

Article

From AI Adoption to Trustworthy Metaverse Ecosystems: A Systematic Review and Metaverse AI Trust–Governance Nexus Framework

Hendi Herdiana¹, Munir Munir¹, Rasim Rasim¹, Komarudin Komarudin^{1,*}, Juliana Juliana^{1,*}, Asep Miftahuddin¹, Maykada Harjono Kurniawan¹, Shafinar Ismail²

¹ Universitas Pendidikan Indonesia, Bandung, Indonesia

² Universiti Teknologi MARA (UiTM), Shah Alam, Malaysia

* **Corresponding author:** Komarudin Komarudin komarudin_pko@upi.edu; Juliana Juliana, julian@upi.edu

CITATION

Herdiana H, Munir M, Rasim R, et al. From AI Adoption to Trustworthy Metaverse Ecosystems: A Systematic Review and Metaverse AI Trust–Governance Nexus Framework. *Metaverse*. 2026; 7(4): m260019. <https://doi.org/10.54517/m260019>

COPYRIGHT



Copyright © 2026 by author(s).

Metaverse is published by Asia Pacific Academy of Science Pte. Ltd. This work is licensed under the Creative Commons Attribution (CC BY) license.

<https://creativecommons.org/licenses/by/4.0/>

Abstract: The rapid integration of artificial intelligence (AI) into metaverse ecosystems has created new opportunities for immersive digital services while simultaneously introducing governance, trust, and accountability challenges. Although studies have examined AI, governance, institutional trust, and technology adoption individually, their relationships remain conceptually fragmented. This study addresses this gap through a Systematic Literature Review (SLR) following the PRISMA 2020 guidelines. A structured search of the Scopus database identified 3,698 records, which were reduced through sequential eligibility filtering to 676 open-access journal articles. Bibliometric mapping was performed on this eligible set, and 183 articles that explicitly addressed the governance–trust–adoption nexus were retained for qualitative thematic synthesis. Included studies were additionally classified as metaverse-native, metaverse-adjacent, or foundational, so that the evidential basis of every thematic claim is made explicit. The analysis reveals four interrelated themes: AI as the enabling capability of metaverse ecosystems, governance mechanisms for trustworthy AI, institutional trust as the foundation of stakeholder confidence, and AI adoption as a socio-technical outcome of governance and trust. A recurring pattern across the literature is the AI Adoption Paradox, whereby advances in AI simultaneously expand opportunities for intelligent and immersive services while intensifying concerns over transparency, privacy, explainability, accountability, and digital identity. Persistence, embodiment, avatar-mediated identity, and cross-jurisdictional virtual economies render these challenges qualitatively different from those of conventional digital platforms. These findings suggest that technological capability alone cannot sustain AI adoption in metaverse ecosystems. Rather, governance mechanisms foster institutional trust, which enables responsible and sustainable adoption. Building on this synthesis, the study proposes the Metaverse AI Trust–Governance Nexus Framework, integrating AI capability, governance, institutional trust, and AI adoption into a coherent socio-technical perspective. Because the framework is derived from qualitative synthesis rather than meta-analytic testing, its relationships are presented as evidence-informed propositions rather than as empirically validated causal effects. By consolidating fragmented research streams, this review contributes to the emerging literature on trustworthy metaverse ecosystems and provides a conceptual foundation for future empirical research and governance strategies supporting responsible, human-centred, and sustainable AI-enabled digital ecosystems.

Keywords: Metaverse ecosystems; Artificial intelligence; AI governance; Institutional trust; Trustworthy AI; AI adoption; Responsible AI; Systematic literature review

1. Introduction

The emergence of the metaverse has accelerated a new stage of digital transformation in which immersive virtual environments are increasingly intertwined

with artificial intelligence (AI), extended reality (XR), blockchain, digital twins, the Internet of Things (IoT), and high-speed communication infrastructures [1]. Rather than serving merely as advanced digital platforms, metaverse ecosystems create persistent environments where individuals, organisations, intelligent agents, and digital assets continuously interact across physical and virtual spaces. As these technologies converge, AI assumes a broader role than conventional computational automation by enabling intelligent virtual agents, adaptive decision-making, personalised services, and dynamic interactions that shape users' experiences within immersive environments [2–7]. Consequently, discussions surrounding the metaverse have gradually shifted from technological innovation alone toward questions of governance, institutional responsibility, and the conditions required to establish trustworthy digital ecosystems.

This rapid technological evolution has simultaneously created new governance challenges. Unlike traditional digital platforms, AI-enabled metaverse ecosystems operate through persistent virtual identities, autonomous intelligent agents, decentralised digital assets, and increasingly complex virtual economies that extend beyond conventional organisational and national boundaries [8–11]. These characteristics have intensified concerns regarding algorithmic transparency, explainability, accountability, privacy protection, fairness, cybersecurity, and regulatory compliance. Recent empirical evidence further indicates that users' willingness to participate in immersive digital environments depends not only on technological capability but also on their confidence that AI systems operate within transparent, accountable, and ethically governed institutional arrangements [12–15]. Technological sophistication therefore represents only one component of sustainable adoption. Without credible governance and institutional trust, even highly capable AI technologies may struggle to achieve broad acceptance.

Four properties distinguish metaverse ecosystems from conventional AI-enabled digital platforms and shape how governance and trust operate within them. First, persistence: the environment and its state continue independently of any individual session, so algorithmic decisions accumulate consequences over time instead of terminating with a discrete transaction. Second, embodiment and real-time social presence: users act through avatars, and AI-driven agents are encountered as social actors rather than as background services, which heightens sensitivity to manipulation, fairness, and emotional influence. Third, identity portability and asset ownership: persistent identities, wallets, and tokenised assets travel across platforms, so a failure in one governance arrangement propagates beyond its originating service. Fourth, jurisdictional ambiguity: immersive interaction and virtual economies fall only partially within existing regulatory perimeters. Taken together, these properties mean that governance in metaverse ecosystems cannot be reduced to conventional platform compliance, and that institutional trust must be sustained continuously rather than established at a single point of adoption. This distinction motivates the present review and is revisited in Section 5.2.

Research addressing these issues has expanded considerably during the past few years. Studies have investigated AI adoption across diverse contexts, including healthcare, education, public administration, manufacturing, tourism, financial services, and digital business ecosystems. Parallel streams of research have also examined responsible AI, institutional trust, digital governance, and metaverse technologies from disciplinary perspectives such as information systems, computer science, management, engineering, and the social sciences [16]. Collectively, these

studies have generated valuable insights into individual aspects of AI adoption and digital transformation. Nevertheless, they have generally evolved as relatively independent conversations. AI adoption research frequently emphasises behavioural acceptance and technology use, governance studies concentrate on regulatory principles and ethical AI, while metaverse research often focuses on technological architecture, immersive interaction, and virtual experiences [5,17–20]. Comparatively fewer studies examine how these perspectives intersect to explain the development of trustworthy metaverse ecosystems.

A closer examination of the literature reveals that this fragmentation extends beyond disciplinary boundaries. Governance studies commonly discuss responsible AI principles such as transparency, accountability, fairness, and explainability, yet they seldom explain how these mechanisms cultivate institutional trust that ultimately supports AI adoption [12,13,15,21–23]. Conversely, studies of technology adoption often recognise the importance of trust but rarely position governance as an institutional capability that shapes stakeholder confidence within immersive digital environments. Meanwhile, metaverse research continues to prioritise technological innovation, digital identity, virtual economies, and user experience, with comparatively limited attention devoted to the governance structures required to sustain long-term adoption. As a result, current scholarship provides only a partial explanation of how governance, institutional trust, and AI adoption interact within emerging metaverse ecosystems.

Evidence across the reviewed studies also suggests that successful AI adoption in metaverse ecosystems should not be viewed as an isolated technological phenomenon. Instead, adoption reflects the combined influence of governance quality, institutional legitimacy, transparency, privacy protection, secure digital identity, and stakeholder confidence [12,13,18,24–26]. These relationships appear consistently throughout the literature, yet they remain dispersed across different theoretical traditions and application domains. The absence of an integrated perspective limits both theoretical development and practical guidance for organisations seeking to deploy trustworthy AI within immersive digital environments.

Against this background, the present review addresses this gap by synthesising evidence from the selected studies to develop the Metaverse AI Trust–Governance Nexus Framework. Rather than proposing a new technology adoption theory or replacing existing governance models, the framework integrates established knowledge into a higher-order socio-technical explanation of how governance mechanisms foster institutional trust and subsequently support AI adoption within trustworthy metaverse ecosystems. By positioning governance as the institutional mechanism linking technological capability with stakeholder confidence, the framework provides a more comprehensive explanation of sustainable AI adoption in immersive digital environments.

The review is guided by three research questions. **RQ1:** examines the current state of research on AI-enabled metaverse ecosystems by identifying publication trends, dominant themes, and emerging research directions. **RQ2:** investigates how existing studies conceptualise the relationships among AI governance, institutional trust, and AI adoption within metaverse ecosystems. **RQ3:** synthesises these findings into an integrated conceptual framework explaining how governance mechanisms cultivate institutional trust and support AI adoption toward trustworthy metaverse ecosystems.

These questions correspond to three research objectives. RO1 is to map the

publication dynamics and intellectual structure of research on AI-enabled metaverse ecosystems. RO2 is to analyse how the reviewed literature conceptualises the relationships among AI capability, governance mechanisms, institutional trust, and AI adoption, and to establish which of these relationships are evidenced in metaverse settings specifically rather than inferred from adjacent domains. RO3 is to consolidate these relationships into an integrative socio-technical framework and to specify the propositions that follow from it. **Table 1** summarises the alignment between the research questions, the research objectives, the analytical procedures used, and the sections in which each is reported.

Table 1. Alignment of research questions, research objectives, analytical procedures, and reporting sections.

Research question	Research objective	Analytical procedure	Reported in
RQ1. What is the current state of research on AI-enabled metaverse ecosystems in terms of publication trends, dominant themes, and emerging directions?	RO1. Map the publication dynamics and intellectual structure of the field.	Descriptive publication analysis of the identified records and keyword co-occurrence mapping of the eligible records using Bibliometrix (Biblioshiny) and VOSviewer.	Sections 4.1–4.2
RQ2. How does existing research conceptualise the relationships among AI governance, institutional trust, and AI adoption within metaverse ecosystems?	RO2. Analyse how the literature relates AI capability, governance, institutional trust, and adoption, and establish which relationships are evidenced in metaverse settings.	Qualitative thematic synthesis of the 183 included studies; classification of each study by metaverse relevance; cross-theme comparison.	Sections 4.1, 4.3–4.4
RQ3. How can these findings be integrated into a conceptual framework explaining how governance cultivates institutional trust and supports AI adoption?	RO3. Consolidate the thematic findings into an integrative socio-technical framework and specify the propositions that follow from it.	Conceptual synthesis and explicit evidence-to-framework mapping, distinguishing evidenced relationships from author-synthesised propositions.	Sections 4.5 and 5

This review makes three principal contributions. First, it systematically maps the rapidly expanding literature on AI-enabled metaverse ecosystems by identifying major publication trends, thematic developments, and unresolved research issues. Second, it brings together previously fragmented streams of research concerning AI governance, institutional trust, and AI adoption, demonstrating how these dimensions collectively shape trustworthy metaverse ecosystems. Finally, it proposes the Metaverse AI Trust–Governance Nexus Framework, an ecosystem-oriented socio-technical framework that provides both a conceptual foundation for future empirical research and practical guidance for policymakers, organisations, and technology developers seeking to promote responsible AI governance in immersive digital environments.

2. Literature Review

2.1 Metaverse Ecosystems

The concept of the metaverse has evolved considerably from its initial interpretation as an immersive virtual space into a broader socio-technical ecosystem where digital and physical environments increasingly converge. Recent developments in artificial intelligence, extended reality, blockchain, digital twins, cloud computing,

and the Internet of Things have accelerated this transformation, enabling persistent environments in which users, intelligent agents, organisations, and digital assets interact continuously across interconnected digital infrastructures [4,8,11,20,27]. Much of the existing literature highlights the technological opportunities created by these developments. Researchers have reported applications spanning education, healthcare, manufacturing, tourism, public services, and business innovation, illustrating the versatility of metaverse ecosystems across multiple sectors [28–33]. At the same time, another body of research has drawn attention to the institutional challenges accompanying this technological progress. Issues such as privacy protection, digital identity management, algorithmic accountability, cybersecurity, transparency, and ethical responsibility have become increasingly central to discussions surrounding sustainable metaverse development [34–38]. Viewed collectively, these studies suggest that the long-term viability of metaverse ecosystems depends on considerably more than technological innovation. As immersive environments become increasingly dependent on autonomous AI and decentralised digital infrastructures, governance mechanisms emerge as equally important conditions for sustaining stakeholder trust and supporting responsible digital interaction.

Recent studies further indicate that AI-enabled conversational agents, intelligent virtual assistants, and adaptive recommendation systems have become integral components of immersive digital ecosystems. Rather than functioning solely as automation technologies, these AI capabilities facilitate personalised interaction, continuous service delivery, and adaptive decision support across multiple organisational contexts [2,39,40].

2.2 Artificial Intelligence in the Metaverse

Within metaverse ecosystems, AI functions as the primary enabling capability that transforms virtual environments into adaptive and intelligent socio-technical systems. Rather than simply automating predefined tasks, AI supports autonomous virtual agents, personalised services, intelligent content generation, computer vision, natural language interaction, and adaptive decision-making that collectively shape immersive user experiences [17,25,27].

The literature consistently demonstrates the growing contribution of AI across diverse application domains. Intelligent virtual assistants improve customer engagement, AI-driven analytics support organisational decision-making, while adaptive learning systems, healthcare applications, and digital public services increasingly rely on AI to deliver responsive and personalised interactions [28,29,31,33,41]. These developments reinforce the position of AI as the cognitive infrastructure underpinning metaverse ecosystems.

Nevertheless, the same technological capabilities that enhance intelligent interaction also introduce new sources of uncertainty. Greater autonomy frequently raises questions concerning explainability, fairness, accountability, privacy, and security. These recurring concerns indicate that technological capability alone cannot guarantee sustainable AI implementation. Instead, its long-term acceptance appears closely associated with the governance mechanisms that regulate how AI systems are designed, deployed, and monitored within metaverse ecosystems [38,42].

Beyond enabling intelligent interaction, recent empirical research demonstrates that conversational AI, generative AI, and intelligent virtual agents increasingly function as socio-technical actors capable of supporting knowledge sharing, personalised interaction, intelligent decision support, and adaptive service delivery

across education, tourism, and digital business environments [2,3,18,39,43].

Recent adoption studies also suggest that users increasingly evaluate AI systems through AI literacy, perceived risk, ethical concerns, explainability, and institutional confidence rather than technological performance alone [19,44–46].

2.3 AI Governance and Responsible AI

The rapid diffusion of AI has stimulated growing scholarly interest in governance frameworks capable of ensuring that intelligent systems remain transparent, accountable, and socially responsible. International initiatives, including the OECD AI Principles, the NIST AI Risk Management Framework, and the European Union AI Act, have established widely recognised principles that emphasise transparency, explainability, accountability, fairness, privacy, human oversight, robustness, and safety [35,39,45]. Recent governance studies emphasise that trustworthy AI implementation depends not only on ethical principles but also on organisational accountability, stakeholder participation, institutional transparency, and responsible implementation throughout the AI lifecycle [26,42,47,48].

Although these initiatives provide an important normative foundation, much of the literature remains oriented toward prescribing governance principles rather than explaining their role within broader socio-technical systems. Existing studies frequently discuss responsible AI as a regulatory or ethical requirement, yet comparatively fewer investigations explore how governance mechanisms shape institutional trust or influence stakeholder acceptance of AI-enabled services.

This distinction becomes particularly relevant within metaverse ecosystems, where governance extends beyond legal compliance. Governance also provides the institutional structures through which users evaluate the credibility, legitimacy, and trustworthiness of increasingly autonomous AI-enabled environments. Accordingly, governance should be understood not merely as a regulatory framework but as an enabling institutional capability that supports sustainable AI adoption [22,49].

2.4 Institutional Trust

Institutional trust has long been recognised as an important determinant of technology acceptance, particularly in situations characterised by uncertainty and perceived risk [50,51]. Within AI-enabled metaverse ecosystems, however, trust assumes a broader institutional dimension because users increasingly depend upon autonomous decision-making, intelligent virtual agents, digital identities, and algorithmic governance.

Across the reviewed studies, institutional trust consistently emerges through perceptions of transparency, accountability, competence, fairness, and responsible governance rather than through technological performance alone. Mechanisms that strengthen privacy protection, secure digital identity management, and effective institutional oversight appear to reduce uncertainty surrounding AI-enabled interactions and increase stakeholder confidence [11,49,52].

Despite its recognised importance, institutional trust is often examined as an independent construct within technology adoption research. Comparatively limited attention has been devoted to explaining how governance mechanisms actively cultivate trust within immersive digital ecosystems. This unresolved relationship represents one of the conceptual issues addressed by the present review.

Recent empirical evidence distinguishes institutional trust from interpersonal trust by emphasising users' confidence in organisations responsible for governing, deploying, and supervising AI systems. Organisational transparency, institutional

accountability, user support, and governance quality consistently emerge as important antecedents of institutional trust [24,26,47,53–55].

2.5 AI Adoption in Metaverse Ecosystems

Research on AI adoption has traditionally been informed by behavioural theories such as the Technology Acceptance Model (TAM), the Unified Theory of Acceptance and Use of Technology (UTAUT), the Diffusion of Innovation Theory (DOI), and Trust Theory [56,57]. These frameworks explain adoption primarily through perceptions of usefulness, ease of use, facilitating conditions, social influence, trust, and perceived risk. Recent AI adoption studies indicate that classical behavioural theories including TAM, UTAUT, TPB, TOE, and the Information Systems Success Model continue to explain behavioural intention effectively. Nevertheless, recent studies increasingly extend these theories by incorporating AI literacy, perceived risk, educational support, stakeholder alignment, organisational capability, and governance mechanisms [2,3,18,19,24,58,59]. Readiness at the level of human capital and supporting digital infrastructure has likewise been identified as a precondition for adoption in cloud, artificial intelligence, and data-centre ecosystems, which supply the computational substrate on which persistent immersive environments depend [60].

The emergence of metaverse ecosystems, however, introduces a considerably more complex adoption context. Users are no longer evaluating isolated technologies but are engaging with persistent environments in which AI, virtual identities, immersive interaction, blockchain, and decentralised governance operate simultaneously. Consequently, adoption increasingly reflects institutional evaluations regarding governance quality, transparency, privacy protection, and organisational legitimacy alongside traditional assessments of technological capability.

This shift suggests that existing behavioural theories remain valuable but may no longer provide a sufficiently comprehensive explanation of AI adoption within metaverse ecosystems when considered independently. A broader socio-technical perspective is therefore required to integrate technological capability, governance arrangements, and institutional trust into a unified explanation of adoption.

Recent systematic reviews similarly conclude that AI adoption research remains fragmented across multiple theoretical perspectives, with comparatively limited integration between governance mechanisms, institutional trust, and post-adoption behaviour [5,27,48].

2.6 Positioning Relative to Existing Metaverse and AI Reviews

The claim that AI adoption, governance, institutional trust, and metaverse research have developed as fragmented streams requires substantiation rather than assertion, because the review literature on AI in the metaverse is already substantial. Existing reviews can be grouped into three broad families. The first comprises technology-convergence surveys that map enabling technologies, architectures, and application domains, including surveys of metaverse fundamentals, security, and privacy and of the convergence of the metaverse with blockchain and AI [61,62]. The second comprises domain-specific systematic reviews that examine AI within a single application field, most prominently education [63]. The third comprises adoption-oriented reviews that synthesise behavioural determinants of AI acceptance but are not situated in immersive or persistent environments [5].

Read together, these reviews establish the technical architecture of AI-enabled immersive environments, catalogue their security and privacy risks, and identify behavioural antecedents of acceptance. What they do not provide is an account of

the institutional pathway that links the two: how governance arrangements translate technical safeguards into stakeholder confidence, and how that confidence in turn conditions adoption. Reviews of metaverse security and privacy explicitly identify governance and regulatory compliance as under-addressed [64], while adoption reviews treat trust as an individual perception rather than as an institutional property. **Table 2** compares these reviews along the dimensions relevant to the present study.

Table 2. Comparison of representative reviews of AI and the metaverse with the present review.

Review	Scope and review type	Coverage	Governance focus	Trust focus	Metaverse focus	Principal contribution and residual gap
Wang et al. [62]	Survey of metaverse fundamentals, security, and privacy	Narrative survey	Partial (technical safeguards)	Indirect	Native	Establishes the threat landscape; institutional and adoption questions remain outside scope
Uddin et al. [61]	Convergence of metaverse, blockchain, and AI; comprehensive survey using SLR techniques	Multi-database narrative survey	Limited	Absent	Native	Maps enabling technologies and applications; does not theorise trust or adoption
Alsharhan et al. [5]	Chatbot adoption; multiperspective systematic review	Multi-database systematic review	Limited	Present, as an individual adoption antecedent	Absent	Synthesises adoption determinants across perspectives; not situated in immersive or persistent environments
Almeman et al. [63]	Integration of AI and the metaverse in education; PRISMA systematic review	Scopus and Web of Science; to August 2024; 465 records identified, 62 studies included	Absent	Limited	Native, education only	Classifies AI techniques and challenges in educational metaverse; governance and institutional trust not addressed
Rahartomo et al. [64]	Metaverse security and privacy; systematic review	114 studies, 2013–2024	Identified by the authors as an under-addressed gap	Indirect	Native	Characterises research methods and validity threats; explicitly calls for governance and regulatory analysis
The present review	AI governance, institutional trust, and AI adoption in metaverse ecosystems; PRISMA 2020 systematic review with qualitative thematic synthesis	Scopus; to 21 July 2026; 3,698 records identified, 676 eligible, 183 included	Central	Central, conceptualised as institutional trust	Central; included studies classified by metaverse relevance	Integrates four previously separate streams into the Metaverse AI Trust–Governance Nexus Framework and specifies propositions for empirical testing

The comparison indicates that the contribution of the present review is neither a further technical survey nor a further adoption review, but the integration

of governance, institutional trust, and adoption into one socio-technical account of trustworthy metaverse ecosystems. Its incremental value lies in making the institutional pathway explicit and testable rather than in extending the coverage of any single existing review.

2.7 Literature Synthesis

Taken together, the reviewed literature demonstrates that scholarship on metaverse ecosystems, artificial intelligence, governance, institutional trust, and technology adoption has expanded rapidly during recent years. Nevertheless, these topics continue to develop largely as parallel streams of research. Studies addressing AI adoption often pay limited attention to governance, governance research frequently overlooks the institutional processes through which trust develops, and metaverse research predominantly emphasises technological innovation while providing comparatively little explanation of how governance influences long-term AI adoption.

Rather than treating these themes as independent research domains, the present review synthesises them into a single socio-technical perspective. The resulting synthesis positions governance as the institutional mechanism that cultivates trust, trust as the process through which governance influences stakeholder confidence, and AI adoption as the cumulative outcome of interactions between technological capability and institutional legitimacy.

Building on this synthesis, the following sections develop the Metaverse AI Trust–Governance Nexus Framework, which integrates the fragmented evidence identified throughout the literature into a coherent conceptual explanation of trustworthy AI adoption within emerging metaverse ecosystems.

Collectively, recent review studies consistently indicate that AI adoption, institutional trust, governance, and metaverse research continue to evolve as relatively independent research streams despite their strong conceptual interdependence. Existing evidence therefore supports the development of integrative socio-technical frameworks capable of explaining how governance mechanisms cultivate institutional trust and subsequently facilitate trustworthy AI adoption [5,27,37,47,48].

3. Methodology

3.1 Research Design

Given the fragmented nature of research on AI, governance, institutional trust, and metaverse ecosystems, a Systematic Literature Review (SLR) was considered the most appropriate research design for this study. Rather than aggregating previous findings descriptively, the review was intended to integrate dispersed evidence into a coherent conceptual explanation of how governance mechanisms contribute to institutional trust and ultimately support AI adoption within trustworthy metaverse ecosystems.

To achieve this objective, the review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) guidelines, which provide a transparent and reproducible framework for identifying, screening, and selecting relevant studies. Adhering to PRISMA also reduced the risk of selection bias by applying explicit inclusion and exclusion criteria throughout the review process [65].

Because the objective of this study was conceptual integration rather than statistical effect estimation, qualitative thematic synthesis was adopted in place of

meta-analysis. This approach enabled recurring concepts, governance mechanisms, trust dimensions, AI adoption factors, and metaverse application contexts to be examined collectively, providing the analytical basis for developing the proposed conceptual framework.

3.2 Search Strategy and Study Selection

The literature search was conducted exclusively through the Scopus database because of its extensive coverage of peer-reviewed publications across information systems, management, computer science, engineering, and related interdisciplinary fields. These disciplines collectively represent the primary knowledge base for research concerning AI-enabled metaverse ecosystems.

To capture studies addressing both technological and institutional dimensions, the search strategy combined keywords related to artificial intelligence, metaverse technologies, digital identity, virtual economy, governance, institutional trust, responsible AI, and technology adoption using Boolean operators. The complete search query is presented below to support the transparency and reproducibility of the review. The search was conducted using Scopus on July 21, 2026 with the following search string.

The search combined three concept blocks with the Boolean operator AND, while synonyms within each block were combined with OR. Block A captured artificial intelligence (“artificial intelligence”, “AI”, “generative AI”); Block B captured the metaverse and its constitutive elements (“metaverse”, “digital twin”, “digital identity”, “virtual economy”); and Block C captured the institutional dimensions of interest (“trust”, “governance”, “adoption”, “responsible AI”). A record was therefore retrieved only if it matched at least one term from each of the three blocks. The grouping is stated explicitly below so that the query can be reproduced without inferring the intended precedence of operators; no restriction was placed on publication year.

(TITLE-ABS-KEY (“artificial intelligence” OR “ai” OR “generative ai”))
 AND (TITLE-ABS-KEY (“metaverse” OR “digital twin” OR “digital identity”
 OR “virtual economy”))
 AND (TITLE-ABS-KEY (“trust” OR “governance” OR “adoption” OR
 “responsible ai”))

The initial search identified 3,698 records. Sequential filtering was subsequently applied by restricting the dataset to peer-reviewed journal articles (1,032 records), English-language publications (962 records), and open-access studies (676 records). These filtering stages were conducted before full-text assessment to ensure that only studies meeting the predefined publication criteria entered the eligibility review.

Because the search was executed in a single database, no cross-database duplicate records were generated; Scopus internal duplicates were removed automatically at export and none required manual reconciliation. The three publication filters applied at the identification stage reduced the pool from 3,698 records to 1,032 journal articles, 962 English-language articles, and 676 open-access articles. The restriction to open access was applied to guarantee full-text availability for every record entering screening; its implications for selection bias are acknowledged in Section 5.4.

The 676 eligible records then entered the eligibility stage. Titles and abstracts were examined first, and records that clearly fell outside the governance–trust–adoption nexus were set aside at that point; all remaining records were assessed in full text where the full text was obtainable, and otherwise against the complete

Scopus record. Across the eligibility stage 493 records were excluded, yielding the final corpus of 183 included studies. Exclusion grounds were recorded against the five mutually exclusive criteria set out in **Table 3** and reproduced in Figure 1, each record being assigned to the first criterion it failed. Because the screening log recorded exclusions in aggregate rather than by category, the distribution of the 493 records across the five grounds is not reported here; this is acknowledged as a reporting limitation in Section 5.4.

Table 3. Eligibility criteria and grounds for exclusion at the eligibility stage.

Order	Criterion	Inclusion requirement	Ground for exclusion
1	Scope	The study addresses artificial intelligence in a digital, immersive, or organisational setting relevant to metaverse ecosystems	Outside research scope
2	AI relevance	AI is a substantive object of study rather than an incidental mention	Limited AI relevance
3	Governance, trust, or adoption content	The study addresses at least one of governance, institutional trust, responsible AI, or technology adoption	No governance, trust, or adoption perspective; engineering optimisation, algorithm development, or hardware implementation only
4	Publication and reporting criteria	Peer-reviewed journal article in English, open access, with sufficient methodological reporting to permit extraction	Did not satisfy the stated criteria
5	Extractable evidence	The study reports evidence or argument from which constructs and findings can be extracted	Insufficient empirical evidence; overlapping coverage with an included study by the same authors
—	Total excluded	—	493 records
—	Included	Meets all five criteria	183 records

Note: Criteria were applied in the order listed and each record was assigned to the first criterion it failed. The five grounds correspond to the exclusion reasons reported in Figure 1.

Following the PRISMA 2020 protocol, the remaining publications were evaluated through four sequential stages: identification, screening, eligibility assessment, and inclusion. During the eligibility assessment, each article was examined against predefined inclusion and exclusion criteria. Studies that primarily addressed engineering optimisation, algorithm development, hardware implementation, or unrelated digital technologies without explicit discussion of governance, institutional trust, responsible AI, or technology adoption were excluded. This process resulted in a final dataset of 183 peer-reviewed journal articles, which formed the basis for the subsequent qualitative synthesis. The complete study selection process is illustrated in **Figure 1**.

The full record set is reported in **Supplementary Table S1**, which lists all 183 included studies with their bibliographic details and Digital Object Identifiers, so that the corpus can be reconstructed independently. The reference list of this article contains the works cited in the text; it is not coextensive with the review corpus, and Supplementary Table S1 should be consulted for the complete set of included studies.

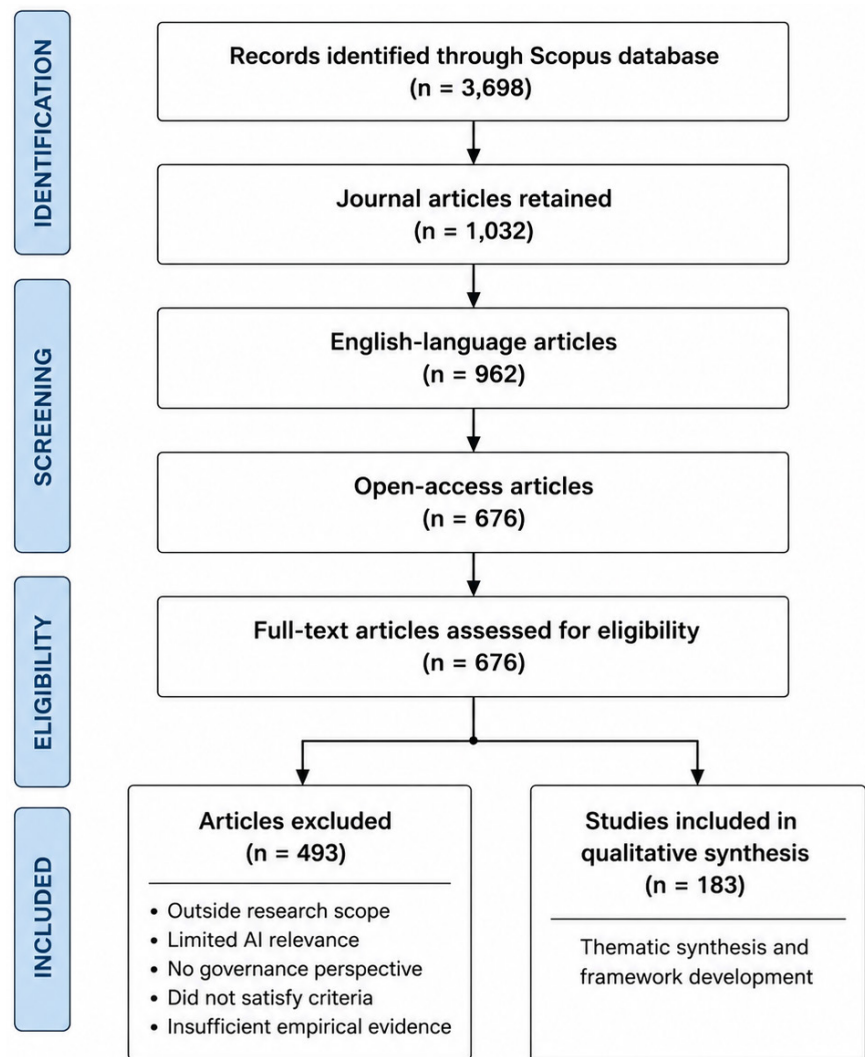


Figure 1. PRISMA flow

3.3 Data Extraction and Analysis

Following study selection, the retained articles were analysed using a structured extraction protocol designed to ensure consistency across the review. For each publication, information relating to bibliographic characteristics, research objectives, methodological approach, theoretical foundation, application context, governance mechanisms, institutional trust, AI adoption determinants, principal findings, and reported research gaps was systematically documented.

The analytical process relied on iterative qualitative coding rather than quantitative synthesis. Open coding was first employed to identify recurring concepts across individual studies, after which related concepts were grouped through axial coding into broader analytical categories. Throughout this process, particular attention was given to recurring relationships among governance mechanisms, institutional trust, AI adoption, and metaverse ecosystems.

The coding procedure was as follows. A preliminary codebook was developed inductively from a pilot subsample of the corpus and specified, for each first-order code, a label, a definition, an inclusion rule, and an anchor example. Two authors coded the pilot subsample independently; discrepancies were listed, discussed, and resolved by consensus, with a third author adjudicating where consensus could not be reached, after which the codebook was finalised and applied to the remaining

articles. Because the analytical objective was interpretive synthesis rather than classification into mutually exclusive categories, agreement was established through documented consensus rather than through a reliability coefficient; this choice and its consequences are acknowledged in Section 5.4. All coding decisions, together with the codebook and the audit trail of code merges, were maintained in a shared extraction workbook.

Extraction depth was not uniform across the corpus, and this is reported rather than smoothed over. The full text was obtainable and read for 19 of the 183 included studies; for the remaining 164, extraction relied on the complete Scopus record, comprising the title, author keywords, index keywords, and abstract. Methodological design and theoretical foundation could therefore be recorded reliably only for the subset with accessible full text, whereas thematic assignment and construct identification for the remainder rest on abstract-level evidence. **Supplementary Table S1** records the extraction depth of every included study so that each entry can be weighed accordingly, and Section 5.4 discusses the consequences for the strength of the synthesis.

First-order codes were consolidated into second-order sub-themes and subsequently into the four aggregate themes through constant comparison. **Table 4** illustrates this trail for one code within each theme, showing how descriptive codes taken from individual studies were abstracted into the categories that structure Section 4.3.

Table 4. Illustrative coding trail from first-order codes to aggregate themes.

Examples of first-order codes	Second-order sub-theme	Aggregate theme
“intelligent virtual assistant improves service quality”; “generative agent produces adaptive content”; “AI chatbot capability”; “real-time decision support in digital twin”	AI as adaptive service capability	Theme 1: AI as the enabling foundation
“algorithmic transparency requirement”; “explainability obligation”; “human oversight of autonomous agents”; “decentralised autonomous organisation as governance form”; “self-sovereign identity as safeguard”	Governance as institutional safeguard	Theme 3: Governance mechanisms
“confidence in the deploying organisation”; “perceived institutional accountability”; “trust calibration under uncertainty”; “willingness to relinquish control to an agent”	Institutional rather than interpersonal trust	Theme 2: Trust formation
“behavioural intention to use”; “continuance intention”; “perceived risk reduces use”; “AI literacy as enabling condition”; “organisational readiness”	Adoption as socio-technical outcome	Theme 4: AI adoption

Rather than evaluating studies independently, the analysis emphasised conceptual convergence and divergence across the literature. This comparative process enabled fragmented findings originating from different disciplines to be synthesised into a unified socio-technical interpretation of trustworthy AI-enabled metaverse ecosystems.

To complement the systematic literature review, bibliometric analysis was conducted using Bibliometrix (Biblioshiny) and VOSviewer. Bibliometrix was employed to generate descriptive bibliometric indicators and summarise publication characteristics, whereas VOSviewer was used to construct keyword co-occurrence networks and overlay visualisations for exploring the conceptual structure and temporal evolution of the research field [66]. The integration of these complementary

tools enables both quantitative bibliometric assessment and graphical visualisation of thematic relationships, thereby providing a robust analytical foundation for subsequent qualitative thematic synthesis. Keyword co-occurrence analysis was selected because it identifies the conceptual relationships among frequently occurring author keywords, while network visualisation reveals thematic clusters and overlay visualisation illustrates the temporal evolution of research topics. This combined approach has been widely adopted in recent bibliometric studies to identify research hotspots, intellectual structures, and emerging research trends across scientific domains. Reviews that combine PRISMA 2020 screening with Bibliometrix and Biblioshiny analysis of Scopus-derived corpora provide a direct methodological precedent for the procedure adopted here [67].

Three analytical layers are distinguished, and each figure and table in Section 4 states the corpus on which it is based. Descriptive publication analysis (annual output, document type, and source distribution; **Figures 2–4**) was performed on the full set of 3,698 identified records in order to characterise the field as a whole, which is why non-journal formats such as conference papers, book chapters, and reviews appear in those figures. Keyword co-occurrence and overlay mapping (**Figures 5–6**) was performed on the 676 eligible open-access journal articles. The qualitative thematic synthesis reported in Sections 4.3 to 4.5, and every substantive claim derived from it, rests exclusively on the 183 included studies. No bibliometric result is used as evidence for a thematic or framework relationship.

3.4 Conceptual Framework Development

The Metaverse AI Trust–Governance Nexus Framework emerged through an iterative process of conceptual synthesis grounded in the reviewed literature rather than through statistical modelling or deductive hypothesis development. Framework construction therefore reflected patterns consistently identified across the selected studies instead of relying on a single theoretical perspective.

The development process consisted of four successive stages. First, recurring concepts were identified through thematic coding. Second, related themes were integrated to establish broader conceptual categories. Third, relationships among these categories were examined to identify recurring governance–trust–adoption patterns across the literature. Finally, these relationships were abstracted into a higher-order socio-technical framework explaining how governance mechanisms foster institutional trust and subsequently facilitate AI adoption within trustworthy metaverse ecosystems.

By grounding the framework directly in evidence synthesised from 183 peer-reviewed studies, the resulting model represents an integrative conceptual interpretation of the current literature rather than a normative or speculative proposition. This framework provides the analytical foundation for the thematic findings presented in the following section.

4. Results

4.1 Overview of the Selected Studies

The systematic literature search was conducted using the Scopus database following the PRISMA 2020 review protocol described in the Methods section. The initial search, based on keywords related to artificial intelligence, metaverse, digital identity, virtual economy, governance, trust, responsible AI, and technology adoption, identified 3,698 records. Successive screening procedures refined the dataset to 1,032

journal articles, 962 English-language publications, and 676 open-access studies before thematic screening produced a final corpus of 183 studies for qualitative synthesis. This sequential filtering ensured that only studies explicitly addressing the intersection of AI, governance, institutional trust, technology adoption, and metaverse ecosystems were included.

As illustrated in Figure 1, the PRISMA-guided screening process resulted in the inclusion of 183 peer-reviewed journal articles from an initial dataset of 3,698 records. These studies formed the basis for the subsequent publication analysis, thematic synthesis, and conceptual framework development. Study selection followed the PRISMA framework through identification, screening, eligibility assessment, and inclusion. Eligibility was determined according to conceptual relevance rather than keyword occurrence alone. Studies focusing primarily on technical topics, such as digital twins, IoT, or sector-specific AI applications without substantive discussion of governance, institutional trust, or AI adoption, were excluded. Publications with insufficient methodological information, overlapping review coverage, or limited relevance to the governance–trust–adoption nexus were also omitted. Consequently, the final dataset represents a focused body of literature directly examining governance, trust, and AI adoption within AI-enabled metaverse ecosystems.

Two features of the corpus require comment before the thematic findings are presented, because they bear directly on how far the review can speak to metaverse ecosystems specifically. The systematic search was constructed around the constitutive elements of metaverse environments rather than the term alone, so that Block B combined “metaverse” with “digital twin”, “digital identity”, and “virtual economy”. This decision was deliberate and follows the convergence view of the metaverse advanced in the reviewed literature, in which persistent virtual representations, portable identity, and tokenised assets are the constitutive infrastructure of immersive ecosystems rather than adjacent technologies [17,27,61,62]. Its consequence, however, is that digital twin research dominates the retrieved corpus, and studies using the word “metaverse” in their titles are comparatively few.

The review therefore distinguishes two evidential functions within its reference base, and states which is doing the work at each point. The 183 included studies constitute the systematically retrieved corpus and are the sole basis of the thematic synthesis reported in Sections 4.3 to 4.5. Alongside these, the Introduction, Literature Review, and Discussion cite contextual literature that was not retrieved by the search protocol: classical adoption and trust theory, the comparator reviews examined in Section 2.6, and metaverse-specific adoption studies drawn from the immersive-environments literature. This is a standard and necessary feature of a systematic review rather than a departure from protocol, since background, theory, and comparison cannot be confined to a corpus assembled for a different purpose. The distinction is nonetheless made explicit here so that no reader mistakes a contextual citation for corpus evidence, and the evidential category of every framework relationship is recorded in Section 4.5 for the same reason.

The metaverse-specific adoption literature is particularly relevant to Theme 4 and to proposition P3, and warrants brief treatment because it is where the constructs of this review have been tested closest to immersive settings. A recent structural study of 200 Spanish university students found perceived usefulness to be the strongest predictor of intention to use the metaverse, with result demonstrability the dominant antecedent of usefulness and perceived enjoyment and perception of external control the antecedents of ease of use [68]. Its finding that demonstrability drives

usefulness is consistent with the account offered here, in which adoption in immersive environments depends on the visible and attributable consequences of institutional arrangements rather than on capability alone. Evidence from smart-city service adoption points in the same direction from the community side, identifying data security, e-safety, and e-literacy as the principal brakes on adoption and institutional incentive design as its principal enabler [69]. Both are consistent with R2 and R3 while testing neither pathway in full.

Two further observations follow. First, the concentration of metaverse adoption research in educational settings, noted in Section 2.6, means that the organisational and public-sector governance questions central to this review remain largely untested in immersive environments; the present framework is directed at that gap. Second, the propositional form adopted here has precedent in the digital adoption literature, where conceptual frameworks are advanced as sets of propositions for subsequent mixed-method testing rather than as validated models [70], and behavioural adoption research continues to rely on the technology acceptance model as the baseline against which extensions are assessed [71]. The framework proposed in Section 4.5 is offered in that tradition.

Because the corpus spans immersive environments and adjacent digital contexts, each included study was classified according to the directness of its relationship to metaverse settings. Three categories were used. Metaverse-native studies investigate persistent immersive environments, avatar-mediated interaction, virtual worlds, virtual economies, or digital twins operated as persistent virtual environments. Metaverse-adjacent studies examine AI in digital platforms that exhibit at least one constitutive metaverse property, such as persistent digital identity, autonomous agents acting on a user's behalf, immersive or embodied interaction, or decentralised asset ownership. Foundational studies examine AI adoption, governance, or trust in conventional organisational settings such as healthcare, banking, education, tourism, or manufacturing.

Applying this classification to the corpus yields 22 metaverse-native studies (12.0%), 106 metaverse-adjacent studies (57.9%), and 55 foundational studies (30.1%). The metaverse-native group comprises research conducted explicitly in immersive or avatar-mediated settings, including work on avatar embodiment and digital selfhood [11,72], metaverse threat and defence architectures [37,73], metaverse governance for smart cities [8,23], and the industrial metaverse [20,30]. The metaverse-adjacent group is dominated by digital twin research, in which persistent virtual representations of physical assets are governed and trusted in ways closely analogous to metaverse environments [4,55,74,75], together with work on decentralised identity, blockchain provenance, and zero-trust architectures [6,76–78]. The foundational group comprises studies of AI adoption, governance, and responsible practice in conventional organisational settings [58,59,79].

The classification governs how evidence is used throughout the review. Claims about metaverse ecosystems specifically are supported by metaverse-native and metaverse-adjacent studies; foundational studies are used only to define constructs, to establish theoretical lineage, and to identify boundary conditions, and are never treated as direct evidence of metaverse phenomena. The distribution of the corpus is reported in **Table 5**, and the evidential category of each proposition is stated explicitly in the evidence-to-framework mapping presented in Section 4.5.

Two features of this distribution warrant comment because they qualify the conclusions that can be drawn. First, only 22 studies are metaverse-native, so the review's metaverse-specific claims rest on a comparatively narrow evidential base

and are strengthened by analogy from digital twin research rather than established independently. Second, extraction depth was uneven: the full text was obtainable and read for 19 studies, whereas for the remaining 164 extraction relied on the Scopus record, comprising the title, author keywords, index keywords, and abstract. Constructs, methodologies, and theoretical foundations could therefore be recorded reliably only for the subset with accessible full text, and thematic assignment for the remainder was made from abstract-level evidence. **Supplementary Table S1** characterises all 183 included studies, reporting author, year, title, source, Digital Object Identifier, study context, metaverse relevance category, thematic assignment, and the depth at which each study could be extracted. Fields that depend on full-text access, principally research design and theoretical foundation, are populated for the studies whose full text was obtainable and are marked as not extractable for the remainder, because reporting them from abstracts alone would misrepresent their provenance.

Table 5. Distribution of the 183 included studies by metaverse relevance and thematic category.

Metaverse relevance	Assignment rule (title and author keywords)	Studies, n (%)	Full text read, n	Use in the synthesis
Metaverse-native	Metaverse, avatar, virtual world, immersive, extended/virtual/augmented/mixed reality, virtual environment, or digital human	22 (12.0)	6	Direct evidence for metaverse-specific claims and for all framework relationships
Metaverse-adjacent	Not native, but exhibits at least one constitutive metaverse property: digital twin, digital identity, self-sovereign identity, blockchain, zero trust, decentralisation, federated learning, autonomous or conversational agent, or virtual asset	106 (57.9)	11	Supporting evidence for metaverse-specific claims, with the shared property stated in each case
Foundational	Neither of the above: AI adoption, governance, or trust in conventional organisational settings	55 (30.1)	2	Construct definition, theoretical lineage, and boundary conditions only
Total	—	183 (100.0)	19	—

Note: Categories are mutually exclusive and were assigned in the order listed. Percentages are of the 183 included studies. “Full text read” reports the number of studies in each category for which the full text was obtainable; for the remainder, extraction relied on the Scopus record (title, author keywords, index keywords, and abstract). Study-level detail is given in Supplementary Table S1.

The characteristics of the selected publications demonstrate the rapid emergence of this research field. Most studies originated from interdisciplinary areas including information systems, computer science, engineering, management, and the social sciences, illustrating the multidisciplinary nature of trustworthy metaverse research. Publication patterns further indicate that scholarly attention has accelerated in parallel with advances in generative AI, immersive digital environments, and responsible AI

4.2 Publication Characteristics

The publication profile demonstrates how rapidly scholarly interest in AI-enabled metaverse ecosystems has expanded during recent years. Although only limited research appeared before 2020, publication activity increased substantially from 2022 onwards, reflecting growing academic attention to generative AI, immersive virtual environments, institutional trust, digital governance, and responsible AI. As shown

in **Figure 2**, this upward trajectory suggests that trustworthy metaverse ecosystems have become an emerging interdisciplinary research domain rather than a niche technological topic.

Documents by year

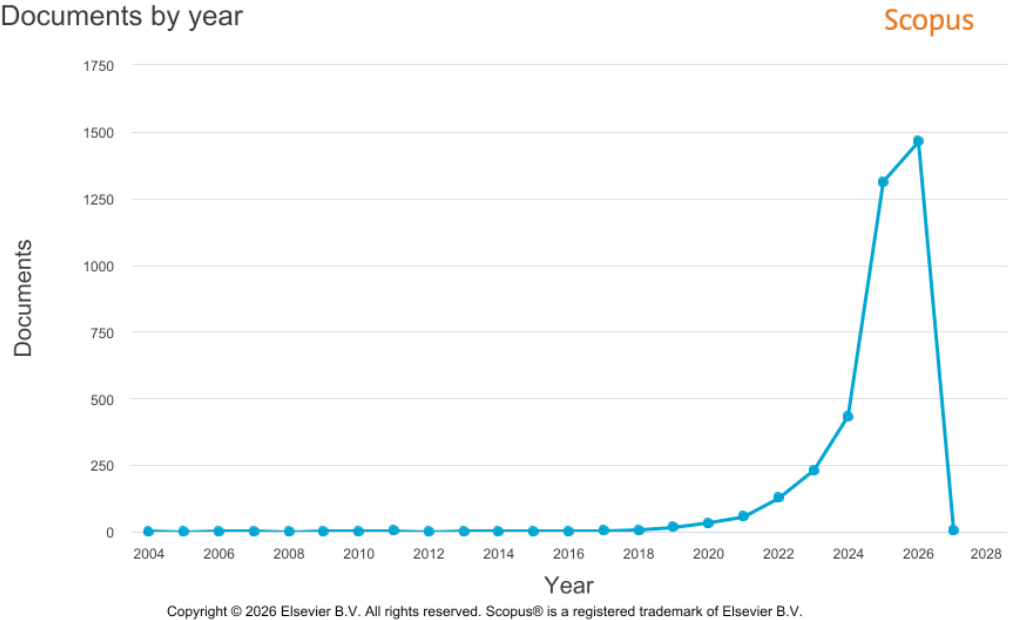


Figure 2. Documents by year (n = 3,698 identified records)

The broader publication landscape further highlights the developmental stage of the field. Journal articles represented approximately 31.3% of the initial search results, followed by conference papers (24.5%), book chapters (18.0%), and review articles (16.1%), with the remaining publications consisting of conference reviews, books, editorials, notes, errata, and short surveys (**Figure 3**). Although the qualitative synthesis focused exclusively on peer-reviewed journal articles, the diversity of publication formats indicates that many new ideas continue to emerge through conferences before reaching archival journals.

Documents by type

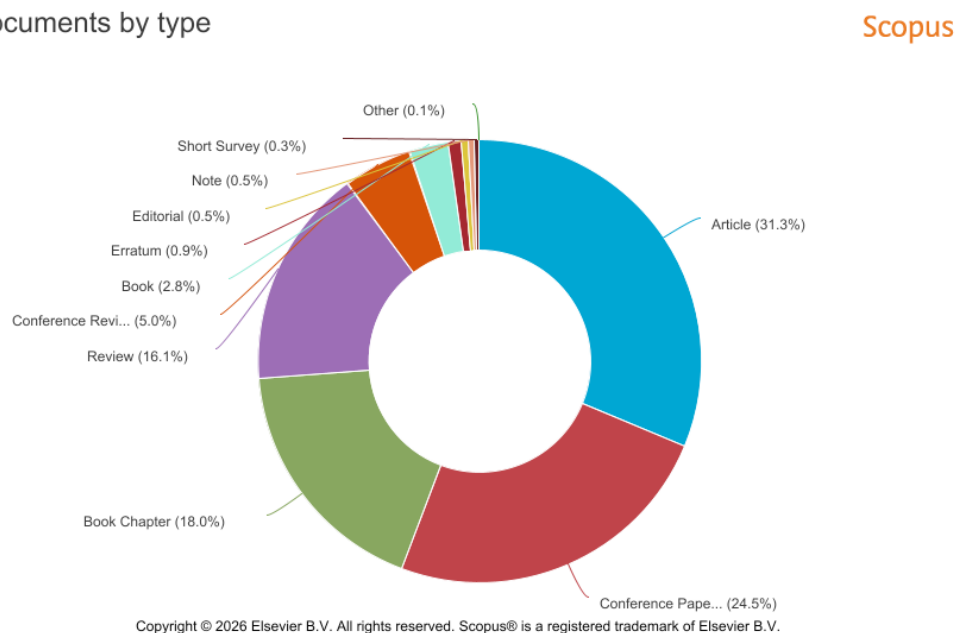


Figure 3. Documents by type (n = 3,698 identified records)

The distribution of publication outlets also reflects the interdisciplinary nature of the field. Major sources include *Sustainability*, *IEEE Access*, *Lecture Notes in Networks and Systems*, *Communications in Computer and Information Science*, and *Lecture Notes in Computer Science* (Figure 4). Collectively, these venues demonstrate that discussions concerning trustworthy metaverse ecosystems increasingly integrate technological innovation with organisational, behavioural, managerial, and governance perspectives.

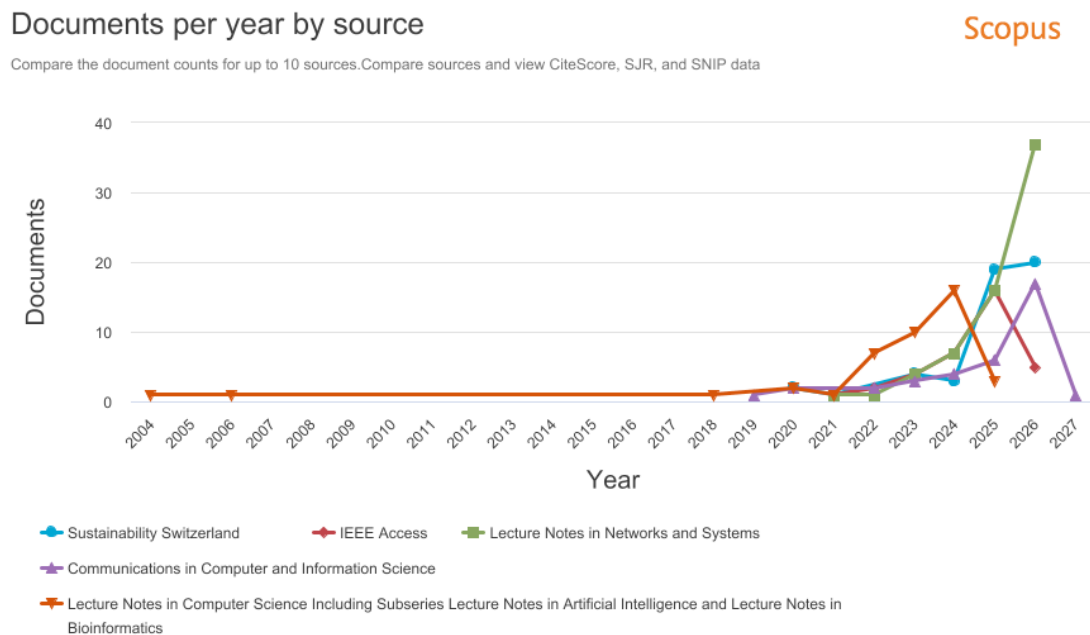


Figure 4. Documents per year by source (n = 3,698 identified records)

To provide an initial overview of the conceptual structure of the selected literature, bibliometric mapping was performed using Bibliometrix (Biblioshiny) and VOSviewer based on author keyword co-occurrence. As illustrated in **Figures 5 and 6**, the network visualisation reveals that the literature has evolved into several interconnected thematic clusters rather than isolated research domains. Among the most influential keywords, *artificial intelligence*, *digital twin*, *machine learning*, *Internet of Things*, *network security*, *blockchain*, *metaverse*, *decision making*, and *sustainable development* occupy central positions within the network, indicating their pivotal role in connecting technological, organisational, and application-oriented research streams. The extensive interconnections among these keywords suggest that contemporary metaverse research has increasingly adopted a multidisciplinary perspective that integrates artificial intelligence with digital infrastructure, cybersecurity, intelligent decision support, and socio-technical applications.

The overlay visualisation further demonstrates the temporal evolution of research themes. Earlier studies predominantly focused on foundational technologies such as the Internet of Things, machine learning, cybersecurity, blockchain, and digital twins, whereas more recent publications increasingly emphasise artificial intelligence, generative AI, digital identity, federated learning, data privacy, predictive maintenance, and intelligent human-centred applications. This transition reflects a broader shift in the literature from technology-oriented innovation toward governance-oriented and application-driven research that addresses the implementation of

trustworthy AI within complex digital ecosystems. These bibliometric findings provide an empirical foundation for the thematic synthesis presented in the following sections by demonstrating how the research landscape has progressively converged around the relationships between AI capability, governance, institutional trust, and AI adoption

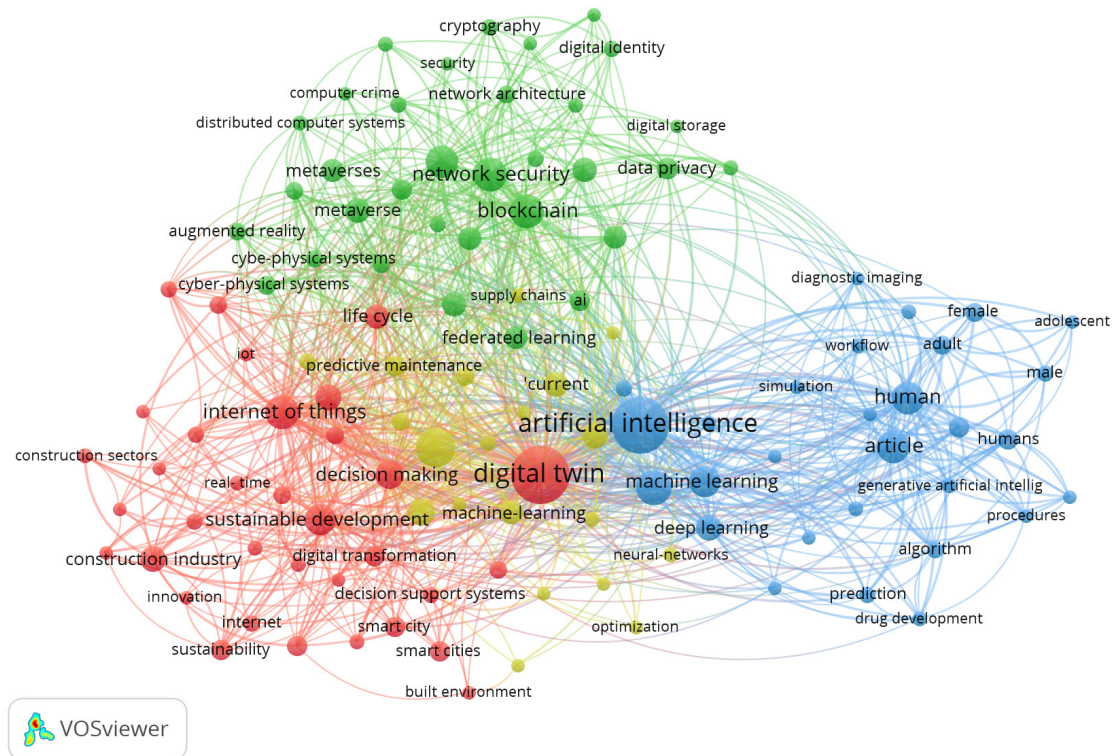


Figure 5. Keyword co-occurrence network (n = 676 eligible articles)

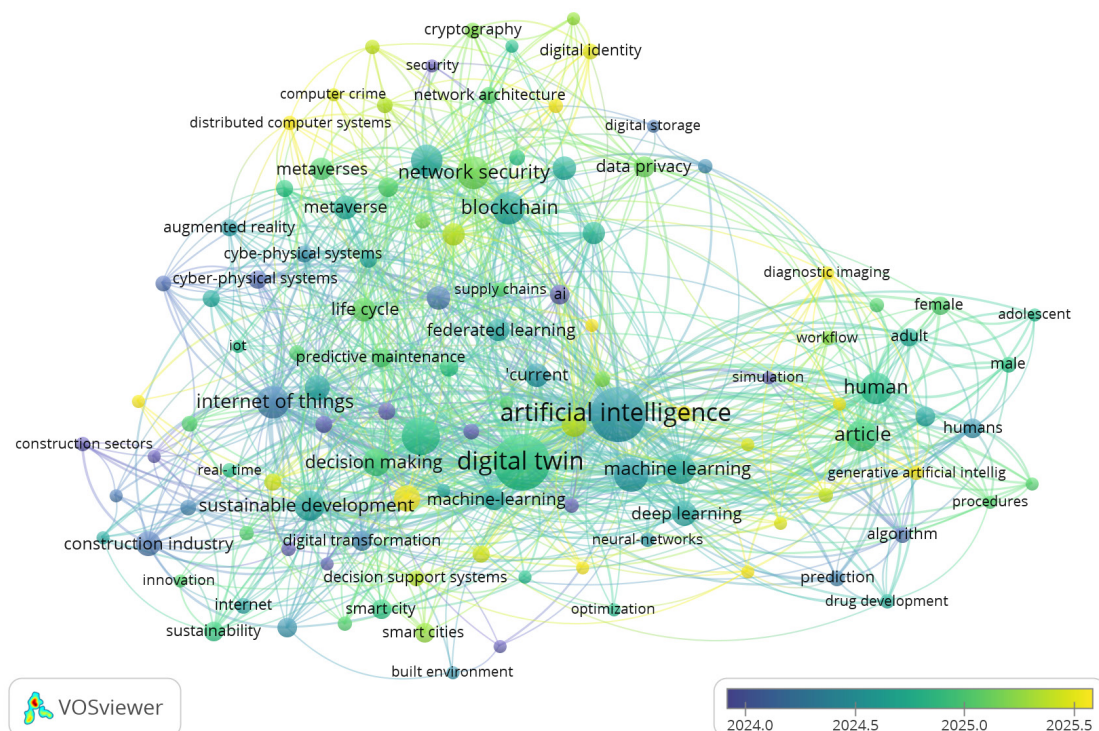


Figure 6. Keyword overlay visualisation (n = 676 eligible articles)

4.3 Thematic Findings

Qualitative coding of the 183 selected studies revealed four closely connected themes that together explain how AI-enabled metaverse ecosystems evolve beyond purely technological environments toward governance-oriented socio-technical systems. Rather than representing isolated research topics, these themes collectively address the relationships among AI capability, governance, institutional trust, and AI adoption.

The four themes are reported in the sequence in which they emerged from the coding process. Section 4.4 and the framework developed in Section 4.5 subsequently arrange them according to the socio-technical logic identified across the reviewed studies, namely AI capability, governance mechanisms, institutional trust, and AI adoption.

Table 6 summarises the principal empirical evidence synthesised from the reviewed studies and illustrates how the four themes identified in this review are consistently supported across different theoretical perspectives, application domains, and methodological approaches. Although the reviewed studies employ diverse theoretical foundations—including TAM, UTAUT, UTAUT2, TOE, TPB, Social Support Theory, and Trust Theory—they collectively indicate a gradual transition from technology-centric explanations toward broader socio-technical perspectives that integrate governance, institutional trust, and organisational capability. This convergence provides empirical support for the Metaverse AI Trust–Governance Nexus Framework proposed in this study.

Table 6. Summary of Empirical Evidence Supporting the Four Thematic Findings

Theme	Representative Evidence	Dominant Theory	Emerging Constructs / Novelty	Implications for This Review
AI as the Enabling Foundation	AI functions as the enabling capability of immersive and persistent environments, supporting intelligent agents and non-player characters, natural language interaction, computer vision, adaptive content generation, and real-time decision support across education, industry, and urban applications [2–4,17,18,27,29,80].	Socio-technical systems perspective; AI taxonomies for immersive environments	Intelligent agents and NPCs; digital twins; multimodal interaction; real-time analytics	AI has evolved from an automation tool into the core enabling capability of metaverse ecosystems.
Institutional Trust	Trust is established through verifiable institutional and technical mechanisms rather than technological capability alone, including reputation and trust-management frameworks for virtual service providers, traceable federated learning, and trust-by-design architectures [26,47,52–55,81].	Trust management theory; zero-trust security paradigm; trust-by-design	Reputation scoring; proof of humanity; traceability; mutual human–machine trust	Institutional trust functions as the mechanism linking governance and AI adoption.
Governance Mechanisms	Governance operates across multiple scales and actors, combining lifecycle accountability for AI models, multi-scale coordination between public institutions and community platforms, and safeguards against corporate-led capture of public authority [22,23,35,40,42,47,48].	Multi-scale governance; stakeholder theory; responsible AI lifecycle management	Multi-scale coordination; corporate-led governance risk; model lifecycle accountability	Governance represents an organisational capability rather than merely regulatory compliance.

Continuation Table:

Theme	Representative Evidence	Dominant Theory	Emerging Constructs / Novelty	Implications for This Review
AI Adoption	Adoption reflects behavioural, organisational, and institutional conditions jointly, including perceived usefulness and risk, avatar appearance and presence, organisational readiness, workforce capability, and public-sector enablers [19,24,44,58,59,82–84].	TAM; UTAUT; TOE; multi-case organisational analysis	Avatar appearance; perceived risk; organisational readiness; workforce capability	AI adoption increasingly reflects a socio-technical process rather than a purely behavioural response.

Theme 1. AI as the Enabling Foundation of Metaverse Ecosystems

Across the reviewed publications, AI consistently appears as the technological capability that enables metaverse ecosystems to operate as intelligent and adaptive environments. Beyond conventional automation, AI supports intelligent virtual agents, natural language interaction, adaptive content generation, computer vision, digital humans, and personalised services that enhance immersive user experiences.

Several studies also emphasise AI's contribution to digital identity management, blockchain-enabled virtual economies, digital twins, and real-time decision support. Together, these capabilities position AI as the technological infrastructure underpinning interactions within emerging metaverse ecosystems.

At the same time, the reviewed evidence highlights an important tension. Increasing AI capability is accompanied by growing concerns regarding algorithmic opacity, privacy protection, digital identity security, and institutional accountability. These recurring concerns suggest that technological advancement alone cannot ensure sustainable ecosystem development.

Theme 2. Trust Formation in AI-enabled Metaverse Ecosystems

A consistent pattern across the reviewed studies is that institutional trust develops through governance quality rather than technological sophistication alone. Users appear more willing to engage with AI-enabled environments when governance arrangements demonstrate transparency, accountability, privacy protection, and responsible AI practices.

Digital identity governance represents one of the strongest trust-building mechanisms identified throughout the literature. Approaches such as Self-Sovereign Identity (SSI), Proof of Humanity, blockchain-based authentication, Zero Trust Architecture (ZTA), and Federated Learning are frequently presented as institutional safeguards that strengthen confidence while reducing uncertainty associated with immersive digital interaction.

Nevertheless, trust remains fragile. Greater AI autonomy, behavioural data collection, and algorithmic decision-making continue to generate concerns regarding fairness, explainability, and long-term institutional accountability. Consequently, governance emerges not simply as a regulatory requirement but as a prerequisite for maintaining stakeholder confidence.

Theme 3. Governance Mechanisms for Trustworthy Metaverse Ecosystems

Governance occupies a central position throughout the reviewed literature by providing the institutional structures through which trustworthy AI-enabled metaverse ecosystems can develop. Rather than focusing exclusively on regulatory compliance,

governance encompasses technical safeguards, organisational arrangements, ethical principles, collaborative institutions, and accountability mechanisms that collectively shape stakeholder confidence.

Responsible AI principles—including transparency, explainability, fairness, accountability, privacy protection, human oversight, and security—appear consistently across different application contexts. The reviewed studies also point to increasing interest in decentralised governance involving governments, technology providers, standards organisations, industry, and users. Emerging mechanisms such as Decentralised Autonomous Organisations (DAOs) illustrate this transition toward more collaborative governance models.

Despite this growing body of work, governance dimensions are often examined independently. Comparatively few studies explain how governance mechanisms jointly cultivate institutional trust that ultimately supports sustainable AI adoption.

Theme 4. AI Adoption as an Outcome of Governance and Trust

The reviewed evidence suggests that AI adoption within metaverse ecosystems is better understood as the outcome of an institutional process than as a simple behavioural response to technological capability. While AI clearly enhances efficiency, automation, personalisation, and intelligent interaction, adoption decisions simultaneously reflect perceptions of governance quality, transparency, privacy protection, explainability, and institutional legitimacy.

Institutional trust consistently emerges as the mechanism connecting governance with adoption. Strong governance arrangements strengthen stakeholder confidence, whereas weaknesses in governance reduce trust and consequently diminish willingness to engage with AI-enabled metaverse services.

An additional insight emerging from the synthesis is the AI Adoption Paradox. Technological progress simultaneously creates opportunities for innovation while increasing concerns regarding surveillance, algorithmic bias, privacy, explainability, and institutional accountability. Sustainable adoption therefore depends not only on continued technological advancement but equally on governance mechanisms capable of maintaining institutional legitimacy.

4.4 Cross-theme Synthesis

Examining the four themes collectively reveals a coherent socio-technical process rather than a collection of independent observations. AI provides the technological capability that enables immersive digital environments, yet increasing technological sophistication also introduces new governance challenges related to privacy, accountability, transparency, digital identity, and ethical responsibility.

Governance mechanisms respond to these challenges by establishing institutional safeguards that reduce uncertainty surrounding AI-enabled interactions. Through these governance arrangements, institutional trust gradually develops and subsequently influences stakeholders' willingness to adopt AI within metaverse ecosystems.

Perhaps the most significant observation arising from the cross-theme comparison is the recurring AI Adoption Paradox. Across multiple contexts, the same technological innovations that expand AI capability also amplify governance concerns. This recurring pattern suggests that governance and technological innovation should not be viewed as competing priorities. Instead, sustainable AI adoption depends on their continuous co-evolution.

4.5 Proposed Metaverse AI Trust–Governance Nexus Framework

The thematic synthesis culminates in the development of the Metaverse AI Trust–Governance Nexus Framework, which integrates the recurring relationships identified across the reviewed studies into a single socio-technical perspective.

Within the framework, AI capability forms the technological foundation supporting intelligent virtual agents, adaptive interaction, digital identity management, personalised services, and automated decision-making. Governance mechanisms constitute the institutional layer that incorporates transparency, explainability, accountability, privacy protection, security architecture, digital identity governance, and human oversight. Institutional trust subsequently operates as the mechanism through which governance arrangements are translated into stakeholder confidence, while AI adoption represents the resulting socio-technical outcome.

As shown in **Figure 7**, the framework also provides a conceptual explanation for the AI Adoption Paradox identified throughout the review. As AI capability expands, governance becomes increasingly important rather than less necessary because greater technological autonomy simultaneously increases institutional expectations regarding transparency, accountability, and responsible AI.

Taken together, the proposed framework integrates previously fragmented streams of research into a coherent explanation of trustworthy AI adoption within metaverse ecosystems. It also establishes the conceptual basis for the theoretical and practical implications discussed in the following section.

Each relationship depicted in **Figure 7** is specified below in terms of its direction, its substantive meaning, and its evidential status. R1 runs from AI capability to governance mechanisms and is a demand relationship: as autonomy, personalisation, and behavioural data capture increase, the institutional requirements placed on the deploying organisation increase with them, so that capability generates governance obligations rather than displacing them. R2 runs from governance mechanisms to institutional trust and is an enabling relationship: transparency, explainability, accountability, privacy protection, identity governance, and human oversight reduce the uncertainty that stakeholders face and thereby render the deploying institution creditworthy. R3 runs from institutional trust to AI adoption and is a conversion relationship: confidence in the governing institution is what permits stakeholders to accept exposure to autonomous systems whose behaviour they cannot fully verify. R4 runs directly from governance mechanisms to AI adoption and captures the compliance and assurance pathway through which certification, audit, and regulatory conformity affect adoption decisions without passing through subjective confidence, and is most visible in organisational rather than individual adoption. R5 runs from AI adoption back to AI capability and is a feedback relationship: sustained use generates the interaction data, revenue, and institutional learning that fund further capability, closing the loop and producing the co-evolution described in Section 4.4.

The AI Adoption Paradox is represented in **Figure 7** as the tension between R1 and R5. Because adoption feeds capability and capability raises governance demand, each cycle of successful adoption increases rather than reduces the governance burden. The paradox is therefore a structural property of the system depicted, not a transitional condition that maturity will dissolve.

Two clarifications are necessary regarding the conceptual status of these relationships. First, the framework distinguishes relationships that are directly evidenced in the reviewed literature from relationships that the authors have

synthesised across studies that did not themselves connect the constructs. R2 and R3 are each supported by studies that examine the relationship directly, whereas the composition of R2 and R3 into a single pathway from governance through trust to adoption, together with the feedback relationship R5, is an authors' synthesis: no reviewed study tests the complete pathway. **Table 7** records this distinction for every relationship. Second, the depiction of institutional trust as the mechanism linking governance to adoption summarises a recurrent interpretive pattern across heterogeneous qualitative, conceptual, and survey-based studies. It is not the outcome of a meta-analytic mediation test, no pooled indirect effect is estimated, and no claim of statistical mediation is intended. The arrow should be read as the statement that the reviewed literature explains the governance-adoption relationship principally through institutional trust, not as a verified causal path. Accordingly, the relationships are stated below as propositions for empirical testing rather than as findings.

P1. Increases in AI capability within metaverse ecosystems increase, rather than diminish, the institutional governance requirements placed on deploying organisations. P2. Governance mechanisms that operationalise transparency, explainability, accountability, privacy protection, and identity governance are positively associated with institutional trust. P3. Institutional trust is positively associated with AI adoption in metaverse ecosystems and accounts for a substantial part of the association between governance and adoption. P4. Governance also relates to adoption directly through compliance and assurance, and this direct pathway is more pronounced for organisational than for individual adoption. P5. Adoption reinforces AI capability over time, so that governance demand and technological capability co-evolve rather than converge to equilibrium.

Table 7. Evidence-to-framework mapping: relationships, evidential basis, and strength of support.

Relationship (proposition)	Representative supporting evidence	Evidential basis	Convergence and divergence	Strength of support
R1: AI capability → governance demand (P1)	Studies reporting that greater autonomy, personalisation, and behavioural data capture intensify concerns over opacity, privacy, and accountability [25,37,38,64]	Metaverse-native (n = 22) and metaverse-adjacent	Convergent in direction; divergent on whether the resulting demand is regulatory, organisational, or architectural	Moderate; consistent across the corpus but largely argued rather than measured
R2: Governance → institutional trust (P2)	Studies relating organisational transparency, accountability, oversight, and identity safeguards to stakeholder confidence [11,15,26,47,49,52]	Predominantly metaverse-adjacent (digital twin and decentralised identity research), with limited metaverse-native support	Convergent on transparency and accountability; divergent on whether governance operates through perceived competence or perceived integrity	Moderate to strong in adjacent domains; requires testing in immersive settings
R3: Institutional trust → AI adoption (P3)	Survey-based studies in which trust and perceived risk predict intention and continued use [12,13,24,44,59,83]	Mixed: one metaverse-native acceptance study, otherwise foundational	Convergent on sign and significance; divergent on whether trust operates as antecedent, mediator, or moderator	Moderate for the association; unverified as a mediation pathway and thinly evidenced in immersive settings

Continuation Table:

Relationship (proposition)	Representative supporting evidence	Evidential basis	Convergence and divergence	Strength of support
R4: Governance → AI adoption, direct (P4)	Studies in which certification, compliance, and institutional legitimacy affect organisational adoption independently of individual perceptions [21,23,48,84]	Foundational and metaverse-native governance studies	Weakly convergent; rarely modelled alongside the trust pathway	Weak; proposition for future testing
R5: AI adoption → AI capability, feedback (P5)	Inferred from longitudinal and co-evolutionary accounts of platform development across the corpus	Authors' synthesis; not tested in any reviewed study	Not applicable	Proposition only; no direct empirical support
Composite pathway: governance → trust → adoption	No reviewed study tests the full pathway; assembled from evidence for R2 and R3 separately	Authors' synthesis	Not applicable	Proposition only; the central claim requiring empirical validation

Note: Evidential basis follows the classification defined in Section 4.1 and reported in Table 5. Relationships marked as authors' synthesis are conceptual propositions derived from the thematic findings and are not empirically established by the reviewed literature.

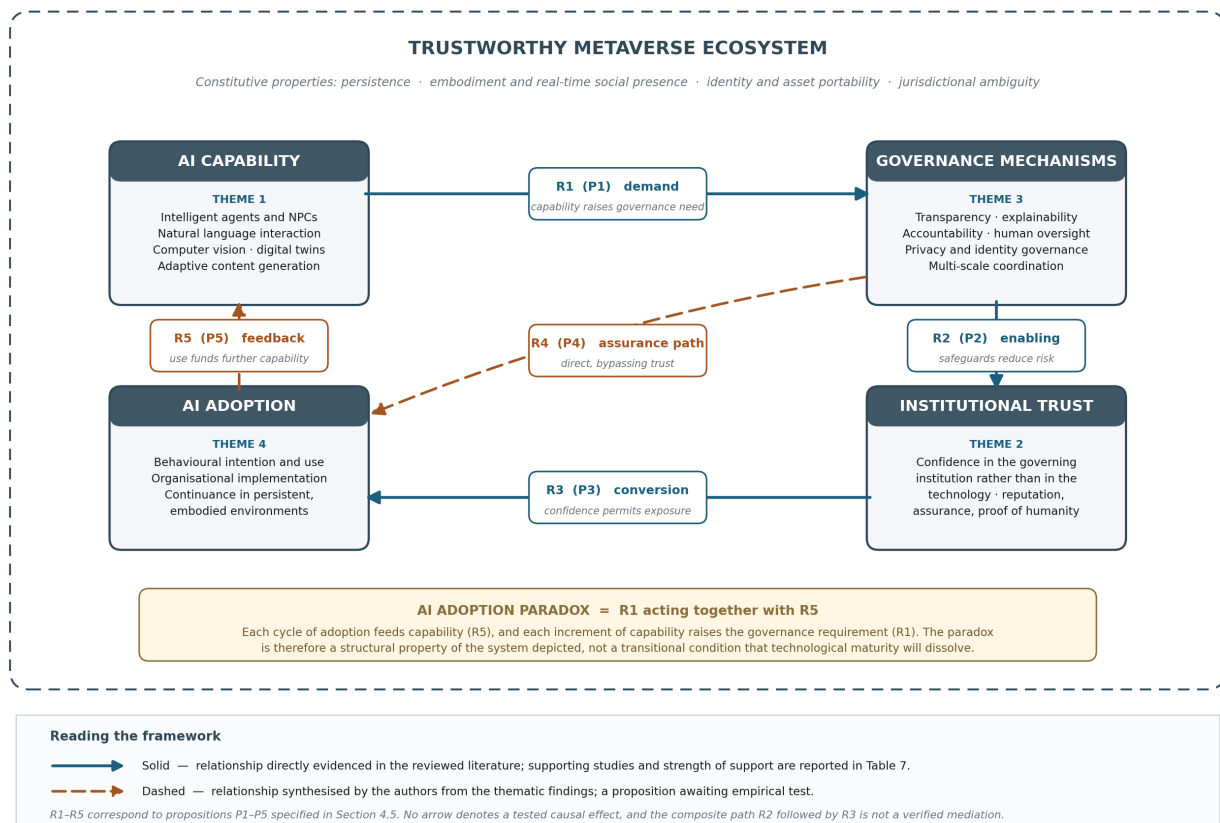


Figure 7. The Metaverse AI Trust–Governance Nexus Framework. Solid arrows denote relationships directly evidenced in the reviewed literature; dashed arrows denote relationships synthesised by the authors and stated as propositions.

R1–R5 correspond to propositions P1–P5 in Section 4.5, and the evidential basis of each is reported in **Table 7**.

5. Discussion

5.1 Theoretical Contributions

The findings of this review reinforce an important shift in contemporary research on AI-enabled digital ecosystems. Rather than portraying AI adoption as a predominantly technological or behavioural phenomenon, the evidence synthesised across the reviewed studies indicates that adoption increasingly reflects institutional conditions surrounding AI deployment. This observation suggests that technological capability alone provides an incomplete explanation of stakeholder acceptance within metaverse ecosystems. As immersive environments become more autonomous and socially embedded, governance quality and institutional legitimacy appear to play progressively greater roles in shaping long-term adoption.

One notable contribution emerging from this review concerns the relationship between governance and institutional trust. Previous studies have frequently acknowledged the importance of trust in technology adoption, yet trust has often been conceptualised primarily as an individual psychological perception or behavioural determinant. Across the reviewed literature, however, institutional trust appears to develop through organisational transparency, accountability, responsible governance, and the perceived legitimacy of AI deployment rather than through technological performance alone. This distinction extends existing adoption research by positioning trust as an institutional outcome rather than merely an individual attitude.

The review also provides a broader interpretation of AI governance. Much of the current literature discusses governance in normative terms, emphasising ethical principles, regulatory compliance, or responsible AI guidelines. While these contributions remain valuable, the present synthesis suggests that governance performs a more fundamental socio-technical function. Governance not only constrains technological development but also enables sustainable innovation by reducing uncertainty, strengthening institutional credibility, and creating conditions under which stakeholders are willing to engage with AI-enabled services. In this sense, governance operates simultaneously as a mechanism of control and a catalyst for technological adoption.

A particularly significant insight emerging from the cross-theme synthesis is the recurring AI Adoption Paradox. Across multiple application domains, technological advancement consistently produces two opposing effects. On one hand, increasingly capable AI systems improve automation, personalisation, intelligent interaction, and service efficiency. On the other hand, these same advances intensify concerns regarding explainability, privacy, algorithmic bias, surveillance, digital identity, and institutional accountability. Interestingly, this paradox appears not to represent a temporary transitional issue but rather a structural characteristic of AI-enabled digital ecosystems. Consequently, future discussions of AI adoption may need to move beyond the assumption that technological maturity naturally leads to higher acceptance. Instead, the evidence reviewed here suggests that technological progress simultaneously increases the importance of governance and institutional trust.

Taken together, these observations provide the foundation for the Metaverse AI Trust–Governance Nexus Framework proposed in this review. Rather than introducing new theoretical constructs, the framework reorganises fragmented knowledge into a coherent socio-technical explanation of how AI capability, governance mechanisms, institutional trust, and AI adoption interact within trustworthy metaverse ecosystems. Its primary contribution therefore lies in conceptual integration rather than theoretical

replacement, offering a platform upon which future empirical research may build. What is distinctive about that integration in immersive settings is set out in Section 5.2.

5.2 Metaverse-Specific Implications

The relationships synthesised above are not confined to immersive environments, and it is therefore necessary to state what changes when governance and institutional trust are situated in metaverse ecosystems rather than in conventional AI-enabled platforms. Five differences follow from the constitutive properties identified in Section 1.

First, persistence alters the temporal structure of governance. On a conventional platform, accountability can be discharged transactionally: a decision is made, contested, and closed. In a persistent environment, an algorithmic decision about an avatar, an asset, or a reputation continues to condition subsequent interaction indefinitely, so governance must be continuous and auditable across sessions rather than exercised at discrete points. This implies obligations that have no clear analogue in conventional platform governance, including model version traceability, retention of decision logs across environment states, and mechanisms for revisiting decisions whose effects persist. Lifecycle approaches to trustworthy AI models and data-and-system traceability frameworks illustrate what such continuity requires in practice [42,85].

Second, embodiment changes what explainability requires. Where users encounter AI as a background service, explanation delivered through policy documents or interface disclosures may be adequate. Where users encounter AI as an embodied social actor, explanation must be available in-world and in-interaction, at the moment of the encounter, and must be legible to a user whose attention is committed to an immersive task. Transparency that is formally available but practically inaccessible during immersive interaction does not generate institutional trust. Research on presence and awareness cues in the metaverse, and on the ontological status of avatar embodiment, indicates how far this departs from conventional interface disclosure [36,72].

Third, identity governance carries the load that credential verification carries elsewhere. Because avatars decouple presentation from personhood and because persistent identities and assets are portable across platforms, mechanisms such as self-sovereign identity, proof of humanity, and blockchain-based authentication are not one safeguard among many but the principal institutional instrument through which trustworthiness is established. This also means that a governance failure at one provider propagates across every environment that recognises the affected identity, so trust in metaverse ecosystems is partly ecosystem-level rather than provider-level, and cannot be secured by any single organisation acting alone [11,86,87]. Zero-trust architectures, which replace perimeter assumptions with continuous verification, are the closest existing analogue [6,88].

Fourth, the data generated in immersive environments differ in kind. Gaze, gait, posture, proximity, and physiological response are captured continuously as a by-product of interaction, are strongly identifying even when de-identified, and reveal attentional and affective states that conventional interaction data do not. Consent frameworks built for discrete data-collection events map poorly onto sensor streams that are inseparable from participation itself, which places particular weight on purpose limitation, on-device processing, and minimisation of granularity and retention. Privacy-preserving techniques developed for sensitive biometric and clinical data indicate the direction such safeguards must take [35,89].

Fifth, harm and therefore trust repair are experienced differently. Harassment, manipulation, and exclusion in embodied environments are experienced with an immediacy that text-based mediation does not produce, and the resulting loss of confidence is correspondingly harder to restore through disclosure or compensation alone. Governance in metaverse ecosystems must therefore include in-world remedy, graduated sanction, and dispute resolution that operate at the speed of interaction, rather than relying on after-the-fact complaint procedures. Analyses of integrity threats in the metaverse, and of the shift from an anthropocentric to a metasubjective conception of the participant, converge on the same requirement [37,90].

Taken together, these differences indicate that the framework proposed here is not a general model of AI governance applied to a new setting. The constructs are shared with the wider literature, but their operationalisation, their temporal structure, and the locus at which they must be enacted are specific to persistent, embodied, and socially interactive environments. This is also why the predominance of foundational studies in the corpus, reported in Table 5, constitutes a research agenda rather than a limitation to be explained away: the constructs transfer, but the mechanisms have yet to be tested where they matter most.

5.3 Practical Implications

Beyond its theoretical implications, the review also offers several practical insights for organisations seeking to deploy AI within emerging metaverse ecosystems. One recurring observation throughout the reviewed studies is that technological sophistication does not automatically translate into stakeholder acceptance. Organisations investing heavily in AI capabilities may therefore achieve only limited adoption if governance arrangements fail to generate institutional confidence. From a managerial perspective, governance should not be regarded solely as a regulatory obligation but rather as a strategic capability that supports sustainable digital transformation.

For policymakers, the findings highlight the need to balance technological innovation with institutional accountability. Current regulatory initiatives increasingly emphasise transparency, explainability, fairness, privacy protection, and human oversight. The present review suggests that these principles should not be implemented merely to satisfy regulatory requirements. Instead, they represent practical mechanisms through which governments cultivate public trust in AI-enabled digital services. Policies that strengthen institutional legitimacy may ultimately prove as important as investments in technological infrastructure.

Technology developers face a similar challenge. Much of the existing innovation effort continues to prioritise computational performance, automation, and user functionality. While these capabilities remain essential, the reviewed evidence consistently indicates that users also evaluate AI systems according to their transparency, explainability, privacy protection, and perceived fairness. Integrating these governance principles during system design rather than after deployment may therefore reduce implementation risk while improving long-term adoption.

The review also carries implications for organisations operating across multiple institutional contexts. As metaverse ecosystems expand internationally, governance expectations are unlikely to remain uniform across jurisdictions. Differences in legal systems, cultural norms, privacy expectations, and institutional maturity may influence how stakeholders interpret trustworthy AI. Consequently, governance frameworks should remain sufficiently adaptive to accommodate contextual variation rather than relying exclusively on universal regulatory principles.

Finally, the proposed framework may serve as a practical reference for evaluating organisational readiness. Instead of assessing readiness exclusively through technological capability, decision-makers may consider whether governance structures, institutional trust, organisational accountability, and stakeholder engagement have developed alongside technological innovation. This broader perspective may help organisations anticipate implementation challenges before large-scale deployment occurs.

5.4 Limitations

Although this review provides a comprehensive synthesis of the current literature, several limitations should be acknowledged when interpreting the proposed framework.

First, the review draws exclusively upon publications indexed in the Scopus database. While Scopus offers broad coverage of peer-reviewed research across multiple disciplines, relevant studies indexed elsewhere may not have been captured. Consequently, the framework should be interpreted as reflecting the current Scopus-indexed evidence base rather than the entirety of global scholarship.

Second, only peer-reviewed journal articles published in English were included. This decision strengthened methodological consistency and ensured that the synthesis was grounded in high-quality academic evidence. At the same time, it inevitably excluded conference proceedings, technical reports, industry documents, policy papers, and other forms of grey literature that frequently introduce emerging technological developments before they appear in journal publications.

A further limitation concerns the conceptual nature of the proposed framework. The relationships presented in the Metaverse AI Trust–Governance Nexus Framework are derived from qualitative synthesis rather than empirical testing. Accordingly, they should be interpreted as theoretically grounded propositions requiring validation through future quantitative, qualitative, or mixed-method investigations.

A fourth limitation concerns the restriction to open-access publications. This criterion was applied to guarantee full-text availability for every screened record, but it is not neutral: open-access publishing is unevenly distributed across regions, disciplines, and publishers, and the resulting corpus may therefore over-represent research funded through article-processing-charge arrangements and under-represent subscription-based venues in which parts of the governance and information systems literature appear. The 493 records excluded between eligibility and inclusion should be interpreted with this in mind.

A fifth limitation concerns critical appraisal. Because the corpus comprises heterogeneous conceptual, qualitative, and survey-based studies, no formal risk-of-bias or quality-appraisal instrument was applied, and included studies were not weighted by methodological strength during synthesis. Relatedly, agreement between coders was established through documented consensus rather than through a reliability coefficient, as explained in Section 3.3. Both choices are defensible for interpretive synthesis but limit the extent to which the relative weight of individual studies can be assessed.

Two reporting limitations should also be recorded. First, the screening log captured exclusions in aggregate rather than by category, so the distribution of the 493 excluded records across the five grounds in Table 3 cannot be reported; future reviews by this team will record exclusion reasons at record level. Second, and more consequentially for the strength of the synthesis, the full text was obtainable for only 19 of the 183 included studies. For the remaining 164, extraction relied on

the Scopus record. This constrains what can be claimed: methodological design, theoretical foundation, and construct operationalisation were verified only for the smaller subset, while thematic assignment for the majority rests on title, keyword, and abstract evidence. Readers should weigh the thematic findings accordingly, and Supplementary Table S1 makes the extraction depth of every study visible for that purpose.

A sixth limitation follows from the composition of the evidence base. As Table 5 reports, a substantial proportion of the included studies examine AI adoption, governance, and trust in conventional organisational settings rather than in immersive environments. The framework is therefore assembled partly from constructs whose behaviour in persistent, embodied, and socially interactive environments has not yet been tested. The classification introduced in Section 4.1 and the mapping in Table 7 make this dependence visible rather than remove it, and the transfer of these constructs to metaverse settings should be treated as a hypothesis rather than an assumption.

Finally, the technological landscape surrounding AI and the metaverse continues to evolve at an exceptional pace. Governance mechanisms, regulatory frameworks, and stakeholder expectations are therefore likely to change alongside continuing advances in generative AI, autonomous intelligent agents, immersive environments, and decentralised digital infrastructures. The framework proposed here should consequently be regarded as an evolving conceptual interpretation rather than a definitive representation of future AI-enabled metaverse ecosystems.

5.5 Future Research Directions

Several promising directions emerge from the present synthesis.

The most immediate priority concerns empirical validation of the proposed framework. While the conceptual relationships identified in this review appear consistently across the literature, future studies should examine whether these relationships remain stable across different organisational settings, technological environments, and institutional contexts. Structural equation modelling, longitudinal designs, comparative case studies, and mixed-method approaches may all contribute to refining the framework.

A further priority follows directly from the propositions stated in Section 4.5. P2 and P3 can be tested with survey-based structural models fielded among users and operators of immersive platforms, with institutional trust measured as confidence in the governing organisation rather than in the technology. P4 requires organisational rather than individual designs, since the compliance and assurance pathway operates at the level of adopting institutions. P5 and the composite pathway require longitudinal or multi-wave designs, because a feedback relationship cannot be identified from cross-sectional data. Studies that compare metaverse-native with conventional platform settings would additionally establish whether the mechanisms proposed here are specific to immersive environments or common to AI-enabled services generally, which is the boundary condition the present corpus cannot resolve.

Future research should also pay greater attention to contextual variation. Governance mechanisms that foster institutional trust in healthcare, public administration, education, financial services, or manufacturing may not operate identically within metaverse ecosystems. Comparative investigations across sectors and countries could therefore provide a richer understanding of how institutional environments shape trustworthy AI adoption.

Another area deserving greater attention concerns the evolution of governance

itself. Most existing studies conceptualise governance as relatively stable, whereas the evidence reviewed here suggests that governance is increasingly adaptive, collaborative, and multi-level. Future investigations may therefore explore how governance evolves in response to advances in generative AI, autonomous agents, decentralised identity management, and cross-platform interoperability.

Perhaps the most intriguing opportunity lies in further examination of the AI Adoption Paradox. Although this review identifies the paradox as a recurring pattern, relatively little empirical evidence explains how organisations successfully balance technological innovation with institutional accountability over time. Longitudinal research examining the co-evolution of governance, institutional trust, and AI adoption could therefore make an important contribution to both information systems theory and responsible AI research.

More broadly, future scholarship should continue moving beyond technology-centred explanations of AI adoption. As AI becomes increasingly embedded within socio-technical ecosystems, understanding adoption will require greater integration of technological capability, governance quality, institutional legitimacy, organisational behaviour, and stakeholder trust. Such integrative perspectives are likely to become increasingly important as trustworthy AI evolves from an organisational objective into a broader societal expectation.

6. Conclusion

This systematic literature review synthesised evidence from 183 open-access peer-reviewed journal articles, retrieved from Scopus and classified by their directness of relationship to metaverse settings, to examine how trustworthy AI adoption develops within metaverse ecosystems. The review indicates that AI adoption is best understood not simply as a technological outcome but as a socio-technical process shaped by the interaction of AI capability, governance mechanisms, and institutional trust. This perspective addresses the fragmented nature of existing research and offers a more integrated understanding of trustworthy metaverse ecosystems.

The three research questions are answered as follows. In response to RQ1, the review shows that research on AI-enabled metaverse ecosystems has expanded rapidly since 2022 and has shifted from technology-oriented to governance-oriented concerns, while remaining institutionally fragmented across disciplines. In response to RQ2, the review shows that the literature conceptualises governance principally as an antecedent of institutional trust and trust principally as an antecedent of adoption, but that these relationships are established largely in conventional organisational settings rather than in immersive environments, and that no reviewed study examines the complete pathway. In response to RQ3, the review consolidates these relationships into the Metaverse AI Trust–Governance Nexus Framework and states them as five propositions whose empirical validation, rather than whose confirmation, is the contribution offered to subsequent research.

The thematic synthesis revealed that advances in AI simultaneously create opportunities for intelligent automation, personalisation, and immersive digital experiences while intensifying concerns regarding transparency, explainability, privacy, accountability, and digital identity. These findings suggest that technological capability alone is insufficient to sustain stakeholder confidence. Instead, governance mechanisms play a critical role in fostering institutional trust, which in turn supports responsible and sustainable AI adoption.

Building on these findings, this review proposes the Metaverse AI Trust–

Governance Nexus Framework, which integrates AI capability, governance, institutional trust, and AI adoption into a unified socio-technical perspective. Rather than introducing new theoretical constructs, the framework consolidates fragmented knowledge and provides a conceptual foundation for future empirical research across diverse organisational and institutional contexts.

The review also identifies the AI Adoption Paradox, whereby increasing AI capability simultaneously strengthens innovation potential and heightens governance challenges. This paradox highlights that the long-term success of AI-enabled metaverse ecosystems depends not only on technological advancement but also on governance arrangements capable of maintaining institutional legitimacy and public trust. Accordingly, the proposed framework offers practical guidance for policymakers, organisations, and technology developers seeking to promote responsible AI governance while supporting the development of trustworthy, sustainable, and human-centred metaverse ecosystems.

Three qualifications define what this review does and does not establish, and they are stated here rather than left to be inferred. The framework is a conceptual synthesis, not a tested model: no reviewed study examines the complete pathway from governance through institutional trust to adoption, and the propositions P1 to P5 are offered as an agenda for empirical work rather than as results. The evidential base is uneven: 22 of the 183 studies are metaverse-native, so the metaverse-specific claims are supported substantially by analogy from digital twin and decentralised identity research, which shares the properties of persistence and portable identity but not those of embodiment and real-time social presence. And the depth of extraction varied, with full text available for 19 studies and Scopus records alone for the remainder.

These qualifications do not diminish the contribution so much as locate it. What the review establishes is that the institutional pathway linking governance, trust, and adoption is treated as self-evident across a literature that has never tested it, and that the properties of persistent, embodied environments give the four constructs a different content from the one they carry in conventional platform research. Making that pathway explicit, specifying it as testable propositions, and stating plainly where the evidence is thin are the contributions this review offers to the study of trustworthy metaverse ecosystems.

Author contributions: Conceptualization, HH, MM; methodology, HH, MM, AM; formal analysis, HH, JJ, AM; investigation, HH, MHK; data curation, HH, MHK; visualization, HH, SI, KK; validation, MM, RR, JJ, AM, SI; writing – original draft, HH; writing – review & editing, MM, RR, JJ, AM, MHK, SI; supervision, MM, RR, KK; project administration, MM, KK.

Funding: This research received no external funding.

Acknowledgment: The authors would like to express their sincere gratitude to Universitas Pendidikan Indonesia (UPI) and Universiti Teknologi MARA (UiTM), Malaysia, for their academic and institutional support throughout this research. The authors also appreciate the constructive comments and suggestions provided by colleagues and anonymous reviewers, which significantly improved the quality and clarity of this manuscript.

Conflicts of interest: The authors declare no conflicts of interest.

Software and tools: VOSviewer and Bibliometrix were used for bibliometric mapping. The use of generative AI assistance is disclosed in the AI Assistance

Statement below. All authors have read and approved the published version of the manuscript.

AI Assistance Statement: During manuscript preparation, the authors used Scopus AI to support literature exploration and thematic visualisation, Vosviewer and Google Gemini AI to assist in refining conceptual visualisations and improving graphical presentation. All literature screening, thematic analysis, interpretation of findings, conceptual framework development, and manuscript writing were conducted, verified, and approved by the authors. The authors take full responsibility for the accuracy, originality, and integrity of the published work.

Supplementary material: Supplementary Table S1 (characterisation of the 183 included studies) is provided as a separate file.

References

- [1] Ramini N, Rahayu A, Disman D, et al. Mapping the evolution of cybersecurity and machine learning research: implications for secure metaverse environments. *Metaverse*. 2026, 7(3):26017. <https://doi.org/10.54517/m26017>
- [2] Abed SS. Acceptance and Use of Artificial Intelligence in Online Tourism Services by Generation Z in Saudi Arabia. *IEEE Access*. 2024, 12:164533–164542. <https://doi.org/10.1109/ACCESS.2024.3492001>
- [3] Almogren AS. Creating a Novel Approach to Examine How King Saud University's Arts College Students Utilize AI Programs. *IEEE Access*. 2025, 13:63895–63917. <https://doi.org/10.1109/ACCESS.2025.3553631>
- [4] Alourani A, Alam M, Ali A, et al. Hybrid AI-IoT Framework with Digital Twin Integration for Predictive Urban Infrastructure Management in Smart Cities. *Computers, Materials and Continua*. 2026, 86(1). <https://doi.org/10.32604/cmc.2025.070161>
- [5] Alsharhan A, Al-emran M, Shaalan K. Chatbot Adoption: A Multiperspective Systematic Review and Future Research Agenda. *IEEE Transactions on Engineering Management*. 2024, 71:10232–10244. <https://doi.org/10.1109/TEM.2023.3298360>
- [6] Swapna MP, Chandran S, Rajeev D, et al. ZTNA: Sector-Specific Analysis on the 'NO TRUST' Security Paradigm. *International Journal of Information Security*. 2026, 25(4). <https://doi.org/10.1007/s10207-026-01286-2>
- [7] Haryanto T, Rahayu A, Istakariana I, et al. Developing an infrastructure-enabled digital human capital framework in emerging digital markets. *Human Resources Management and Services*. 2026, 8(4). <https://doi.org/10.18282/hrms6110>
- [8] Allam Z, Sharifi A, Bibri SE, et al. The Metaverse as a Virtual Form of Smart Cities: Opportunities and Challenges for Environmental, Economic, and Social Sustainability in Urban Futures. *Smart Cities*. 2022, 5(3):771–801. <https://doi.org/10.3390/smartcities5030040>
- [9] Singh SK, Kumar S, Attar RW, et al. Leveraging AI and Machine Learning for Enhanced Data Analytics and Visualization in Database Management With Digital Twins. *Journal of Database Management*. 2025, 36(1):1–29. <https://doi.org/10.4018/JDM.388135>
- [10] Smith DP, Sokoya D, Moore S, et al. Embedding Generative AI as a digital capability into a year-long skills program. *Journal of University Teaching and Learning Practice*. 2025, 22(4):1-24. <https://doi.org/10.53761/fh6q4v89>
- [11] Tauber S. Constructing and evolving with the virtual self: exploration of digital twins and experiential living. *AI and Society*. 2026, 41(3):2619–2636. <https://doi.org/10.1007/s00146-025-02594-2>
- [12] Love B, Ghosh C, Shi W, et al. Trust and Technology Acceptance: Comparing Traditional Search Engines and Artificial Intelligence for Colorectal Cancer Information Seeking. *Cancer Control*. 2026, 33:1–23. <https://doi.org/10.1177/10732748261454474>
- [13] Qian Y, Fu Y, Chen Y, et al. Public acceptance of LLM-driven healthcare chatbots in China: An empirical study. *Digital Health*. 2026, 12:1–16. <https://doi.org/10.1177/20552076261437614>
- [14] Rizzo A, Mozgai S, Sigaras A, et al. Expert Consensus Best Practices for the Safe, Ethical, and Effective Design and Implementation of Artificially Intelligent Conversational Agent (Chatbot or Virtual Human) Systems in Health Care Applications. *Journal of Medical Extended Reality*. 2025, 2:209–222. <https://doi.org/10.1177/29941520251369450>
- [15] Sides T, Kbaier D. Bridging trust gaps: Stakeholder perspectives on AI adoption in the United Kingdom NHS primary care. *Digital Health*. 2025, 11:1–11. <https://doi.org/10.1177/20552076251386706>
- [16] Nugraha RA, Ryadi Y, Juliana J, et al. Metaverse-oriented civic learning in civic education: integrating augmented reality,

- GCE, and Project Citizen to strengthen global diversity character. *Metaverse*. 2026, 7(3):26015. <https://doi.org/10.54517/m26015>
- [17] Bhattacharya P, Saraswat D, Savaliya D, et al. Towards Future Internet: The Metaverse Perspective for Diverse Industrial Applications. *Mathematics*. 2023, 11(4). <https://doi.org/10.3390/math11040941>
- [18] Bansal C, Pandey KK, Goel R, et al. A PLS-SEM mediation analysis of factors influencing the adoption intention of banking chatbots. *Issues in Information Systems*. 2025, 26(3). https://doi.org/10.48009/3_iis_2025_2025_116
- [19] Lazarte-aguirre A, Fern R. Determinants of Generative AI Adoption Through the UTAUT Model: Insights From Postgraduate. *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje*. 2025, 20:391–400. <https://doi.org/10.1109/RITA.2025.3613135>
- [20] Waehrens BV, Berger U, Ducholm BC, et al. Operationalizing the Industrial Metaverse: Strategies, Challenges, and Opportunities for the Sustainable Factory of the Future. *Sustainability (Switzerland)*. 2026, 18(6). <https://doi.org/10.3390/su18062941>
- [21] Kolf Schooten H. AI chatbots for promoting healthy habits: Legal, ethical, and societal considerations. *Digital Health*. 2025, 11:1–9. <https://doi.org/10.1177/20552076251390004>
- [22] Pastor-Escuredo D, Gardezabal A, Koo J, et al. Multi-scale governance and data for sustainable development. *Frontiers in Big Data*. 2022, 5. <https://doi.org/10.3389/fdata.2022.1025256>
- [23] Sharifi A, Amirzadeh M, Khavarian-Garmsir AR. Responsible metaverse-powered smart cities can contribute to sustainable development goals. *Computational Urban Science*. 2025, 5(1). <https://doi.org/10.1007/s43762-025-00201-0>
- [24] Aldraiweesh AA. The Influence of Social Support Theory on AI Acceptance: Examining Educational Support and Perceived Usefulness Using SEM Analysis. *IEEE Access*. 2025, 13:18366–18385. <https://doi.org/10.1109/ACCESS.2025.3534099>
- [25] Jin SV, Ryu E. Unraveling the dynamics of digital equality and trust in AI-empowered metaverses and AI-VR-convergence. *Technological Forecasting and Social Change*. 2025, 210. <https://doi.org/10.1016/j.techfore.2024.123877>
- [26] Vasiliu-Feltes I, Zhang CY, Dennis S, et al. Trust By Design: Enabling Responsible Precision Health Through Blockchain-Powered Digital Twins And Trusted AI. *Blockchain in Healthcare Today*. 2025, 8(2). <https://doi.org/10.30953/bhty.v8.453>
- [27] Fadhel MA, Duham AM, Albahri AS, et al. Navigating the metaverse: unraveling the impact of artificial intelligence—a comprehensive review and gap analysis. *Artificial Intelligence Review*. 2024, 57(9). <https://doi.org/10.1007/s10462-024-10881-5>
- [28] Ikhsan RB, Fernando Y, Prabowo H, et al. An empirical study on the use of artificial intelligence in the banking sector of Indonesia by extending the TAM model and the moderating effect of perceived trust. *Digital Business*. 2025, 5(1):100103. <https://doi.org/10.1016/j.digbus.2024.100103>
- [29] Chamola V, Peelan MS, Mittal U, et al. Metaverse for Education: Developments, Challenges, and Future Direction. *Computer Applications in Engineering Education*. 2025, 33(3). <https://doi.org/10.1002/cae.70018>
- [30] De Giovanni P. Sustainability of the Metaverse: A Transition to Industry 5.0. *Sustainability (Switzerland)*. 2023, 15(7). <https://doi.org/10.3390/su15076079>
- [31] Rahman M, Pow H, Kaium A, et al. Examining the dynamics of mobile banking app. Adoption during the COVID-19 pandemic: A digital shift in the crisis. *Digital Business*. 2024, 4(2):100088. <https://doi.org/10.1016/j.digbus.2024.100088>
- [32] Thor H, Aaltonen K, Snaer A, et al. Personalised learning in project management education: Insights from an artificial intelligence-driven chatbot. *Project Leadership and Society*. 2025, 6:100193. <https://doi.org/10.1016/j.plas.2025.100193>
- [33] Vidovszky AA, Fisher CK, Loukianov AD, et al. Increasing acceptance of AI-generated digital twins through clinical trial applications. *Clinical and Translational Science*. 2024, 17(7). <https://doi.org/10.1111/cts.13897>
- [34] Alhamam AA. Investigating the role of AI tool interactions in enhancing English language acquisition among Saudi college students. *Arab World English Journal*. 2025, 16(3):426–440. <https://doi.org/10.24093/awej/vol16no3.26>
- [35] Almomani A, Al-Qerem A, Alauthman M, et al. Ethical Foundations of AI-Driven Avatars in the Metaverse for Innovation and User Privacy. *IEEE Access*. 2025, 13:130610–130628. <https://doi.org/10.1109/ACCESS.2025.3589714>
- [36] Arnold R, Schneider A. Overlooked – Presence and awareness cues in the metaverse. *Telecommunications Policy*. 2025, 49(8). <https://doi.org/10.1016/j.telpol.2025.103020>
- [37] Jung YS, Lee JD, Kim M, et al. Systematic Analysis of Integrity Threats and Defense Strategies in the Metaverse. *KSII Transactions on Internet and Information Systems*. 2025, 19(11):4093–4118. <https://doi.org/10.3837/tiis.2025.11.018>
- [38] Plevris V, Hosamo H. Responsible AI in structural engineering: a framework for ethical use. *Frontiers in Built Environment*.

- 2025, 1–22. <https://doi.org/10.3389/fbuil.2025.1612575>
- [39] Carmona-Martínez A, Jara AJ, Asín A. A Sovereign Conversational Assistant Powered by ALIA and Mistral for the AI Act Age: Architecture, Governance, and Evaluation. *Machine Learning and Knowledge Extraction*. 2026, 8(6). <https://doi.org/10.3390/make8060155>
- [40] Lopez DL, Francesc J, Gascón F, et al. Aligning stakeholders in AI - enabled customer service: Toward human - centric adoption in electronic marketplaces. *Electronic Markets*. 2026, 1–19. <https://doi.org/10.1007/s12525-026-00892-1>
- [41] Sustaningrum R, Haldaka M. Student utilization and perceptions of AI technology for academic purposes: a mixed- method analysis. *Cogent Education*. 2025, 12(1). <https://doi.org/10.1080/2331186X.2025.2553835>
- [42] Parra-Ullauri J, Zhou X, Moazzeni S, et al. Lifecycle Management of Trustworthy AI Models in 6G Networks: The Reason Approach. *IEEE Wireless Communications*. 2025, 32(2):42–51. <https://doi.org/10.1109/MWC.001.2400288>
- [43] Al-emran M, Alqudah AA, Abbasi GA, et al. Determinants of Using AI-Based Chatbots for Knowledge Sharing: Evidence From PLS-SEM and Fuzzy Sets (fsQCA). *IEEE Transactions on Engineering Management*. 2024, 71:4985–4999. <https://doi.org/10.1109/TEM.2023.3237789>
- [44] Lauser T, Weinberger M. Does Appearance Matter? A Technology Acceptance Study of Mixed Reality Avatars in Citizen Services. *Future Internet*. 2026, 18(3). <https://doi.org/10.3390/fi18030169>
- [45] Li X. Public Trust in Global AI Governance Across Geopolitical Rivals. *Global Policy*. 2025, 1–11. <https://doi.org/10.1111/1758-5899.70105>
- [46] Skalka JN, Przybyła-Kasperek M. Artificial Intelligence Literacy Structure and the Factors Influencing Student Attitudes and Readiness in Central Europe Universities. *IEEE Access*. 2025. <https://doi.org/10.1109/ACCESS.2025.3573575>
- [47] Heinisuo E, Kuoppakangas P, Stenvall J. Navigating AI Implementation in Local Government: Addressing Dilemmas by Fostering Mutuality and Meaningfulness. *Information Systems Frontiers*. 2025. <https://doi.org/10.1007/s10796-025-10599-x>
- [48] Khanfar AA, Kiani R, Mohammad M, et al. Determinants of artificial intelligence adoption: research themes and future directions. *Information Technology and Management*. 2026, 27(1):31–51. <https://doi.org/10.1007/s10799-024-00435-0>
- [49] Gkinko L, Elbanna A. Designing trust: The formation of employees' trust in conversational AI in the digital workplace. *Journal of Business Research*. 2023, 158:113707. <https://doi.org/10.1016/j.jbusres.2023.113707>
- [50] Gefen D, Karahanna E, Straub DW. Trust and TAM in online shopping: an integrated model. *MIS Quarterly*. 2003, 27(1):51–90. <https://doi.org/10.2307/30036519>
- [51] McKnight DH, Choudhury V, Kacmar C. Developing and validating trust measures for e-commerce: an integrative typology. *Information Systems Research*. 2002, 13(3):334–359. <https://doi.org/10.1287/isre.13.3.334.81>
- [52] Abed AZM, Abdelkader T, Hashem M. Fuzzy logic trust management framework for Cyber Physical-Social Systems (FTM-CPSS). *PeerJ Computer Science*. 2025, 11. <https://doi.org/10.7717/peerj-cs.3164>
- [53] Agostinelli S, Bisconti P, Casciani A, et al. Exploring Human Trust in Software Robots for Robotic Process Automation. *Business and Information Systems Engineering*. 2026. <https://doi.org/10.1007/s12599-026-01007-y>
- [54] Li J, Zhu E, Lin W, et al. A novel digital twins-driven mutual trust framework for human–robot collaborations. *Journal of Manufacturing Systems*. 2025, 80:948–962. <https://doi.org/10.1016/j.jmsy.2025.04.009>
- [55] Romandini N, Roberta Costagliola A, Bujari A, et al. TrustFlow: A traceable federated learning framework to enable trustworthy digital twins. *Future Generation Computer Systems*. 2026, 178. <https://doi.org/10.1016/j.future.2025.108267>
- [56] Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*. 1989, 13(3):319–340. <https://doi.org/10.2307/249008>
- [57] Venkatesh V, Morris MG, Davis GB, et al. User acceptance of information technology: toward a unified view. *MIS Quarterly*. 2003, 27(3):425–478. <https://doi.org/10.2307/30036540>
- [58] McPhee S, Saravana A, Tahmasebinia F, et al. Education and Training for Emerging Technology Adoption and Expertise: Insights from Australian Construction. *Sustainability (Switzerland)*. 2026, 18(12). <https://doi.org/10.3390/su18125855>
- [59] Turan H. Determinants of AI-Enabled Quality Control Adoption Intention in Manufacturing SMEs: An Integrated TOE–TAM Analysis Using PLS-SEM, IPMA, and fsQCA. *Journal of Manufacturing and Materials Processing*. 2026, 10(6). <https://doi.org/10.3390/jmmp10060212>
- [60] Noor R, Rahayu A, Hurriyati R, et al. Human capital readiness for cloud–AI–data center ecosystems: a bibliometric review. *Human Resources Management and Services*. 2026, 8(3). <https://doi.org/10.18282/hrms6018>
- [61] Uddin M, et al. Exploring the convergence of metaverse, blockchain, and AI: a comprehensive survey of enabling

- technologies, applications, challenges, and future directions. *WIREs Data Mining and Knowledge Discovery*. 2024, e1556. <https://doi.org/10.1002/widm.1556>
- [62] Wang Y, Su Z, Zhang N, et al. A survey on metaverse: fundamentals, security, and privacy. *IEEE Communications Surveys & Tutorials*. 2023, 25(1):319–352. <https://doi.org/10.1109/COMST.2022.3202047>
- [63] Almeman K, EL Ayeb F, Berrima M, et al. The integration of AI and metaverse in education: a systematic literature review. *Applied Sciences*. 2025, 15(2):863. <https://doi.org/10.3390/app15020863>
- [64] Rahartomo A, Merino L, Ghafari M. Metaverse security and privacy research: a systematic review. *Computers & Security*. 2025. <https://doi.org/10.1016/j.cose.2025.104602>
- [65] Page MJ, Mckenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021, 372:n71. <https://doi.org/10.1136/bmj.n71>
- [66] Marlina L, Juliana J. Halal industry development strategy: a bibliometric analysis. *Journal of Islamic Marketing*. 2025, 324–348. <https://doi.org/10.1108/JIMA-12-2024-0591>
- [67] Damayanti A, Miftahuddin A, Rofaida R, et al. Two decades in five years: mapping digital HR, workflow automation, and innovation ecosystem research (2020–2025). *Daengku: Journal of Humanities and Social Sciences Innovation*. 2026, 6(2):259–275. <https://doi.org/10.35877/454RI.daengku4894>
- [68] de Matías Batalla D, Nicolás Sans R. Student's perception of the use of the metaverse in higher education on the employability of Spanish talent market. *Metaverse*. 2025, 6(4):8270. <https://doi.org/10.54517/m8270>
- [69] Vinodan A, Meera S, Shreansh M. Incentivized adoption of smart tourism services: host community perspective. *Smart Tourism*. 2023, 4(2):2454. <https://doi.org/10.54517/st.v4i2.2454>
- [70] Tirdasari NL, Dhewanto W, Arief NN. Developing a model for digital adoption in family business: a proposed research framework. *Sustainable Social Development*. 2024, 2(3):2622. <https://doi.org/10.54517/ssd.v2i3.2622>
- [71] Yu K. Research on the influencing factors of the e-commerce adoption behaviour of new agricultural business entities under the TAM model. *Value, Function, Cost*. 2023, 3(2):2609. <https://doi.org/10.54517/vfc.v3i2.2609>
- [72] Valera L, Alamos F, Ramos P, et al. Bodies in the metaverse: Is there "someone" out there?. *AI and Society*. 2026, 41(1):99–110. <https://doi.org/10.1007/s00146-025-02439-y>
- [73] Han J, Yang M, Chen X, et al. ParaDefender: A Scenario-Driven Parallel System for Defending Metaverses. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*. 2023, 53(4):2118–2127. <https://doi.org/10.1109/TSMC.2022.3228928>
- [74] Karthikeyan J, Saad NMN, Lim CK, et al. Human-Centric Urban Planning Using Behavioural Digital Twins and Predictive Analytics. *International Journal of Artificial Intelligence and Machine Learning*. 2026, 6(1S):25–33. <https://doi.org/10.51483/IJAIML.6.1s.2026.25-33>
- [75] Raiththa K, Shabaz M, Kumar G. Trust-Aware Intrusion Detection Framework for AI-Augmented Vehicular Digital Twin. *IEEE Open Journal of Vehicular Technology*. 2026, 7:1603–1613. <https://doi.org/10.1109/OJVT.2026.3694672>
- [76] Egho-Promise E, Ekereuke U, Gashi E, et al. AI-Driven Intrusion Detection and Threat Mitigation in Zero Trust Architectures. *Indonesian Journal of Electrical Engineering and Informatics*. 2026, 14(1):211–227. <https://doi.org/10.52549/ijeei.v14i1.7706>
- [77] Han HC, Huang DC. The Blockchain Neural Network Superior to Deep Learning for Improving the Trust of Supply Chain. *CMES - Computer Modeling in Engineering and Sciences*. 2025, 143(3):3921–3941. <https://doi.org/10.32604/cmcs.2025.065627>
- [78] Kumar A H, Venkatram C P, Saran N, et al. Decentralised digital health ecosystems: a unified architecture for AI-enhanced medical record management. *Frontiers in Digital Health*. 2025, 7. <https://doi.org/10.3389/fdgth.2025.1685628>
- [79] Weber-Lewerenz B. Corporate digital responsibility (CDR) in construction engineering—ethical guidelines for the application of digital transformation and artificial intelligence (AI) in user practice. *SN Applied Sciences*. 2021, 3(10). <https://doi.org/10.1007/s42452-021-04776-1>
- [80] Rahman MK, Ismail NA, Hossain A, et al. Students' mindset to adopt AI chatbots for effectiveness of online learning in higher education. *Future Business Journal*. 2025. <https://doi.org/10.1186/s43093-025-00459-0>
- [81] Oc Y, Gonsalves C. Generative AI in Higher Education Assessments: Examining Risk and Tech-Savviness on Student's Adoption. *Journal of Marketing Education*. 2025. <https://doi.org/10.1177/02734753241302459>
- [82] Acikgul K. Educators' Acceptance and Use of AI Software: UTAUT2 Model. *SAGE Open*. 2026, 1–19. <https://doi.org/10.1177/2158244025561111>

- org/10.1177/21582440251413449
- [83] Al Humaid Alneyadi MRM, Normalini MK. FACTORS INFLUENCING USER'S INTENTION TO ADOPT AI-BASED CYBERSECURITY SYSTEMS IN THE UAE. *Interdisciplinary Journal of Information, Knowledge, and Management*. 2023, 18:459–486. <https://doi.org/10.28945/5166>
- [84] Alshahrani A. Adopting Emerging Technologies in Digital Government: A Multi-Case Analysis of Drivers, Enablers, and Challenges in Saudi Arabia. *Digital Government: Research and Practice*. 2025, 6(3). <https://doi.org/10.1145/3719297>
- [85] Kalokyri V, Tachos NS, Kalantzopoulos CN, et al. AI Model Passport: Data and system traceability framework for transparent AI in health. *Computational and Structural Biotechnology Journal*. 2025, 28:386–404. <https://doi.org/10.1016/j.csbj.2025.09.041>
- [86] Chakranarayan V, Hussain F, Jaber FA, et al. Safeguarding Brand and Platform Credibility Through AI-Based Multi-Model Fake Profile Detection. *Future Internet*. 2025, 17(9). <https://doi.org/10.3390/fi17090391>
- [87] Roijen DV. The European Digital Identity Wallet: A Healthcare Perspective. *Blockchain in Healthcare Today*. 2024, 7(2). <https://doi.org/10.30953/bhty.v7.344>
- [88] Joshi H. Emerging Technologies Driving Zero Trust Maturity Across Industries. *IEEE Open Journal of the Computer Society*. 2025, 6:25–36. <https://doi.org/10.1109/OJCS.2024.3505056>
- [89] Yang Y, Lyu J, Wang R, et al. A digital mask to safeguard patient privacy. *Nature Medicine*. 2022, 28(9):1883–1892. <https://doi.org/10.1038/s41591-022-01966-1>
- [90] Kostenko O, Dniprov O, Zhuravlov D, et al. A New Paradigm of Metaverse Philosophy: From Anthropocentrism to Metasubjectivity. *Philosophies*. 2025, 10(6). <https://doi.org/10.3390/philosophies10060117>