Kasu Journal of Computer Science Vol.3 No.4 [Dec, 2026], pp. 029-037
https://doi.org//10.47514/kjcs/2026.3.4.003

Online ISSN: 1597-2216
Print ISSN: 1597-2178
www.kjcs.edu.ng

Artificial Intelligence and Blockchain for Social Media Security: Authentication, Scam Detection and Federated Learning

Ngene Samuel[1] and Ugah John Otozi[2]

[1] Department of Computer Science, David Umahi Federal University of Health Sciences, Uburu, Ebonyi State, Nigeria
[2] Department of Computer Science/Informatics Department, Alex Ekwueme Federal University Ndufu-Alike, Ikwo, Ebonyi State, Nigeria
ngenedad@gmail.com[1], ugahjohn@gmail.com[2]

Abstract

Background: Social-media security increasingly involves threats that remain difficult to address through login controls alone, particularly fake accounts, impersonation, phishing, account takeover and behaviour that becomes malicious after successful authentication. Aim: This review examines how Artificial Intelligence (AI), blockchain, sequential behavioural models and Federated Learning (FL) can be combined as complementary security mechanisms for social-media environments. Method: A structured narrative review was undertaken using peer-reviewed journal articles, conference papers, surveys and foundational studies relevant to social-media cybersecurity, decentralized identity, AI-based fraud detection, behavioural modelling and blockchain-enabled FL. Recent literature from 2021–2026 was emphasized, while earlier seminal studies were retained where required to explain RNN/LSTM and FL foundations. The evidence was synthesized by security function, technical approach, reported contribution and limitation. Results: The reviewed studies show a consistent division of strengths. AI is effective for adaptive fraud and behavioural detection; blockchain strengthens credential verification, integrity and accountability; sequential models capture temporal behaviour; and FL reduces the need to centralize raw behavioural data. Recent blockchain-enabled FL studies improve coordination and accountability but continue to face scalability, poisoning, privacy leakage and deployment overhead. The principal gap is the limited empirical evaluation of end-to-end architectures that join wallet-based authentication, post-authentication behavioural analysis, privacy-preserving collaborative learning and blockchain-based accountability.

Keywords: Artificial Intelligence, Blockchain, Cybersecurity, Federated Learning, Recurrent Neural Network, Scam Detection, Social Media Security.

1. Introduction

Social media now supports communication, commerce, education, professional networking and public information on a scale that makes it equally attractive to legitimate users and cybercriminals. The resulting security problem is broader than unauthorized login. Fake profiles, impersonation, phishing, malicious links and account takeover may originate from newly created fraudulent accounts or from legitimate accounts that have been compromised. This distinction matters because a system can authenticate a credential correctly and still fail to recognize that the authenticated

Kasu Journal of Computer Science Vol.3 No.4 [Dec, 2026], pp. 029-037
https://doi.org//10.47514/kjcs/2026.3.4.003
Online ISSN: 1597-2216
Print ISSN: 1597-2178
www.kjcs.edu.ng

session has become harmful. Current research approaches this problem from several directions. AI and machine-learning models examine profile attributes, content, interaction patterns and temporal behaviour that fixed rules may miss. Ensemble, deep-learning and hybrid models have reported strong results for fake-profile and fraudulent-account detection (Alharbi et al., 2024; Goyal et al., 2025; Shukla et al., 2025). Sequential models are especially relevant where risk emerges from a pattern of actions rather than a single event; for example, an unusual login followed by rapid profile changes, repeated friend requests and mass distribution of external links. The use of behavioural data for security, however, introduces a second problem. For example, the data needed to improve detection may itself be sensitive. According to McMahan et al *(*2017), Federated Learning (FL) offers one response by allowing local training and exchanging model updates rather than complete raw datasets. This reduces raw-data centralization, but it does not remove privacy or integrity risks because model updates may leak information and malicious participants may attempt poisoning or backdoor attacks. Blockchain addresses a different part of the problem. A wallet-based authentication process can verify control of a cryptographic key without transmitting the private key to the platform, while smart contracts and tamper-evident records can support selected identity, reputation and audit functions (Chen & Omote, 2023; Petcu et al**.,** 2023). These properties do not make blockchain a substitute for behavioural detection. They instead suggest a layered design in which authentication, behavioural intelligence, privacy-preserving learning and accountability perform distinct but related roles. A related KJCS study demonstrates that this complementary use of AI and blockchain extends beyond social media. Oduntan et al. (2024) reviewed their integration in modern telehealth systems and showed how AI-supported analysis can be combined with blockchain-based security, authentication and data protection. Although the application domain differs, the study reinforces a principle relevant to the present review: AI and blockchain address different security functions and can be more useful when their roles are deliberately integrated rather than treated as substitutes. Against this background, the review asks a practical question: what is gained when these technologies are considered together rather than as isolated solutions? It synthesizes recent work on social-media threats, blockchain-based identity, AI and sequential behavioural detection, FL and blockchain-enabled FL (BFL). The review also compares representative studies to identify where current approaches stop short of an end-to-end social-media security lifecycle.

2. Related Works

2.1 Social-Media Threats and Fake Accounts

Fake accounts support spam, impersonation, phishing, fraudulent advertising and coordinated manipulation. Machine-learning studies commonly distinguish such accounts using profile completeness, account age, follower relationships, posting frequency, engagement and content features. Kerrysa and Utami (2023) show that both profile and behavioural attributes recur across the literature, while Alharbi et al. (2024) demonstrate the usefulness of LSTM-based modelling for fake-account classification across Instagram and Twitter datasets. More recent work has moved toward hybrid and multimodal representations. Goyal et al. (2025) combined Random Forest, XGBoost and SMOTE to address class imbalance in Instagram fake-profile detection. Goyal et al**.** (2026) later combined textual, visual and metadata information through LSTM, CNN and ANN

Online ISSN: 1597-2216
Print ISSN: 1597-2178
www.kjcs.edu.ng

components, reporting about 97.21% accuracy, 0.9874 precision, 0.9682 recall and 0.9777 F1-score. These results are useful, but they should not be read as proof that fake-account classification solves account takeover: a previously legitimate account may become malicious without acquiring the static characteristics of a fabricated profile.

2.2 Impersonation, Phishing and Post-Authentication Risk

Impersonation and phishing exploit technical weaknesses and human trust at the same time. Attackers can imitate trusted persons or organizations, distribute malicious URLs or request money and credentials. Rathod et al. (2025) describe social engineering as an adaptive problem that requires more than static access control. For social media, this supports combining content evidence with account history and temporal behaviour, particularly after authentication has already succeeded.

2.3 Blockchain-Based Identity and Authentication

Blockchain contributes tamper-evident records, cryptographic signatures and programmable smart contracts. In a social-media setting, these properties are most useful when applied selectively to identity, authentication evidence, reputation state and audit records. Chen and Omote (2023) examined blockchain for social-network impersonation prevention, while Petcu et al. (2023) demonstrated decentralized Web3/Ethereum authentication. Both strands strengthen proof of credential control, but neither removes the need to examine what the user does after access is granted. A wallet-login flow typically issues a fresh nonce, asks the wallet to sign it and verifies that the signature corresponds to the claimed address. The private key remains with the user, and the nonce limits replay of an old signature. The limitation is conceptual as well as technical: control of a wallet proves control of a cryptographic credential, not necessarily the civil identity of the person holding it or the legitimacy of every later action.

2.4 Artificial Intelligence and Sequential Behavioural Analysis

AI-based social-media security ranges from conventional classifiers to deep, sequential and multimodal models. The attraction of RNN/LSTM models is temporal context. Ten messages sent gradually during a normal conversation and ten messages broadcast within seconds after an unusual login may have the same static count but very different security meanings. RNNs and LSTMs are designed to retain information from preceding events, making them suitable for modelling such sequences (Hochreiter & Schmidhuber, 1997). For deployment, a risk score is often more defensible than an irreversible malicious/legitimate label. A moderate score could request renewed authentication, whereas repeated high-risk behaviour could temporarily restrict sensitive actions or trigger moderator review. This keeps the model in an evidential role and reduces the operational harm of false positives.

2.5 Federated Learning and Blockchain-Enabled Federated Learning

FL changes where learning takes place rather than eliminating the need for coordination. In the Federated Averaging model, a global model is distributed, trained locally and updated through

Kasu Journal of Computer Science Vol.3 No.4 [Dec, 2026], pp. 029-037
https://doi.org//10.47514/kjcs/2026.3.4.003
Online ISSN: 1597-2216
Print ISSN: 1597-2178
www.kjcs.edu.ng

aggregation (McMahan et al., 2017). For social-media security, the attraction is clear: behavioural histories can remain closer to the user or participating organization. The difficulty is that heterogeneous data, malicious clients and information leakage from updates remain possible. BFL adds a blockchain coordination and accountability layer to this distributed-learning process. Qammar et al. (2023) and Oktian and Lee (2023) show that blockchain can support contribution records, reputation, incentives and decentralized coordination. Salim et al. (2025) are particularly relevant because BFL-SC applies blockchain, FL and smart contracts to social-media-integrated IoT security. Ying et al. (2025), De Chaudhury et al. (2024), Jain et al. (2026) and Kamran et al. (2026) further illustrate how Byzantine resilience, differential privacy, reputation or trust mechanisms can be incorporated into secure FL. Taken together, these studies also expose a boundary that is easy to overlook: a blockchain can prove that an update was submitted by a recognized participant and preserve an audit trail, but transaction validity does not establish that the update is statistically benign. BFL therefore still requires robust aggregation, update validation and privacy protection.

3. Methodology

3.1 Review Design and Scope

This study used a structured narrative review to examine the intersection of social-media cybersecurity, AI-based threat detection, blockchain identity and FL. The approach was selected because the literature spans heterogeneous application domains, datasets and technical architectures that are better compared by security function and design contribution than by pooling a single common effect size.

3.2 Search and Selection Strategy

The search was organized around combinations of the following concepts: social media security, fake or fraudulent account detection, blockchain authentication, decentralized identity, AI cybersecurity, behavioural anomaly detection, RNN/LSTM fraud detection, Federated Learning security and blockchain-enabled Federated Learning. Priority was given to peer-reviewed work published from 2021 to 2026 and directly connected to at least one of the review themes. Earlier seminal studies were retained only where they establish foundational concepts. Sources were excluded from the comparative synthesis when their primary focus had no meaningful connection to cybersecurity, identity, fraud detection, distributed learning or social-media environments.

3.3 Data Extraction and Synthesis

For each retained study, the synthesis recorded the application domain, security problem, learning technique, blockchain role, FL support, privacy mechanism, reported performance where available, and stated limitations. The evidence was then grouped into four analytical streams: AI-based threat detection, blockchain identity/authentication, privacy-oriented FL, and blockchain mechanisms for securing FL. This thematic comparison was used to identify overlap, missing links and the research gap discussed in Section 4.

Kasu Journal of Computer Science Vol.3 No.4 [Dec, 2026], pp. 029-037
https://doi.org//10.47514/kjcs/2026.3.4.003
Online ISSN: 1597-2216
Print ISSN: 1597-2178
www.kjcs.edu.ng

4. Results and Discussion

4.1 Comparative Findings

Three patterns dominate the reviewed literature. First, AI-centred studies are strongest at classification and behavioural detection but generally do not address decentralized credential ownership. Second, blockchain-centred studies strengthen identity, integrity and accountability but do not by themselves determine whether authenticated behaviour is malicious. Third, BFL research improves distributed learning and accountability but is rarely designed around the full social-media path from login to post-authentication risk assessment. Table 1 summarizes representative examples.

Table 1. Comparative synthesis of representative studies (Source: Authors' synthesis of reviewed literature)

Study	**Main focus**	**AI / sequence**	**Blockchain**	**FL**	**Main contribution / limitation**
Alharbi **et al**. (2024)	Fake accounts	LSTM	No	No	Strong sequential detection; platform/dataset dependence
Shukla **et al**. (2025)	Fraudulent accounts	Hybrid temporal DL	No	No	High-capacity hybrid detection; no decentralized identity
Chen & Omote (2023)	Impersonation	No	Yes	No	Blockchain identity evidence; limited behavioural intelligence
Petcu **et al**. (2023)	Authentication	No	Yes	No	Web3/Ethereum authentication; post-login behaviour unresolved
Salim **et al.** (2025)	SM-IoT security	Yes	Yes	Yes	Smart-contract BFL; integration/overhead complexity
Ying **et al.** (2025)	Secure FL	Yes	Yes	Yes	Incentives, Byzantine resilience and privacy; not social-media end-to-end
Jain **et al.** (2026)	Secure FL	Yes	Yes	Yes	Reputation, incentives and differential privacy; deployment overhead

Online ISSN: 1597-2216
Print ISSN: 1597-2178
www.kjcs.edu.ng

Premkumar & Nachimuthu (2026)	Decentralized social-media threats	Cross-modal graph model	No/limited	Yes	Multimodal federated detection; synthetic/experimental validation
Kamran **et al**. (2026)	IoT intrusion detection	Yes	Yes	Yes	Trust-assisted BFL; application domain differs from social media

Table 1 makes the fragmentation visible. The Artificial Intelligence studies concentrate on detection; the blockchain identity studies concentrate on authentication and impersonation; and the BFL studies concentrate on trustworthy collaborative learning. Salim et al. (2025) come closest to the combined domain because their framework explicitly addresses social-media-integrated IoT, yet its main problem remains distributed learning security rather than an end-to-end conventional social-media authentication and behavioural-monitoring lifecycle.

4.2 Recent Research Direction

The 2025–2026 literature nevertheless shows clear convergence. Hasnain et al. (2025), reviewing 158 studies, identify privacy, scalability, smart-contract deployment, platform support and trust as persistent problems in blockchain-based social media. Anter et al. (2026) similarly emphasize the need to combine blockchain with complementary security mechanisms. At the same time, Premkumar and Nachimuthu (2026) apply federated cross-modal graph transformers to decentralized social-media threat detection, while Jain et al. (2026), Kamran et al. (2026) and Yun and Liu (2026) advance secure or asynchronous BFL designs. The direction of travel is therefore from isolated mechanisms toward layered systems.

4.3 Identified Research Gap

The central research gap is not the absence of AI, blockchain or FL techniques individually. It is the limited empirical evaluation of architectures that connect them across the same social-media security lifecycle. A user can pass authentication and later behave abnormally; an AI detector can flag that behaviour without establishing decentralized credential ownership; FL can reduce raw-data centralization without guaranteeing trustworthy coordination; and blockchain can provide accountability without judging whether a model update is malicious. The reviewed studies suggest that an integrated social-media security architecture should operate across successive but interconnected security functions. Identity establishment and authentication provide the initial verification of users, while behavioural monitoring extends security beyond the point of login. Suspicious activity identified through sequential AI models can subsequently inform threat assessment, while Federated Learning enables participating nodes to improve detection models without routinely centralizing raw behavioural data. Blockchain can provide verifiable identity records, integrity and accountability across these processes, with the resulting evidence supporting proportionate security responses such as additional authentication, moderation or temporary restrictions.

Kasu Journal of Computer Science Vol.3 No.4 [Dec, 2026], pp. 029-037
https://doi.org//10.47514/kjcs/2026.3.4.003
Online ISSN: 1597-2216
Print ISSN: 1597-2178
www.kjcs.edu.ng

4.4 Challenges and Open Issues

Scalability is the first practical constraint. Social-media platforms process high-frequency events, while blockchain consensus and FL both add communication and computation. The literature therefore supports selective rather than indiscriminate blockchain use. For example, large models, private messages and detailed behavioural data should remain off-chain, while compact hashes, validation outcomes or carefully designed reputation states may be recorded on-chain (Hasnain et al., 2025; Anter et al., 2026). A second constraint is trust in collaborative learning. Poisoning, backdoor, Sybil and free-riding attacks remain possible even when contributions are recorded immutably. BFL designs consequently need robust aggregation and update validation alongside blockchain accountability (Ying et al., 2025; De Chaudhury et al., 2024). Privacy also remains incomplete because gradients or model parameters can reveal information, motivating differential privacy and secure aggregation. Operational deployment adds further concerns such as concept drift, false positives, smart-contract vulnerabilities, wallet-key recovery and interoperability. These issues argue for risk-based rather than purely automatic enforcement. Consequential decisions should permit human review or appeal, and wallet-based identity should be paired with a recovery design that does not simply recreate the weaknesses of centralized password reset.

5. Conclusion

This review considered AI, blockchain, sequential behavioural modelling and FL as complementary security mechanisms rather than competing technologies. Across the literature, AI contributes adaptive fraud and anomaly detection; blockchain strengthens cryptographic credential verification, integrity and accountability; RNN/LSTM models add temporal context; and FL reduces the need to centralize raw behavioural data. Recent BFL studies show that these capabilities can be coordinated, but they also introduce new costs and attack surfaces. The main conclusion is that stronger social-media security is likely to come from carefully composed layers rather than from any single technology. The unresolved research problem is to demonstrate, under realistic conditions, an end-to-end architecture that links wallet-based authentication with post-authentication behavioural assessment, privacy-preserving collaborative learning and accountable security response. Future studies should therefore prioritize empirical validation, explainability, robust aggregation, privacy protection, scalability and user-centred recovery mechanisms. Future work should move beyond adding technologies side by side and test how they interact under realistic attack conditions. Personalized federated behavioural models could combine global scam patterns with user-specific baselines; multimodal and graph-based models could integrate text, images, metadata and relationships; and explainable AI could show which behavioural signals contributed to a security decision. The multimodal results of Goyal et al. (2026) and the federated graph approach of Premkumar and Nachimuthu (2026) indicate the value of these directions. Research should also examine asynchronous BFL where participants have unequal connectivity and computing resources (Yun & Liu, 2026), together with Byzantine-robust aggregation and hybrid on-chain/off-chain storage. Evaluation should report not only accuracy but also precision, recall, F1-score, false-positive rate, latency, communication overhead, convergence behaviour and blockchain cost. Those measures are necessary to judge whether an integrated architecture is operationally useful rather than merely technically feasible.

Online ISSN: 1597-2216
Print ISSN: 1597-2178
www.kjcs.edu.ng

References

Akhtar, N., & Feng, C. (2021). An overview of the applications of artificial intelligence in cybersecurity. EAI Endorsed Transactions on Security and Safety.

Alharbi, N., Alkalifah, B., Alqarawi, G., & Rassam, M. A. (2024). Countering social media cybercrime using deep learning: Instagram fake accounts detection. Future Internet, 16(10), 367. https://doi.org/10.3390/fi16100367

Anter, F., Elmendili, F., Benamar, N., Mrani, N., & Fattah, M. (2026). A survey of blockchain-based solutions for secure social networks. Scientific African, 31, e03233. https://doi.org/10.1016/j.sciaf.2026.e03233

Chen, Z., & Omote, K. (2023). Preventing SNS impersonation: A blockchain-based approach. IEICE Transactions on Information and Systems, E106-D(9), 1354–1363. https://doi.org/10.1587/transinf.2022ICP0003

De Chaudhury, S., Morreddigari, L. R., Varun, M., Sengupta, T., Chakraborty, S., Sural, S., Vaidya, J., & Atluri, V. (2024). Blockchain based secure federated learning with local differential privacy and incentivization. IEEE Transactions on Privacy, 1, 31–44. https://doi.org/10.1109/TP.2024.3487819

Goyal, B., Gill, N. S., Gulia, P., Alduaiji, N., Shukla, P. K., et al. (2025). Instagram fake profile detection using an ensemble learning method. Scientific Reports, 15, 26464. https://doi.org/10.1038/s41598-025-03973-x

Goyal, B., Gill, N. S., Gulia, P., Alduaiji, N., Shukla, P. K., & Vishnoi, S. (2026). Deep learning with optimization for social media fake profile detection. Social Network Analysis and Mining, 16, 71. https://doi.org/10.1007/s13278-026-01586-4

Hasnain, M., Ghani, I., Smith, D., Daud, A., & Jeong, S. R. (2025). Cybersecurity challenges in blockchain-based social media networks: A comprehensive review. Blockchain: Research and Applications, 6(3), 100290. https://doi.org/10.1016/j.bcra.2025.100290

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8

Jain, C., Chaudhari, P., & Das, M. L. (2026). A blockchain-infused framework for secure federated learning. Discover Computing, 29, 265. https://doi.org/10.1007/s10791-026-10139-4

Kamran, M., Akhtar, S. M., Gilani, A., Alhashmi, A. A., Kanwal, S., Darem, A. A., & Alofairi, A. A. (2026). A blockchain-assisted secure federated learning architecture for intrusion detection in Internet of Things networks. Scientific Reports, 16, 26072. https://doi.org/10.1038/s41598-026-53053-x

Kerrysa, N. G., & Utami, I. Q. (2023). Fake account detection in social media using machine learning methods: Literature review. Bulletin of Electrical Engineering and Informatics, 12(6), 3790–3797. https://doi.org/10.11591/eei.v12i6.5334

Lim, S. Y., Musa, O. B., Al-Rimy, B. A. S., & Almasri, A. (2022). Trust models for blockchain-based self-sovereign identity management: A survey and research directions. In Y. Maleh, L. Tawalbeh, S. Motahhir, & A. S. Hafid (Eds.), Advances in blockchain technology for cyber physical systems (pp. 277–302). Springer. https://doi.org/10.1007/978-3-030-93646-4_13

McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, 54, 1273–1282.

Online ISSN: 1597-2216
Print ISSN: 1597-2178
www.kjcs.edu.ng

Oktian, Y. E., & Lee, S.-G. (2023). Blockchain-based federated learning system: A survey on design choices. Sensors, 23(12), 5658. https://doi.org/10.3390/s23125658

Oduntan, O. E., Bature, H. J., Wisdom, D. D., Okechukwu, U. C., Isaac, S., & Falayi, C. F. (2024). Artificial Intelligence & Blockchain Technology in Modern Telehealth Systems-A Systematic Review. KASU Journal of Computer Science, 1(3), 607–625. https://doi.org/10.47514/kjcs/2024.1.3.0015

Petcu, A., Pahontu, B., Frunzete, M., & Stoichescu, D. A. (2023). A secure and decentralized authentication mechanism based on Web 3.0 and Ethereum blockchain technology. Applied Sciences, 13(4), 2231. https://doi.org/10.3390/app13042231

Premkumar, D., & Nachimuthu, S. K. (2026). Privacy-preserving cyberthreat detection in decentralized social media with federated cross-modal graph transformers. Scientific Reports, 16, 3608. https://doi.org/10.1038/s41598-025-33596-1

Qammar, A., Karim, A., Ning, H., & Ding, J. (2023). Securing federated learning with blockchain: A systematic literature review. Artificial Intelligence Review, 56, 3951–3985. https://doi.org/10.1007/s10462-022-10271-9

Rathod, T., Jadav, N. K., Tanwar, S., Alabdulatif, A., Garg, D., & Singh, A. (2025). A comprehensive survey on social engineering attacks, countermeasures, case study, and research challenges. Information Processing & Management, 62(1), 103928. https://doi.org/10.1016/j.ipm.2024.103928

Salim, S., Moustafa, N., & Turnbull, B. (2025). BFL-SC: A blockchain-enabled federated learning framework, with smart contracts, for securing social media-integrated Internet of Things systems. Ad Hoc Networks, 169, 103760. https://doi.org/10.1016/j.adhoc.2025.103760

Shukla, P. K., Veerasamy, B. D., Alduaiji, N., Addula, S. R., Pandey, A., et al. (2025). Fraudulent account detection in social media using hybrid deep transformer model and hyperparameter optimization. Scientific Reports, 15, 38447. https://doi.org/10.1038/s41598-025-24326-8

Ying, C., Xia, F., Wei, D. S. L., Yu, X., Xu, Y., Zhang, W., Jiang, X., Jin, H., Luo, Y., Zhang, T., & Tao, D. (2025). BIT-FL: Blockchain-enabled incentivized and secure federated learning framework. IEEE Transactions on Mobile Computing, 24(2), 1212–1229. https://doi.org/10.1109/TMC.2024.3477616

Yun, J., & Liu, T. (2026). BCAFL: A secure and efficient blockchain framework for asynchronous federated learning. Scientific Reports, 16, 27584. https://doi.org/10.1038/s41598-026-57780-z