PUXplore: Multidisciplinary Journal of Engineering , Vol 2 , Issue 1, Jan - Mar 2026 ||
Online (ISSN 3108-2106)

PUXplore Multidisciplinary Journal of Engineering

Parul® University NAAC GRADE A++

A Review of Security, Privacy, and Authentication Mechanisms in Social Media Web Applications

Dr. Ashwini Kumar Jha[1], Akil Khatri[2], Kavi Kanda[3], Areeb Haider[4], Raunak Shah[5]

[1]*Associate Professor, AI-DS Department, Parul University, India*

[2&3&4&5]*Student, B.Tech AI, Parul University, India*

E-mail: [1]ashwini.jha34918@paruluniversity.ac.in, [2]2303031240529@paruluniversity.ac.in,
[3]2303031240523@paruluniversity.ac.in, [4]2303031240220@paruluniversity.ac.in, [5]2303031241132@paruluniversity.ac.in

ORCID: [1]https://orcid.org/0000-0002-6607-2168

Abstract - Social media was never supposed to become this complicated. What started as a way to stay connected has turned into one of the most contested spaces in modern life where fake identities thrive, personal data gets harvested in ways users never agreed to, and the systems meant to verify who we are can barely keep up. Researchers have thrown a lot at these problems over the years, and there has been genuine progress. But an uncomfortable pattern keeps repeating itself, solutions that look great on paper tend to wobble once they hit the real world. The privacy paradox still has no clean answer. Detection models still break when moved between platforms. Authentication still trades security for convenience in ways nobody is fully happy with. Until the technical side of this work starts taking human behaviour and policy seriously not as afterthoughts but as core design constraints, the same gaps will keep reappearing.

Keywords: Social Media Security; Social Media Privacy; Social Media Authentication; Fake Account Detection; Bot Detection; Phishing Attacks; Machine Learning in Cybersecurity; Deep Learning for Social Networks; Privacy Paradox; AI-driven Privacy Risks; Inference Attacks; Data Governance; Multi-Factor Authentication (MFA); Single Sign-On (SSO); OAuth 2.0; PKCE; Identity Verification; Socio-Technical Security Frameworks; Cross-Platform Security Evaluation; User Behavior in Cybersecurity

I. INTRODUCTION

Nobody mapped out the security implications of social media before building it. The platforms scaled fast, the users came faster, and the problems arrived almost immediately. Early researchers did what they could documenting the threats, naming the patterns, trying to build a shared vocabulary for something genuinely new (Manoj, 2021). It was useful foundational work, but it was mostly observational. It described what was happening without offering much to actually stop it.

That changed as the attacks got worse. Fake account networks became more coordinated, phishing grew more convincing, and the old manual approaches to detection stopped scaling. The research community responded by building automated tools like Random Forest and Naïve Bayes classifiers that could flag suspicious accounts with surprising accuracy [1]. Deep learning followed, bringing the ability to catch behavioral signals too subtle for earlier models [2]. Layering in natural language processing made things more robust still [3]. These were genuine advances. The frustrating part is that nearly all of them were built on data from one platform and quietly struggled the moment anyone tried applying them somewhere else.

Privacy research has wrestled with its own version of this problem. The central question for a long time was deceptively simple: why do people who say they care about privacy keep sharing so freely? That tension, the privacy paradox, generated enormous amounts of research and still does not have a satisfying resolution [4]. What the field has gradually accepted is that privacy is not a purely individual choice. How people think about and protect their information is shaped by culture, race, and social context in ways that matter enormously for how any solution gets designed [5]. And while researchers were working through those questions, AI was quietly making the problem worse like pulling sensitive inferences from photos, location data, and behavioral traces that users had no reason to think of as revealing [6], [7]. Regulation has tried to respond, but it has mostly been playing catch-up with technology that moves faster than legislation can [8].

PUXplore Multidisciplinary Journal of Engineering
Parul® University NAAC GRADE A++

Authentication has probably changed more visibly than either of the other two areas. The slow death of the password has been underway for years, replaced gradually by Single Sign-On, Multi-Factor Authentication, and federated identity frameworks built on protocols like OAuth 2.0 and PKCE [9], [10]. MFA in particular has earned its reputation as it genuinely reduces unauthorised access in ways that are hard to argue with [11]. But it also introduced new problems nobody fully anticipated. The friction is real, and users find ways around it. Phishing attacks have adapted to specifically target MFA flows. Rollout across platforms has been inconsistent enough to undermine the gains [12], [13]. And tighter identity verification carries a cost that rarely gets discussed honestly like it can quietly exclude the people least able to meet its demands, while raising legitimate questions about surveillance [14].

The honest read across all three areas is that the research tends to treat these as separate technical problems, when the thing limiting progress in each of them is usually something human like behavior, incentives, institutional failure, or the gap between how systems are designed and how people actually use them.

II. METHODOLOGY

A. *Research Design*

The approach taken in this study was always going to be a literature review, the research questions didn't call for primary data collection, and there's already a reasonable body of work out there on security, privacy, and authentication in social media that was worth pulling together and making sense of. The review was kept deliberately focused on those three areas. Honestly, there were moments where it felt like the scope could easily creep outward, algorithmic accountability kept coming up in searches, and data ethics is hard to ignore entirely, but pulling on those threads would have turned this into a very different piece of work. So the decision was made early on to hold the line.

On the PRISMA question, it came up during the planning stage and was considered, but it didn't feel like the right fit. PRISMA works well when you're dealing with clinical trials or health interventions where every excluded paper needs a documented reason. This review isn't that kind of project, and forcing that structure onto it would have been more performance than substance. A structured keyword search with careful manual screening did the job without the unnecessary overhead.

B. *Search Strategy*

Google Scholar was used as the main search tool throughout. Part of that was practical as it's accessible and indexes a genuinely wide range of academic material, from journal articles and conference papers to theses and edited volumes. For a topic that cuts across computer science, law, and social science the way this one does, that breadth mattered. To keep results manageable and actually relevant, Boolean operators and quotation marks were used in every search. Dropping those constraints in early test searches produced results that were all over the place (loosely related at best) so the structure was necessary.

The searches themselves were split across three areas. For security, terms like "social media platform security" AND "web application vulnerabilities" came up alongside "fake accounts" AND "social networks", "phishing attacks" AND "social networking sites", "social media bots" AND "security threats", "malware propagation" AND "social media", "hate speech detection" AND "social media security", and "cyberbullying" AND "social network analysis". Privacy searches drew on strings like "personal data collection" AND "social media platforms", "user tracking" AND "social networking sites", "GDPR" AND "social media compliance", "third party data sharing" AND "social media", "privacy settings" AND "social networking sites", and "right to be forgotten" AND "online platforms".For authentication, searches covered "user authentication" AND "social media", "two factor authentication" AND "social networks", "multi factor authentication" AND "social media", "OAuth 2.0" AND "security vulnerabilities", "single sign on" AND "social media platforms", and "online anonymity" AND "social media".

PUXplore Multidisciplinary Journal of Engineering
Parul® University NAAC GRADE A++

Only the first few pages of results were looked at for each query. Going further than that rarely turned up anything new, past a certain point it's mostly tangential work that had been ranked low for a reason.

C. *Inclusion Criteria*

Peer review was the baseline, if a paper hadn't gone through that process, it wasn't considered regardless of how relevant it looked. Conference proceedings from established venues were included alongside journal articles. Everything had to be in English, and had to be engaging with social media or online social networking in a direct way, not just referencing it as background context. Papers where security, privacy, or authentication was a side note rather than a central concern were dropped. The fifteen-year window covered most of the relevant literature comfortably, though a handful of older papers were kept because the field kept citing them, it made sense to go back to those directly rather than rely on second-hand accounts of what they said.

D. *Exclusion Criteria*

Opinion pieces, editorials, and anything blog-adjacent were out from the start. The same went for industry reports and white papers that hadn't been peer reviewed, they might be useful in a different kind of project, but not here. Papers that were only loosely connected to social media, or that were really about something else entirely and just mentioned social media in passing, were also removed. Any duplicates that appeared across different searches were caught and filtered out during screening.

E. *Screening Process*

Screening happened across three passes. The first was just titles, anything obviously off-topic got dropped quickly, which cleared out a significant chunk of the initial results without much deliberation. Abstract screening was where the bulk of the decisions got made. A lot of papers that looked fine by title turned out to be only loosely connected to the topic once the abstract was read properly, they might mention social media once or frame authentication as a minor variable in a much broader study. Those got cut. If a paper's abstract left any doubt about whether it was genuinely dealing with one of the three domains, it didn't go through to full-text review.

The final stage was reading the remaining papers properly. Some looked promising in the abstract but turned out to be only tangentially relevant once read in full, or their methodology was too weak to justify inclusion. Only papers that had something genuinely useful to contribute like technically, empirically, legally, or analytically were kept.

F. *Data Extraction and Thematic Classification*

Once a paper was confirmed as relevant, the same information was pulled from each one: who wrote it and when, what research question it was working with, what method it used, what it actually found, and which domain it sat in. Keeping that consistent across every paper made the comparison stage far less painful, patterns started becoming visible once everything was laid out in the same format.

Papers were then grouped into more specific sub-themes within each domain. Security covered infrastructure protection, bots and fake accounts, malware and phishing, and behavioural threats like harassment and hate speech. Privacy broke down into data collection practices, third-party data sharing, user-facing privacy controls, and regulatory compliance, particularly around GDPR. Authentication covered identity verification approaches, multi-factor and two-factor systems, single sign-on mechanisms, and questions around online anonymity. Organising things this way made the write-up much more coherent and helped bring out where research is clustered and, just as importantly, where it isn't.

G. *Limitations*

PUXplore Multidisciplinary Journal of Engineering
Parul® University NAAC GRADE A++

A few honest admissions here. Sticking to Google Scholar meant that some work published in databases like Scopus or ACM Digital Library might have been missed, that's a real gap, even if it's a manageable one. Not using PRISMA means someone trying to replicate this review exactly would have some gaps to fill in. Manual screening, however carefully done, inevitably involves judgement calls that another reviewer might make differently.

There's also a citation bias baked into the search method, papers that are newer simply haven't had the time to build up the kind of citation record that pushes them up the rankings, which means some recent work probably slipped through the net. That's a real limitation, not just a theoretical one. But it's the kind of trade-off that comes with any relevance-sorted search, and acknowledging it feels more useful than pretending it didn't affect the results.

H. *Summary*

Looking back at the process, the method was never going to be perfect, literature reviews rarely are, but the decisions made at each stage were deliberate and the reasoning behind them was sound. The search process was systematic enough to be defensible, the screening criteria were clear, and the thematic classification gave the review a shape that made the findings easier to discuss. The aim throughout was to cover the three domains thoroughly enough to say something meaningful — not just to compile a long list of papers and call it a review.

III. SECURITY

The rise of social media platforms has significantly changed how we communicate digitally (See Fig. 1.). These platforms facilitate not only the rapid distribution of information, social interaction, and economic activities on an unprecedented scale but also create a complex security challenge where bad actors take advantage of the open nature of social media platforms, trust, and amplification mechanisms. The literature studied for this paper clearly indicates that social media security has been a subject of intense scholarly interest, with a clear call for addressing various security threats, including fake accounts, bots, phishing, identity deception, and data misuse, among many others. Gaps still exist despite significant advances in methodology, especially through machine learning methods.

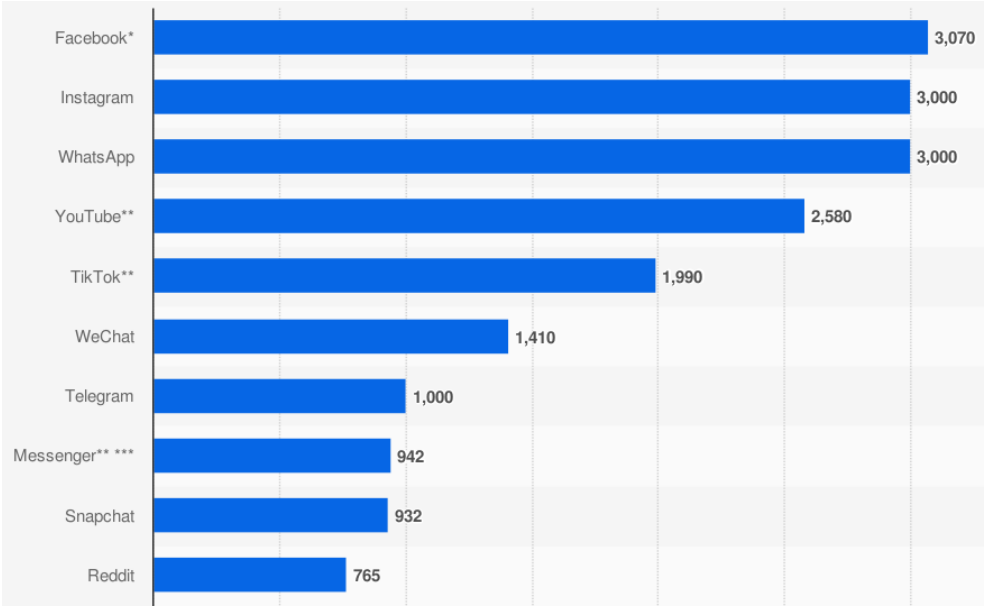

Fig. 1. Most popular social networks worldwide as of February 2025, by number of monthly active users (in millions) [15]

Early research in social media security has primarily been conceptual and descriptive. Target of these studies was to define the threat landscape and raise awareness of growing cyber risks. These studies have identified identity deception, privacy breaches, malware propagation, and unauthorized data exploitation as fundamental challenges

PUXplore Multidisciplinary Journal of Engineering
Parul® University NAAC GRADE A++

faced by both users and platforms [16], [17].

Such work has played an important foundational role by setting the context for social media threats within broader cybersecurity discussion and the need to put an emphasis on user education, regulatory oversight, and platform responsibility. Although, it is important mention that these studies typically lack empirical validation, automated detection mechanisms, and measurable performance indicators, which limits their utility for operational security systems.

These has been a shift toward data-driven and algorithmic approaches in recent research due to the increasing scale and automation of social media threats. Machine learning is increasingly used to detect fake accounts and malicious behaviour, with supervised classification models trained on profile attributes, activity patterns, and textual content. High detection accuracy is reported in studies using traditional classifiers such as Random Forest and Naïve Bayes, especially for straightforward fake account behaviours [1]. Interpretability and relatively low computational cost of these approaches make them more attractive as that makes them suitable for initial filtering stages in security pipelines.

Deep learning architectures including convolutional and deep neural network are being used in more advanced work to capture complex, non-linear relationships within social media data [18]. Deep models are shown to have impressive performance even when they are trained on minimal profile information, this indicates their ability to infer hidden indicators of deception which are not easily captured by handcrafted features [2]. Similarly, hybrid ensemble models that integrate profile-based features with natural language processing techniques improve robustness in bot detection tasks which reflects a growing trend towards multi-feature fusion [3]. Put together these studies show an improvement in algorithmic defence against social media threats.

Even after these methodological advances, the literature has shown a noticeable platform-centric bias. Most studies have tested their models using data from a single social media platform (most commonly Facebook, Twitter, or VK) while not investigating if the same model can be used on different platforms. This limitation plays critical role, as bad actors often adapt their strategies across platforms with different affordances, user demographics, and moderation policies. Models which are optimized for one platform may fail when used elsewhere, this reflects broader machine learning generalization challenges observed in other high-risk domains [19]. widespread use of custom, non-public datasets worsen the problem, which makes it harder to reproduce the results and prevents meaningful benchmarking across studies [2].

The detection of advanced and evolving threats creates another significant challenge. While reported accuracy of many models is often high, closer examination reveals that many models struggle with sophisticated fake accounts which are designed to mimic legitimate user behaviour over extended periods. For example, ARU-type fake accounts show low activity levels and realistic interaction patterns, multiple classifiers struggle to detect such patterns [1]. This shows a broader limitation of static detection approaches that rely on snapshot-based features rather than analyzing behaviour over a long period of time. As bad actors are increasingly using human-assisted automation and adaptive strategies, current detection models risk rapid obsolescence.

Feature selection trends show that there is an over-reliance on content and profile metadata, this data can easily be manipulated by attackers. Although such features are accessible and widely available, they provide limited protection against bad actors who deliberately create realistic profiles and content. Despite the potential of behavioural, temporal, and network-level features (such as interaction dynamics, community structure, and temporal posting patterns) to offer stronger resilience against deception, such features remain underexplored. Advanced relational modeling approaches such as spatiotemporal graph networks show how interaction dynamics can be captured more holistically in other domains [20], yet similar architectures remain underutilized in social media threat detection. This gap shows the need for holistic modelling approaches that move beyond surface-level indicators.

Research on phishing and social engineering threats exposes additional vulnerabilities in current security strategies. Rule-based phishing detection systems and keyword-driven algorithms are common, yet they struggle with short-form language, multilingual expressions, and code-mixed text common in social media communication. Despite

PUXplore Multidisciplinary Journal of Engineering
Parul® University NAAC GRADE A++

growing use of images, videos, and audio to deceive users, most phishing detection approaches focus exclusively on textual content. This exposes a blind spot in contemporary social media security research where multimodal and multilingual attacks are under-researched.

Operational considerations such as scalability and real-time deployment receive limited attention across the reviewed literature. While many studies demonstrate strong offline performance, few evaluate computational efficiency, latency, or system integration challenges. Practical constraints (such as API rate limits, data access restrictions, and platform-specific policies) are acknowledged but rarely incorporated into model design or evaluation [1], [21]. As a result, there remains a significant gap between experimental performance and deployable, large-scale security solutions. Edge–fog architectures have been explored in other domains such as healthcare to support latency-sensitive applications [22], yet similar architectural thinking is rarely incorporated into social media threat detection systems.

Finally, the reviewed studies reveal a narrow view of authentication and identity security. Most approaches treat authentication as a binary classification task (distinguishing real from fake accounts at a single point in time) rather than as a continuous, lifecycle-based process. This perspective fails to account for gradual identity evolution, long-term infiltration strategies, and delayed malicious activation. Moreover, technical solutions are often developed in isolation from broader socio-technical considerations, such as user trust, platform governance, legal frameworks, and ethical implications. Conceptual works emphasize these dimensions [17], but integration with algorithmic systems remains limited.

In summary, the social media security literature reflects a clear evolution from conceptual threat awareness to sophisticated machine learning-based detection techniques. While current approaches achieve high accuracy under controlled conditions, they remain constrained by platform specificity, limited datasets, static modeling assumptions, and insufficient consideration of adversarial adaptation, multimodal content, and real-world deployment. Addressing these challenges will require interdisciplinary approaches that combine advanced AI techniques with behavioral analysis, cross-platform evaluation, and policy-aware system design to ensure robust and sustainable social media security.

Table. 1. Comparison table of research on Social Media Security

Author(s)	Year	Method	Accuracy	Key Limitations
Manoj [17]	2021	Conceptual and descriptive analysis of cyber threats (identity deception, malware, data theft) on social media platforms; review-based discussion without automated detection	Not reported	No empirical dataset; no machine learning or automated detection model; lacks quantitative evaluation, benchmarks, and performance metrics; largely advisory
Frunze & Frolov [1]	2021	Feature-based supervised ML using profile and activity features; classifiers include Random Forest and Gaussian Naïve Bayes on VK social network data	**97%** (Random Forest)	Poor detection of advanced fake accounts (ARU-type); high false positives for Naïve Bayes; dataset limited by VK API constraints; weak cross-platform generalization
Amankeldin et al. [2]	2023	Deep Neural Network (CNN/DNN) using minimal profile and content-based features for fake Facebook	Up to **99.4%** (avg. ~99–	Custom, non-public dataset limits reproducibility; excludes behavioral and network features; scalability and

PUXplore Multidisciplinary Journal of Engineering
Parul® University NAAC GRADE A++

		profile detection	99.8%)	adversarial robustness not evaluated
Talha [3]	2024	Hybrid ensemble ML model combining profile-based, content-based, and NLP features for Twitter bot detection	Not explicitly reported (improved over baselines)	Focused only on Twitter; lacks cross-platform validation; real-time deployment and scalability issues not addressed

IV. PRIVACY

Social media privacy research had gone way more beyond simple studying about why users had shared their personal information,over the years. Earlier the studies were majorly focused on individual behaviour and awareness. Nowadays,privacy is considered a complex social technical issue caused by user psychology, platform algorithms, regulatory systems, AI and even emerging digital environments. The literature has clearly shown that the privacy on social media is no longer a single issue of "protecting data", but it's a multi-layered challenge influenced by both tech design and governance structures [8], [23].

A. *Behavioral Trends and the Privacy Paradox*

The most common observation in privacy research is the so-called privacy paradox. Many users show their concern towards privacy but yet they continue to share their personal information while harming the privacy on social media and other platforms. As per Gruzd and Hernández-García (2024) had shown how users usually prioritize social connection, visibility, and maintaining relationships over potential privacy risks. In other words, being actively online is more significant for them rather than the risk of their data getting misused.

At the same time, privacy behaviour is not similar for all users. Wang and Metzger (2024) had shown that privacy management practices vary among different racial and ethnic groups due to structural inequalities and their experiences of discrimination (See Fig. 2). This shows that privacy decisions are not only based on individual thinking about risk but are also affected by social and cultural factors. However, most of the existing studies are based on surveys and cross-sectional data. Because of this, it becomes difficult to understand how privacy behaviours change over time or how changes in social media platforms affect users' decisions [4], [5].

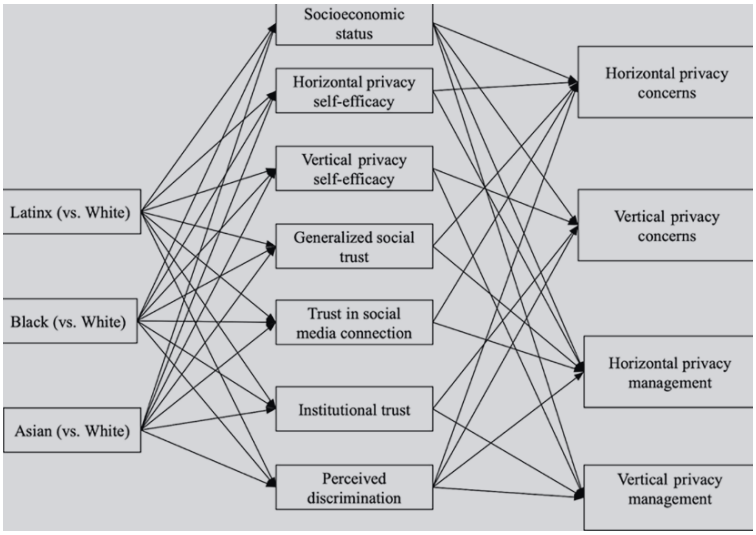

Fig. 2. Racial/ethnic divides in social media privacy concerns and privacy management behaviours.

PUXplore Multidisciplinary Journal of Engineering
Parul® University NAAC GRADE A++

B. *AI-Driven Privacy Risks and Inference Attacks*

As social media platforms are using more artificial intelligence, privacy risks have become more complex. Now threats are not only about directly sharing data but also through inference. Cheng et al. (2022) explain that images shared on social media can reveal more information than users intend, including their location, facial identity, and other personal details. Even simple or harmless posts may lead to privacy leakage when they are processed by advanced AI systems.

To handle this issue, Liu et al. (2022) suggest AI-based systems that can detect privacy risks in social networks. But even though these systems are useful, they have problems in managing large amounts of data and working properly across different platforms. As AI is developing very fast, privacy protection methods are not able to keep up, which creates a gap between new threats and security solutions [6], [7].

C. *Summary*

People keep sharing personal info on social media even if they worry about privacy [4]. Privacy get more tricky as AI can also get hidden info from pictures and what people do online [6], [7]. Blockchain and other solutions exist, but they are not fully practical or easy to use yet [23], [24].

Table. 2. Comparison table of research on Social Media Privacy

Author(s)	Year	Method	Accuracy / Key Findings	Limitations
Cheng et al. [6]	2022	Comprehensive survey on image privacy in Online Social Networks (OSNs); taxonomy-based analytical review	Identified privacy leakage risks in image sharing (face recognition, location inference, metadata extraction); proposed privacy intelligence frameworks for automated protection.	Review-based study; lacks large-scale empirical validation; rapidly evolving AI techniques may outdate taxonomy.
Gruzd & Hernández-García [4]	2024	Empirical survey study testing privacy paradox and Information Privacy Management (IPM) strategies (N ≈ national sample)	Confirmed privacy paradox; benefits of social media drive self-disclosure more than privacy concerns; IPM strategies and literacy influence behavior.	Single-country (Canada) context; self-reported data; cross-sectional design.
Jain et al. [25]	2021	Comprehensive survey/review of OSN security and privacy threats; comparative analysis of defense mechanisms	Identified major OSN threats and gaps between proposed solutions and real-world deployment; emphasized hybrid frameworks and improved detection.	Survey-based; no experimental validation; limited depth in specific mechanisms.
Kumar et al. [24]	2023	Blockchain-based hybrid framework using Ethereum smart contracts (SCSGI, SCSTI) + Graph Attention Network (BC-GAT)	Demonstrated improved confidentiality, anonymity, and malicious node detection; enhanced trust via blockchain transparency.	Dataset accessibility limitations; implementation complexity; scalability and real-world deployment challenges.
Liu et al. [26]	2022	AI/ML-based privacy leakage detection model; automated risk assessment framework	Achieved improved detection of sensitive attribute inference and privacy risks using deep learning techniques.	Model performance dataset-dependent; computational overhead; limited cross-platform validation.

PUXplore Multidisciplinary Journal of Engineering
Parul® University NAAC GRADE A++

Wang & Metzger [5]	2024	Large-scale survey (N = 1,401) testing resource-based and identity-based explanations for online privacy divide	Found racial/ethnic differences in privacy concerns and management behaviors; resource inequality and perceived discrimination explain variations.	U.S.-focused sample; self-reported measures; cross-sectional design limits causal inference.
Yeung et al. [8]	2023	Policy and regulatory analysis of digital privacy governance frameworks	Identified structural gaps in platform accountability and data governance; emphasized need for adaptive regulatory mechanisms.	Normative/legal analysis; lacks technical evaluation; regulatory impact not empirically measured.

V. AUTHENTICATION

A. *Evolution of Authentication on Social Media*

Authentication on social media has evolved from basic password-based systems focused on strength and attack prevention to layered, identity-centric approaches. Early emphasis was on complex passwords to counter simple guessing or brute-force attacks. Modern research prioritizes integrating Single Sign-On (SSO) with Multi-Factor Authentication (MFA) for centralized yet layered verification and better risk management. MFA adoption has shown clear reductions in unauthorized access, with principles transferable from banking to social platforms. OAuth 2.0 remains dominant but is strengthened via PKCE and encrypted tokens to reduce theft, reflecting incremental improvements rather than radical changes. The usability-security balance persists as a core challenge. [9], [10], [11], [27]

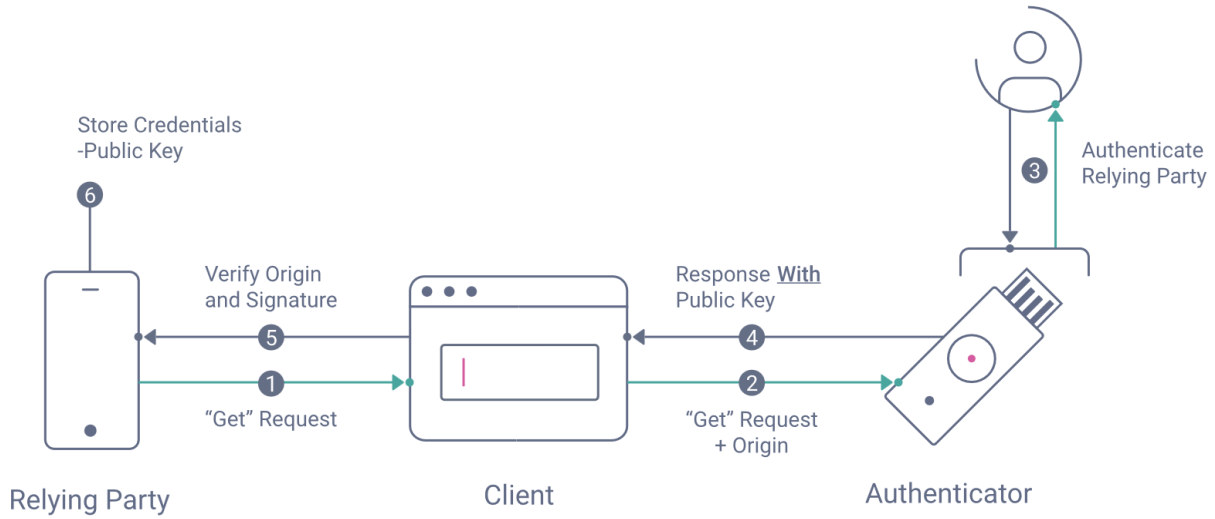

Fig. 3. W3C, Web Authentication: An API for accessing Public Key Credentials Level 3, W3C Recommendation, 2026 [28]

B. *Security vs. Usability Trade-off*

Stronger mechanisms like enhanced OAuth flows and MFA significantly reduce exploits such as token misuse. Additional layers increase technical resistance but introduce user friction through extra steps, impacting engagement. Frequent SSO users show higher vulnerability to phishing via spoofed prompts due to familiarity lowering caution. Awareness gaps exist, with groups like students exhibiting lower SSO risk perception than faculty. Technical fixes

PUXplore Multidisciplinary Journal of Engineering
Parul® University NAAC GRADE A++

close protocol holes, yet human behavior and interface cues heavily influence real-world success. The security-usability tension remains unresolved. [10], [12], [13]

C. *Lingering Weak Spots in OAuth and SSO*

OAuth 2.0 is the standard for social platforms, but vulnerabilities arise mainly from inconsistent implementations. Safeguards like PKCE and secure token storage are essential yet variably adopted, leading to differing security levels across platforms. Enterprise SSO-MFA integrations accelerate, but configuration and update inconsistencies cause fragmentation. Clear verification signals boost user trust significantly, though over-reliance amplifies risks if flaws occur. Security depends more on developer discipline and consistent enforcement than protocol design alone. [9], [10], [29]

D. *Human Behavior and Psychology in Authentication*

Technical controls validate credentials effectively but often ignore user psychology. SSO risk awareness varies widely; students score lower than faculty despite tech familiarity. Perceived anonymity correlates with increased aggression and negative online behaviors, especially under privacy concerns. Systems rarely design for user interpretations of visibility or accountability, allowing social factors to undermine technical strengths. Effective security requires balancing robust design with understanding human nature. [13], [30], [31]

E. *Identity Verification: Trust and Ethical Concerns*

Stronger verification boosts trust via credibility cues, encouraging interactions and potentially reducing anonymity-driven harms like misinformation. However, it raises ethical issues including privacy erosion, surveillance risks, group exclusion, and reduced free expression. Research shows benefits but highlights trade-offs without a unified framework reconciling trust gains with privacy and inclusion. Technical and ethical aspects are often discussed separately, leaving a design gap. [14], [29]

F. *Core Issue: Inconsistent Implementation*

The main challenge is uneven application of proven technologies rather than lack of tools. SSO and MFA adoption grows, yet rollout quality varies greatly across platforms and organizations. PKCE and secure token handling reduce risks substantially when done correctly, but inconsistencies fragment the ecosystem and hinder interoperability. Standardized, disciplined implementation is now the priority for reliable protection. [9], [10]

G. *Research Gaps and Methodological Limits*

Many studies use controlled experiments, narrow samples, or simulations, limiting real-world applicability. Lab setups and single-institution surveys provide strong stats but miss adversarial social media dynamics. Self-reported data on behavior and awareness lacks long-term observation. Technical papers often remain conceptual or simulation-focused without large-scale validation. The field needs more longitudinal, cross-platform, and ecosystem-level adversarial testing for better generalizability. [9], [10], [13], [29], [30], [31]

H. *Remaining Challenges and Future Direction*

Post-2020 progress includes widespread federated logins, MFA effectiveness, OAuth enhancements (PKCE/token encryption), and trust-building verification. Yet challenges persist: SSO phishing susceptibility, uneven security literacy, anonymity-driven negative behavior, and ethical risks from stronger verification (privacy/surveillance). The deepest gap is the lack of an integrated socio-technical framework aligning technical tools, behavioral insights, and ethical considerations for holistic authentication design. Emerging 6G architectures and advanced multiple-access schemes introduce new security considerations at the network layer that will influence authentication robustness in social ecosystems. [9], [10], [11], [12], [13], [14], [29], [30], [31]

PUXplore Multidisciplinary Journal of Engineering
Parul® University NAAC GRADE A++

Author(s)	Year	Method / Technique	Effectiv	Key Limitation
Khurshid [32]	2025	Controlled user study on SSO-based phishing	56.65% participants clicked fake SSO prompts	Small sample size; controlled setting
Ahmad et al. [14]	2024	Digital identity verification framework proposal	Conceptual security enhancement	Adoption feasibility not evaluated
Technische Universität Darmstadt et al. [29]	2020	Online experiment (SEM) on identity verification and trust	Significant indirect trust effect ($\beta = 0.310$, $p = 0.019$)	Artificial experimental environment
Kovalan et al. [27]	2021	Systematic literature review (PRISMA)	MFA identified as most secure practice	Limited large-scale empirical studies
Pratama et al. [13]	2022	Survey-based SSO security awareness study	Significant awareness disparities	Single-institution sample
Hossain & Raza [11]	2023	Mixed-method evaluation of MFA effectiveness	~88% reduction in unauthorized access	Limited generalizability
Rehman et al. [31]	2025	Quantitative survey on anonymity and aggression	Strong statistical association identified	Context-specific population
Nguyen et al. [30]	2025	PLS-SEM analysis of anonymous NeWOM	Anonymity significantly amplifies negative behavior	Limited contextual scope
Deepka Pandey [9]	2025	Integrated SSO–MFA architecture analysis	Highlights industry adoption trends	Conceptual; no experimental validation
Matcha & Kumar [10]	2025	OAuth 2.0 vulnerability mitigation analysis	Token theft reduced to near-zero with safeguards	Performance overhead

Table. 3. Comparison table of research on Social Media Authentication

VI. CONCLUSION

Progress has been real. That is worth saying clearly before picking at the edges of it. Detection systems have become genuinely sophisticated, authentication has improved in measurable ways, and privacy research has expanded well beyond where it started. But the recurring pattern across all three domains is hard to ignore — strong results in controlled settings that do not always survive in practice.

Security models are a good example of this. The benchmark numbers look impressive, but those numbers come from datasets that were never designed to reflect adversarial conditions or cross-platform variation [1], [2]. Someone determined to fool these systems usually can, because the features they rely on are visible and replicable. The field has known this for a while without fully solving it.

Privacy sits in a stranger position because the problem resists purely technical fixes. People share more than they intend to partly because the value of connection feels immediate and the risks feel abstract — and that calculus is not irrational, even if it produces bad outcomes [4]. What has changed is the scale of what can be inferred from ordinary-seeming content. AI can now reconstruct surprisingly detailed pictures of a person's life from data that nobody would have considered sensitive ten years ago [6], [7]. The governance response has been real but slow, and the structural inequalities shaping who bears the cost of weak privacy protections have barely entered the mainstream policy conversation [5], [8].

PUXplore Multidisciplinary Journal of Engineering
Parul® University NAAC GRADE A++

Authentication is probably where the concrete wins are clearest, and also where the remaining frustrations feel most avoidable. MFA works when people use it properly, and the protocol-level improvements have closed genuine vulnerabilities [10], [11]. But systems that work in theory only matter if people use them consistently, and friction has a way of turning good security into a checkbox that gets ticked once and then worked around [12], [14].

What all three areas need, and what the research has been slow to fully embrace, is a willingness to treat technical design and human reality as inseparable. The problems that remain are not mostly engineering problems. They are problems about how people behave, how institutions respond, and how solutions get built for the average case while the edge cases bear the cost. Research that takes all of that seriously, rather than controlling for it, is what the next phase of this work actually requires.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to Dr. Ashwini Kumar Jha for his continuous guidance, insightful feedback, and unwavering support throughout the development of this review paper. His expertise in research and his constructive suggestions significantly strengthened the conceptual clarity and analytical depth of this work. We are deeply thankful for his mentorship, encouragement, and valuable academic direction.

REFERENCES

[1] A. D. Frunze and A. A. Frolov, “Methods for Detecting Fake Accounts on the Social Network VK,” in *2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus)*, Jan. 2021, pp. 342–346. doi: 10.1109/ElConRus51938.2021.9396670.

[2] D. Amankeldin, L. Kurmangaziyeva, A. Mailybayeva, N. Glazyrina, A. Zhumadillayeva, and N. Karasheva, “Deep Neural Network for Detecting Fake Profiles in Social Networks,” *Comput. Syst. Sci. Eng.*, vol. 47, no. 1, pp. 1091–1108, 2023, doi: 10.32604/csse.2023.039503.

[3] Z. Talha, “Enhancing Social Network Security: Machine Learning-Based Bot Detection,” University of Guelma, Working Paper, 2024. Accessed: Feb. 02, 2026. [Online]. Available: https://dspace.univ-guelma.dz/jspui/handle/123456789/16472

[4] A. Gruzd and Á. Hernández-García, “A balancing act: how risk mitigation strategies employed by users explain the privacy paradox on social media,” *Behav. Inf. Technol.*, vol. 43, no. 1, pp. 21–39, Jan. 2024, doi: 10.1080/0144929X.2022.2152366.

[5] L. H. Wang and M. J. Metzger, “The Online Privacy Divide: Testing Resource and Identity Explanations for Racial/Ethnic Differences in Privacy Concerns and Privacy Management Behaviors on Social Media,” *Commun. Res.*, p. 00936502241273157, Aug. 2024, doi: 10.1177/00936502241273157.

[6] X. Cheng, L. Qiao, B. Yang, and X. Zhang, “Investigation on users’ resistance intention to facial recognition payment: a perspective of privacy,” *Electron. Commer. Res.*, pp. 1–27, Nov. 2022, doi: 10.1007/s10660-022-09588-y.

[7] Y. Liu, W. K. Tse, P. Y. Kwok, and Y. H. Chiu, “Impact of Social Media Behavior on Privacy Information Security Based on Analytic Hierarchy Process,” *Information*, vol. 13, no. 6, p. 280, May 2022, doi: 10.3390/info13060280.

[8] C. A. Yeung, I. Liccardi, K. Lu, O. Seneviratne, and T. Berners-Lee, “Decentralization: The Future of Online Social Networking,” in *Linking the World’s Information*, 1st ed., O. Seneviratne and J. Hendler, Eds., New York, NY, USA: ACM, 2023, pp. 187–199. doi: 10.1145/3591366.3591383.

[9] Deepak Pandey, “Enhancing Digital Security through SSO and MFA Integration: A Technical Perspective,” Aug. 2025, doi: 10.5281/ZENODO.16790057.

[10] S. Matcha and M. Kumar, “Enhancing Software Security with OAuth 2.0:Implementation Strategies and Vulnerability,” vol. 12, Mar. 2025.

[11] M. A. Hossain and A. Raza, “EXPLORING THE EFFECTIVENESS OF MULTIFACTOR AUTHENTICATION IN PREVENTING UNAUTHORIZED ACCESS TO ONLINE BANKING SYSTEMS,” vol. 01, no. 01, 2023.

[12] N. Khurshid, “Single Sign-On (SSO) and its Intersection with Phishing Attacks: An Investigation,” 2025.

[13] A. R. Pratama, F. M. Firmansyah, and F. Rahma, “Security awareness of single sign-on account in the academic community: the roles of demographics, privacy concerns, and Big-Five personality,” *PeerJ Comput. Sci.*, vol. 8, p. e918, Mar. 2022, doi: 10.7717/peerj-cs.918.

[14] W. Ahmad, R. Berg, and S. Kim, “Combating Fake News with Digital Identity Verification,” 2024.

[15] “Biggest social media platforms by users 2025,” Statista. Accessed: Feb. 12, 2026. [Online]. Available: https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/

[16] A. Kumar, “Sensing and Supervising through IOT,” *Int. J. Comput. Appl.*, vol. 152, no. 9, pp. 7–9, Oct. 2016, doi: 10.5120/ijca2016911723.

[17] D. K. S. Manoj, “CYBER-SECURITY: DETECTING IDENTITY DECEPTION ON SOCIAL MEDIA PLATFORMS,” *Int. J. Electr. Eng. Technol. IJEET*, vol. 12, no. 1, Jan. 2021.

[18] V. Soni and A. Jha, “IoT botnet attacks detection using deep learning approaches: a review,” *IET Conf. Proc.*, vol. 2025, no. 7, pp. 253–260, Sep. 2025, doi: 10.1049/icp.2025.1303.

[19] S. Agal, K. Raulji, and N. D. Odedra, “A machine learning approach to risk based asset allocation in portfolio optimization,” *Sci. Rep.*, vol. 15, no. 1, p. 42263, Nov. 2025, doi: 10.1038/s41598-025-26337-x.

[20] S. Agal, K. Raulji, N. Bhavsar, and P. Bhatt, “Spatiotemporal Graph Networks for Relational Reasoning in Campus Infrastructure Management,” *Int. J. Adv. Comput. Sci. Appl. Ijacsa*, vol. 16, no. 10, Oct. 2025, doi: 10.14569/IJACSA.2025.0161085.

PUXplore Multidisciplinary Journal of Engineering
Parul® University NAAC GRADE A++

[21] A. K. Jha, M. Patel, and T. Pawar, "Fog offloading: Review, Research Opportunity and Challenges," in *2019 International Conference on Smart Systems and Inventive Technology (ICSSIT)*, Tirunelveli, India: IEEE, Nov. 2019, pp. 1224–1227. doi: 10.1109/ICSSIT46314.2019.8987905.

[22] A. K. Jha and T. Pawar, "Computation Offloading for Smart Healthcare Applications," in *IoT Applications for Healthcare Systems*, R. K. Kher, C. Paunwala, F. Thakkar, H. Kher, M. Paunwala, P. K. Sahoo, and L. Ladid, Eds., in EAI/Springer Innovations in Communication and Computing. , Cham: Springer International Publishing, 2022, pp. 121–136. doi: 10.1007/978-3-030-91096-9_7.

[23] A. K. Jain, S. R. Sahoo, and J. Kaubiyal, "Online social networks security and privacy: comprehensive review and analysis," *Complex Intell. Syst.*, vol. 7, no. 5, pp. 2157–2177, Oct. 2021, doi: 10.1007/s40747-021-00409-7.

[24] A. Kumar, T. Vyas, S. Ahmed, N. Girdharwal, E. Vijayakumar, and A. Thangavelu, "Security and Privacy Enabled Framework for Online Social Networks using Blockchain," in *2023 4th International Conference on Electronics and Sustainable Communication Systems (ICESC)*, Jul. 2023, pp. 641–647. doi: 10.1109/ICESC57686.2023.10193119.

[25] A. K. Jain, S. R. Sahoo, and J. Kaubiyal, "Online social networks security and privacy: comprehensive review and analysis," *Complex Intell. Syst.*, vol. 7, no. 5, pp. 2157–2177, Oct. 2021, doi: 10.1007/s40747-021-00409-7.

[26] C. Liu, T. Zhu, J. Zhang, and W. Zhou, "Privacy Intelligence: A Survey on Image Privacy in Online Social Networks," *ACM Comput Surv*, vol. 55, no. 8, p. 161:1-161:35, Dec. 2022, doi: 10.1145/3547299.

[27] K. Kovalan *et al.*, "A Systematic Literature Review of the Types of Authentication Safety Practices among Internet Users," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 7, 2021, doi: 10.14569/IJACSA.2021.0120792.

[28] "Web Authentication: An API for accessing Public Key Credentials - Level 3." Accessed: Feb. 25, 2026. [Online]. Available: https://www.w3.org/TR/webauthn-3/

[29] Technische Universität Darmstadt, Information Systems & E-Services, N. Siegfried, J. Löbbers, and A. Benlian, "The Trust-Building Nature of Identity Verification in the Sharing Economy: An Online Experiment," in *WI2020 Zentrale Tracks*, GITO Verlag, 2020, pp. 1506–1521. doi: 10.30844/wi_2020_n5-siegfried.

[30] M. H. Nguyen, T. M. H. Dam, P. H. Pham, M. N. Pham, T. N. Nguyen, and H. T. Nguyen, "Behind the Digital Mask: Unveiling the Drivers of Anonymous Negative Word-Of-Mouth in Education," 2025.

[31] S.- Rehman, N. Rehman, S. Saleem, and Y. A. Jaffri, "An Examination of the Impact of Social Media Anonymity and Intensity of Online Conflict and Aggressive Behavior," *Rev. Appl. Manag. Soc. Sci.*, vol. 8, no. 1, pp. 279–290, Feb. 2025, doi: 10.47067/ramss.v8i1.457.

[32] N. Khurshid, "Single Sign-On (SSO) and its Intersection with Phishing Attacks: An Investigation," 2025.