

Journal of Multidisciplinary Studies and Research

Volume 2, Issue 1, 2026

Indexed Journal
Peer Reviewed Journal

Social Media in the Age of Generative AI: Trust, Misinformation, and the Transformation of Social Reality

Tanjuma Mohima

Lecturer, Department of Sociology, Dhaka International University

Orcid Id: 0009-0002-7850-1537

E-mail: tanjuma.mohima31@gmail.com

ABSTRACT

The rapid emergence of Generative Artificial Intelligence (GenAI) has significantly changed the ways information is produced, distributed, and understood within social media. Previously a platform for personal interactions and user-generated content, social media is becoming a space where both human- and machine-generated information is produced without any visible distinction between the two types of information. It has become possible to produce, modify, and distribute texts, images, sounds, and videos at a remarkable speed, offering new forms of communication along with new concerns related to misinformation, credibility, manipulation, and trust. This paper will analyze how the merger of social media and generative AI changes the way social reality is constructed and how the basis for public trust is changing. Based on the integrative review of the existing scientific literature, institutional reports, and interdisciplinary studies in artificial intelligence, misinformation, social media, trust, and digital communication, this research reveals some interrelated changes. First, generative AI has significantly lowered the costs and requirements for producing realistic synthetic content. Second, the growing challenge of distinguishing between authentic information and information that was produced by AI leads to an extended “epistemic uncertainty,” in which individuals might become unsure not only of whether a particular piece of information is true or false, but also of whether they can recognize authentic information. Third, amplification and engagement techniques used on the platform could lead to faster dissemination of content that might evoke strong emotions or be misleading in nature. Fourth, exposure to misleading and synthetic information could lead to decreasing trust among people and in institutions, as well as result in “reality apathy,” when individuals become demotivated to verify the authenticity of content. Fifth, the effects of GenAI communication differ across individuals depending on their digital literacy, socio-economic status, education, and institutional capacity. The paper suggests a multi-dimensional approach linking generative AI, misinformation, algorithmic amplification, digital literacy, trust, and the social construction of reality. The study finds that technological measures alone will not suffice and recommends improvements in AI literacy, platform accountability, provenance systems, media education, transparency in institutions, and governance.

Keywords: generative artificial intelligence, social media, misinformation, social trust, digital literacy, algorithmic communication, synthetic media, social reality, AI-generated content, digital society.

Introduction

The arrival of social media changed the structure of human communications (Benkler et al., 2018). The information which used to be generated and disseminated by governments, newspapers, television channels, and publishers could all of a sudden be generated and disseminated by regular people (Benkler et al., 2018). Platforms such as Facebook, YouTube, Instagram, TikTok, and X facilitated a space where people acted not only as the receivers of information but also as its producers and distributors (Benkler et al., 2018). However, digital transformation of communication has also raised a number of social issues. The speedy distribution of false information, misleading, conspiracy theories, altered images, fake news

stories, and sensational content is increasingly complicating the process of individuals' discernment of credibility (Wardle & Derakhshan, 2017; Molina et al., 2021). It was shown that disinformation can circulate quickly on social networks and that emotionally salient and novel information may receive more attention (Vosoughi et al., 2018). It is not merely that people are exposed to false information. The digital communication environment increasingly affects how people understand what information is reliable, whom people trust, and how they interpret their social reality (Giddens, 1990; Luhmann, 1979). The advent of generative artificial intelligence brings a new dimension into this situation. Generative AI technologies can create sophisticated text, images, audio, video, and synthetic content in general (OpenAI, 2023; NIST, 2023). The creation of large language and image generation models has drastically decreased technical requirements for the production of credible digital content (OpenAI, 2023; NIST, 2023). An individual lacking in technical skill can today create an image of an event that never took place, write an article that lacks original reportage, copy someone else's voice, and produce a video of a politician saying things that were never said before. Hence, the link between representation and reality has become quite intricate. The problem is especially hard to tackle in social media since these networks feature speed, scale, personalization, networking, and algorithmic recommendation (Benkler et al., 2018; UNESCO, 2023). Generative AI is capable of creating content in bulk which could not have been done by humans alone. Meanwhile, the algorithms of platforms decide what content will reach the user and in what frequency. Therefore, automated content creation and automated content dissemination lead to the emergence of an information ecosystem where misinformation may potentially be generated and propagated at the same time. Misinformation affects more than just an individual since trust plays an essential role in any kind of human interaction (Luhmann, 1979; Giddens, 1990). Individuals need trust in order to relate to others, institutions, governments, media outlets, science, and markets. In his work, Luhmann (1979) characterized trust as a way in which people handle complexity and uncertainty. Trust was equally highlighted by Giddens (1990) in relation to modern social systems. When people are not able to clearly differentiate between true and false information, there is the risk that the problems will be experienced at a larger scale than just certain pieces of information. Here arises a critical sociological question of what will happen to social reality when individuals are unable to tell whether what they see and hear and read is real. Social reality, according to Berger and Luckmann (1966), is produced through interactions, institutionalization, and common understanding. In the modern world, where social media are becoming one of the environments for the process, synthetic information, when used in such a way, turns into an actor in the process of producing socially significant representations. It relates to computer science since AI produces synthetic content; communication studies, since synthetic content is spread in the media environment; sociology, since information shapes societal beliefs and relations; psychology, since cognitive bias affects people's perception of information; political science, since misinformation can affect public opinions and the process of democracy; education, since digital literacy affects an individual's capability of analyzing information; and ethics and law, since issues of responsibility, accountability, privacy, and regulation cannot be avoided.

Background and Conceptual Context

Social Media and the Transformation of Communication

There has been an unprecedented shift in the traditional balance between information producers and audiences due to social media (Benkler et al., 2018). Traditionally, information would move via clear channels that were easily identifiable. This is because newspapers, broadcasting companies, publishers, and journalists were gatekeepers of the information.

Social Media Weakened These Traditional Boundaries

The modern-day consumer is able to create, comment on, distribute, recontextualize, and shape the visibility of content (Benkler et al., 2018). There is significant democratic potential in this model. In addition, it complicates the division among professional journalism, individual opinions, advertisements, political propaganda, satires, and fake news.

The consequence of this dynamic is the creation of a more hybrid informational sphere that contains not only content from professional media sources but also content from influencers, anonymous profiles, political entities, corporations, algorithms, and common users (Wardle & Derakhshan, 2017).

It creates the broader ecosystem of the information disorder described by Wardle and Derakhshan (2017). Misinformation, disinformation, and malinformation differ from each other both in their intention and form, but at the same time can interact in the same communicative space.

Generative AI introduces an additional layer, as the author of the content does not have to be a human anymore.

Generative AI and Synthetic Media

Generative AI is defined as the term for artificial intelligence technologies that can create new content through learned patterns (NIST, 2023; OpenAI, 2023). While large language models can create text, generative technologies can create imagery, sound, videos, and many other forms of media. There are important ramifications for communication in the development of such technologies. Previously, the production of even a single realistic image or professionally written article necessitated a range of technologies and know-how (NIST, 2023; OpenAI, 2023).

The outcome of this is an unprecedented revolution in the economics of content creation:

Low technical barrier → rapid production → large-scale dissemination → increased information volume

Such a technology has many positive uses. Generative AI may be employed in education, accessibility, creative industries, journalism, translation, research, and communication (NIST, 2023; UNESCO, 2023). However, the very same technology may be abused to create fake information. The key issue, thus, is not the fact that AI technologies allow generating false information. False information could be generated by humans even before the advent of AI. What is unique about such a technology is its potential to accelerate information manipulation.

Objectives of the Study

The general objective of this paper will be to study the social aspects of generative AI-produced information in social media, especially misinformation, trust, and the construction of social reality.

Specific objectives are:

1. To study the changes in the production and distribution of information due to generative AI on social media platforms.
2. To study the correlation between AI-generated information and misinformation and disinformation.

3. To study the impact of synthetic information on interpersonal, institutional, and public trust.
4. To study the impact of algorithms of social media platforms in promoting AI-generated information and misleading information.
5. To understand the potential impact of generative AI on the social construction and perception of reality.
6. To study the role of digital and AI literacy in navigating a complex information environment.
7. To study the larger issues of social inequality emerging from the new digital information environment.
8. To give recommendations to individuals, educational institutions, tech firms, governments, media houses, and researchers.

Research Questions

The study addresses the following research questions:

Q1: How is generative AI transforming the process of information generation and dissemination on social media platforms?

Q2: How can artificial intelligence-generated content be instrumental in spreading dis/misinformation?

Q3: How is one's exposure to synthetic media impacting social and organizational trust?

Q4: How do social media algorithms impact the visibility and spread of AI-generated content?

Q5: How is the prevalence of synthetic media shaping the socially constructed reality?

Q6: How can digital/ AI literacy help counteract the negative effects of AI-mediated dis/misinformation?

Methodology

Research Design

The research makes use of a qualitative integrative literature review methodology. An integrative review is an ideal approach for research conducted in a new interdisciplinary research domain since it enables the integration of results and theories from various disciplines to detect wider connections.

The current research is not meant to present any novel survey or experimental results. The results are a description of what has already been researched regarding generative AI, social media, disinformation, trust, digital literacy, algorithmic communication, and social reality.

The methodology chosen was geared towards answering a conceptual and interdisciplinary research question, and not estimating a statistical effect.

Sources of Literature

Relevant academic literature was identified from research addressing:

- generative artificial intelligence;
- large language models;
- social media;
- misinformation and disinformation;
- synthetic media and deepfakes;
- social and institutional trust;
- digital literacy and media literacy;
- algorithmic recommendation;
- political communication;

- digital inequality;
- social construction of reality;
- AI ethics and governance.

Peer-reviewed articles in journals, academic books, important reports from institutions, and policy documents were given preference. Important sociological theories have been considered in order to build a theoretical base for studying the connection between information technology and social reality.

Analytical Approach

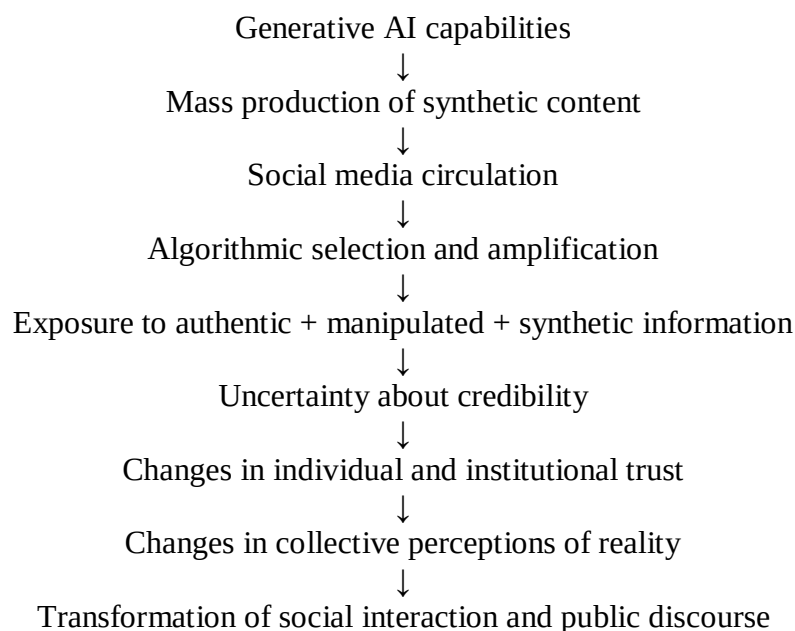
The reviewed literature was analyzed thematically. Five major analytical themes were developed:

1. AI-enabled content production
2. Misinformation and information disorder
3. Algorithmic amplification
4. Trust and epistemic uncertainty
5. Transformation of social reality

A sixth cross-cutting theme, digital inequality and literacy, was incorporated to examine how the consequences of these transformations differ across social groups.

Conceptual Framework

The study proposes that the relationship between generative AI and social trust is not linear. Instead:



This process is influenced by digital literacy, education, socioeconomic status, political polarization, platform governance, media institutions, and regulatory environments.

Findings

The Democratization of Content Creation

One of the most important facts that have come out of the literature review is that the emergence of generative AI has resulted in the democratization of the creation of digital

media (NIST, 2023; OpenAI, 2023). The creation of convincingly fake digital media usually involved highly developed technical abilities. However, the emergence of generative AI has lowered this requirement significantly. This results in two effects at once. First, people can now express themselves better. The technology can be used to provide personalized educational assistance; create multilingual media; conduct research for journalists; provide people with disabilities with assistive technologies; and help artists explore new artistic horizons. Second, the same technological accessibility can result in manipulation because the process of creation is no longer limited by time, skills, and funding.

The Industrialization of Misinformation

Conventional misinformation strategies have been limited by the requirement for human effort. An individual or group needed to draft an article, design an image, edit a video, or manage various social media accounts. However, generative AI could turn this paradigm on its head. Large amounts of content can be produced rapidly and tailored for different populations (NIST, 2023; OECD, 2024). The narrative that misleads could be drafted in various languages and different cultural settings through various formats. This could signal a shift from:

individual misinformation → scalable misinformation

This is a critical implication, where the danger lies not in one amazing deepfake but in many tiny fragments of fabricated content that together shape perceptions.

A single fabricated image may draw attention for just one day. But a steady stream of misleading stories will eventually shape perceptions over weeks or months.

Therefore, the phenomenon of misinformation cannot be considered simply as an individual lie but rather a whole communication system (Wardle & Derakhshan, 2017; Molina et al., 2021).

Deepfakes and the Problem of Visual Evidence

For many years, visual information carried significant evidentiary weight. A video or photograph used to serve as proof of the occurrence of an event. Deepfake technology contradicts such reasoning. Facial expressions, voiceovers, body movements, and even entire scenes can now be modified by deepfake technologies (Chesney & Citron, 2019; NIST, 2023). Realistic fake visual content poses significant problems for traditional techniques of visual verification (Chesney & Citron, 2019; NIST, 2023). As a result, we face what can be referred to as an authenticity deficit.

When an individual faces a questionable video, they may wonder:

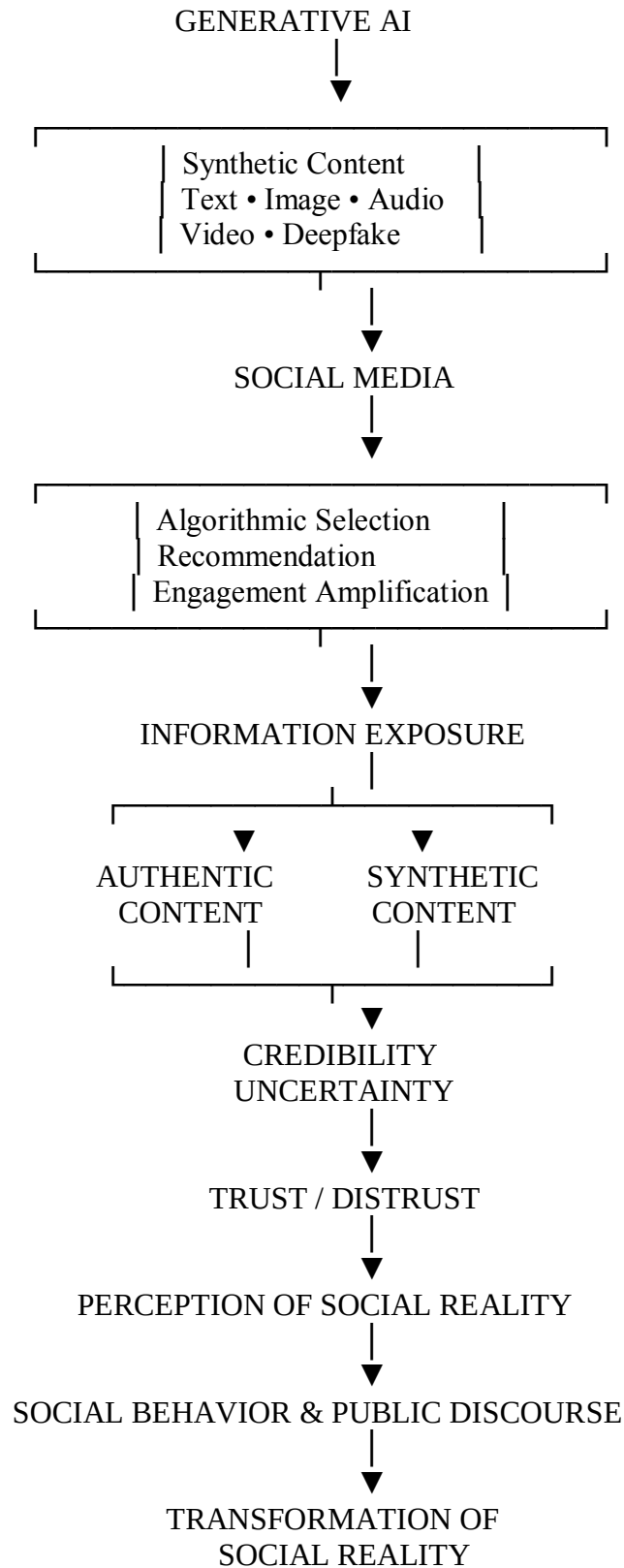
"Is this real?"

But there is another question one needs to pose:

"Can I still trust what I see?"

Such a difference matters sociologically.

The presence of fake visual information not only means that people have a chance to believe something untrue but also to doubt something true.

Figure 1. Proposed Model of AI-Mediated Transformation of Social Reality**Moderating factors:**

- Digital literacy
- Education
- Inequality

- Political polarization
- Platform governance
- Institutional credibility
- Regulation

The Crisis of Epistemic Trust

Trust plays a key role in social coordination. It is impossible for individuals to confirm all of the information they come across independently (Luhmann, 1979). Instead, they depend on institutions, experts, journalists, friends, family members, scientific organizations, and others. The digital information age makes these relations complicated. If one encounters contradictions, fakes, manipulated images, artificial videos, machine-generated articles, and deceptive profiles, doubt grows.

This can be illustrated as:

More information ≠ more knowledge

Actually, too much information can make one confused (McIntyre, 2018; OECD, 2024). The use of generative AI might exacerbate the situation by generating even more material. The person thus faces an epistemic issue. He or she has to identify not just information, but which of it should be believed.

The “Liar’s Dividend” and the Collapse of Evidentiary Confidence

One of the most important implications of this new media form is the chance of real evidence being treated as a fraud. Chesney and Citron (2019) refer to a similar concept as "the liar's dividend." Once people realize that deepfakes and other manipulative technologies can be used for malicious purposes, one can use the excuse that whatever evidence was presented is actually made up. As a result, we have a paradox in which:

False information makes people believe lies, but the knowledge of false information allows people to disbelieve the truth.

This may be even more dangerous for society than deception alone.

For instance, if there is a real video that proves wrongdoing, the person can say that the video is a product of artificial intelligence. Given that audiences will already know that "anything can be faked," the chances of verifying information become much harder.

Algorithmic Amplification

The use of generative AI does not automatically decide what is shown to the users. Social media companies apply recommendation and ranking systems to decide which content will get visibility (UNESCO, 2023). Clicks, shares, comments, watching time, and other forms of engagement may affect how the content is distributed. Studies of misinformation on social media have shown that false information spreads quickly in the network environment (Vosoughi et al (Vosoughi et al., 2018), 2018). This gives a very dangerous feedback mechanism:

GenAI generated content → engagement → algorithmic amplification → exposure → more engagement → further amplification

It does not necessarily mean that algorithms always amplify misinformation. It means only that research shows that content generation technology and content distribution technology cannot be separated.

Table 1. Major Transformations Associated with Generative AI on Social Media

Dimension	Earlier Digital Environment	Generative AI Environment	Social Implication
Content production	Primarily human-generated	Human + AI-generated	Blurred authorship
Production speed	Relatively limited	Extremely rapid	Information abundance
Content personalization	Human effort required	Automated at scale	Targeted persuasion
Visual manipulation	Specialized skills	Increasingly accessible	Authenticity concerns
Information verification	Source/content assessment	Source + provenance + AI detection	Greater complexity
Misinformation	Often labor-intensive	Potentially scalable	Increased volume
Trust	Source-based	Increasingly verification-based	Epistemic uncertainty
Social reality	Mediated by humans/institutions	Mediated by humans, platforms, and AI	New forms of reality construction

The Emergence of “Reality Apathy

Another issue is the potential for repeated exposure to misinformation to create not just skepticism but indifference (Pennycook & Rand, 2021; McIntyre, 2018). When people constantly hear contradictory claims, see manipulated media, and witness contradictory stories, they might eventually conclude:

"Nothing online can be trusted."

Such an approach can be called reality apathy, a situation where people feel demotivated to find out whether the information is true since the cost of verification is too high. Reality apathy differs from healthy skepticism.

Healthy skepticism questions:

"What proof is there for this statement?"

Reality apathy questions:

"Who cares? It could all be lies anyway."

The latter question can harm public debate because democracies rely not only on the right to free speech but also on shared norms regarding evaluation of evidence.

Transformation of Social Reality

The sociological implications of generative AI become more obvious in the context of the theory of the social construction of reality proposed by Berger and Luckmann (1966). Humans perceive reality not just through the prism of perception of what is happening, but also through socially constructed meanings, language, institutions, and interaction. In modern societies, one such socially constructed space is social media (Berger & Luckmann, 1966; Benkler et al., 2018). Let us think about a significant political event.

The person does not necessarily have to see it themselves. They might come across:

- photographs;
- videos;
- news reports;
- comments;
- memes;
- influencer interpretations;
- political statements;
- AI-generated images;
- AI-generated summaries;
- algorithmically recommended posts.

As such, the individual's comprehension of the event is framed by a complicated information ecosystem. In this case, generative AI joins this ecosystem. It means that the AI-created content can become an indirect part of the meaning-making process within society. The key point, therefore, is not so much whether the AI-created content is "real," but rather how synthetic representations shape the beliefs about what actually happened, what is happening now, and what will happen next.

Trust Becomes a Scarce Social Resource

Trust has traditionally enabled people to cope with complexity.

We trust doctors without knowing everything about all their procedures. We trust engineers without solving all the structural equations ourselves. We trust journalists without investigating everything ourselves.

Digital disinformation puts pressure on the trust systems. As institutional credibility wanes, people might be forced to depend on:

- friends;
- influencers;
- online communities;
- ideological groups;
- celebrities;
- anonymous accounts;
- algorithmically recommended content.

This could lead to a fragmented trust environment. Rather than one general information environment, society could have several parallel universes, each having its own network of trust.

Political and Democratic Consequences

The connection between misinformation and democratic societies has been a subject of extensive scholarly research. Democracies rely on citizens' access to reliable information. The information space must not be systematically distorted since, without it, citizens will not be able to make an informed choice. The application of generative AI can affect political communication in several ways.

First, it can lower the costs of creating political content.

Second, it can enable very personalized messages.

Third, synthetic media can mimic political figures.

Fourth, automation can enable massive creation of comments and posts.

Fifth, fake information can potentially influence perceptions about political figures or events. Thus, the problem is not confined to misinformation on election day but extends to the change of political communication as a whole.

Social Polarization and Fragmented Realities

Misinformation impacts not all social groups in the same way. People tend to understand information based on their beliefs, identities, and social networks. Hence, political polarization will amplify the effect of misleading information. Misleading information that supports a person's existing worldview is more likely to be accepted than information that contradicts it.

Thus, we have a feedback loop:

Belief → selective exposure → algorithmic recommendation → supporting information → even stronger belief

Generative AI might further intensify this trend since the information will be generated specifically for certain ideological groups. As a result, people will live in social realities that are more divided, where they not only differ in values but in facts as well.

Digital Inequality in the Age of AI

The impact of AI-driven misinformation on society should also be evaluated with respect to the idea of digital inequality.

Digital inequality cannot be reduced to the mere question of internet availability (van Deursen & van Dijk, 2014). The issue is about discrepancies between:

- digital skills;
- information literacy;
- educational resources;
- technological familiarity;
- access to reliable devices;
- ability to verify information;
- language proficiency;
- socioeconomic resources.

It is evident from van Deursen and van Dijk (2014) that digital skill inequalities can lead to an uneven distribution of results despite equal availability of technology (van Deursen & van Dijk, 2014). This factor is particularly relevant for developing nations.

Those who have more education and technological opportunities will be in a better position to spot dubious materials, verify sources, compare information, and understand AI-generated texts. On the other hand, others may be at a disadvantage. In effect, while democratizing the production of content, AI may perpetuate existing social inequalities.

The Language Dimension of AI Misinformation

There is also the issue of linguistic inequality with generative AI. Historically, many AI models have been shown to perform better on languages that have wider coverage than languages that do not (NIST, 2023). Misinformation detection is one of the factors affected by linguistic inequalities. It might happen that an AI model capable of detecting manipulation in English might not be able to detect manipulation in other languages such as Bangla. It is

important for the country of Bangladesh and others like it to consider this challenge when implementing AI governance.

Discussion

The results show that generative AI is not merely a new tool in the social media world (NIST, 2023; UNESCO, 2023). Rather, it is an evolution in the communication setting.

There are three principal arguments.

From Information Scarcity to Verification Scarcity

In the past, one of the most important problems of communication was access to information. With the digital revolution, access to information has increased tremendously. However, the new problem is very different – it is less the access to information that poses a problem and more the access to credible information (OECD, 2024; UNESCO, 2023). Generative AI increases this trend even further. When there is an opportunity to create millions of pieces of content fast, there is no chance for individuals to verify every piece of information. Verification becomes the scarce resource.

They need to understand:

- who created the content;
- when it was created;
- whether the source is credible;
- whether the image is manipulated;
- whether other reliable sources confirm the claim;
- whether the content has been generated or altered by AI.

From Human-Mediated Communication to Hybrid Communication

Traditional The conventional form of communication had people communicating with people. With digital communication, there is platform-mediated communication.

The new age of communication has generative AI, which entails machine-generated communication.

One may thus interpret the current information landscape as:

Human → Platform → Human

becoming more:

Human → AI → Platform → Human

and sometimes:

AI → Platform → Human

There are several questions raised in this case.

In the case of harmful misinformation through AI, who is liable?

The person who prompted the AI?

- The developer?
- The platform?
- The person who shared the content?
- The algorithm that amplified it?

There is no simple answer. This indicates the need for governance frameworks that distribute responsibility according to different stages of the information lifecycle.

From Trusting Sources to Verifying Provenance

Evaluating information traditionally was done using source authority. There was trust placed in well-known media sources, academic institutions, government organizations, and specialists. In the case of synthetic information, establishing source authority becomes problematic since the source might not be trustworthy. Hence, in the future, the information environment might rely on content provenance – the process of establishing the origin of the information, detecting any changes, and verifying whether artificial intelligence had been used for its generation (NIST, 2023; UNESCO, 2023). Technical methods such as digital watermarking, metadata, provenance standards, and content credentials might help with establishing content provenance (NIST, 2023; UNESCO, 2023). Yet, provenance systems alone will not replace social trust.

Theoretical Interpretation

The social construct can be understood from different sociological perspectives and others.

Social Construction of Reality

Sociological theories of Berger and Luckmann (1966) state that reality depends upon shared meanings and institutions (Berger & Luckmann, 1966). Nowadays, social media provides an effective platform for such activities. Generative AI contributes to this reality by adding synthetic actors and their images to the meaning creation process.

Self-Presentation in the View of Goffman

Another theory of Goffman (1959) about self-presentation can also be used in the context of AI and social media (Goffman, 1959).

Previously, people had the opportunity to create their virtual identities. Generative AI makes it possible for them to develop more sophisticated and high-quality images, voice, biography, and stories that may not match their real identity at all. Thus, the gap between presentation and fabrication becomes blurred.

Bourdieu and Digital Capital

The concept of capital by Bourdieu (1986) also offers an important lens (Bourdieu, 1986). Persons who have more cultural and educational capital can be better at analyzing information. With the advent of AI, new forms of capital related to knowledge of AI technologies, verification, sources, and algorithms could arise. Thus, the changes in technology can lead to new social stratifications.

Public Sphere and Habermas

The theory of the public sphere developed by Habermas (1989) stresses the importance of rational critical debate (Habermas, 1989). The information society flooded with lies and synthetic information can destroy the very conditions that are needed for such debates. If citizens cannot agree on basic facts, any reasonable discussion will become hard. Thus, the problem of misinformation is not only the presence of misinformation but the destruction of the communicative conditions.

A Multidimensional Model of AI-Mediated Trust

According to the literature reviewed above, this paper suggests that trust in an AI-mediated social media environment includes no less than five interrelated components:

Content trust

Does the individual trust the information itself?

Source trust

Does the individual trust the source that provides the information?

Platform trust

Does the individual trust the platform's ability to manage information?

Institutional trust

Does the individual trust institutions that have to verify and regulate the process and communicate it publicly?

Technological trust

Does the individual trust in the proper use of AI technologies? These components relate to each other. An individual may trust the journalist but distrust the platform where the journalist's article is placed. The other individual might distrust traditional media but trust a social media influencer. This means that "trust" must not be regarded as one variable.

Table 2. Dimensions of Trust in the Generative AI Era

Dimension	Central Question	Potential Threat
Content trust	Is the information accurate?	AI fabrication
Source trust	Who produced it?	Fake accounts/impersonation
Platform trust	Why am I seeing it?	Algorithmic opacity
Institutional trust	Who verifies it?	Political/institutional distrust
Technological trust	Can AI be used responsibly?	Unexplained AI systems

Recommendations

It is not possible to solve the issue of the evolution of social media via generative AI through any one measure.

Strengthen AI and Media Literacy

Educational institutions should introduce AI literacy as an extension of media literacy (UNESCO, 2023; NIST, 2023).

Students should learn:

- how generative AI works;
- how synthetic content is produced;
- how misinformation spreads;
- how to identify unreliable sources;
- how to verify images and videos;
- how to cross-check information;
- how algorithms influence exposure.

AI literacy should begin at school and continue through higher education.

Introduce Verification Education

Users should be encouraged to adopt a simple verification process:

STOP → CHECK → CROSS-VERIFY → SHARE

Before sharing potentially controversial information, users should ask:

1. Who created it?
2. What is the original source?
3. Is the date correct?
4. Do reliable sources report the same event?
5. Is there evidence of AI manipulation?
6. Is the headline consistent with the actual content?

Improve Platform Transparency

Social media companies should provide greater transparency regarding:

- recommendation systems;
- AI-generated content;
- political advertising;
- automated accounts;
- content moderation;
- synthetic media;
- algorithmic amplification.

Users should have meaningful information about why particular content is recommended to them.

Expand Content Provenance Mechanisms

Technology firms need to develop tools that can assist in establishing the origin and the modification of digital information (NIST, 2023). The use of provenance technology may help in establishing whether there has been any modification in the information. Nonetheless, the use of provenance does not negate human verification.

Develop Stronger Policies against Malicious Synthetic Media

Governments should develop proportionate legal frameworks addressing malicious uses of synthetic media, particularly:

- non-consensual intimate deepfakes;
- identity impersonation;
- financial fraud;
- election manipulation;
- targeted harassment;
- fabricated emergency information;
- deliberate political deception.

Regulation should distinguish harmful deception from legitimate creative, artistic, educational, or satirical uses.

Support Independent Fact-Checking

Independent fact-checking entities should get more institutional backing (OECD, 2024; UNESCO, 2023).

Fact-checking should be made multilingual and more culturally aware as well.

Where information is being spread in native languages, the verification process should run in these languages as well.

Strengthen Journalistic AI Standards

News organizations should establish clear guidelines regarding:

- AI-generated text;
- AI-generated images;
- synthetic voices;
- automated translation;
- AI-assisted reporting;
- disclosure of AI use.

Transparency can help preserve public trust.

Encourage Responsible AI Design

AI developers should embed safety systems into the development process, instead of considering misinformation an afterthought (NIST, 2023).

A responsible AI system should take into consideration:

Accuracy + Transparency + Accountability + Fairness + Privacy + Human Oversight

Address Digital Inequality

AI literacy initiatives should prioritize socially vulnerable populations.

Programs should target:

- rural communities;
- older adults;
- low-income populations;
- individuals with limited formal education;
- linguistically underserved communities;
- first-time internet users.

Otherwise, AI literacy itself may become a new form of social inequality.

Establish Multidisciplinary AI Governance

No single discipline can adequately address the social consequences of generative AI (NIST, 2023; UNESCO, 2023).

Governance frameworks should involve:

- sociologists;
- computer scientists;
- psychologists;
- journalists;
- educators;
- legal scholars;
- policymakers;
- ethicists;
- civil society organizations.

This interdisciplinary approach is essential because technological risks are simultaneously social, political, psychological, economic, and ethical.

Implications for Bangladesh and Other Developing Countries

The above discussion has special significance for developing countries like Bangladesh. Bangladesh has seen a rapid growth in internet and social media usage (UNESCO, 2023). Social media sites, including Facebook, YouTube, TikTok, among others, have become major forums for the purpose of news intake, politics, entertainment, shopping, and personal interactions.

Accessibility to generative AI comes with its pros and cons.

AI can contribute to:

- education;
- translation;
- entrepreneurship;
- public services;
- research;
- creative industries;
- accessibility;
- digital employment.

However, misinformation generated or modified through AI may also affect:

- elections;
- communal relations;
- public health communication;
- financial decision-making;
- reputation;
- gender-based harassment;
- political polarization.

The linguistic dimension is particularly important.

AI technology in Bangla and disinformation detectors must be developed further. A governance model tailored mainly to the English language might not fit the information landscape of Bangladesh well.

Finally, digital literacy initiatives have to go beyond mere computer literacy (van Deursen & van Dijk, 2014; UNESCO, 2023). The future calls for AI literacy. People will have to learn not just how to use technology but also how it affects what they see and believe.

Future Research Directions

There are various avenues that need further empirical study.

Firstly, longitudinal studies can be conducted to assess whether long-term exposure to AI-generated misinformation leads to any change in trust.

Secondly, there can be experimental studies carried out to assess the reaction of people to the same information when generated by humans and AI, respectively.

Thirdly, cross-cultural studies can be undertaken to assess whether there are any differences in the level of trust towards AI-generated information in different cultures.

Fourthly, there must be further study regarding the impact of AI misinformation in low-resource languages like Bangla.

Fifthly, there needs to be further study to find out if AI literacy makes people less vulnerable to AI-generated misinformation.

Lastly, there should be interdisciplinary studies conducted on the long-term effects of pervasive uncertainty about digital evidence.

Limitations of the Study

There are several limitations associated with this research.

Firstly, it is an example of an integrative review and thus lacks empirical information. This implies that the article cannot demonstrate a causal relationship between exposure to AI-generated content and trust decline.

Secondly, the fast pace at which generative AI technology evolves implies that there is more technological capability than academic research.

Thirdly, most of the current literature revolves around the English-speaking and Western population. This implies that the results are less likely to reflect the experiences of people from developing nations.

Finally, the article focuses on the social consequences of generative AI technology and lacks a detailed technical assessment of AI detection systems.

Future empirical research should consider these limitations.

Conclusion

Generative AI is changing the social media platform from an environment where most content is created by humans into a hybrid ecosystem where both human- and machine-created information exist side-by-side (NIST, 2023; UNESCO, 2023). These changes have massive potential. Generative AI can make creativity more democratic, improve accessibility, enhance learning, enable communication, and open up new economic and cultural possibilities. At the same time, these positive possibilities come along with major challenges. It is not just that AI can spread misinformation (OECD, 2024; NIST, 2023). Humans have always been doing so. What makes the current change unique is the unprecedented speed, scale, realism, personalization, and accessibility of synthetic information production and dissemination (Chesney & Citron, 2019; NIST, 2023). In combination with social media algorithms, misinformation becomes more scalable and harder to detect. Even more importantly, the development of synthetic media introduces an even more profound epistemic challenge. People will increasingly find themselves unable to answer the simple question:

“Can I trust what I see?”

The implications go far beyond mere individual deception. Uncertainty surrounding the authenticity of digital content may undermine interpersonal relationships, institutions' legitimacy, political communication, journalism, democratic discourse, and the joint process of constructing social reality (Chesney & Citron, 2019; UNESCO, 2023; OECD, 2024).

One needs to be careful when talking about the “crisis of trust.” It does not mean that society will no longer trust anything, but become fragmented in its approach to trust, personalizing it as well as depending on closed communities and identity networks.

In other words, the development of generative artificial intelligence can lead to an interesting paradox whereby the more society realizes that digital content can be fake, the easier it will become to dismiss real evidence as being made up (Chesney & Citron, 2019). It is clear then that the ability to build shared factual grounds becomes crucial in such conditions.

The solution is neither fear nor rejection of the technology. Instead, society requires a new social contract for digital information that would be founded on transparency, accountability, verification, literacy, provenance, institutional responsibility, and ethical technology design. Educational institutions need to train people not only to be consumers of information but also to question that information. Technology companies need to be more responsible about the context in which synthetic content is circulating. Governments require proportional regulatory systems. Journalists need to work on transparency and verification. Researchers need to study the social implications of AI in different languages and cultural contexts. In the end, the future of social reality does not depend only on artificial intelligence. It depends on how societies decide collectively which technologies to trust, which institutions to hold responsible, what information to verify, and which kinds of knowledge should constitute the basis of their existence. The problem posed by generative AI is no longer “Who created this information?” but:

“Who—or what—created it, how was it transformed, why was it shown to me, and why should I believe it?”

The ability of individuals and institutions to answer these questions will become one of the defining challenges of the digital society.

References

1. Benkler, Y., Faris, R., & Roberts, H. (2018). *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*. Oxford University Press. <https://doi.org/10.1093/oso/9780190923624.001.0001>
2. Berger, P. L., & Luckmann, T. (1966). *The social construction of reality: A treatise in the sociology of knowledge*. Anchor Books.
3. Bourdieu, P. (1986). The forms of capital. In J. G. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241–258). Greenwood.
4. Chesney, R., & Citron, D. K. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review*, 107(6), 1753–1820. <https://doi.org/10.15779/Z38RV0D15J>
5. European Commission. (2022). *The 2022 Code of Practice on Disinformation*. European Commission.
6. Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2024). Generative AI. **Business & Information Systems Engineering*, 66*, 111–126. <https://doi.org/10.1007/s12599-023-00834-7>
7. Giddens, A. (1990). *The consequences of modernity*. Stanford University Press.
8. Goffman, E. (1959). *The presentation of self in everyday life*. Doubleday.
9. Habermas, J. (1989). *The structural transformation of the public sphere: An inquiry into a category of bourgeois society* (T. Burger, Trans.). MIT Press. (Original work published 1962)

10. Luhmann, N. (1979). *Trust and power*. Wiley.
11. Kaplan, A. M., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. (*Business Horizons*, 62*(1), 15–25. <https://doi.org/10.1016/j.bushor.2018.08.004>
12. McIntyre, L. (2018). *Post-truth*. MIT Press.
13. Molina, M. D., Sundar, S. S., Le, T., & Lee, D. (2021). “Fake news” is not simply false information: A concept explication and taxonomy of online content. *American Behavioral Scientist*, 65(2), 180–212. <https://doi.org/10.1177/0002764219878224>
14. National Institute of Standards and Technology. (2023). *Artificial intelligence risk management framework (AI RMF 1.0)*. U.S. Department of Commerce. <https://doi.org/10.6028/NIST.AI.100-1>
15. OECD. (2024). *Facts not fakes: Tackling disinformation, strengthening information integrity*. OECD Publishing. <https://doi.org/10.1787/d909ff7a-en>
16. Open AI. (2023). GPT-4 technical report. *arXiv*. <https://doi.org/10.48550/arXiv.2303.08774>
17. Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences*, 25(5), 388–402. <https://doi.org/10.1016/j.tics.2021.02.007>
18. UNESCO. (2023). *Guidelines for the governance of digital platforms: Safeguarding freedom of expression and access to information through a multistakeholder approach*. UNESCO.
19. Van Deursen, A. J. A. M., & van Dijk, J. A. G. M. (2014). The digital divide shifts to differences in usage. *New Media & Society*, 16(3), 507–526. <https://doi.org/10.1177/1461444813487959>
20. Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>
21. Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policymaking*. Council of Europe.
22. Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Public Affairs.