderation mechanisms. By improving access to such information, government can better hold platforms accountable and ensure that they operate in a manner that respects public interest.

The same human rights protections afforded to offline conduct must be extended to online activities. The government should ensure that any legal limitations on online expression are enforced in a consistent, transparent, and fair manner, upholding due process while respecting freedom of speech. As individuals encounter an increasing array of information sources, from traditional media to social media, it is crucial to equip them with the tools necessary to navigate this complexity. Lastly, there is no single solution to tackling misinformation and disinformation. Rather, a multi-faceted approach that integrates regulatory measures, educational initiatives, and collaborative efforts across sectors is key to effectively addressing these challenges.