Fake News on Social Media Platforms: A Survey Paper

Cathrine Shalby
College of Computing and Information Technology
University of Doha for Science and Technology
Doha, Qatar

Alaa Al-Halabi
College of Computing and Information Technology
University of Doha for Science and Technology
Doha, Qatar

Suzana Zaid
College of Computing and Information Technology
University of Doha for Science and Technology
Doha, Qatar

Khaled Khalafi
College of Computing and Information Technology
University of Doha for Science and Technology
Doha, Qatar

Abdelrahman Attili
College of Computing and Information Technology
University of Doha for Science and Technology
Doha, Qatar

Abstract—The rapid dissemination of information, particularly on social media, has resulted in a rise in fake news, which has weakened public confidence in media sources. This survey paper This survey paper enhances the understanding of the potential and challenges in combating the fake news plague by combining current studies and offering insights into the evolving nature of disinformation. It explores the causes, types, incentives, and historical context of fake news. It investigates how social media features including verification processes recommender systems, and concerns about censorship and centralization contribute to the propagation of false information. In addition, Proposed solutions are also covered within the area of blockchain and artificial intelligence. And finally, innovative strategies that big platforms have used to combat false information is outlined.

I. INTRODUCTION

In an era dominated by the rapid evolution of information dissemination, the rapid increase of misinformation and fake news has emerged as a formidable challenge, especially on social media platforms. With an estimated 5.3 billion internet users, including 4.95 billion active on social media, one of the reasons for 50.9% of them is staying up to date with news and current events [1]. In the United States, 19% of adults report that they often obtain news through social media, with an additional 31% indicating occasional consumption [2]. This emphasizes the critical role of these platforms in shaping the public's opinion and how the prevalence of fake news has increased with the recent rise of social media.

II. FAKE NEWS

Fake news refers to inaccurate or deceptive information that is presented in the form of legitimate news. The term is a neologism, which refers to a newly coined or repurposed expression that is being introduced into the language as a result of cultural or technological changes. Fabricated information, commonly known as fake news, lacks any factual foundation but is deceitfully portrayed as being true. Although many definitions primarily emphasize the accuracy of its content and its format, recent research suggests that the rhetorical structure of the content may have a notable impact on how fake news is perceived.

This section dissects the complex issue of fake news, tracing its origins, categorizing its types, understanding the motivations behind its spread, and examining its societal impacts. Initially, we provide a historical perspective to understand how misinformation has evolved with technological advancements. We then identify and describe the various forms of fake news, from unintentional misinformation to malicious disinformation, and explore the psychological and social reasons behind their propagation. Lastly, we assess the consequences of fake news on public opinion, democratic integrity, and individual behaviors. This analysis aims to offer insights into the pervasive challenge of fake news.

A. The History of Fake News

The phenomenon of fake news has an extremely long and complex history stretching back centuries. While the term "fake news" gained a widespread recognition in the digital world, its roots can be traced to ancient times where false narratives, rumours, propaganda, and misinformation were spread through word of mouth, handwritten manuscripts, and other very early forms of communication. It was mostly a tactic employed by political leaders in order to manipulate the public opinion. For instance, in the 13th century BC, the Battle of Kadesh was portrayed by Ramses the Great as a spectacular victory for the Egyptians. The walls of almost all his temples outlined scenes of him smiting his enemies, when in fact the Egyptian-Hittite pact shows that the battle

was a stalemate [3]. Another notable event was during the 1st century BC when a misinformation campaign was run by Octavian against his rival Mark Anthony, which painted him in a negative light and portrayed him as a puppet of the Egyptian queen Cleopatra VII. Their battle of fake information kept going until a message stating that Cleopatra has killed herself was sent to Mark Anthony, contributing to his tragic end when he committed suicide [4].

Moving forward to the 19th and 20th century, the advent of modern media introduced new dimensions of falsehoods. This period saw a rise in journalism, newspapers, articles, magazines, and even radios or televisions, which lead to the popularity and spread of fake news. In the 1890s, "Yellow Journalism" became popular as a style of reporting that highlight sensationalism over facts using exaggerated headlines, it was particularly associated with certain newspapers owned by publishers who competed for readership [5], the concept is very similar to today's clickbait's which are meant to attract attention online by using sensationalized language, misleading images, and incomplete or exaggerated headlines. Yellow Journalism played a pivotal role in the Spanish-American War of 1898 by sensationalizing events such as the sinking of the USS Maine, stoking anti-Spanish sentiment, and influencing public opinion to support war efforts [5]. During World War II, both Axis and Allied powers used propaganda to manipulate public opinion and boost morale, using radio broadcasts and printed media to spread false stories and undermine enemy morale [6]. Additionally, the United States saw disinformation campaigns like "Red Scare" during two periods in history, the first was after World War I, and the second during the Cold War between the US and the Soviet Union after World War II, the purpose was to fuel paranoia and cause social and political upheaval [7].

Now, during this digital age, fake news has truly flourished with the rise of the internet and social media platforms which serve as a ground for the dissemination of false information. The notable event where fake news reached its peak was during the 2016 U.S elections, where disinformation and misinformation circulated widely on social media to cause conflicts and sway public opinion [8]. Additionally, during the COVID-19 period information about the virus' origins, treatment, number of cases, or the vaccine misinformation proliferated online resulting in what is called an "infodemic" [9]. Furthermore, recently the Israeli-Palestinian conflict has also caused a surge in fake news, weaponizing the social media to propagate narratives and garner international support, fuelling tensions between conflicted parties. Aside from spreading fake news, the spread of deepfakes which are realistic AI-generated images audios, and videos, pose a new frontier in the battle against fake news, further complicating the efforts of distinguishing the truth from fiction [3].

B. Types of Fake News

False information comes in various forms, influenced by the intentions behind its dissemination. The inadequacy of the term "fake news" in accurately reflecting our current information landscape is one of the reasons its usage is being avoided. Given the importance of language, it is more appropriate to employ alternative terms like misinformation, disinformation, and malinformation [10]. This approach aids in distinguishing the various shapes of false information to gain a deeper understanding of the challenges posed by each.

- Misinformation: This type of false information is often spread without the intention to deceive. It can arise from various sources, usually they are inaccurate, incomplete, vague, or ambiguous information in a context, on various topics. Misinformation can occur due to factors like cognitive biases, lack of access to reliable information, human errors, etc. These factors affect how the entities perceive this information and its accuracy.
- Disinformation: Unlike misinformation, disinformation is intentionally spread with the purpose of deceiving and manipulating the audience, it is often intended to influence opinions. Disinformation can be spread by individuals, organizations, or governments seeking to advance their agendas.
- Malinformation involves the dissemination of true information but with the intent to cause harm or damage to individuals, organizations and their reputations. Unlike misinformation and disinformation, malinformation relies on genuine facts or private information that's weaponized for malicious purposes. An example of this would be doxing which is done by sharing private information online.

Misinformation and disinformation encompass a diverse array of deceptive content. Here, we delve into seven primary types of misinformation and disinformation, shedding light on their distinct characteristics and implications [11].

1) Fabricated content: This is the content that is completely false and new, such as quotes or statements that were allegedly said by someone, fake news articles and made-up stories, or hoaxes. All this meant to mislead audiences and push particular agendas.
2) Manipulated Content: Involves content that has been manipulated to deceive people, usually pictures or videos that were altered in a way that they seem real. The most used method nowadays is deepfakes, a form of synthetic media that uses powerful AI and machine learning techniques to generate and manipulate audio and visuals with the potential to deceive people and trick them into thinking it's real [12]. Speech synthesis is another form used to produce artificial human-like speech, where speech fragments from recordings are concatenated, or models/vocal traits are used to create entirely synthetic voices [12]. This adds more concerns into the realm of fake news.
3) Imposter Content: Refers to the impersonation of legitimate users and news sources. Examples include mimicking legitimate sources to spread false information or creating fake profiles online to pose as legitimate individuals/organizations to deceive the other users and

take advantage of the trust these specific entities have.

4) Misleading Content: The information that misleads users. This may include selective quotes or statistics picked out of context to distort their meaning and misrepresent the original intent, as well as cropping pictures or using biased language to influence and shape opinions or support a specific argument.

5) False Context: Factually accurate content that is combined and shared with false contextual information. Such as clickbait headlines which attract clicking by using sensational headlines when it has only little resemblance to the actual content, or misleading thumbnails that misrepresent the content of an article or video.

6) Satire or Parody: The usage of humorous but fake stories shared to fool, or the creation of parody accounts on social media used to mock and satirize individuals, organizations, or events. Also, the use of memes or images that are not intended to be taken seriously but may be misleading and taken as genuine news.

7) False Connection: Using headlines, captions, or pictures that are not related to the content, to present information in a way that implies a connection/relationship that does not exist.

C. Motivations for Sharing Fake News

In this section, we delve into the different motivations that drive the dissemination of false information on social media platforms. The spread of false information can arise from both unintentional and intentional motives, each influenced by various psychological and social factors.

Unintentional spread of fake news might be motivated when individuals want to express themselves and connect with others, seeking different viewpoints [13]. Conversely, intentional sharing could stem from wanting to fit in with existing beliefs or maintain good relationships. Fake news sharing has been positively associated with social media use behaviors. People who tend to disclose more information on social media generally, those who receive fake news from a trusted source, individuals who seek popularity and a sense of belonging, and those who experience information overload are more likely to pass fake news on to others in their social networks [13]. Moreover, those inclined towards entertainment, provocation, or debate may also engage in the dissemination of fake news.

Beyond these motivations, other factors contribute to the spread of misinformation. Collective fact-checking, sensationalism, and ridicule play roles in amplifying fake news content. Recipients of such stories may share the suspected misinformation with their friends and followers on social media to crowd-source its truthfulness and reduce uncertainty about the information's credibility. Furthermore, sharing misinformation for educational purposes functions to debunk misinformation and serves as a warning to others not to believe the information. This altruistic act of alerting recipients to potential misinformation may also stem from a desire to demonstrate moral or intellectual superiority [13].

Understanding these diverse motivations provides insights into the complex dynamics underlying the dissemination of fake news on social media platforms.

D. Consequences of Fake News

Following the exploration of various types of false information and the underlying motivations behind their spread, this subsequent section will discuss the potential consequences and outcomes of spreading such information, regardless of where it originated from. This segment will analyze how fake news can impact individuals, organizations, and even countries, illuminating the wide-ranging consequences of misinformation in our digital landscape.

- Financial Impact: The stock market may suffer greatly from the spread of false information. This was demonstrated by the events of November 10, 2022, when a fabricated tweet impersonating Eli Lilly falsely claimed the distribution of free insulin, resulting in a 4.37% decline in Eli Lilly's shares [14].
- Reputation damage: Maliciously spread fake news can tarnish the reputation of individuals or companies. For instance, Tesla's safety record was defamed by false information in 2022 [15]. Similarly, in 2017, a well-known member of the pseudoscientific media David Wolfe published an article claiming that ingredients in Reese's Peanut Butter Cups could negatively impact fertility, cause gastrointestinal problems in kids, and cause symptoms like dizziness, nausea, tinnitus, and fainting. The article was mainly made up of false material because the scientific foundation for these claims was outdated and lacked thorough accuracy [16].
- Public Deception: False narratives are frequently created utilizing propaganda and fake news in order to mislead the public. One example is the bombing of Gaza's Al-Ahli Baptist Hospital on October 17, 2023, where Israel's digital spokesperson Hananya Naftali first acknowledged their responsibility on X. However, the post was then removed, and followed by denials [17], deceiving the public and obscuring the actual truth.
- Political Conflict: Fake news can increase political conflicts and create an unstable environment for political discourse and decision-making by spreading misinformation, widening rifts, and weakening public trust in authorities. An instance of this occurred on May 23, 2017, when the official website of the Qatar News Agency released fabricated quotes written by an unidentified hacker attributed to Emir Tamim Bin Hamad Al-Thani telling the crowd that Qatar has "tensions" with the new US administration and that "there is no wisdom in harboring hostility toward Iran" and relations with Israel are "good", which sparked international hashtags calling for cutting ties with Qatar. This particular incident almost started a war and led to a blockade [18].

III. The Current Technological Enablers of Fake News in Social Media

In this section, the several elements present in existing social media platforms that facilitate the extensive spread of false information will be discussed, and how they all work together to actively promote false information in the digital era.

A. Recommender Systems

Recommender systems, employed by various online platforms, utilize sophisticated algorithms to analyze user data and engagement patterns, aiming to deliver personalized content recommendations [19]. While these systems enhance user experience, they inadvertently contribute to the formation of filter bubbles, echo chambers, and rabbit holes.

Filter bubbles occur when the recommender systems prioritize content that aligns with the user's existing interests and viewpoint, limiting the exposure to diverse perspectives. This phenomenon fosters confirmation bias, reinforcing users' pre-existing beliefs and insulating them for dissenting opinions. Similarly, echo chambers emerge when users are predominantly exposed to content that validates their opinions, leading to the amplification of partisan discourse and ideological polarization in online communities. Moreover, the proliferation of rabbit-holes which are narrow and repetitive content loops, further complicates this problem by preventing users from accessing a wide range of information and encouraging intellectual isolation [20].

These algorithm-driven recommendation systems lead to either overrepresentation or underrepresentation of certain groups, which results in the spread of misinformation, societal divisions, and discrimination [19]. Addressing the unintended consequences of recommender algorithms is paramount to fostering a more inclusive and informed digital environment.

B. Paid Verification

In the context of identity verification, recent shifts in popular platforms' policies such as X and Meta regarding the verification check mark highlight the developing nature of how online identities are verified, raising questions about its relevance. Originally the verification check mark was given to maintain the integrity of notable accounts to prevent threats like impersonation, but in March 2023, Meta started offering U.S. users paid subscriptions for a verified account, making it possible for regular users to obtain it [21]. Concurrently In 2022, Elon Musk's X (previously Twitter) updated their verification process entirely, making it possible for users to obtain that check mark through a monthly subscription to Twitter Blue, now X premium. In 2023 they decided to put an end to the legacy verification program and continue with Twitter Blue verified accounts [22].

However, this shift allowed various corporations and notable figures to be impersonated once their legacy verification check mark vanished due to the policy changes. Following a concerning incident where a Washington Post reporter managed to create a fabricated verified account impersonating Massachusetts Democrat Senator Ed Markey, the senator addressed Elon

Musk directly through X. Markey urged Musk to provide an explanation for the incident and to outline preventive measures to ensure such impersonation attempts are limited [23]. This was only one of the many impersonation cases during that period, even though such accounts got suspended afterward, they still underscore the challenges of paid verification.

C. Centralization and Censorship

The problem of centralization and censorship on social media platforms poses significant threats to free speech and open discourse. Centralized platforms have the power to censor or moderate content, thereby constraining the ability of users to freely express themselves. Recent accusations of Instagram purposefully censoring posts or stories in support of Palestine highlight this issue [24]. Such actions not only silence individuals advocating for social justice, but also raise concerns about the fairness and neutrality of platform moderation procedures.

Another case of such censorship was when X (Twitter then) removed over 50 tweets in India, criticizing the government's ways of handling the COVID-19 pandemic after the Indian government asked Twitter to take them down [25]. As centralized platforms continue to play a dominant role in shaping online discourse, addressing the challenges of centralization and censorship becomes increasingly urgent to safeguard the principles of free expression and democratic participation in the digital age.

D. Automated Manipulation

In the realm of automated manipulation, social media bots serve as pivotal weapons for spreading false information, wielding the ability to artificially amplify the popularity of an individual or a movement, with follower bots available for purchase on the black market, thereby influencing perceptions of influence and importance. These bots are frequently employed during elections and political campaigns to sway public opinion, undermining the integrity of democratic processes, as seen in the significant percentage of political discussions generated by bots during the 2016 U.S. presidential election [26].

Bots extend their reach to financial markets, where they flood social media with false information to manipulate stock prices [27], leading to market volatility. Detecting and mitigating the influence of social media bots pose significant challenges due to their increasing capability to mimic human behaviour and evade detection algorithms.

IV. State of the Art

In the dynamic and ever-evolving fight against the spread of misinformation on social media, platforms are increasingly adopting sophisticated strategies and technologies. These efforts are geared towards not only mitigating the dissemination of fake news but also enhancing the information ecosystem for users worldwide. This section delves into the nuanced approaches of major platforms like X, Meta (Facebook and Instagram), Google (YouTube), and TikTok, alongside innovative

solutions like Birdwatch, to provide a comprehensive overview of the current landscape in combating misinformation.

X has taken a leading role with its innovative Birdwatch feature, which utilizes both crowd wisdom and advanced algorithms to combat misinformation. The 2022 paper, "Birdwatch: Crowd Wisdom and Bridging Algorithms can Inform Understanding and Reduce the Spread of Misinformation" highlights an algorithm that employs matrix-factorization and a "bridging-based ranking" approach. The aim is to pinpoint annotations that are not only accurate but also resonate with diverse audiences, particularly on polarizing topics. By surpassing traditional voting methods, this algorithm effectively selects notes that enhance comprehension and positively impact social media sharing habits by reducing unnecessary retweets and likes. This initiative underscores X's commitment to fostering a more informed community and mitigating the spread of fake news through strategic, crowd-sourced content moderation [28].

Meta Platforms, encompassing Facebook and Instagram, have implemented robust strategies to counter misinformation. These include advanced algorithmic modifications to demote false content and partnerships with global fact-checking organizations to verify the accuracy of information. Meta's approach emphasizes user engagement tools, such as warning labels on misinformation and contextual information panels, to encourage critical thinking and responsible content sharing among its users [29].

Google and YouTube prioritize credible content and authoritative sources through algorithmic updates and the introduction of information panels. These efforts are particularly focused on sensitive and high-stakes topics, aiming to guide users toward accurate information and away from misleading content [29].

TikTok addresses misinformation through its fact-checking program and by labeling unverified content, attempting to strike a balance between informing users and preserving freedom of expression. This approach reflects a broader industry trend toward empowering users to critically evaluate the content they encounter online [29].

V. PROPOSED SOLUTIONS

This section collects several studies offering a variety of solutions to the dissemination of false information grouping them by the technology used.

A. Blockchain

In 2023, Xiaowan Wang, Huiyin Xie, Shan Ji, Liang Liu, and Ding Huange introduced a mechanism leveraging a blockchain database cloud server to store complete news articles. Meanwhile, metadata about these articles is stored on the blockchain and linked together through polynomial commitment. This approach enhances blockchain storage efficiency while ensuring the integrity and traceability of news content by detecting any alterations or updates made to the news subsequent to its publication [30].

In 2019, M. Torky, Emad Nabil, and Wael Said explored the potential development of a blockchain-based Proof of Credibility (PoC) consensus protocol as a solution to the issue of fake news on current social media platforms, where artificial intelligence is misused to spread misinformation instead of preventing it. They utilize a Peer-To-Peer (P2P) social network model where each peer shares a distributed ledger and is responsible for validating the shared information. If identified as fake, the information is added to the blockchain, in order to detect fake news on social networks [31].

In 2017, Antorweep Chakravorty and Chunming Rong proposed Ushare, a blockchain-based social media network. On Ushare, users have the ability to determine the sharing limit for their posts and utilize a Personal Certificate Authority (PCA) to control who can view it. This approach empowers users by enabling them to control, trace, and claim ownership of their content, thus securing it from unauthorized use and view [32].

In 2019, Iago Sestrem Ochôa, Gabriel de Mello, Luis Augusto Silva, Abel JP Gomes, Anita Maria Rocha Fernandes, and Valderi Reis Quietinho Leithardt proposed FakeChain, a centralized blockchain architecture where each block is a news item that contains all other sources related to it with validation based on users consensus to detect and alert fake news on social media, punish those who spread misinformation, and reward those who publish correct information on the network. [33].

B. Artificial Intelligence (AI)

In 2023, Claudio Marche, Ilaria Cabiddu, Christian Giovanni Castangia, Luigi Serreli, and Michele Nitti proposed a web browser plugin responsible for tracking user web page activity, evaluating the news, and providing the user with feedback if it detects the news is fake using machine learning to verify the reliability of the news and its provider [34].

C. Blockchain and AI

In their 2021 research, Akash Dnyandeo Waghmare and Dr. Girish Kumar Patnaik proposed an innovative model for Fake News Detection using a combination of custom blockchain and machine learning classification algorithms. The system utilizes supervised machine learning algorithms, particularly focusing on Natural Language Processing (NLP), to address the challenge of identifying false information. This model is designed to detect fake news generated by reporters as well as malicious bots [35].

In 2023, Sakshi Kalra, Yash Bansal, Yashvardhan Sharma, and Gajendra Singh Chauhan propose a blockchain-based solution for fake news detection called "FakeSpotter: A blockchain-based trustworthy idea for fake news detection in social media". The system involves the voting of various users, including News Agencies, Journalists, Editors, Crowd Annotators, and Expert Annotators. Each user registers on a portal, and their data is hashed and stored in a block. All posts get stored in the blockchain [36].

In 2021, Zeinab Shahbazi and Yung-Cheol Byun proposed a system that integrated machine learning by using Natural Language Processing (NLP) and reinforcement learning to classify and predict fake accounts and news with blockchain

to add a layer of security to store the collected news to detect fake news on social media and reduce its re-usage [37].

In 2023 FP Team published this article that discusses how blockchain technology can be a powerful tool to combat fake news. The article also explores the use of smart contracts in combating fake news, which are self-executing contracts stored on a blockchain network. A case study on Fact Protocol also was made, which is a decentralized fact-checking system using blockchain that illustrates how a two-factor approach involving news registrars and validators helps reduce bias and increase trust in information. Fact Protocol believes that blockchain technology can effectively combat fake news by ensuring the validity and integrity of news content through cryptographic techniques and processes [38]

In 2020, WhistleBlower was proposed. A decentralized fake news detection platform leveraging blockchain, Artificial Intelligence (AI), and Machine Learning (ML). It employs a Token Curated Registry (TCR) where the community uses AI and ML algorithms for fake news detection. Blockchain ensures transparency and security. Users can challenge algorithmic scores, encouraging community engagement and enhancing the accuracy of detection. The platform uniquely combines text analysis, fact-checking, source evaluation, and community feedback, providing a decentralized solution to combat misinformation [39].

VI. Reflections

To address this pressing issue of fake news, various strategies and technologies have been proposed as seen in the previous sections, each with its unique approach to mitigate the spread of misinformation. Table I presents an examination of these solutions, and reveals a nuanced landscape where technological innovations, community engagement, and transparency intersect in the fight against fake news.

Some of the approaches observed use advanced algorithms and crowd wisdom, such as the one demonstrated by X's Birdwatch. By harnessing the collective knowledge of users and employing sophisticated algorithms, platforms aim to enhance understanding and reduce the dissemination of false information. Similarly, Meta Platforms, including Facebook and Instagram, leverage warning labels and contextual information panels to encourage critical thinking and combat misinformation. However, challenges like the reliance on centralization introduces the risk of biases and may not capture all instances of misinformation.

Blockchain technology emerges as another promising approach for combating fake news, some approaches utilized blockchain to offer decentralized solutions that prioritize transparency and integrity such as Ushare and Fact Protocol. Yet, challenges persist in the cost, scalability, and adoption of blockchain networks, highlighting the need for further refinement.

Additionally, solutions that prioritize community engagement play a vital role in the fight against fake news. Platforms like WhistleBlower and FakeSpotter's user voting approach engage users in the detection and verification of fake news, fostering transparency and accountability within online communities. However, reliance on community participation may introduce biases and scalability concerns.

Addressing the spread of fake news on social media platforms demands a multifaceted approach. While each solution discussed offers unique advantages in combating misinformation, it is essential to acknowledge their limitations and potential unintended consequences.

VII. Ethical Considerations

After examining the numerous problems within today's social media platforms and considering the solutions that are being developed and implemented, it's crucial to discuss the significant ethical responsibilities that social media companies bear. These responsibilities demand proactive actions and strategies to ensure responsible management of social media platforms.

Maintaining the integrity of circulating content on their platform stands as one of the most important duties. Social media companies must prioritize the accuracy of the shared information on their platforms, which requires employing solutions including human moderators and advanced algorithms, to mitigate the spread of misleading content. However, this must be balanced with upholding principles of free speech and open dialogue, avoiding censorship of different voices, and implementing recommender systems [40] while still safeguarding the users from the harms of false information.

Transparency emerges as another critical consideration in the efforts to fight fake news. Providing users with clear indicators of the reliability and sources of information enables them to make informed judgements about the encountered content, fostering a more insightful community. Moreover, platforms should prioritize educating users and promoting critical thinking skills and digital literacy, resulting in empowered individuals capable of independently evaluating the truthfulness of the content they see. This can create a more informed and engaged community that can lessen the impact of fake news.

Another foundational ethical responsibility for social media platforms is verification. Ensuring the authenticity of shared information and its sources is crucial in fostering trust and reliability among users. A robust verification process needs to be set in place for both content and users, leveraging technologies, such as artificial intelligence, human fact-checkers, or both, to authenticate the accuracy of news and their sources. Clearly labelling verified accounts and news sources enables users to distinguish between credible information and potentially misleading content. A proper verification mechanism in place can mitigate the spread of fake news by limiting the visibility and reachability of unverified sources. However, it's essential to approach the verification with transparency and inclusivity, without allowing users to pay for a check mark, to ensure reliability and accessibility to diverse voices and perspectives [41].

Furthermore, in the battle against fake news, protecting users' privacy on platforms is an important aspect that needs

TABLE I: Comparison table

Solution	Approach	What is Addressed	Drawbacks
Birdwatch (Twitter) [28]	Crowd wisdom and advanced algorithms.	Enhances understanding of the context in misleading posts and reduces spread of viral disinformation posts.	Heavily relies on user participation which may be inaccurate or biased in some cases, and it's also centralized.
Meta Platforms (Facebook and Instagram) [29]	Implements warning labels and contextual information panels encouraging critical thinking.	Combats various forms of misinformation.	Relies on external fact-checking organizations. The centralization of content moderation can lead to bias and lack of transparency.
Google and YouTube [29]	Prioritizes accurate information through algorithmic updates, information panels, video fact-checking.	Emphasizes credible content and authoritative sources to combat misinformation.	Algorithmic updates may not always effectively demote false content, and reliance on authoritative sources can lead to echo chambers and the problem of centralization.
TikTok [29]	Labelling unverified content and fact-checks to balance informing users and preserving expression.	Targets various forms of misinformation and promotes transparency.	Labelling unverified content may not always be effective, it is centralized, and there could be challenges in distinguishing between satire and misinformation.
Wang et al. [30]	Uses blockchain to store news articles and metadata for integrity and traceability.	Enhances integrity and traceability of news content, reduces centralization, and addresses the challenges of the limited blockchain storage.	Scalability issues with blockchaina and the off-chain solutions introduce dependency on centralization, adoption challenges.
Torky et al. [31]	Utilizes blockchain for PoC consensus protocol.	Detect and adds rumours and fake news to blockchain through peer validation, promotes transparency.	Relies on peer validation which may introduce biases and scalability concerns.
Ushare [32]	Utilizes blockchain to empower users to control sharing limits and claim ownership of content.	Ownership and sharing control of user-generated content, defeats centralization.	Complexity in user adoption and potential limitations in scalability.
FakeChain [33]	Uses blockchain and data mining as consensus for detecting fake news.	Detects and alerts fake news, punishes the entities who share fake news, and rewards the ones who do otherwise.	Uses a centralized blockchain and consensus mechanisms which may introduce biases and scalability concerns.
Marche et al. [34]	Employs machine learning to verify news reliability and their provider using web browser plugin.	Text-based fake news detection using machine learning.	Relies on training data which may introduce biases and challenges in real-time detection.
Waghmare and Patnaik. [35]	Utilizes blockchain and ML classification for fake news detection.	Detects fake news through supervised ML, promotes transparency.	Integration complexities, the need of many computational requirements, and cost concerns.
FakeSpotter [36]	Involves various user groups in voting for news authenticity stored in the blockchain.	Detects various types of fake news through user voting, promoting transparency.	Reliance on user participation, may suffer from scalability issues.
Shahbazi and Byun. [37]	Integrates NLP and reinforcement learning with blockchain for fake news detection.	Detects and predicts fake accounts and news, promoting transparency and accountability.	Relies on machine learning models which may not capture all forms of fake news and could introduce biases.
Fact Protocol [38]	Utilizes blockchain and smart contracts for decentralized fact-checking.	Ensures validity and integrity of news content, promotes transparency.	Relies on community validators and may introduce biases.
WhistleBlower [39]	Utilizes community engagement and AI algorithms for fake news detection.	Engages community in detecting fake news, promotes transparency.	Challenges in algorithmic accuracy and community participation, potential biases.

consideration. Social media companies must ensure that verification processes do not compromise individuals' privacy, employing robust data protection measures and transparency regarding data usage. This enhances user trust and inclusivity while maintaining the effectiveness of the efforts made to combat fake news. Integrating privacy considerations into

strategies is vital for promoting a healthy online environment [41].

Additionally, collaboration with regulators, researchers, civil society organizations, and other stakeholders is necessary to address challenges posed by fake news and develop comprehensive strategies, leveraging the expertise and resources of different actors to identify and counter fake news while minimizing consequences. Social media companies should understand the complexity of the information ecosystem and acknowledge the role of human behaviour in fake news dissemination. They must foster online communities where users feel a sense of responsibility for shared content. Encouraging positive norms such as fact-checking before sharing information can help mitigate the spread of misinformation within these online communities [40], and investing in research and development of innovative technologies, such as blockchain and decentralized networks, can offer alternative approaches to address the challenges of misinformation while preserving privacy and decentralizing the control over content dissemination.

Embracing approaches that combine technological innovations, user empowerment, and collaboration with external stakeholders will be essential in effectively combating the spread of fake news on social media platforms.

VIII. CONCLUSION

In conclusion, the widespread problem of fake news has gained more attention in the current digital era, especially with the rise of social media platforms as the primary means for the circulation of information. As highlighted throughout this survey paper, there are several ways that false news might appear intentionally or unintentionally, ranging from fabricated content to satire and parody, all of which provide different challenges for the accuracy of the information that is shared online with significant impacts on people, organizations, and even the dynamics of geopolitics.

In addition, the multifaceted nature of social media dynamics—such as verification processes, recommender systems, automated ways of manipulation, and concerns about censorship and centralization makes the issue more complicated by contributing to the dissemination of false information.

However, despite these difficulties, there are several approaches to countering fake news, from artificial intelligence to blockchain technology. Stakeholders can contribute to the development of a more reliable digital information ecosystem by utilizing novel approaches like crowd-sourced fact-checking projects, AI-powered detection algorithms, and blockchain-based verification systems.

Furthermore, a rising dedication to preventing the spread of misinformation can be seen in the approaches used by major social media platforms, such as Twitter's Birdwatch function and Meta's fact-checking collaborations where through their emphasis on promoting reliable content and enabling users to assess information critically, they will play a vital role in reducing the influence of fake news.

Finally, numerous barriers hinder the propagation of fake news, but cooperation and creativity are also required to safeguard the truthfulness of information on the internet. By conducting additional research and implementing practical solutions, we may strive toward a digital society that is more informed, resilient, and dependable. In addressing the complex challenges posed by fake news, it's crucial to consider not only technological solutions but also societal and regulatory measures to foster a more trustworthy digital information landscape.

REFERENCES

[1] "Digital 2023 october global statshot report — datareportal – global digital insights." https://datareportal.com/reports/digital-2023-october-global-statshot, 2024. Accessed: 2024-01-19.

[2] "Social media and news fact sheet — pew research center." https://www.pewresearch.org/journalism/fact-sheet/social-media-and-news-fact-sheet/, 2024. Accessed: 2024-01-19.

[3] A. C. Doiul, "History of fake." https://www.thedeepfake.report/en/history-of-fake-copy.

[4] "Cleopatra's death." https://penelope.uchicago.edu/~grout/encyclopaedia_romana/miscellanea/cleopatra/rixens.html.

[5] "Milestones: 1866–1898 - office of the historian." https://history.state.gov/milestones/1866-1898/yellow-journalism.

[6] "World war ii propaganda — american experience — official site — pbs." https://www.pbs.org/wgbh/americanexperience/features/goebbels-propaganda/.

[7] "Red scare: Cold war, mccarthyism & facts — history." https://www.history.com/topics/cold-war/red-scare, 2010.

[8] "Fake news in the 2016 us elections - evaluating information - library guides at victoria university." https://libraryguides.vu.edu.au/evaluating_information_guide/fakenews2016, 2024.

[9] M. Ries, "The covid-19 infodemic: Mechanism, impact, and counter-measures—a review of reviews," *Sustainability*, vol. 14, p. 2605, Feb. 2022.

[10] "Understanding information disorder - first draft." https://firstdraftnews.org/long-form-article/understanding-information-disorder/.

[11] Antje, "Disinformation and 7 common forms of information disorder - the commons." https://commonslibrary.org/disinformation-and-7-common-forms-of-information-disorder/, 2024.

[12] Respeecher, "Deepfake or synthetic voice. what's the difference? — by respeecher — medium." https://respeecher.medium.com/deepfake-or-synthetic-voice-whats-the-difference-43d84895b0a3, 2024.

[13] M. J. Metzger, A. J. Flanagin, P. Mena, S. Jiang, and C. Wilson, "From dark to light: The many shades of sharing misinformation online," *Media and Communication*, vol. 9, p. 134–143, Feb. 2021.

[14] "How a fake tweet cost eli lilly billions." https://www.advisory.com/daily-briefing/2022/11/14/eli-lilly-twitter, 2024.

[15] "Why brand impersonation is increasing on twitter & how to combat it - corsearch." https://corsearch.com/content-library/blog/why-brand-impersonation-is-increasing-on-twitter-how-to-combat-it/, Dec. 2024. Accessed: 2024-01-19.

[16] "Three reasons you should stop eating peanut butter cups — snopes.com." https://www.snopes.com/fact-check/peanut-butter-cups/, 2017.

[17] "Who's to blame for the massacre at gaza's baptist hospital?." https://www.middleeastmonitor.com/20231018-whos-to-blame-for-the-massacre-at-gazas-baptist-hospital/, Oct. 2023.

[18] D. N. Team, "Hackers take aim at qatar by publishing fake news on qna - doha news — qatar." https://dohanews.co/hackers-take-aim-at-qatar-by-publishing-fake-news-on-qna-2/, 2017.

[19] "The impact of social media algorithms on content distribution." https://aicontentfy.com/en/blog/impact-of-social-media-algorithms-on-content-distribution, 2024. Accessed: 2024-01-19.

[20] "Echo chambers, rabbit holes, and ideological bias: How youtube recommends content to real users — brookings." https://www.brookings.edu/articles/echo-chambers-rabbit-holes-and-ideological-bias-how-youtube-recommends-content-to2022.

[21] N. Pettijohn, "Everyone is verified. now what?." https://www.forbes.com/sites/nathanpettijohn/2023/04/21/everyone-is-verified-now-what/?sh=1827e0643b95, 2024. Accessed: 2024-01-19.

[22] "How to get verified on twitter in 2023." https://buffer.com/library/how-to-get-verified-on-twitter/, 2024.

[23] A. Capoot, "Sen. ed markey hits back at elon musk after his response to questions about impersonation." https://www.cnbc.com/2022/11/13/sen-ed-markey-hits-back-at-elon-musk-after-his-response-to-questions-about-impersonation-.html, 2022.

[24] "Instagram users accuse platform of censoring posts supporting palestine — instagram — the guardian." https://www.theguardian.com/technology/2023/oct/18/instagram-palestine-posts-censorship-accusations, 2024.

[25] "At government request, twitter takes down some tweets critical of official covid handling." https://thewire.in/tech/as-covid-19-crisis-deepens-twitter-takes-down-tweets-criticising-modi-government, 2024. Accessed: 2024-01-19.

[26] "View of social bots distort the 2016 u.s. presidential election online discussion — first monday." https://firstmonday.org/ojs/index.php/fm/article/view/7090/5653, 2016.

[27] "What is a social media bot? — social media bot definition — cloudflare." https://www.cloudflare.com/learning/bots/what-is-a-social-media-bot/.

[28] S. Wojcik, S. Hilgard, N. Judd, D. Mocanu, S. Ragain, M. B. F. Hunzaker, K. Coleman, and J. Baxter, "Birdwatch: Crowd wisdom and bridging algorithms can inform understanding and reduce the spread of misinformation," 2022.

[29] C. Asselin, "Social media giants fight against fake news: Moving forward or stagnant?." https://blog.digimind.com/en/insight-driven-marketing/social-media-giants-fight-fake-news, May 2023.

[30] X. Wang, H. Xie, S. Ji, L. Liu, and D. Huang, "Blockchain-based fake news traceability and verification mechanism," *Heliyon*, vol. 9, p. e17084, July 2023.

[31] M. Torky, E. Nabil, and W. Said, "Proof of credibility: A blockchain approach for detecting and blocking fake news in social networks," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 12, 2019.

[32] A. Chakravorty and C. Rong, "Ushare: user controlled social media based on blockchain," in *Proceedings of the 11th International Conference on Ubiquitous Information Management and Communication*, IMCOM '17, ACM, Jan. 2017.

[33] I. S. Ochoa, G. de Mello, L. A. Silva, A. J. P. Gomes, A. M. R. Fernandes, and V. R. Q. Leithardt, *FakeChain: A Blockchain Architecture to Ensure Trust in Social Media Networks*, p. 105–118. Springer International Publishing, 2019.

[34] C. Marche, I. Cabiddu, C. G. Castangia, L. Serreli, and M. Nitti, "Implementation of a multi-approach fake news detector and of a trust management model for news sources," *IEEE Transactions on Services Computing*, vol. 16, no. 6, pp. 4288–4301, 2023.

[35] A. D. Waghmare and G. K. Patnaik, "Fake news detection of social media news in blockchain framework," *Indian Journal of Computer Science and Engineering*, 2021.

[36] S. Kalra, Y. Bansal, Y. Sharma, and G. S. Chauhan, "Fakespotter: A blockchain-based trustworthy idea for fake news detection in social media," *Journal of Information and Optimization Sciences*, vol. 44, no. 3, p. 515–527, 2023.

[37] Z. Shahbazi and Y.-C. Byun, "Fake media detection based on natural language processing and blockchain approaches," *IEEE Access*, vol. 9, pp. 128442–128453, 2021.

[38] F. Team, "Blockchain technology as a tool to detect and combat fake news - fact protocol." https://fact.technology/learn/blockchain-technology-to-combat-fake-news/, 2024.

[39] G. Ramachandran, D. Nemeth, D. Neville, D. Zhelezov, A. Yalçin, O. Fohrmann, and B. Krishnamachari, "Whistleblower: Towards a decentralized and open platform for spotting fake news," in *2020 IEEE International Conference on Blockchain (Blockchain)*, pp. 154–161, 2020.

[40] B. Dhiman, "Ethical issues and challenges in social media: A current sce- nario," Mar. 2023.

[41] K. Keane, "Ethical issues in social media and digital journalism." https://www.linkedin.com/pulse/ethical-issues-social-media-digital-journalism-kayla-k-keane-hpage, Nov. 2023.