Int. J. Media. Mas. Com. Vol. 06, No. 02, 173-193 (2024)
DOI: 10.46988/IJMMC.06.02.2024.07
ISSN: 2536-0035
ISBN: 2535-9797
Arab Impact Factor (1.04) (2021).

International Journal of Media and Mass Communication المجلة الدولية للإعلام والاتصال الجماهيري

Artificial Intelligence and Trust: Analyzing the Impact of AI-Generated Fake Content on Social Media Platforms

Dr. Wassim Alsahafi [1], Dr. Mohamed Allaymoun[2], Asmaa Dakheel Alla Alsalmi, Rawan Khaled Aljamaee, Manar Abdulrahman Alharbi, Noora Salim Alahmadi, Hadeel Thabet Alawi [3]

[1] Assistant Professor and Head of the Department of Communication and Media Technology, University of Jeddah

[2] Assistant Professor at the Department of Administrative Science, Gulf University, Bahrain

[3] Graduates of the Digital Media Bachelor's program, Department of Communication and Media Technology, at the University of Jeddah.

Email: waalsahafi@uj.edu.sa, dr.mohammed.allaymoun@gulfuniversity.edu.bh

Received: 25 March. 2024 Revised: 26 April. 2024 Accepted: 25 May. 2024 Published: 01 July 2024

Abstract:

This study discusses the impact of artificial intelligence (AI) technologies on social media platforms through the creation and distribution of fake content and how this affects user trust. One of the key technologies highlighted in the study is “Deepfake,” which uses AI to replace faces or voices in images or videos with fake ones, making it difficult to distinguish between real and fabricated content. This technology poses a significant threat to the credibility of information shared on digital media. Fake content relies on complex algorithms, trained to mimic human facial expressions, voices, and movements with extreme accuracy, causing users to question the reliability of the information they see online. The study explores how these technologies influence user trust in the content they encounter on social media platforms. With the rapid spread of false information and its increasing use, these platforms have become fertile ground for the proliferation of fake and misleading news. The research notes that users vary in their ability to detect this fake content, with some becoming more skeptical and cautious, while others are quick to believe information without verifying its authenticity. Furthermore, the study examines the social and psychological aspects of AI’s impact on user behavior and how these technologies can erode trust in digital information sources, contributing to social divisions and general misinformation. The study suggests technical and managerial solutions to counter this phenomenon, including enhancing the use of AI technologies to detect, rather than solely produce, fake content. It also emphasizes the importance of establishing ethical standards

for the use of AI in media production, as well as increasing awareness of these technologies and how to identify fake content. The significance of this study lies in its in-depth understanding of AI's impact on digital society and identifying mechanisms to combat the spread of fake content. It also provides a set of recommendations for media institutions and communities to enhance the credibility of information and improve public awareness of the dangers posed by these technologies. Based on the study's findings, there is an urgent need to develop new tools and technologies to detect fake content and improve user trust in social media platforms. The impact of these technologies is not limited to individuals but extends to influence society by distorting public opinion and manipulating information. The study highlights the need for coordinated efforts between technological, media, and legal sectors to mitigate the negative effects of this growing phenomenon.

Keywords: Artificial Intelligence; Fake Content; Social Media Platforms; User Trust; Deep Fake.

1 Introduction

Digital media has transformed the creation, sharing, and consumption of information in the digital landscape. Its integration with social media platforms has streamlined content dissemination, creation, and delivery. Social media platforms play a pivotal role in influencing and shaping public opinion on a range of issues. However, the instantaneous nature of social media raises concerns about the accuracy of information, the potential for misinformation, and the reliability of user feedback. Understanding the influence of digital media is essential for navigating the complexities of the digital environment (Dolega, Rowe, & Branagan, 2021).

The incorporation of artificial intelligence (AI) in social media has significantly enhanced platform efficiency, benefiting startups and communities by creating spaces for interaction and service exchange. AI serves as a foundational component of social media, enabling the delivery of more efficient and effective services. Furthermore, AI contributes to safeguarding user and organizational data privacy through advanced monitoring and machine learning techniques (Al-Ghamdi, 2021).

The communications and internet revolution has further amplified the application of AI across various domains, including government, social, and economic services. AI has been leveraged for data analysis, decision-making, automation, education, training, enhancing user experiences, and even prediction and strategic planning. In government services, for instance, AI has the potential to drive development, especially in low-income nations, by improving efficiency and increasing initial salaries (Alhashmi, 2024).

However, AI's advancements have also introduced challenges, such as the rise of deep-fake technology. This involves the use of AI to merge images, videos, and audio to create highly realistic but fabricated content. These technologies can mimic facial expressions, voices, and behaviors with a level of realism that makes distinguishing between fabricated and genuine content increasingly difficult. Deepfake technology allows the alteration of faces or voices to depict individuals saying or doing things that never occurred (Al-Babli, 2023).

This paper explores the use of artificial intelligence in generating fake content on social media platforms and examines its impact on users' trust. By delving into the ethical, technical, and social dimensions of this phenomenon, the study aims to provide insights into mitigating the negative implications of AI-generated misinformation.

2 Using artificial intelligence in SNS

Artificial intelligence (AI) is widely utilized across various fields, with a significant presence in the realm of media, particularly on social networking platforms. AI technologies are leveraged in numerous ways, including analyzing user behavior, processing platform data, and identifying emerging trends by employing predictive tools. Prominent social networking companies such as Facebook, X (formerly Twitter), and LinkedIn have integrated AI applications to enhance their operations. AI technologies on these platforms are implemented through various mechanisms, such as chatbots, blockchain technology, big data analysis, and image recognition tools. While these AI-driven technologies offer benefits like improved efficiency, automation, round-the-clock availability, and personalized user experiences, they also raise concerns about reduced human interaction, unequal access to technology, privacy and security risks, and potential job displacement (Kemp, 2024; Rabie & Abdel Fattah, 2024).

Deepfake Technology Using Artificial Intelligence

Deepfake technology employs AI to create deceptive content by altering visual and auditory elements. This process involves three primary stages: alignment, training, and merging, using sophisticated algorithms. For instance, deepfakes can replace faces or voices in videos to depict individuals performing actions or speaking words that never occurred, or they can generate entirely artificial faces and bodies.

Alignment:

This stage focuses on replacing the subject's face in a video using targeted images. It involves precisely identifying the face’s location and dimensions, often utilizing techniques like the Histogram of Oriented Gradients (HOG), which analyzes each pixel to determine its orientation (Al-Babli, 2023).

Training:

In this stage, facial image extracts are used to train an autoencoder. The encoder compresses facial images into reduced data, while the decoder reconstructs the original image from this compressed information (Mahjoub, 2023).

Merging:

The final stage modifies the original images using the extracted data, resulting in a manipulated video. This process generates high-quality fake videos that are difficult to differentiate from authentic ones (Lujain & Abdel Nasser, 2023).

Deepfake techniques span various forms, including face swaps in videos, voice manipulations, still image falsifications, and 3D deepfakes. They also integrate audio with video, producing highly convincing fake content (Abdelsalam, 2023; Al-Obaidi, 2021).

Detecting and Responding to Fake Content

Despite their sophistication, deepfakes can be detected through advanced techniques. These include analyzing videos and images for visual inconsistencies, assessing audio for subtle tonal variations, applying biological markers, and using temporal and spatial analysis. Additional tools, such as generative network techniques, further enhance detection capabilities (Al-Bayoumi, 2023; Passos, 2024).

User Behavior Towards Deepfake Content

User responses to deepfake content vary based on awareness and prior experience. Some users quickly believe high-quality fake content due to its realistic production, making it difficult to discern its authenticity. Conversely, a growing segment of users has become more cautious, particularly in politically or media-sensitive contexts. These users often rely on trusted sources to verify the credibility of content presented to them (Kaate et al., 2023).

This section underscores the dual role of AI in social media platforms, both as a tool for enhancement and a potential source of manipulation. The varying user responses highlight the need for greater awareness and reliable detection methods to mitigate the impact of AI-generated fake content on trust.

3 Literature Review

3.1 How Artificial Intelligence Facilitates Deepfake Technology

The study by Hasan Abir et al. (2022) examined the effectiveness of deep learning techniques in detecting fake images and evaluated the accuracy of interpretable artificial intelligence methods in this domain. Using a deep learning approach and quantitative analysis, the study applied Convolutional Neural Networks (CNNs) alongside interpretation tools such as SHAP and LIME. The study analyzed a dataset of fake and real images collected from multiple sources. The results demonstrated that deep learning techniques effectively identified fake images, while interpretable AI methods enhanced transparency by clarifying the model’s decision-making processes. This, in turn, increased users' trust in automated systems for detecting image manipulation.

Hanif and Dave (2022) assessed the risks and benefits associated with deep-fake technology using a descriptive analytical approach, with data collected via questionnaires and interviews. Their study targeted media, technology, and social specialists, comprising a sample of 100 participants. Results highlighted the growing necessity for legislation to protect individuals from deepfake risks and the importance of educating the public on identifying manipulated content.

Al-Suwaileh (2023) investigated the role of the internet and cyberspace in propagating psychological warfare and rumors, relying on a descriptive analytical approach. By analyzing the spread of electronic rumors and their effects through social media, the study revealed that rumors and misinformation have become tools of fifth-generation warfare, harming societies through digital platforms.

Mahjoub (2023) explored deepfake methods and the efforts of Sudanese media

professionals to combat fake content. Using a descriptive survey approach, the study focused on media institutions in Khartoum. Results showed that Sudanese media organizations use digital tools to counter fake content but are hindered by the lack of strict laws and media professionals' limited responses to such content.

Amr (2024) employed a descriptive approach to investigate how artificial intelligence facilitates media image manipulation through deep-fake technologies. The study revealed that the misuse of AI weakens the credibility of media imagery, promotes the spread of rumors, and normalizes the denial of authentic events.

3.2 Fake Content on Social Media Platforms and Its Impact on Users' Trust

The study by Femi Olan et al. (2022) analyzed how fake content influence's public opinion and social behaviors. Relying on a descriptive analytical approach, the study involved 200 social media users. Results highlighted that fake content exacerbates social divisions, undermines trust in media sources, and negatively impacts decision-making processes.

Similarly, the study by Al-Naghi and Al-Naghi (2022) explored social media users' awareness of deepfake threats and their safe usage of such platforms. Using a media survey approach and digital questionnaires, the study sampled 600 users. Findings revealed a direct relationship between awareness of deepfake threats and safe social media usage, while confidence in detecting deepfakes inversely correlated with users' safety practices. Additionally, perceptions of threat varied based on demographic factors such as geography, gender, and education level.

Hatamleh et al. (2023) investigated how trust in social media influences the quality of social relationships. Using quantitative analysis and questionnaires, the study surveyed 300 users in Riyadh. Results demonstrated that trust in social platforms fosters stronger social ties and enhances positive interactions among users.

Ahmed Shehata et al. (2023) examined how misinformation on social media during the COVID-19 pandemic affected the Omani public's behavior and the government's response. The study used a quantitative survey approach with data collected from 15 WhatsApp users across Oman. Findings revealed that misinformation heightened anxiety and impaired users' decision-making, demonstrating their exposure to false information while browsing social media.

Tung and Lan (2024) analyzed university students' awareness of content authenticity and trust in social media, focusing on TikTok users. Using a quantitative approach, the study surveyed 300 students. Results indicated low trust in TikTok content but growing awareness of the importance of verifying information sources.

Commentary on Studies

The reviewed studies provide valuable insights into deep-fake technology and its implications for media, social platforms, and user behavior. Combining qualitative and quantitative approaches, the research addresses technical aspects of detecting fake content and examines the social and psychological risks posed by its proliferation.

A recurring theme across studies is the need to address the multifaceted challenges posed by deep-fake technology. These challenges extend beyond technical detection to encompass

social and psychological dimensions, such as eroding trust in digital media and influencing societal behaviors. The studies emphasize the urgency of developing robust fake content detection methods, with a focus on transparency and understanding algorithmic processes.

Additionally, the research highlights the importance of cross-sector collaboration between technology, media, and legal institutions to combat the spread of fake content. Efforts should also focus on raising public awareness about identifying and distinguishing real from manipulated information.

The studies employed diverse methodologies and theoretical frameworks, such as deep learning for AI-based detection, fifth-generation warfare theory for understanding rumors, and media framing theory for addressing fake content. The use of both quantitative and qualitative tools, including surveys and interviews, enriched the findings by evaluating both technical capabilities and social impacts. Collectively, these studies underscore the critical need for comprehensive strategies to address the technical, legal, and societal challenges posed by fake content on social media platforms.

4 Training problems

Information Overload Theory

The information overload theory describes a state where individuals are unable to process the vast amount of information presented to them due to its rapid and continuous flow. This concept originated in the late 20th century when researchers began examining the effects of information abundance on decision-making clarity (Gare, 2024). The theory gained prominence with the advent of modern information and communication technologies, particularly the rise of the Internet and social media. These platforms inundate individuals with more information than they can comprehend or retain, leading to distraction and a decline in decision-making quality. Alvin Toffler, a writer and thinker, was the first to introduce this concept in his book Future Shock (Roetzel, 2019).

Key Theoretical Assumptions:

- The sheer volume of available information surpasses individuals' cognitive capacity, resulting in mental fatigue.
- Overload information diminishes the ability to make decisions, understand, and recall information effectively.
- Personal factors, such as an individual’s cognitive abilities and capacity to verify information influence the extent of susceptibility to information overload (Mustafa, 2024).

The information overload theory emphasizes how excessive information impacts individuals’ capacity to make informed decisions. It suggests that when the influx of information exceeds an individual's ability to process it, distraction and confusion occur. This aligns with findings that the vast amount of information on social media platforms—often including fake content—makes it challenging to differentiate between accurate and misleading information.

Deep-fake technology exacerbates this issue by eroding users’ trust in content. The rapid flow of information often compels individuals to make hasty decisions without sufficient time to verify their accuracy. The theory and supporting research demonstrate that the high

volume and velocity of information not only facilitate the spread of fake content but also encourage quick, often uninformed decisions, reducing the overall quality of information processing and trust in digital platforms.

Theoretical Framework

Study Problem

Social media platforms have become a fertile ground for leveraging artificial intelligence (AI) technologies to create and disseminate misleading information. Techniques such as deep-fake videos, altered texts, and manipulated images are increasingly used to produce fake content, making it difficult for users to discern genuine and fabricated information. As the sophistication of these falsehoods grows, user confidence in these platforms has eroded due to rising uncertainty. This study addresses the impact of AI technologies on user behavior and the factors contributing to the proliferation of misinformation. Furthermore, it aims to fill the research gap by exploring how these technologies are used to generate and distribute fake content and their subsequent effects on users' trust. The central problem of the study is to "identify the methods through which artificial intelligence technologies are employed to present fake content on social media platforms and their influence on users' confidence."

Study Importance

1. Expanding the scientific understanding of AI technologies and their societal effects in the digital era.

2. Highlighting the mechanisms behind the production and dissemination of fake content.

3. Examining the influence of fake content on collective perceptions and individual trust in information sources.

4. Proposing recommendations and solutions to combat fake content on social media and improve detection mechanisms.

5. Developing public policies to regulate AI in the digital sphere and guiding innovations toward enhancing information transparency.

Study Objectives

1. To identify how AI is used to produce content on social media platforms.

2. To examine the impact of deepfake technology on user behavior on these platforms.

3. To analyze the relationship between advancements in AI technologies and the rising quality of deepfake content.

4. To investigate the correlation between exposure to fake content and trust levels among users.

Study Questions

1. How is artificial intelligence utilized in producing content on social media platforms?

2. What is the impact of deepfake technology on user behavior on these platforms?

Study Hypotheses

1. There is a relationship between users' exposure to fake content and their trust in the information provided by social media platforms.

2. The quality of content produced through AI technologies affects user awareness of fake content.

3. Verifying information sources before adoption or sharing increases trust in the content.

Study Terms

- **Media Misinformation:**

A deliberate effort to mislead or falsify facts with the goal of influencing public opinion and directing it in specific ways. This is achieved through propaganda techniques, psychological strategies, and modern technologies to distort reality and spread misinformation. Social media platforms are often used to amplify this content, enabling its rapid and wide dissemination, which enhances its impact on the public (Abdulaziz, 2022).

- **Deep-fake:**

A process that uses artificial intelligence and machine learning to replace faces in images or videos. AI algorithms are trained on datasets from various sources to reconstruct new faces and create manipulated video or audio clips. These techniques are also used to produce realistic fake audio (Muharram, 2022).

- **Artificial Intelligence (AI):**

The development of technology that can think and act in a human-like manner, going beyond merely simulating specific behaviors. AI relies on a knowledge base composed of algorithms and codes that enable computers to perform tasks autonomously and effectively (Azibi, 2024).

This theoretical framework establishes the foundational understanding of AI's role in producing and distributing fake content on social media platforms, as well as its implications for users' trust and behavior. It underscores the need to explore the interplay between AI-driven misinformation and public trust to develop effective mitigation strategies.

Study Type and Methodology

This study falls under the category of descriptive and exploratory research, which extends beyond mere description to delve deeper into the phenomenon (Tawaher & Othman, 2021). The research adopts the survey methodology, a structured scientific approach used by researchers to collect data on the study subject. Surveys can encompass either the entire target population or a large representative sample (Lamya, 2021).

Study Tool

The study utilized an electronic questionnaire as the primary tool for collecting data and information from the sample population. This method is one of the most used tools in survey research for gathering necessary data related to the study problem (Raqam,

BouSaha, & Haroush, 2022). The questionnaire was designed using Google Forms and consisted of a set of questions and statements organized into thematic sections. The final version of the questionnaire included 19 questions.

Study Population and Sample

The study population comprised users of social media platforms in Saudi Arabia. Data was collected by distributing the electronic questionnaire between September 1, 2024, and November 1, 2024. To ensure the validity of responses, the questionnaire's settings were adjusted appropriately. A simple random sampling method was employed, resulting in a sample size of 210 participants who completed the questionnaire.

Validity and Reliability

The questionnaire form was reviewed by a panel of media experts (*) to provide feedback and suggestions regarding the questions and their structure. The researchers incorporated the suggested revisions from these evaluators.

Table 1: Cronbach's Alpha Reliability Coefficient

Number of Items or Questions	Cronbach's Alpha Coefficient for Reliability
19	0.802

The results in the table above indicate that Cronbach’s Alpha coefficient for the questionnaire items is 0.802, which reflects a high level of reliability.

* Dr. Hassan Faraj Hassan, Assistant Professor, Department of Communication and Media Technology, University of Jeddah.

Prof. Abdullah Al-Wazzan, Associate Professor, Department of Communication and Media Technology, University of Jeddah.

Dr. Majed Diab Al-Zubair, Associate Professor, Department of Communication and Media Technology, University of Jeddah.

Statistical Methods Used in the Study

The study relied on the SPSS software (version 23) for data processing and analysis, employing the following statistical methods:

1. Percentages.

2. Frequencies.

3. Pearson's correlation coefficient to measure the strength of associations for descriptive data that can be presented ordinally.

This theoretical framework outlines the study's robust methodology, which leverages validated tools and statistical techniques to investigate the use of artificial intelligence technologies in generating fake content on social media platforms and their impact on user trust.

5 Search Results

The study presents its findings by clarifying the correlations between the independent and

dependent variables, as outlined in the study's hypotheses.

Section 1: Basic Data

Table 2: Age

Age Group	**Frequency (n)**	**Percentage (%)**
18–24	129	60.6
25–34	36	16.9
35–45	48	22.5
Total	**213**	**100**

Table 2 indicates the age distribution of the study sample. It shows that 60.6% of participants are aged 18–24, 22.5% are aged 35–45, and 16.9% are aged 25–34. This suggests that the majority of the sample consists of young social media users who are likely more aware of the content published on these platforms.

Table 3: Gender

Gender	**Frequency (n)**	**Percentage (%)**
Male	30	14.1
Female	183	85.9
Total	**213**	**100**

Table 3 shows the gender distribution of the study sample, with 85.9% of the participants being female and 14.1% male.

Table 4: Educational Level

Educational Level	**Frequency (n)**	**Percentage (%)**
Intermediate	8	3.8
Secondary	46	21.6
University	159	74.6
Total	**213**	**100**

Table 4 indicates that 74.6% of the participants have a university-level education, 21.6% have a secondary-level education, and 3.8% have an intermediate-level education. This highlights that the majority of participants have a higher educational level, particularly university students, who are active followers of digital media.

Table 5: Most Used Social Media Platforms

Platform	**Frequency (n)**	**Percentage (%)**
Instagram	25	11.7
X (Twitter)	28	13.1

Platform	Frequency (n)	Percentage (%)
Snapchat	67	31.5
TikTok	93	43.7
Total	**213**	**100**

Table 5 reveals that TikTok is the most used platform (43.7%), followed by Snapchat (31.5%), X (13.1%), and Instagram (11.7%).

Section 2: Trust in Content

Table 6: Verification of Information Sources Before Adoption or Sharing

Verification Frequency	Frequency (n)	Percentage (%)
Always	98	46.0
Sometimes	102	47.9
Rarely	13	6.1
Total	**213**	**100**

Table 6 shows that 47.9% of users sometimes verify information sources before adopting or sharing, 46% always verify, and 6.1% rarely verify sources.

Table 7: Trust in Information on Social Media

Trust Level	Frequency (n)	Percentage (%)
1 - Completely Untrustworthy	20	9.4
2 - Largely Untrustworthy	39	18.3
3 - Somewhat Trustworthy	123	57.7
4 - Largely Trustworthy	25	11.7
5 - Completely Trustworthy	6	2.8
Total	**213**	**100**

Table 7 reveals that 57.7% of users find information on social media somewhat trustworthy, while 18.3% find it largely untrustworthy. Only 2.8% of participants find social media completely trustworthy.

Table 8: Impact of Design and Appearance on Trust

Impact Level	Frequency (n)	Percentage (%)
1 - No Impact at All	34	16.0
2 - Slight Impact	22	10.3
3 - Moderate Impact	43	20.2
4 - Significant Impact	46	21.6

Impact Level	**Frequency (n)**	**Percentage (%)**
5 - Very Significant Impact	68	31.9
Total	**213**	**100**

Table 8 shows that 31.9% of users believe the design and appearance of content have a very significant impact on trust, while 21.6% reported a significant impact and 20.2% indicated a moderate impact.

Table 9: Most Trusted Sources on Social Media

Source	**Frequency (n)**	**Percentage (%)**
Friends	84	22.9
Influencers	23	6.3
Official Pages	162	44.1
Traditional Media	98	26.7
Total	**367**	**100**

Table 9 reveals that official pages are the most trusted source on social media (44.1%), followed by traditional media (26.7%), friends (22.9%), and influencers (6.3%).

These findings provide a comprehensive understanding of user demographics, behaviors, and levels of trust in content on social media platforms, highlighting the role of design, source reliability, and user verification habits in shaping trust and interaction.

Table 10: User Perceptions on Mentions Increasing Trust in Content

Response	**Frequency (n)**	**Percentage (%)**
Yes	72	33.8
No	61	28.6
Neutral	80	37.6
Total	**213**	**100**

Table 10 reveals that 37.6% of users remained neutral about mentions increasing trust in content, while 33.8% believed mentions enhance trust, and 28.6% disagreed.

Table 11: Methods Users Use to Identify Fake Content

Method	**Frequency (n)**	**Percentage (%)**
Using reliable sources	158	53.0
Reading comments and opinions	138	46.3
Others	2	0.7
Total	**298**	**100**

Method	Frequency (n)	Percentage (%)
(*n = 213*)		

Table 11 indicates that 53% of users identify fake content by using reliable sources, while 46.3% rely on comments and opinions, and 0.7% use other methods such as analyzing links or activated comments.

Table 12: Impact of Social Media Opinions and Discussions on Awareness of Fake Content

Response	Frequency (n)	Percentage (%)
Yes	120	56.3
No	48	22.5
Neutral	45	21.1
Total	**213**	**100**

Table 12 shows that 56.3% of users stated that social media opinions and discussions significantly influence their awareness of fake content, while 22.5% believed they do not, and 21.1% remained neutral.

Table 13: Previous Exposure to Fake Content

Response	Frequency (n)	Percentage (%)
Yes	88	41.3
No	125	58.7
Total	**213**	**100**

Table 13 reveals that 58.7% of users had not encountered fake content, while 41.3% had prior exposure.

User Reactions to Identifying Fake Content:

- Surprise at the realistic nature of fake content.
- Warning acquaintances and blocking the content if possible.
- Disbelief that the source appeared highly credible.
- Reporting the content.
- Feelings of disappointment for believing the fake content.
- Learning a lesson to only trust content from verified sources.
-

Table 14: Preference for Educational Courses on Identifying Fake Content.

Response	Frequency (n)	Percentage (%)
Yes	103	48.4
No	52	24.4

Response	Frequency (n)	Percentage (%)
Neutral	58	27.2
Total	**213**	**100**

Table 14 indicates that 48.4% of users prefer educational courses on identifying fake content, while 27.2% remained neutral, and 24.4% did not express interest.

Table 15: Awareness of AI Technologies

Response	Frequency (n)	Percentage (%)
Yes	62	29.1
No	84	39.4
Neutral	67	31.5
Total	**213**	**100**

Table 15 shows that 39.4% of users felt they lacked sufficient knowledge of AI technologies, 31.5% remained neutral, and 29.1% believed they were knowledgeable.

Table 16: Use of AI-Based Applications or Services in Daily Life

Response	Frequency (n)	Percentage (%)
Yes	111	52.1
No	72	33.8
Neutral	30	14.1
Total	**213**	**100**

Table 16 reveals that 52.1% of users use AI-based applications or services, 33.8% do not, and 14.1% remain neutral.

Table 17: Perceived Impact of AI on Content Quality on Social Media

Impact Level	Frequency (n)	Percentage (%)
1 - Extremely Negative	14	6.6
2 - Largely Negative	25	11.7
3 - No Impact	93	43.7
4 - Largely Positive	42	19.7
5 - Extremely Positive	39	18.3
Total	**213**	**100**

Table 17 shows that 43.7% of users believe AI has no impact on content quality, while

19.7% perceive a largely positive impact, and 18.3% view it as extremely positive.

Table 18: Knowledge of AI Technologies Used in Content Creation

Response	Frequency (n)	Percentage (%)
Yes	109	51.2
No	104	48.8
Total	**213**	**100**

Table 18 reveals that 51.2% of users are aware of AI technologies used in content creation, while 48.8% are not.

Table 19: Concerns About Use of AI in Content Creation

Concern	Frequency (n)	Percentage (%)
Loss of credibility	125	33.0
Spread of misinformation	135	35.6
Reduced human interaction	117	30.9
Others	2	0.5
Total	**379**	**100**
(*n = 213*)		

Table 19 indicates that 35.6% of users are most concerned about the spread of misinformation, 33% about loss of credibility, and 30.9% about reduced human interaction.

These results provide detailed insights into user behavior, awareness, and perceptions of AI and its implications for content quality and trust on social media platforms.

Hypothesis Validity Testing

Hypothesis 1: There is a relationship between users' exposure to fake content and trust in the information provided by social media platforms.

Table 20: Correlation Between Users' Exposure to Fake Content and Trust in Social Media Information

Variable	Pearson Correlation Coefficient	Significance Level (p-value)	Sample Size (n)
Exposure to Fake Content	-0.745	0.004	213

Table 20 shows a strong negative correlation that is statistically significant between users' exposure to fake content and their trust in the information provided by social media platforms. The Pearson correlation coefficient is -0.745, with a significance level of 0.004.

This indicates that greater exposure to fake content reduces users' trust in social media information, while limited exposure to fake content tends to increase trust.

Hypothesis 2: There is a relationship between the effects of artificial intelligence technologies on content quality and awareness of fake content.

Table 21: Correlation Between AI Effects on Content Quality and Awareness of Fake Content

Variable	**Pearson Correlation Coefficient**	**Significance Level (p-value)**	**Sample Size (n)**
Effects of AI Technologies on Content Quality	0.697	0.001	213

Table 21 indicates a moderate positive correlation that is statistically significant between the effects of AI technologies on content quality and awareness of fake content. The Pearson correlation coefficient is 0.697, with a significance level of 0.001. This suggests that AI technologies' influence on content quality increases users' awareness of fake content.

Hypothesis 3: There is a relationship between verifying information sources before adoption or sharing and increased trust in content.

Table 22: Correlation Between Verifying Information Sources and Increased Trust in Content.

Variable	**Pearson Correlation Coefficient**	**Significance Level (p-value)**	**Sample Size (n)**
Verifying Information Sources	0.683	0.002	213

Table 22 demonstrates a moderate positive correlation that is statistically significant between verifying information sources before adoption or sharing and increased trust in content. The Pearson correlation coefficient is 0.683, with a significance level of 0.002. This indicates that verifying sources significantly enhances trust in content. Conversely, a lack of verification reduces trust in content substantially.

These findings validate the proposed hypotheses, providing statistical evidence of the relationship between exposure to fake content, the influence of AI technologies, source verification, and trust in social media information. These results underline the importance of combating fake content and enhancing user practices, such as source verification, to maintain trust in social media platforms.

6 Discussion

Based on a comprehensive discussion of the study's findings and hypothesis testing, several

key conclusions were drawn:

The **first hypothesis** revealed a strong, statistically significant negative correlation between users' exposure to fake content and their trust in the information provided by social media. The Pearson correlation coefficient was -0.745, with a significance level of 0.004. This finding suggests that frequent and repeated exposure to false content diminishes trust in social media platforms. As the prevalence of false content increases, users become more skeptical about the credibility of the information sources, reducing their willingness to trust the content they encounter. Conversely, limited exposure to fake content can bolster users' trust. When users encounter fake content rarely, they may view it as an anomaly, reinforcing their confidence that most social media content is reliable.

The **second hypothesis** demonstrated a moderate, statistically significant positive correlation between the effects of artificial intelligence (AI) technologies on content quality and users' awareness of fake content. The Pearson correlation coefficient was 0.697, with a significance level of 0.001. AI technologies utilize advanced algorithms to analyze text, images, and videos, detecting signs of manipulation, such as inconsistencies or modifications. This enables AI systems to alert users to potential fake content, thereby increasing their awareness and equipping them to distinguish between authentic and misleading information.

The **third hypothesis** revealed a moderate, statistically significant positive correlation between verifying information sources before adopting or sharing them and increased confidence in the content. The Pearson correlation coefficient was 0.683, with a significance level of 0.002. Verification is a critical step in ensuring the accuracy of information and demonstrates users' or transmitters' commitment to sharing reliable content. This process strengthens public trust in the content, as it highlights a commitment to accuracy and authenticity. Conversely, neglecting source verification increases the likelihood of spreading inaccurate information, undermining public trust. Without effective verification mechanisms, users may perceive content as unreliable, contributing to the spread of false information.

Public trust in social media content is a significant focus in media and communication research. This trust is influenced by various factors, including the spread of false content and the effectiveness of AI technologies in detecting it. Previous studies have shed light on these dynamics:

- The study by Olan, Jayawickrama, Arakpogun, Suklan, and Liu (2022) highlighted how the prevalence of false content exacerbates distrust in media, heightens social divisions, and raises concerns about its impact on public decision-making.
- Tung and Lan (2024) examined university students' awareness of the need to verify content on TikTok. Despite increased awareness, many students expressed a lack of trust in the platform's content, underscoring the role of verification in fostering trust between audiences and the content they consume.
- Mahjoub (2023) explored the efforts of Sudanese media institutions to combat fake content using digital tools, emphasizing the necessity of strict laws to curb its spread.
- Amr (2024) discussed the impact of deep-fake technology powered by AI on media credibility. The study indicated that while deepfake technology can enhance

content manipulation, it also significantly increases the spread of rumors, weakening public trust.

From these findings, it can be concluded that trust in social media is heavily influenced by the prevalence of fake content and users' ability to identify it. The effective application of AI technologies can play a crucial role in raising users' awareness of fake content, ultimately fostering greater trust and promoting more responsible use of digital platforms.

7 Conclusion and Recommendations

This research highlights the significant impact of artificial intelligence (AI) techniques in creating and disseminating fake content on social media platforms and their subsequent effects on users' trust. The findings reveal critical relationships between exposure to false content, the role of AI in content verification, and the importance of user practices such as source verification.

The study concludes that excessive exposure to fake content undermines users' trust in social media, emphasizing the need for platforms to employ effective mechanisms for identifying and limiting the spread of false information. AI technologies, with their advanced capabilities to analyze and detect manipulated content, play a pivotal role in enhancing users' awareness and enabling them to differentiate between real and fake content. This underscores the value of integrating AI-driven tools in content moderation strategies to foster a trustworthy digital environment.

Moreover, the research demonstrates that verifying information sources before adoption or sharing significantly enhances trust. This highlights the importance of promoting media literacy and educating users about verification practices to mitigate the impact of misinformation.

The study also underscores the role of social media platforms in addressing the challenges posed by fake content. Collaborating with AI developers, policymakers, and educators to create robust frameworks for detecting and managing misinformation can strengthen public trust and ensure responsible use of social media.

In conclusion, building trust in social media platforms requires a multifaceted approach that includes leveraging AI technologies, enhancing user awareness, and fostering a culture of verification and accountability. These efforts are crucial for combating the spread of fake content and creating a more reliable and transparent digital information ecosystem.

Recommendations:

1. **Enhance Verification Mechanisms**

   It is recommended that social media platforms adopt advanced technologies to ensure content accuracy before publication or promotion. Leveraging sophisticated AI algorithms can effectively detect and flag fake content, improving overall content credibility.

2. **Promote AI Ethics**

   Establish strict ethical standards governing the use of AI in content creation to mitigate the risk of exploiting these technologies for spreading false information. Ethical guidelines will help ensure that AI is used responsibly.

3. **Conduct Studies on User Perception**

   It is recommended to conduct studies that explore the gaps in AI awareness among different user groups. Such research can inform the design of targeted awareness strategies, effectively educating each demographic on AI technologies and their implications.

4. **Develop Digital Trust Metrics**

   Develop standardized metrics to measure digital trust levels among social media users. By analyzing user experiences with content and monitoring interactions with both reliable and misleading materials through AI tools, platforms can better understand and address trust issues.

5. **Study Cross-Cultural Impacts**

   Investigate the effects of AI-generated fake content on user trust across different countries and cultures. Understanding how cultural factors influence trust levels can inform the development of tailored strategies to mitigate the negative impact of fake content globally.

References

A. N. Häuselmann. (April 23, 2024). EU Privacy and Data Protection Law Applied to AI: Unveiling the Legal Problems for Individuals.

Abdelaziz, M. (2022). Media Manipulation Techniques in Addressing Religious Issues Through YouTube TV Videos and Their Relationship to Adolescents' Attitudes Toward Religious Renewal.

Abdel-Fattah, H. (2024). The Relationship Between AI Applications and the Economics of Media Professions and Industries. Egypt.

Abdel-Salam, M. (2023). Criminal Liability for the Misuse of Deepfake Technology: A Comparative Study, pp. 1–67.

Ahlam Ail Baabdullah & Yasmeen Ansari. (June 3, 2020). Factors Influencing Online Buying Behavior of Customers in Saudi Arabia. Journal of Economic, Administrative and Legal Sciences, pp. 146–159.

Ahmed Shehata et al. (2023). Behavior in Dealing with Misinformation on Social Media Platforms in Omani Society During the COVID-19 Crisis. Retrieved from https://www.qscience.com/content/journals/10.5339/jist.2023.4#html_fulltext

Ajay Agrawal, Joshua S. Gans, & Avi Goldfarb. (February 7, 2017). What to Expect From Artificial Intelligence. Retrieved from Sloan Review: https://sloanreview.mit.edu/article/what-to-expect-from-artificial-intelligence/

Al-Babli, A. (2023). Security Risks of Deepfake and Mechanisms to Counter Them. Police Thought, pp. 167–222.

Al-Bayoumi, R. (2023). Legal Protection Against the Risks of Deepfake Applications in Islamic Jurisprudence and Secular Law. The Spirit of the Law.

Al-Naghi, Y., & Al-Naghi, W. (2022). Social Media Users' Perception of Deepfake Threats and Its Relationship to Safe Usage Practices. Scientific Journal of Journalism Research, pp. 383–425.

Al-Obaidi, H. (2021). Deepfake Technology and Its Impact on the Credibility of Electronic Media: A Descriptive Study.

Al-Qahtani, L. A., Al-Mutairi, S. M., & Al-Juhani, A. S. (2024). The Role of Saudi Digital Media in

Raising Awareness About Cybercrime Technologies Among Saudi Citizens, pp. 283–327.
Al-Rasheed, I. (2023). Artificial Intelligence and the Transformation of the Concept of Creativity in Digital Fine Arts Photography.
Al-Suwaileh, S. (2023). The Role of Artificial Intelligence Systems in Combating Electronic Rumors. Arab Journal of Security Studies, pp. 80–97.
Amr, M. (2024). The Credibility of Media Imagery Amid the Evolution of AI and Deepfake Technology. Retrieved from https://journals.ekb.eg/article_330263_836a81b53e910101e83028f4801c1d76.pdf
Ángel Fernández Gambín, Anis Yazidi, Athanasios Vasilakos, Hårek Haugerud, & Youcef Djenouri. (February 19, 2024). Deepfakes: Current and Future Trends. Retrieved from https://link.springer.com/article/10.1007/s10462-023-10679-x
Arran Gare. (2024). The Philosophy of Anti-Dumping as the Affirmation of Life. Biosemiotics.
Ayenda Kemp. (July 1, 2024). Competitive Advantage Through Artificial Intelligence: Toward a Theory of Situated AI. Retrieved from https://journals.aom.org/doi/full/10.5465/amr.2020.0205
Azibi, Y. (2024). The Role of Artificial Intelligence in Enhancing Health Services: A Case Study of Jazan Health Hospitals.
Barakat Karine Aoun, Amal Dabbous, & Abbas Tarhini. (March 19, 2021). 1468-4527. Retrieved from Emerald: https://www.emerald.com/insight/content/doi/10.1108/OIR-08-2020-0333/full/html
Femi Olan, Uchitha Jayawickrama, Emmanuel Ogiemwonyi Arakpogun, Jana Suklan, & Shaofeng Liu. (2022). Fake News on Social Media: The Impact on Society. Retrieved from Springer: https://link.springer.com/article/10.1007/s10796-022-10242-z
Hatamleh, I., Safori, A., Habes, M., Tahat, O., Ahmad, A., Abdallah, R.-Q., & Aissani, R. (July 20, 2023). Trust in Social Media: Enhancing Social Relationships. Retrieved from MDPI: https://www.mdpi.com/2076-0760/12/7/416
Hossam, M. (2022). Digital Media: Its Concept, Tools, and Theories, pp. 87–104.
Ilkka Kaate, Joni Salminen, Joao Santos, Soon-Gyo Jung, Rami Olkkonen, & Bernard Jansen. (June 19, 2023). The Realness of Fakes: Primary Evidence of the Effect of Deepfake Personas. Retrieved from https://www.sciencedirect.com/journal/international-journal-of-human-computer-studies
Jaron Mink, Licheng Luo, Natã M. Barbosa, Olivia Figueira, Yang Wang, & Gang Wang. (August, 2022). DeepPhish: Understanding User Trust Towards Artificially Generated Profiles in Online Social Networks. Retrieved from https://www.usenix.org/conference/usenixsecurity22/presentation/mink
Khodor, L., & Wanos, A. (2023). Applications of Deepfake Technology in Contemporary Marketing and Promotion, pp. 166–187.
Khourara, A. (2021). Social Media and the Spread of Rumors During the Corona Pandemic: Facebook as a Model – A Field Study on Professors at the Faculty of Humanities and Social Sciences.
L. A., Jodas, D., Costa, K. A. P., Souza Júnior, L. A., Rodrigues, D., Del Ser, J., Camacho, D., & Papa, J. P. Passos. (2024). A Review of Deep Learning-Based Approaches for Deepfake Content Detection. Expert Systems.
Lamia, M. (2021). Research Methods. Retrieved from https://dspace.univ-guelma.dz/jspui/bitstream/123456789/12736/1/%D9%85%D8%B7%D8%A8%D9%88%D8%B9%D8%A9%20%D8%A7%D9%84%D8%A3%D8%B3%D8%AA%D8%A7%D8%B0%D8%A9%20%D9%85%D8%AC%D8%AF%D9%88%D8%A8.pdf
Les Dolega, Francisco Rowe, & Emma Branagan. (February 23, 2021). Retrieved from

https://www.sciencedirect.com/science/article/abs/pii/S0969698921000679

Lina Muhammad Al-Ghamdi. (January 1, 2021). Towards Adopting AI Techniques for Monitoring Social Media Activities, pp. 15–22.

Mahjoub, S. (2023). The Use of Artificial Intelligence and Deepfake Technologies in Information Manipulation: A Survey of Sudanese Journalists' Practices in Detecting Fake Content on Social Media Networks. Journal of Arts.

Moharram, A. (2022). Applications of Artificial Intelligence (AI) Using Deepfake Technology in Defamation: A Comparative Jurisprudence Study.

Mustafa, M. (2024). The Impact of Information Overload on Egyptian Audiences' Perceptions of International Crises: The Russian-Ukrainian War as a Model – A Semi-Experimental Field Study. Media Research and Studies.

Peter Gordon Roetzel. (2019). Information Overload in the Information Age: A Review of the Literature from Business Administration, Business Psychology, and Related Disciplines with a Bibliometric Approach and Framework Development. Retrieved from Springer: https://link.springer.com/article/10.1007/s40685-018-0069-z

Qazadri, H. (2022). Sampling and Its Applications in Media Studies. Retrieved from Mandumah Database.

Rabie, E., & Abdel-Fattah, L. (2024). The Pros and Cons of Artificial Intelligence in Specialized Education: An Analytical Study. Retrieved from https://journals.ekb.eg/article_369570_1efa2955edbc01e55356fd2da4903d8f.pdf

Raqam, S., BouSaha, K., & Haroush, M. (2022). The Role of Digital Promotion for Algerian Tourism Agencies in Enhancing Customer Loyalty. Retrieved from http://dspace.univ-guelma.dz/jspui/handle/123456789/13396

Saiyad Mahmmadakram Hanif & Vivek Dave. (2022). Deepfakes Technology Using AI. Retrieved from IJSRSET: https://ijsrset.com/home/issue/view/article.php?id=IJSRSET229522

Tawahir, A., & Othman, B. (2021). Research Methods and Selection Techniques. Retrieved from Mandumah Database.

Tung, T., & Lan, D. (2024). Exploring Fake News Awareness and Trust in the Age of Social Media Among University Student TikTok Users. Retrieved from Taylor & Francis: https://www.tandfonline.com/doi/full/10.1080/23311886.2024.2302216

Wahidul Hasan Abir, Faria Rahman Khanam, Kazi Nabiul Alam, Myriam Hadjouni, Hela Elmannai, Sami Bourouis, & Mohammad Monirujjaman Khan. (July 19, 2022). Retrieved from Tech Science Press: https://www.techscience.com/iasc/v35n2/48928