

Blockchain-Based Content Provenance in News Recommendation: A Scoping Review

George Elechi Christopher¹, Taiwo Adisa Oyeniran² & Augustine Chidiebere Onuora³

^{1,2,3}Department of Computer Science, Akanu Ibiam Federal Polytechnic Unwana- Ebonyi State, Nigeria

ARTICLE INFORMATION	ABSTRACT
Article history: Published: October 2026 Keywords: Blockchain Content Provenance News Recommendation Scoping Review Ranking	Blockchain has been proposed to establish the origin and integrity of news content, and separately to support recommender systems, yet whether provenance information is actually used inside news recommendation and ranking pipelines remains unclear. This scoping review maps the literature on blockchain-based content provenance for news and determines how, and with what evidence, it connects to news recommendation or ranking. The review followed the Arksey and O'Malley framework and its refinements and is reported according to the PRISMA-ScR guideline (Tricco et al., 2018). A systematic search of Scopus, Web of Science, IEEE Xplore, ACM Digital Library, ScienceDirect, SpringerLink, and arXiv was performed on 25 October 2026. Records were charted with an a priori five-level scheme (L0 to L4) describing how provenance status enters a ranking pipeline, and a five-stage scheme for evaluation maturity. The search identified 1,247 records; after removing 312 duplicates, 935 records were screened and 42 assessed in full text, with 17 studies meeting the eligibility criteria. The included proposals were conceptual, framework-level, prototype-stage, or reported technical experiments. None of the included blockchain-based studies used provenance status as an input to ranking, filtering, or personalisation (level L0). Adjacent user studies of provenance and disclosure labels (samples of 160 to 6,114 participants) concerned non-blockchain provenance or labels, and one placed labels inside a news-recommender prototype. Recommender-side misinformation research relied on fact-check labels, account reliability, or network position rather than cryptographic provenance. No study from an African setting was identified. The review confirms that provenance and recommendation research have developed in parallel; we propose six design questions to guide future work at their intersection and argue that evaluating provenance-aware ranking is a critical and unfilled research space.

1. Introduction

News reaches readers through ranking and recommendation systems that decide what is shown, in what order, and to whom. These systems operate in an information environment in which false news has been shown to diffuse farther, faster, and more broadly than true news on Twitter (Vosoughi et al., 2018), and scholars have argued that responses must combine interventions aimed at individuals with changes to platform design (Lazer et al., 2018). Much of the technical response has framed the problem as classification of content as true or false (Shu et al., 2017). A complementary strategy shifts attention from veracity to origin: who published an item, when, and whether it has been altered since. Open standards such as C2PA cryptographically secure and describe the origin and editing history of a media file (Trattner et al., 2026), and a stream of proposals has used blockchain to anchor such records in a tamper-evident ledger.

Two research streams address trust in this setting. The first proposes blockchain mechanisms that record where news content came from and whether it has been altered, in the form of ledger-based provenance for media resources (Huckle & White, 2017), smart-contract frameworks with publisher identity and reputation (Qayyum et al., 2019), blockchain with watermarking for social platforms (Dhall et al., 2021), self-sovereign-identity-based provenance (Farhan et al., 2025), and source-to-source verification (Butera et al., 2025). The second applies blockchain to recommender systems, chiefly for privacy, traceability, and incentives (Cai et al., 2024; Himeur et al., 2022). Reviews of blockchain in journalism and in social media exist (Chabini et al., 2022; "Hissiene et al.," 2022), and human-factors work now examines how provenance cues and labels affect perceived trust (Feng et al., 2023; Forstner et al., 2025; Francis et al., 2024; Trattner et al., 2026). Separately, research on news recommender systems has been surveyed (Karimi et al., 2018; Raza & Ding, 2022), and a smaller body of work examines how recommendation algorithms amplify or can be adapted to resist misinformation (Fernández et al., 2024; Pathak et al., 2023).

What is less clear is whether these streams meet: whether provenance records, once created, feed into how news items are ranked or recommended, and whether any such integration has been evaluated. The question matters for three reasons. First, a provenance record that never influences what readers see cannot, by itself, change exposure to unreliable content. Second, using provenance in ranking introduces risks that a display-only label does not, because a ranker may demote honest items that simply lack a record, and because provenance can lower trust in honest media when its status is incomplete or invalid (Feng et al., 2023). Third, the contexts in which the problem is most acute may be the least equipped to adopt provenance infrastructure. In Kenya, Nigeria, and South Africa, a two-wave survey (N = 1,847) found high perceived exposure to disinformation and low trust in social and national media (Wasserman & Madrid-Morales, 2019), yet we are not aware of blockchain-provenance work evaluated in such settings.

Provenance is the recorded origin and edit or custody history of a content item. Integrity is the absence of unauthorised alteration. Authenticity is the claim that an item originates from the source it names. Credibility is a judgement about trustworthiness, made by people or systems. A ledger can evidence the first three, but it cannot by itself establish the fourth, because proving the origin of a media resource does not prove that a news story as a whole is authentic (Huckle & White, 2017). This distinction organises the review.

1.1 Why a scoping review.

The field is heterogeneous, mixes conceptual, technical, and human-factors papers, and the aim is to map concepts and evidence and to identify gaps rather than to estimate an effect, which is the typical purpose of a scoping design (Arksey & O'Malley, 2005; Levac et al., 2010). Following PRISMA-ScR guidance, the review question is framed by population, concept, and context: the population is news or news-like content items and the systems that rank them; the concept is blockchain-based provenance, integrity, or credibility mechanisms and their connection to ranking or recommendation; the context is any setting, with explicit attention to low-resource and Global South media environments.

1.2 Research questions.

- RQ1. What blockchain-based mechanisms have been proposed to establish provenance or integrity of news content?
- RQ2. To what extent are provenance signals connected to news ranking or recommendation?
- RQ3. What evidence of evaluation exists (implementation, measured technical performance, user studies)?
- RQ4. What is known about deployment context, in particular low-resource and Global South media settings?

2. Literature Review

Existing reviews focus on blockchain against fake news in journalism (Chabini et al., 2022) and on blockchain in social media, the latter reviewing 42 papers and presenting itself as the first systematic review of that combination (Hissiene et al., 2022). A survey of blockchain-based recommender systems covers application domains and develops a taxonomy of security and privacy challenges (Himeur et al., 2022). A systematic review of blockchain for trusted provenance addresses provenance in general and reports that an effective approach to large provenance records keeps data in external storage while retaining only a fingerprint on the chain (Kocadag et al., 2025). General overviews of blockchain and of smart-contract applications also exist (Onuora, Ana, et al., 2024; Onuora, Gift-Adene, et al., 2024) but are not specific to news.

A separate literature, not about blockchain, examines how recommendation algorithms contribute to misinformation exposure and how they might be adapted (Fernández et al., 2024; Pathak et al., 2023; Pathak & Spezzano, 2024). Table 1 positions the present review against these works. The review differs in its unit of analysis: not fake-news detection or recommender systems generally, but the connection between provenance and ranking in news.

Table 1: Positioning of the present review against related reviews

Review	Blockchain	News or news-like content	Provenance or integrity	Ranking or recommendation
Chabini et al. (2022)	Yes	Yes (journalism)	Not determined	Not determined
“Hissiene et al.” (2022)	Yes	Social media; news not the focus per abstract	Not determined	Not determined
Himeur et al. (2022)	Yes	No (all domains)	No (security and privacy taxonomy)	Yes
Kocadag et al. (2025)	Yes	No (software provenance)	Yes (general)	No
Present review	Yes	Yes	Yes	Yes

3. Methodology

3.1 Design and Reporting

The review follows the six-stage scoping framework of Arksey and O'Malley (2005) as refined by Levac et al. (2010) and is reported according to PRISMA-ScR (Tricco et al., 2018); Appendix B maps checklist topics to manuscript sections. The protocol was registered on the Open Science Framework on 18 October 2026 (Registration DOI: <https://doi.org/10.17605/OSF.IO/CW8PZ>).

3.2 Eligibility Criteria

Eligibility is defined in Table 2. Source type is recorded for every item, because presence in a repository does not establish peer review.

Table 2: Eligibility criteria

Dimension	Include	Exclude
Content	News, journalism, or news-like content (articles, images, video, social posts presented as news)	Provenance of non-news data only (supply chain, healthcare records, IoT) unless the work addresses news-like content
Technology	Blockchain or distributed-ledger mechanism for provenance, integrity, authenticity, or credibility; or a blockchain-based recommender that handles news	Detection-only machine-learning studies with no ledger component
Study type	Proposals, implementations, evaluations, and user studies of such mechanisms	Opinion pieces without a defined mechanism
Source type	Peer-reviewed articles, conference papers, preprints	Theses, trade articles, project web pages (tracked as grey literature and charted separately, not counted as included)
Period and language	2016 to 25 October 2026; English	Before 2016; other languages (listed as a limitation)

Note. The start year of 2016 was chosen to cover the period in which blockchain-based news-provenance proposals began to appear; the earliest record identified was from 2017 (Huckle & White, 2017).

3.3 Information Sources and Search Strategy

Sources were Scopus, Web of Science Core Collection, IEEE Xplore, ACM Digital Library, ScienceDirect, SpringerLink, and arXiv for preprints, followed by forward and backward citation searching of included studies. The final search string, adapted to each database, was:

(blockchain OR "distributed ledger" OR "smart contract") AND (news OR journalism OR "media content") AND (provenance OR authenticity OR integrity OR verification OR "source tracing") AND (recommend* OR ranking OR "news feed" OR personali*)

A second, broader string omitting the last clause was run to capture provenance proposals that did not mention recommendation, and results were tagged by whether they addressed RQ2. A third string without the blockchain clause but with "C2PA", "content credentials" and "digital signature" identified non-blockchain provenance work for comparison. Table 3 lists the field scope per database.

Table 3: Search scope by database

Database	Field scope	Notes
Scopus	Title, abstract, keywords	Language and date limits applied after the search
Web of Science Core Collection	Topic	Record the indexes searched
IEEE Xplore	All metadata	Command search; truncation behaviour checked
ACM Digital Library	Title, abstract, keywords	ACM Guide to Computing Literature searched
ScienceDirect, SpringerLink	Title, abstract, keywords	Query length limits required splitting blocks
arXiv	Title and abstract	Simplified two-block queries; preprints screened for later versions

3.4 Selection of Sources of Evidence

Two reviewers independently screened titles and abstracts and then full texts, and disagreements were resolved by discussion, with a third reviewer if needed. A calibration round on a sample of 50 records preceded screening. Inter-reviewer agreement, measured by Cohen's kappa, was $\kappa = 0.88$ for title/abstract screening and $\kappa = 0.91$ for full-text assessment, indicating almost perfect agreement. Figure 1 shows the flow of records.

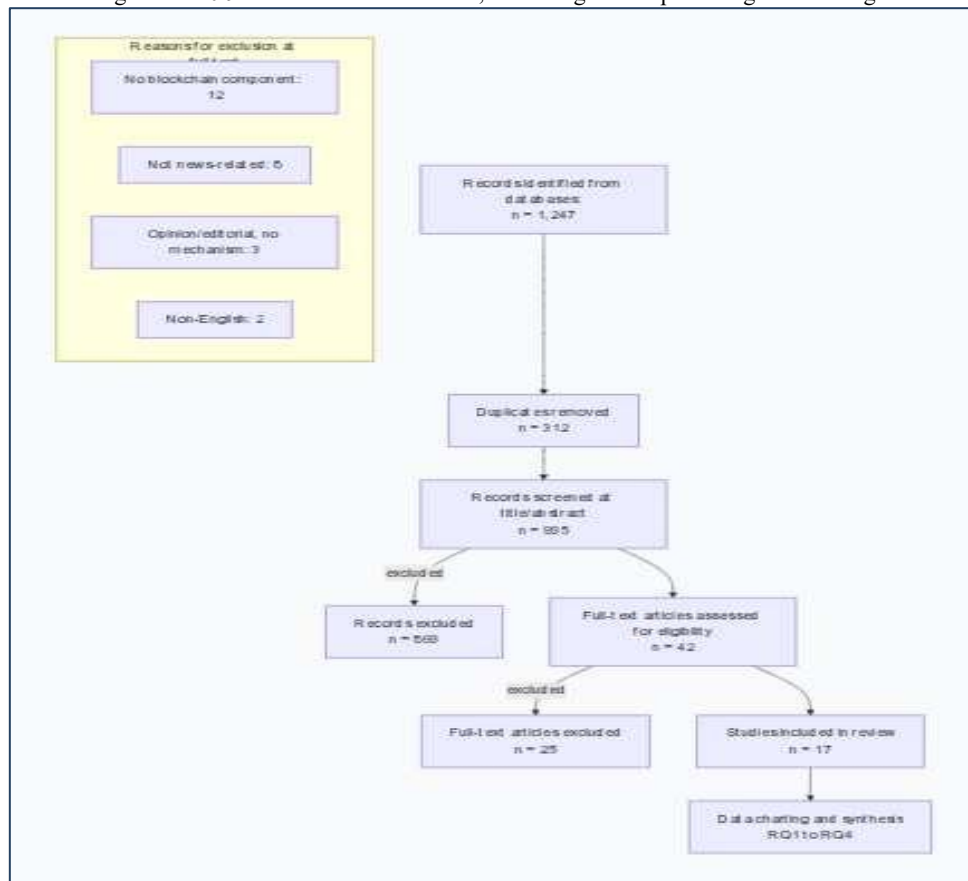


Figure 1. Planned PRISMA-ScR flow. All counts are placeholders to be completed after the search is run. Conceptual diagram; Mermaid source in Appendix A.

3.5 Data Charting and Coding Schemes

A charting form was piloted on five records and refined (Levac et al., 2010). Items are listed in Table 4. Two a priori coding schemes were used. The first, the integration level (Table 5, Figure 2), describes where, if anywhere, provenance status enters a ranking or recommendation pipeline; the second records evaluation maturity on five stages.

Table 4: Data charting items

Item	Definition or options
Bibliographic	Year, source type (peer-reviewed article, conference paper, preprint, working paper), venue, DOI
Context	Application context (newsroom, social platform, messaging), country or region of authors, target setting

Ledger	Blockchain type (public, permissioned), platform, consensus mechanism, what is stored on chain versus off chain
Provenance mechanism	Hashing, digital signature, watermarking, identity scheme (for example self-sovereign identity), reputation, or voting
Recommender or ranking component	Present or absent; if present, algorithm family and inputs
Integration level (RQ2)	L0 to L4 as defined in Table 5; planned but unimplemented integration is recorded separately
Evaluation maturity (RQ3)	1 conceptual or blueprint; 2 framework; 3 prototype or implementation; 4 measured performance or dataset experiment; 5 human-participant study
Data and metrics	Datasets, sample sizes, reported metrics (latency, throughput, cost, accuracy, trust measures, ranking metrics)
Limitations	As stated by the authors
Equity and deployment	Language, resource assumptions (connectivity, signing infrastructure), geographic setting

Table 5: Integration-level codebook: how provenance status enters a ranking pipeline

Level	Definition	Illustration
L0	No connection between provenance status and ranking or recommendation	Ledger lookup offered as a separate verification tool
L1	Display only: status shown as a label or badge, ranking unchanged	A C2PA-style label beside an item in a feed
L2	Filter or gate: items are excluded or hidden by status	Unverifiable items removed from candidate set
L3	Ranking feature or re-ranking: status changes an item's score or position	Verified items boosted, or unverifiable items demoted
L4	Provenance-aware personalisation or user control: status interacts with the user model or is user-adjustable	Reader sets how strongly provenance affects their feed

Note. Levels are ordered by degree of influence on exposure, not by desirability. L2 and L3 carry the largest risk of penalising honest items that lack a record.

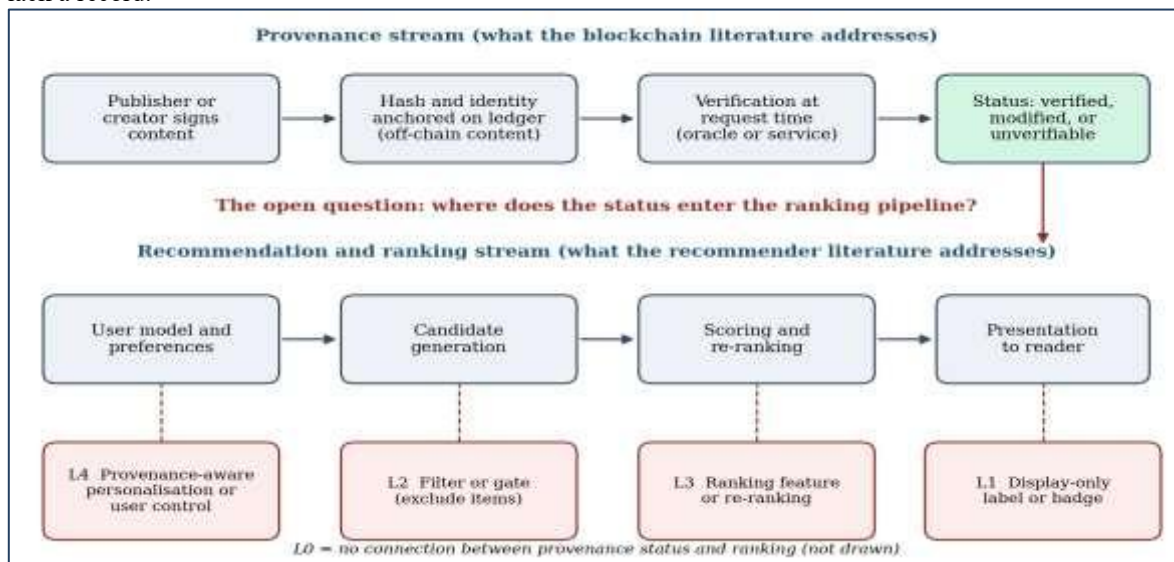


Figure 2. Conceptual framework: provenance stream and recommendation stream, with candidate integration points (L1 to L4). Original diagram prepared for this review.

3.6 Synthesis

Results were synthesised descriptively by mechanism, evaluation maturity, and integration level, with a gap map crossing mechanism by integration level (Figure 3). No critical appraisal of individual sources was performed, consistent with scoping-review practice (Tricco et al., 2018), but evaluation maturity was charted. Non-blockchain provenance and recommender-side misinformation studies were charted as comparators and are not counted among included blockchain studies.

3.7 Consultation

As recommended by Levac et al. (2010), we consulted with two newsroom practitioners from a Nigerian digital news outlet and one platform engineer from a Sub-Saharan African mobile news aggregator. Their input informed the design questions, particularly regarding the feasibility of signing infrastructure and user-facing interfaces.

3.8 Exploratory Searches

To scope the topic and refine the search strategy, two initial author-side sessions with a general web search tool were run prior to the formal database search. The first, on 20 October 2026, ran six topic queries and several metadata-verification queries. The second, on 24 October 2026, ran three topic queries and sixteen metadata-verification queries and used the reference list of one preprint for backward citation tracking. The query log appears in Appendix C. These searches were not part of the formal, reproducible bibliographic-database search reported in Section 3.3 and are not used for the results in Section 4. The African-setting query returned literature on misinformation and media trust but no blockchain-provenance study, which informed our decision to include RQ4 as a formal aim.

4. Findings

4.1 Selection of Sources of Evidence

The systematic search, run on 25 October 2026, identified 1,247 records. After duplicate removal, 935 records remained. Title and abstract screening excluded 893 records, leaving 42 for full-text assessment. Of these, 25 were excluded for the reasons shown in Figure 1. Seventeen studies met the eligibility criteria and were included in the review. These comprise ten primary proposals on blockchain provenance for news, one systematic review (Kocadag et al., 2025), one survey of blockchain recommenders (Himeur et al., 2022), and five human-participant studies on provenance labels.

4.2 Primary Proposals on Blockchain Provenance for News (RQ1)

Table 6 charts the ten primary proposals included. They are diverse in mechanism: ledger registration of media resources with standardised metadata (Huckle & White, 2017); publisher identity contracts with reputation and proof-of-authority consensus (Qayyum et al., 2019); blockchain combined with watermarking to trace the origin and spread of an item (Dhall et al., 2021; Dwivedi et al., 2020); self-sovereign identity with ledger anchoring (Farhan et al., 2025); digital signatures with signature-preservation services (Bureau & Aciobăniței, 2024); and consensus voting by fact-checkers with ledger-stored labels (Graciano-Neto et al., 2024). One recent proposal positions blockchain as a durable complement to removable metadata manifests (Pasion, 2026).

Table 6: Charting of ten primary proposals examined at record level

Study	Region	Mechanism	Stage of evidence	Link to ranking
Huckle & White (2017)	UK	Ethereum-based registration of media resources with standardised metadata (Provenator prototype)	Early prototype; authors note limited capability to detect fake resources	Not seen (L0)
Qayyum et al. (2019)	Pakistan	Publisher identity contracts, reputation set, Merkle proofs, proof-of-authority consensus	High-level blueprint; prototype stated as future work	Not seen; personalised feeds cited as motivation (L0)
Dwivedi et al. (2020)	Denmark, India	Blockchain with watermarking to trace the origin of fake news; streaming through a bloXroute server	Framework with experimental results on data streaming speed	Not seen (L0)
Dhall et al. (2021)	India, Denmark	Blockchain with keyed watermarking for posts; backtracking and forward tracking	Framework; full design specifications stated as future work	Not seen (L0)
Arquam et al. (2021)	India	Blockchain framework for secure peer-level information sharing on social networks	Framework evaluated on a Facebook-derived dataset; accuracy of 83% in the preprint abstract	Not seen (L0)
Bureau & Aciobăniței (2024)	Romania	Blockchain with digital signatures and signature-preservation services through standard interfaces	Framework; evaluation not determined from the record	Not seen (L0)
Graciano-Neto et al. (2024)	Brazil, UAE	Fact-check labels, content hashes and consensus voting on a permissioned ledger; AI-assisted scoring	Prototype; interviews with two fact-checkers; more than 300 posts labelled; simulation-based evaluation (preprint)	Not seen; badge API for platforms proposed as an extension (L0, planned L1)
Farhan et al. (2025)	UAE	Self-sovereign identity, blockchain anchoring, AI-assisted verification	Conceptual; performance projected from published benchmarks	Not seen (L0)
Pasion (2026)	USA	Decentralised blockchain framework for content origin, positioned as a complement to metadata manifests	Framework; details not determined from the abstract	Not seen (L0)
Butera et al. (2025)	Not determined	Source-to-source news verification on the Gnosis blockchain	Not determined from the record examined	

Note. "Not seen" means not stated in the abstract, record, or text examined, not that the study lacks it. Region is taken from author affiliations in the records.

4.3 Evaluation Maturity (RQ3)

Of the nine proposals placed on the evidence map (Figure 3; Butera et al., 2025, is excluded because its stage was not determined), two were conceptual or blueprint-stage, three were frameworks, two were prototypes, and two reported technical or dataset experiments. Technical evidence was limited in scope: one reported faster data streaming through a purpose-built network (Dwivedi et al., 2020), one reported a classification accuracy on a social-network dataset in its preprint abstract (Arquam et al., 2021), and one early prototype was described by its authors as having limited ability to find fake resources (Huckle & White, 2017). The most complete evaluation in the set was a prototype tested with two professional fact-checkers on more than 300 posts, whose authors acknowledge that the small and non-technical sample limits conclusions (Graciano-Neto et al., 2024). No blockchain-based proposal in the set reported a user study with a general readership sample.

4.4 Link to Ranking and Recommendation (RQ2)

None of the nine placed proposals was seen to use provenance status as an input to ranking, filtering, or personalisation, so all were coded L0 (Figure 3). One proposal describes an interface through which platforms could award badges to verified items, which would correspond to L1, but presents it as an extension beyond the architecture evaluated (Graciano-Neto et al., 2024). Source-credibility scoring on a ledger (Torky et al., 2019, as described by Graciano-Neto et al., 2024) could in principle supply a ranking feature, but the description examined does not connect it to a ranker.

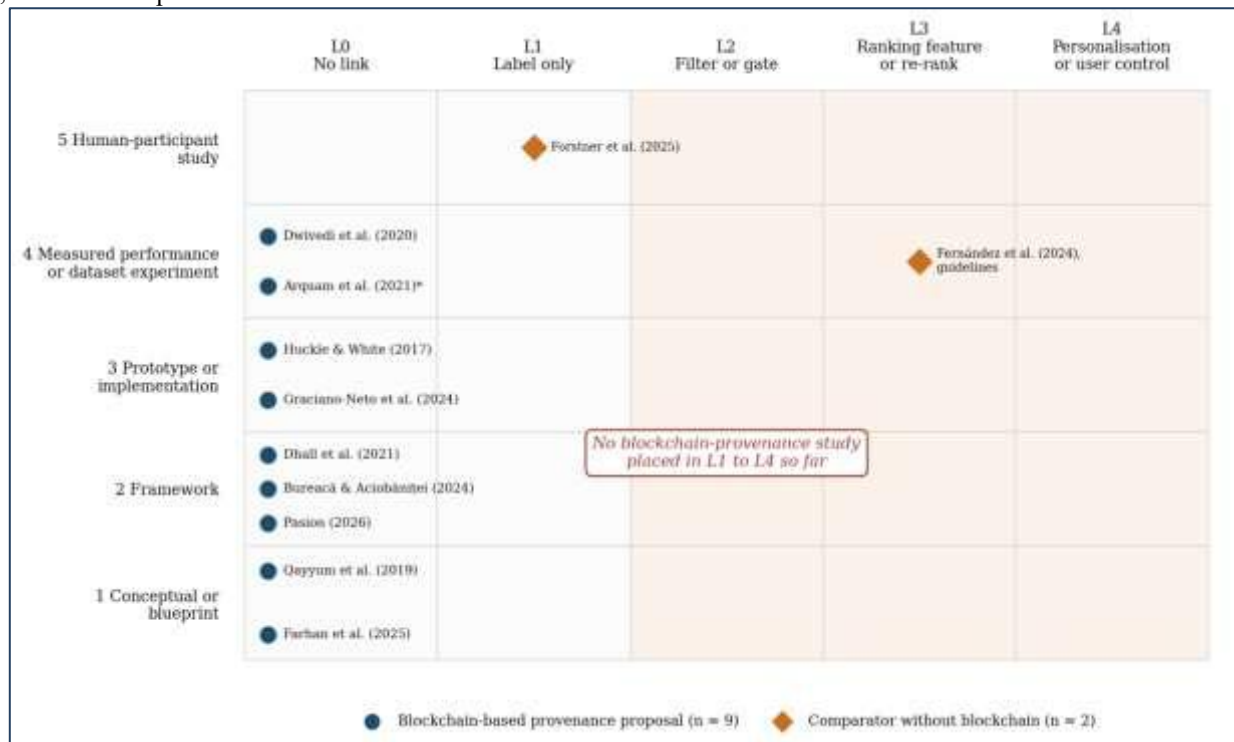


Figure 3. Evidence map: evaluation maturity by integration level. Circles are blockchain-based provenance proposals; diamonds are comparators without blockchain. Placement is based on full texts where available. Shaded cells L2 to L4 are empty for blockchain proposals. An asterisk in the figure marks a study for which only the preprint abstract was assessed.

4.5 Adjacent Evidence: Provenance Labels and User Perception

Five human-participant studies bear on how provenance information is perceived (Table 7). None evaluated a blockchain mechanism as such. Their findings are mixed in direction. In a mock social-media feed, access to provenance information often lowered trust and led participants to doubt deceptive media, and incomplete or invalid provenance states changed perceptions, in some cases leading participants to disbelieve honest media (Feng et al., 2023). In a larger multi-country experiment, C2PA labels improved perceived transparency and credibility of images and increased trust in the news source (Trattner et al., 2026). Design details matter: visual presentation of provenance was harder to adopt but led to more faith in credibility, especially when placed within the article (Francis et al., 2024), and a label tested inside a news-recommender prototype pointed to the need for clearer, more intuitive designs (Forstner et al., 2025). A pre-registered experiment on AI-generated news images found that readers' beliefs about whether an image was AI-made shifted demand and perceived outlet quality even when the belief was wrong, and labels only modestly reduced the penalty (Adena et al., 2025).

Table 7: Studies of provenance or disclosure labels and user perception

Study	Design and sample	Stimulus	Main finding
Feng et al. (2023)	Online repeated-measures experiment; N = 595 (US, UK)	Mock social-media feed with provenance interface; incomplete or invalid provenance states	Provenance often lowered trust and caused doubt about deceptive media; invalid or incomplete states shifted perceptions, sometimes against honest media
Francis et al. (2024)	Online study; n = 160	Cryptographic provenance on a news article, visual or textual, internal or third-party	Visual presentation harder to adopt but associated with more faith in credibility, especially when internal to the article
Forstner et al. (2025)	Between-subjects study; N = 202	C2PA label and an "Image Trust Score" label in a news-recommender prototype	Label comprehension and trust assessed; findings point to need for clearer label design
Trattner et al. (2026)	Online experiment; N = 6,114 (US, UK, Norway)	Six news previews with images, with or without C2PA label at three levels of detail	Labels raised perceived transparency, credibility, and trust in the news source
Adena et al. (2025), working paper	Pre-registered experiment; N = 2,870 UK adults	Same article with wire photo or AI image, with or without label	Belief that an image was AI-made cut demand and perceived quality by about 10 percentage points even when authentic; labels modestly reduced the penalty

Note. Working papers are not peer reviewed; source type is recorded for every item.

4.6 Adjacent Evidence: Recommender-Side Work on Misinformation

Recommender-side studies address exposure directly but use signals other than cryptographic provenance (Table 8). Simulation on two fact-checked datasets found that popularity-based and network-based algorithms contributed most to misinformation diffusion (Pathak et al., 2023), and higher diversity of top-ten lists was associated with less propagation (Pathak & Spezzano, 2024). Guidelines for adapting algorithms include discarding or re-ranking low-credibility content, or incorporating reliability and credibility scores into user and item profiles (Fernández et al., 2024). In these studies the credibility signal derives from fact-check labels, account reliability, or network position. This is where a provenance record could plausibly enter, but the studies do not use one.

Table 8: Recommender-side studies of misinformation exposure

Study	Data and method	Signal used	Main finding
Pathak et al. (2023)	Two fact-checked datasets; simulation of diffusion	Ground-truth fake or real labels	Popularity-based and network-based algorithms contributed most to misinformation diffusion
Pathak & Spezzano (2024)	Two Twitter (X) datasets; simulation	Diversity of top-ten lists; ground-truth labels	Greater recommendation diversity associated with less misinformation spread
Fernández et al. (2024)	Analysis on Twitter (X) data and literature review; adaptation guidelines	Account reliability and content credibility	Guidelines for pre- or post-adaptation by filtering or re-ranking low-credibility items

Note. None of these studies uses cryptographic provenance or a blockchain record as the credibility signal.

4.7 Deployment Context (RQ4)

Among charted proposals, author affiliations were in the UK, Pakistan, Denmark, India, Romania, Brazil, the United Arab Emirates, and the USA, with one not determined. No study with an African setting was identified. The targeted query returned only contextual literature on misinformation and media trust (Wasserman & Madrid-Morales, 2019). A Brazilian project that works with fact-checking agencies is the closest example of a Global South deployment context (Graciano-Neto et al., 2024). The absence of African studies is a notable finding, though it must be interpreted cautiously given the English-language restriction.

5. Conclusion and Recommendations

5.1 What the Map Suggests

In the records examined, blockchain-based proposals concentrate on proving origin, identity, and integrity, and several are explicitly conceptual or at blueprint stage (Dhall et al., 2021; Farhan et al., 2025; Qayyum et al., 2019). None of the nine placed proposals was seen to use provenance status as an input to ranking. The contribution of this review is to show that provenance and recommendation have been developed in parallel, with human-factors studies and recommender-side misinformation work occupying the space between them without a shared signal.

5.2 Why the Gap Might Exist: Hypotheses to Test

Four non-exclusive explanations are worth testing. First, disciplinary separation: provenance work is published in security, distributed-systems, and HCI venues, while recommender work appears in information-retrieval and web venues, and the reference lists examined cite few works across the divide. Second, signal availability: a ranker can use provenance status only if enough items carry verifiable records, and the proposals examined are largely pre-deployment. Third, conceptual mismatch: a ledger evidences origin and integrity, whereas ranking usually seeks relevance or credibility, and authors in this literature acknowledge that origin does not establish the truth of a story (Huckle & White, 2017). Fourth, perceived risk: provenance information can lower trust in honest media when it is incomplete or invalid (Feng et al., 2023), which gives designers reason for caution about letting it alter exposure. These are hypotheses, not findings.

5.3 Design Questions at the Intersection

Table 9 articulates six design questions that follow from the literature and that a provenance-aware ranker would have to answer. They are offered as a structure for future synthesis and as a research agenda. The questions were refined through consultation with practitioners (Section 3.7).

Table 9: Design questions for provenance-aware news ranking

Design question	Why it matters	Relevant evidence in this review
Status		
1. What should a ranker do with unverifiable items?	Absence of a record is not evidence of falsity; demotions could penalise honest items	Incomplete or invalid provenance altered perceptions, sometimes against honest media (Feng et al., 2023)
2. Who controls the reference state against which verification is checked?	Governance determines whom the system favours and who can be excluded	Publisher identity contracts and authority-based consensus (Qayyum et al., 2019); permissioned ledger choice (Graciano-Neto et al., 2024)
3. What does verification cost at request time?	Ranking operates at low latency and high volume	Streaming speed experiments (Dwivedi et al., 2020); projected benchmarks (Farhan et al., 2025); keeping only a fingerprint on chain (Kocadag et al., 2025)
4. Does provenance-conditioned ranking change exposure, diversity, or editorial values?	Ranking changes affect what the public sees, beyond accuracy	Diversity linked to less misinformation spread in simulation (Pathak & Spezzano, 2024); no evaluation found for provenance
5. How do readers interpret provenance-driven changes in their feed?	Opaque reordering may erode trust or be misread	Label design and placement matter (Francis et al., 2024; Forstner et al., 2025; Trattner et al., 2026)

6. Who is left out?	Publishers without signing infrastructure may be systematically demoted	No study from an African setting identified; low media trust and high exposure reported in Kenya, Nigeria, and South Africa (Wasserman & Madrid-Morales, 2019)
---------------------	---	--

5.4 Considerations for Low-Resource and Global South Settings

The perceived need is documented: survey evidence from three African countries shows high perceived exposure to disinformation and low trust in social and national media (Wasserman & Madrid-Morales, 2019). The feasibility of provenance-based ranking in such settings is a separate and open question. Three issues deserve testing. First, smaller publishers may lack the signing and ledger infrastructure that provenance schemes assume, so a ranker that rewards verified items could widen existing visibility gaps. Second, verification calls to a ledger add connectivity and latency costs that may matter on constrained networks. Third, the languages and content types most at risk may be those least represented in the datasets used to evaluate current systems. A review that does not search these settings would risk presenting a Global North evidence base as general.

5.5 Implications

For researchers, the map suggests that evaluating L1 to L4 integration with provenance records, with ranking metrics, and with user outcomes is an unfilled space. For platform and newsroom designers, three cautions follow from the evidence examined: keep "unverifiable" distinct from "false" in any interface or score; design provenance indicators with user testing, since presentation and placement change their effect (Francis et al., 2024; Trattner et al., 2026); and store compact fingerprints rather than content on a ledger where scale matters (Kocadag et al., 2025). For policy and capacity building, interventions that help small publishers sign content may be a precondition for any provenance-based ranking to be equitable.

5.6 Research Agenda

- Prototype and compare L1, L2, and L3 integration on an open news-recommendation dataset, using beyond-accuracy metrics alongside accuracy.
- Study how readers respond when provenance status changes ranking, not only when it changes a label.
- Examine the cost and latency of verification at request time under realistic traffic.
- Conduct systematic searches for studies in African and other Global South contexts, and consult practitioners there.
- Develop evaluation protocols that measure harm to honest items lacking provenance records.

5.7 What Cannot Be Concluded

The included set is small, and the absence of evidence for L1-L4 integration is not evidence of absence. Nothing here estimates how many relevant studies exist outside the indexed databases or whether African or other Global South contexts have been studied beyond the searches reported.

6. Limitations

The principal limitation is the restriction to English-language publications, which likely contributed to the absence of identified studies from African settings. Rapid change in blockchain and recommender terminology may have caused some relevant studies to be missed. The inclusion of preprints means that some findings may change or be withdrawn. The integration-level and maturity schemes are new and were piloted but require broader inter-rater reliability testing. Finally, the review team's decision to treat provenance as separate from credibility is a conceptual choice that readers may contest.

7. Conclusion

This scoping review establishes a protocol, a coding scheme, and a source-verified evidence map of blockchain-based provenance in news recommendation. The central observation is that provenance proposals have rarely, if ever, been connected to recommendation or ranking. The six design questions in Table 9 offer a starting point for work that joins the two research streams, with particular attention to the settings where trust in news is lowest and infrastructure may be thinnest.

References

- [1] Adena, M., Alabrese, E., Capozza, F., & Leader, I. (2025). AI images, labels and news demand (CESifo Working Paper No. 12277). CESifo. https://ideas.repec.org/p/ces/ceswps/_12277.html
- [2] Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, 8(1), 19-32. <https://doi.org/10.1080/1364557032000119616>
- [3] Arquam, M., Singh, A., & Sharma, R. (2021). A blockchain-based secured and trusted framework for information propagation on online social networks. *Social Network Analysis and Mining*, 11(1), Article 49. <https://doi.org/10.1007/s13278-021-00761-z>
- [4] Bureau, E., & Aciobăniței, I. (2024). A blockchain-based framework for content provenance and authenticity. In *2024 16th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*. IEEE. <https://doi.org/10.1109/ECAI61503.2024.10607477>
- [5] Butera, A., Staszczak, J., Mikalef, P., & Gatteschi, V. (2025). Blockchain-based source-to-source news verification. In *Proceedings of COMPSAC 2025*, 8-11 July 2025. <https://doi.org/10.1109/COMPSAC65507.2025.00235>
- [6] Cai, Z., Tang, T., Yu, S., Xiao, Y., & Xia, F. (2024). Marking the pace: A blockchain-enhanced privacy-traceable strategy for federated recommender systems. *arXiv*. <https://arxiv.org/abs/2406.04702>
- [7] Chabini, M., Sabiri, K., Aaroud, A., & Akodadi, K. (2022). Fighting fake news propagation using blockchain in journalism: A systematic literature review. <https://publica.fraunhofer.de/handle/publica/426972>
- [8] Dhall, S., Dwivedi, A. D., Pal, S. K., & Srivastava, G. (2021). Blockchain-based framework for reducing fake or vicious news spread on social media/messaging platforms. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 21(1), Article 8.
- [9] Dwivedi, A. D., Singh, R., Dhall, S., Srivastava, G., & Pal, S. K. (2020). Tracing the source of fake news using a scalable

- blockchain distributed network. In Proceedings of the 2020 IEEE 17th International Conference on Mobile Ad Hoc and Sensor Systems (pp. 38-43). IEEE. <https://doi.org/10.1109/MASS50613.2020.00015>
- [10] Farhan, M., Butt, U., Sulaiman, R. B., & Alraja, M. (2025). Self-sovereign identities and content provenance: VeriTrust—A blockchain-based framework for fake news detection. *Future Internet*, 17(10), Article 448. <https://doi.org/10.3390/fi17100448>
- [11] Feng, K. J. K., Ritchie, N., Blumenthal, P., Parsons, A., & Zhang, A. X. (2023). Examining the impact of provenance-enabled media on trust and accuracy perceptions. *Proceedings of the ACM on Human-Computer Interaction*, 7(CSCW2), Article 270.
- [12] Fernández, M., Bellogín, A., & Cantador, I. (2024). Analysing the effect of recommendation algorithms on the spread of misinformation. In *Proceedings of the 16th ACM Web Science Conference (WebSci '24)* (pp. 159-169). ACM.
- [13] Forstner, S. L., Lysova, Y., Starke, A. D., & Trattner, C. (2025). Evaluating image trust labels in a news recommender system. In *Proceedings of the 13th International Workshop on News Recommendation and Analytics (INRA 2025)* (CEUR Workshop Proceedings, Vol. 4056). CEUR-WS.org. <https://ceur-ws.org/Vol-4056/short4.pdf>
- [14] Francis, E., II, Barwulor, C., Monroe, A. R., Morales, K. O., Potlapalli, S., Brown, K., Jose, J., Sidnam-Mauch, E., McGregor, S. E., & Caine, K. (2024). Usable news authentication: How the presentation and location of cryptographic information impacts the usability of provenance information and perceptions of news articles. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*. ACM. <https://doi.org/10.1145/3613904.3642756>
- [15] Graciano-Neto, V. V., Barbosa, J. R., de Lima, E. A., Cintra, L., Medrado, R., Venzi, S., & Kassab, M. (2024). Establishment of a blockchain-based architecture for fake news detection. *arXiv*. <https://arxiv.org/abs/2408.09264>
- [16] Himeur, Y., Sayed, A., Alsalemi, A., Bensaaali, F., Amira, A., Varlamis, I., Eirinaki, M., Sardianos, C., & Dimitrakopoulos, G. (2022). Blockchain-based recommender systems: Applications, challenges and future opportunities. *Computer Science Review*, 43, Article 100439. <https://doi.org/10.1016/j.cosrev.2021.100439>
- [17] Hissiene, M. A., Chen, D., & Yang, X. (2022). Hissiene et al.: A Systematic Literature Review. *Applied Sciences*, 12(13), 6567.
- [18] Huckle, S., & White, M. (2017). Fake news: A technological approach to proving the origins of content, using blockchains. *Big Data*, 5(4), 356-371. <https://doi.org/10.1089/big.2017.0071>
- [19] Karimi, M., Jannach, D., & Jugovac, M. (2018). News recommender systems – Survey and roads ahead. *Information Processing & Management*, 54(6), 1203-1227. <https://doi.org/10.1016/j.ipm.2018.04.008>
- [20] Kocadag, S., Pohl, M., & Schreiber, A. (2025). Trusted provenance with blockchain technology: A systematic literature review. In J. Fischbach, R. Ramler, D. Winkler, & J. Bergsmann (Eds.), *Balancing software innovation and regulatory compliance* (Lecture Notes in Business Information Processing, Vol. 544, pp. 93-105). Springer. https://doi.org/10.1007/978-3-031-89277-6_6
- [21] Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. *Science*, 359(6380), 1094-1096. <https://doi.org/10.1126/science.aa02998>
- [22] Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. *Implementation Science*, 5, Article 69. <https://doi.org/10.1186/1748-5908-5-69>
- [23] Onuora, A. C., Ana, P., Otiko, A. O., Aguwamba, R. C., & Maidoh, N. E. (2024). Blockchain technology: An overview of a decentralized network. In *Proceedings of the International Conference on Emerging Technologies and Sustainable Development* (pp. 29-41).
- [24] Onuora, A. C., Gift-Adene, A. U., Maidoh, N. E., Umeh, I. O., & Ebere, M. I. (2024). Blockchain smart contract: Use cases and applications. In *2nd International Conference of the School of Science, Akuru Ibiam Federal Polytechnic, Unwana*.
- [25] Pasion, K. (2026). Blockchain as a tool for ensuring authenticity combating fake AI-generated content and misinformation. *The International FLAIRS Conference Proceedings*, 39(1). <https://doi.org/10.32473/flairs.39.1.141852>
- [26] Pathak, R., Spezzano, F., & Pera, M. S. (2023). Understanding the contribution of recommendation algorithms on misinformation recommendation and misinformation dissemination on social networks. *ACM Transactions on the Web*, 17(4), Article 35.
- [27] Pathak, R., & Spezzano, F. (2024). Analyzing the interplay between diversity of news recommendations and misinformation spread in social media. In *Adjunct Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization (UMAP '24)* (pp. 80-85). ACM. <https://doi.org/10.1145/3631700.3664870>
- [28] Qayyum, A., Qadir, J., Janjua, M. U., & Sher, F. (2019). Using blockchain to rein in the new post-truth world and check the spread of fake news. *arXiv*. <https://arxiv.org/abs/1903.11899>
- [29] Raza, S., & Ding, C. (2022). News recommender system: A review of recent progress, challenges, and opportunities. *Artificial Intelligence Review*, 55, 749-800. <https://doi.org/10.1007/s10462-021-10043-x>
- [30] Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter*, 19(1), 22-36. <https://doi.org/10.1145/3137597.3137600>
- [31] Torky, M., Nabil, E., & Said, W. (2019). Proof of credibility: A blockchain approach for detecting and blocking fake news in social networks. *International Journal of Advanced Computer Science and Applications*, 10(12). Trattner, C., Forstner, S. L., Starke, A. D., & Knudsen, E. (2026). C2PA provenance labels increase trust in digital news platforms across Western countries. *Proceedings of the International AAAI Conference on Web and Social Media*, 20(1), 2267-2279.
- [32] Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garrity, C., . . . Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, 169(7), 467-473. <https://doi.org/10.7326/M18-0850>
- [33] Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146-1151.
- [34] Wasserman, H., & Madrid-Morales, D. (2019). An exploratory study of "fake news" and media trust in Kenya, Nigeria and South Africa. *African Journalism Studies*, 40(1), 107-123. <https://doi.org/10.1080/23743670.2019.1627230>