


# Chapter 18

## Thread and Threat of AI: Analyzing Social Implications of Fake News and Hate Speech on Social Media Platforms in India

Aiswaria G. Shajan

 <https://orcid.org/0000-0002-0230-5989>

Central Institute of Indian Languages, India

### ABSTRACT

*The rapid expansion of social media has transformed global communication and information sharing, fostering unprecedented connectivity. However, these platforms have also become hubs for the spread of fake news and hate speech, which threatens social cohesion. This article explores AI's dual role in combating fake news and hate speech on social media platforms in India. While AI offers promising solutions for detecting and reducing harmful content, it raises ethical and social concerns that may aggravate existing societal tensions. The analysis emphasizes the need for a balanced approach that effectively manages harmful content while preserving freedom of speech and avoiding excessive censorship. It also discusses the potential for AI to evolve into a more ethical, inclusive, and culturally sensitive tool within India's dynamic digital landscape, and highlights the importance of interdisciplinary collaboration among AI developers, policymakers, and social scientists to ensure that AI technologies serve both technological innovation and social equity.*

DOI: 10.4018/979-8-3693-9904-0.ch018

## 1. INTRODUCTION

Artificial intelligence (AI) is a general-purpose technology that has existed since the early 1950s refers to a system designed with the help of various programs and electronic circuits, aimed at solving complex real-world problems more easily and efficiently. AI is expected to behave like a human, demonstrating intelligent behavior and responding spontaneously to different situations. A *social system* defines a structured set of relationships and interactions among individuals, groups, and institutions within a society, encompassing the roles, norms, values, and practices that govern social behavior and maintain societal order. *Social media*, as a digital platform, facilitates interactive communication, networking, and community building, enabling users to engage in real-time discussions and share personal or public information. Together, AI, social systems, and social media play integral roles in shaping modern communication, societal behaviors, and technological advancements. Artificial Intelligence (AI) and social media are increasingly intertwined in personal life, providing greater convenience and contributing to safer online spaces. However, while AI helps with content moderation by detecting harmful material such as hate speech and misinformation, it can sometimes go beyond convenience. The algorithms that power these platforms can lead to addictive behaviors, making social media usage difficult to avoid and, at times, causing negative effects on users' well-being. This dual impact—both beneficial and harmful—raises concerns about finding a balance between the positive aspects of AI and the potential negative consequences it may have on individuals and society.

Major among these potential negative consequences are the issues of fake news and hate speech, which can have significant societal impacts. Although fake news and hate speech have existed throughout human history, their nature and reach have changed drastically. According to Lea Bander and Jochen Bender (2022), *the emergence of social media has changed how, where and with what effects fake news and hate speech occur*. Nowadays, fake news and hate speech are becoming more prevalent on social media. Usually, *fake news* is defined as false or misleading information presented as news, often with the intention of deceiving or manipulating public opinion. It is harmful because it distorts the truth, misinforms the public, and undermines trust in credible sources of information. As for *hate speech*, it refers to any form of communication that incites or promotes hatred, violence, or discrimination against individuals or groups based on characteristics such as race, religion, ethnicity, gender, or sexual orientation. Hate speech is dangerous because it fosters intolerance, encourages discrimination, and can lead to social unrest or violence.

Understanding the phenomena of fake news and hate speech within India's diverse socio-political landscape is crucial, particularly when considered in the context of Artificial Intelligence (AI). With its vast and varied population—comprising

16 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: [www.igi-global.com/chapter/thread-and-threat-of-ai/371747](http://www.igi-global.com/chapter/thread-and-threat-of-ai/371747)

## Related Content

---

### Wrapper Induction Programs as Information Extraction Assistants

Klaus Jantke and Carsten Müller (2007). *Intelligent Assistant Systems: Concepts, Techniques and Technologies* (pp. 35-63).

[www.irma-international.org/chapter/wrapper-induction-programs-information-extraction/24172](http://www.irma-international.org/chapter/wrapper-induction-programs-information-extraction/24172)

### Detecting Distributed Predicates Using Genetic Algorithms

Eslam Al Maghayreh, Iyad Abu Doushand and Faisal Alkhateeb (2013). *International Journal of Intelligent Information Technologies* (pp. 56-70).

[www.irma-international.org/article/detecting-distributed-predicates-using-genetic/75546](http://www.irma-international.org/article/detecting-distributed-predicates-using-genetic/75546)

### Applications of Machine Learning (ML) and Artificial Intelligence (AI) in Marketing

Shikha Mittal, Shivani Tyagi and Mani Dubish (2025). *AI and Machine Learning Applications in Supply Chains and Marketing* (pp. 23-54).

[www.irma-international.org/chapter/applications-of-machine-learning-ml-and-artificial-intelligence-ai-in-marketing/359823](http://www.irma-international.org/chapter/applications-of-machine-learning-ml-and-artificial-intelligence-ai-in-marketing/359823)

### Human Action Recognition Using Median Background and Max Pool Convolution with Nearest Neighbor

Bagavathi Lakshmi and S. Parthasarathy (2019). *International Journal of Ambient Computing and Intelligence* (pp. 34-47).

[www.irma-international.org/article/human-action-recognition-using-median-background-and-max-pool-convolution-with-nearest-neighbor/225769](http://www.irma-international.org/article/human-action-recognition-using-median-background-and-max-pool-convolution-with-nearest-neighbor/225769)

### A Study of Different Color Segmentation Techniques for Crop Bunch in Arecanut

Siddesha S, S K. Niranjana and V N. Manjunath Aradhya (2016). *Handbook of Research on Advanced Hybrid Intelligent Techniques and Applications* (pp. 1-28).

[www.irma-international.org/chapter/a-study-of-different-color-segmentation-techniques-for-crop-bunch-in-arecanut/140449](http://www.irma-international.org/chapter/a-study-of-different-color-segmentation-techniques-for-crop-bunch-in-arecanut/140449)