



AI-Powered Leader Digital Twins as Identity-Governed Augmentation Systems: A Quantum-Entangled Framework for Leadership, XR, and Human–AI Coupling

Giammarco Tosi^{a,*}

^a“Aurel Vlaicu” University of Arad, Faculty of Exact Sciences, Arad, Romania

Abstract

This work develops a conceptual framework for AI-powered leader digital twins, defined as digital systems designed to emulate a specific leader’s knowledge, communication style, and decision logic. The paper addresses a central theoretical problem: how such systems can remain meaningfully linked to a leader’s identity without reducing leadership to functional simulation. To address this problem, the study integrates research on human digital twins, digital twin typologies, decentralized identity infrastructures, extended mind theory, trust and transparency, and quantum entanglement as a bounded metaphor for tightly coupled human–AI systems. Extended reality (XR) is incorporated as the embodiment layer through which a leader digital twin becomes socially present in organizational settings. The resulting Quantum-Entangled Leadership Digital Twin (QELDT) framework conceptualizes the twin as an identity-governed augmentation system anchored by verifiable linkage, closed-loop synchronization, bounded delegation, and human accountability. The framework is interpreted against exploratory survey evidence from 38 leaders, which suggests greater willingness to delegate analytical and monitoring activities than identity-laden tasks such as culture shaping, inspiration, and ethically sensitive judgment. The paper contributes a governance-oriented foundation for theorizing leader digital twins as extensions of leadership capacity rather than autonomous substitutes.

Keywords: Leader digital twin, human digital twin, digital identity, quantum entanglement, extended reality, AI governance.

2020 MSC: : 68T42, 68T05, 93C40, 91B06.

1. Introduction

Artificial intelligence is making it increasingly feasible to create digital systems that emulate how particular individuals communicate, reason, and respond. Within this broader evolution, the idea of an AI-powered leader digital twin has become increasingly plausible: a digital system trained to reflect a leader’s knowledge base, communication style, priorities, and decision tendencies so that it can assist, advise, or interact in ways associated with that leader. Recent scholarship on human digital twins suggests that such systems may support prediction, simulation, and augmentation, while also introducing significant questions about autonomy, authenticity, and ethics (Kabalska & Wagner, 2025; Zhu & Sulaiman Khail, 2025; Zhu & Yang, 2025).

However, a leader digital twin is not simply another AI assistant. It is a representation that claims to operate in relation to a specific human identity. Once the discussion moves from generic AI augmentation to

*Corresponding author

E-mail address: gmtosi@gmail.com (Giammarco Tosi)

leader-specific emulation, the key problem is no longer performance alone. It becomes a question of whether the system can remain meaningfully linked to the leader whose authority, values, and symbolic legitimacy it appears to extend. In leadership settings, functional competence is not sufficient. Trust, attribution, authenticity, and accountability are also constitutive of legitimacy.

This paper addresses that gap by proposing a Quantum-Entangled Leadership Digital Twin (QELDT) framework. The term quantum-entangled is used analogically, not literally. The paper does not claim that organizational leadership or AI systems operate through quantum mechanisms. Rather, the language of entanglement is used as a bounded metaphor for strong coupling, coherence, and interdependence between a human leader and a digital counterpart. The metaphor is intentionally constrained by the logic of the no-signaling principle: unlike physical entanglement, human–AI coupling remains mediated by data flows, interfaces, feedback loops, and governance mechanisms rather than any instantaneous transmission of intention or agency (De Angelis *et al.*, 2007; Horodecki *et al.*, 2009).

The paper makes four contributions. First, it reframes leader digital twins as identity-governed systems rather than merely personalized AI artifacts. Second, it integrates digital identity infrastructures into the framework, showing why verifiable linkage matters for accountability and representational legitimacy. Third, it treats XR as an organizational embodiment layer, not a peripheral interface choice. Fourth, it uses exploratory survey evidence from 38 leaders to interpret the likely boundaries of acceptable delegation in leader-twin systems.

2. Review of scientific literature

2.1. Human digital twins and leader-specific representation

The literature on human digital twins increasingly defines them as detailed, personalized digital representations that may include physical, cognitive, behavioural, psychological, and social dimensions (Zhu & Yang, 2025). Compared with conventional industrial digital twins, human digital twins introduce additional complexity because they involve not only measurable states but also values, intentions, tacit knowledge, and socially situated conduct. Kabalska & Wagner (2025) argue that human digital twins may augment human capability and even serve as crisis-prevention tools, while simultaneously foregrounding a paradox between efficiency and authenticity. That paradox is especially important for leadership, where authority is not reducible to information processing.

This is precisely why leader digital twins deserve distinct theorization. A leader does not merely perform analytical tasks. Leadership also involves symbolic authority, relational trust, meaning-making, and the ability to embody institutional values in socially recognizable ways. A digital system may therefore replicate selected aspects of a leader's communicative or decision style without being equivalent to leadership in its fuller organizational sense. The concept becomes meaningful only when questions of identity, legitimacy, and bounded delegation are brought to the forefront.

2.2. Digital twin typologies and the identity problem

The digital twin literature shows that not all twins are alike. Kritzinger *et al.* (2018) distinguish among digital models, digital shadows, and digital twins according to the degree of integration and feedback between physical and digital entities. Van der Valk *et al.* (2022) similarly show that digital twins differ significantly by use case, data linkage, and function. These distinctions matter because they suggest that a leader digital twin should not automatically be treated as a full reciprocal twin merely because it imitates language or judgment.

A leader digital twin may be highly capable as a model or shadow while still falling short of what would justify calling it a genuine twin in a stronger organizational sense. The difference lies in whether the system

is dynamically coupled, continuously updated, and governable as a representation of a specific person. In this paper, that problem is captured through the distinction between entity and identity. The entity is the computational artifact itself. The identity is the legitimized claim that this artifact is authorized to represent a named leader within defined organizational boundaries.

2.3. *Digital identity and verifiable linkage*

For a leader digital twin, identity cannot be reduced to behavioural resemblance. A system may mimic a leader's style convincingly and still fail to qualify as that leader's authorized digital counterpart. The identity problem is therefore distinct from the modelling problem. The modelling problem concerns fidelity. The identity problem concerns legitimacy, traceability, and authorization.

This distinction is clarified by the literature on decentralized identity and self-sovereign identity. The W3C Decentralized Identifiers Recommendation defines DIDs as verifiable identifiers that can refer to any subject and be resolved to documents containing cryptographic material, verification methods, and services ([World Wide Web Consortium, 2022](#), hereafter W3C). The W3C Verifiable Credentials Data Model v2.0 extends this architecture by providing a standardized way to express and verify claims about subjects through issuer–holder–verifier ecosystems ([World Wide Web Consortium, 2025](#)). The systematic review by [Schardong & Custódio \(2022\)](#) further shows that self-sovereign identity research emphasizes control, persistence, consent, portability, and interoperability as central dimensions of digital identity architecture.

Applied to leader digital twins, these identity infrastructures do not solve the philosophical question of selfhood. They do, however, provide a governance substrate for answering practical organizational questions: Who authorized the twin? What role is it allowed to play? Which communications or decisions fall within its scope? How are its outputs attributed and audited? Without such mechanisms, a leader twin remains only an advanced simulation. With them, it becomes a governable representational system.

It is important to distinguish four related but distinct dimensions:

- 1) Identity verification: establishing who the leader is and who controls the twin's identifier.
- 2) Authorization: specifying which tasks, roles, and communications the twin is permitted to perform.
- 3) Behavioural fidelity: the degree to which the twin's outputs resemble the leader's knowledge, style, and priorities.
- 4) Leadership legitimacy: stakeholders' acceptance of the twin as an appropriate and credible representative of the leader.

Verifiable credentials can establish authorized linkage and provenance; they cannot, by themselves, guarantee behavioural fidelity, authenticity, or stakeholder acceptance.

2.4. *Extended mind, trust, and human–AI coupling*

The idea that a leader digital twin could become more than a tool has theoretical precedent in the extended mind thesis. [Clark & Chalmers \(1998\)](#) argue that cognitive processes may extend into external artifacts when those artifacts are reliably integrated into ongoing activity. This insight is highly relevant here. If a leader twin becomes a persistent memory structure, scenario engine, communication proxy, and analytic companion, the relationship is not well described as occasional tool use. Instead, the twin begins to function as a socio-technical extension of cognition.

The entanglement metaphor adds a second layer. In quantum theory, entanglement refers to nonclassical correlations between subsystems, but those correlations do not authorize faster-than-light signalling or magical direct control ([De Angelis et al., 2007](#); [Horodecki et al., 2009](#)). That limitation is analytically useful in the present context. It forces a disciplined interpretation: leader–twin “entanglement” should be

understood as bounded coupling through designed channels of synchronization, correction, and feedback. The metaphor is useful only insofar as it highlights coherence and interdependence while preserving the organizational reality that data, interfaces, oversight, and governance mediate the relationship.

Because leader digital twins are also trust-sensitive organizational systems, the literature on trust and transparency is directly relevant. Transparency influences trust through trustworthiness perceptions rather than automatically generating trust on its own (Tomlinson & Schnackenberg, 2022). Recent evidence further suggests that AI disclosure can erode trust by lowering legitimacy perceptions, even when disclosure is normatively appropriate (Schilke & Reimann, 2025). For leader digital twins, this means disclosure is necessary, but not sufficient. Stakeholders must also regard the system as operating within legitimate scope, under clear human accountability, and with safeguards that preserve the human centre of leadership. Research on algorithm appreciation and AI trust complements this view by showing that people may value algorithmic input under some conditions, while still remaining sensitive to context, framing, and role appropriateness (Gerlich, 2024; Logg *et al.*, 2019).

2.5. XR as the embodiment and interaction layer

If identity is the governance problem, XR is the interaction problem. Research on digital twin–XR integration shows that XR can make twins more visible, interactive, collaborative, and operationally useful by allowing users to visualize, manipulate, and engage with twin representations in immersive environments (Mayer *et al.*, 2025; Stacchio *et al.*, 2022). In industrial settings, this improves coordination and interaction with complex systems. In leadership settings, however, XR has deeper implications because embodiment changes perception.

A text-based leader twin may be interpreted as a tool or assistant. An XR-mediated leader twin, particularly one rendered through voice, avatar, or immersive presence, may be interpreted as a socially present proxy of leadership. This increases the twin’s practical usefulness for training, continuity, and scenario-based interaction, but it also heightens risks of over-attribution, emotional misrecognition, and symbolic substitution. XR should therefore be treated not as a neutral interface layer, but as a design feature that shapes legitimacy, authenticity, and perceived authority.

3. Methodology

This paper adopts a conceptual theory-building design. Such a design is appropriate when the objective is not to test causal relationships through primary empirical data, but to clarify constructs, integrate fragmented literature streams, and propose a framework that can guide subsequent operationalization and empirical inquiry. In line with methodological guidance on conceptual articles, the present study is grounded in an explicit research design and makes transparent the role played by the theories and literature streams brought into the analysis (Jaakkola, 2020).

More specifically, the paper combines theory synthesis and theory adaptation. It synthesizes insights from the literature on human digital twins, digital twin typologies, decentralized identity and verifiable credentials, trust and transparency in AI-mediated interaction, XR-enabled embodiment, and the extended mind perspective. It then adapts these streams to the specific organizational problem of leader digital twins. This combination is necessary because the phenomenon under investigation is emerging, interdisciplinary, and not yet supported by a stable, consolidated theoretical core. A purely single-stream treatment would therefore risk either technological reductionism or conceptual under-specification.

The conceptual development proceeded in three stages. First, the paper identified foundational literature capable of clarifying the ontological and functional status of digital twins, with particular attention to distinctions among models, shadows, and twins, as well as recent work on human digital twins. Second, it

introduced governance-oriented literature on digital identity, traceability, transparency, and trustworthy AI to address the problem of authorized representation. Third, it incorporated the extended mind perspective and quantum entanglement as a bounded metaphor to theorize the human–AI relationship not as replacement, but as tightly coupled augmentation. XR was then added as the embodiment and interaction layer through which the twin becomes socially encounterable within organizational settings.

The outcome of this process is the Quantum-Entangled Leadership Digital Twin (QELDT) framework. The framework is intended as an analytical architecture, not as a claim of empirical confirmation. Its purpose is to specify the dimensions through which leader digital twins may be meaningfully theorized: entity–identity duality, verifiable linkage, closed-loop coupling, scenario plurality under human judgment, and XR-mediated embodiment.

The survey evidence from 38 leaders is used in an exploratory and interpretive manner rather than as confirmatory hypothesis testing. Its role is to assess whether the proposed framework resonates with plausible user boundaries regarding delegation, legitimacy, authenticity, and trust. Accordingly, the survey is not treated as a basis for statistical generalization. Instead, it serves as an empirical touchpoint that helps evaluate whether the distinctions proposed by the QELDT framework, particularly those between entity-intensive and identity-intensive leadership activities, are consistent with how potential users make sense of leader digital twins in practice.

4. Results and discussion

The QELDT framework conceptualizes the leader digital twin as a governed socio-technical extension of leadership rather than an autonomous replacement. It consists of five pillars, summarized in Table 1.

Table 1. The five pillars of the QELDT framework

Pillar	Core idea	Observable signals in organizational practice
1. Duality of entity and identity	The twin has a computational entity dimension (model fidelity, memory, performance) and an identity dimension (legitimacy, authorization, accountability).	Clear separation between “the system performs well” and “the system is authorized to act as the leader.”
2. Verifiable identity anchoring	Persistent, auditable linkage to the leader via DIDs and verifiable credentials.	Metadata specifying identity, role, scope, and accountability; machine-verifiable authorization records.
3. Bounded entanglement through closed-loop coupling	Ongoing, governable synchronization between leader and twin via feedback, corrections, and updates.	Logs of corrections, versioning of preferences, visibility of twin outputs to the leader.
4. Scenario plurality under human judgment	The twin generates multiple scenarios; the leader selects and authorizes action.	Documentation showing multiple options generated, with final decision explicitly attributed to the human leader.
5. XR-mediated embodiment and transparent social presence	XR as the layer through which the twin is socially encountered, paired with disclosure.	Clear labeling of XR interactions as twin-mediated; stakeholders can distinguish leader, twin, and hybrid forms.

Pillar 1 – Duality of entity and identity. A leader digital twin has both an entity dimension and an identity dimension. The entity dimension concerns model fidelity, memory architecture, domain knowledge, behavioural consistency, and interface performance. The identity dimension concerns legitimacy: the

claim that the system is authorized to represent a particular leader in specific tasks and interactions. This distinction prevents a common conceptual slippage from “the system sounds like the leader” to “the system legitimately acts as the leader.”

Pillar 2 – Verifiable identity anchoring. A legitimate leader digital twin requires persistent and auditable linkage to the leader it represents. DIDs and verifiable credentials provide a relevant governance architecture because they support machine-verifiable identifiers, controller relationships, and cryptographically supported claims about authorization and scope ([World Wide Web Consortium, 2022, 2025](#)). In organizational contexts, this implies that every operational twin should carry explicit metadata regarding identity, role, domain boundaries, and accountability. The stronger the system’s representational role, the stronger the need for auditable provenance.

Pillar 3 – Bounded entanglement through closed-loop coupling. The leader and the twin should operate through closed-loop synchronization rather than static imitation. The twin learns from corrective feedback, historical decisions, contextual constraints, and updated preferences. The leader, in turn, receives visibility into the twin’s recommendations, activities, and scenario outputs. The point is not machine sovereignty, but coherence. Entanglement is therefore defined here as structured, ongoing, and governable coupling. Drift is reduced not through absolute autonomy but through continuous calibration.

Pillar 4 – Scenario plurality under human judgment. One of the twin’s most defensible advantages lies in generating multiple plausible interpretations, scenarios, or action paths. The quantum metaphor of superposition is useful here as a metaphor for structured plurality. In this paper, “superposition” refers only to the temporary coexistence of alternative analytical scenarios generated by the system; it does not imply a physical, quantum, or probabilistic equivalence between organizational decisions and quantum states. The twin can hold open and explore several candidate courses of action in parallel, but legitimate action still requires human judgment. The leader “collapses” the scenario space into an actual course of action by selecting, revising, or rejecting candidate outputs. This preserves intentionality and distinguishes augmentation from abdication.

Pillar 5 – XR-mediated embodiment and transparent social presence. When the twin appears through dashboards, coaching tools, avatars, or immersive environments, XR becomes the layer through which the twin is socially encountered. This embodiment can support training, continuity, and interaction, but only if paired with clear disclosure. Stakeholders must be able to distinguish among the human leader, the digital twin, and hybrid forms of AI-assisted interaction. Without that clarity, XR risks converting technical mediation into social ambiguity or deception.

These five pillars operate under a cross-cutting ethical condition: human accountability remains primary. Human accountability should be allocated explicitly across system ownership, model maintenance, authorization of delegated tasks, review of outputs, and final decision authority. The NIST AI Risk Management Framework and its Generative AI Profile provide a practical governance orientation for lifecycle risk management, documentation, and oversight ([Autio et al., 2024; Tabassi, 2023](#)). At the regulatory level, the EU AI Act reinforces the relevance of transparency, risk classification, documentation, and human oversight in high-impact AI deployments ([European Union, 2024](#)).

4.1. Interpreting the framework against exploratory survey evidence

The exploratory survey evidence analyzed in this study, based on 38 managers and executives, is compatible with the QELDT framework’s proposed distinction between entity-intensive and identity-intensive leadership work. The pattern is not simply that some tasks are delegable and others are not. More importantly, the boundary appears to map onto a deeper distinction between entity-intensive and identity-intensive leadership work.

Analytical, monitoring, and informational tasks appear more compatible with delegation to a digital twin because they rely more heavily on structured information processing and can be bounded by procedural or review mechanisms. By contrast, strategic vision, culture shaping, inspiration, conflict mediation, and ethically sensitive judgment appear less delegable because they are closely tied to symbolic authority, relational legitimacy, and authenticity. In other words, leaders seem more willing to extend their cognitive reach than to delegate their identity-bearing presence.

This pattern aligns with the framework's core argument. A leader digital twin is most acceptable when it functions as an augmentation system rather than a substitute authority. It also illustrates the distinction between modelling fidelity and identity legitimacy. Even if the twin performs well as a computational entity, leaders and stakeholders may still resist its use when the task is experienced as inherently human, morally loaded, or socially constitutive of leadership itself.

At the same time, the survey evidence should be interpreted cautiously. The sample is small, and the findings are exploratory rather than confirmatory. The value of the survey lies not in proving the framework statistically, but in showing that the QELDT architecture resonates with plausible user boundaries: analytical support is more acceptable than symbolic replacement, and legitimacy concerns are central rather than secondary.

4.2. *What QELDT adds to existing discussions*

QELDT extends existing discussions in four ways:

- From simulation quality to identity legitimacy. Current debates often focus on how well an AI can imitate a leader. QELDT shifts the focus to whether the system can be governed as a legitimate extension of leadership without eroding accountability or authenticity.
- From productivity tool to representation system. Leader digital twins are repositioned as systems of authorized representation, not merely personalized assistants. This shift matters because leadership is symbolic and relational, not only informational.
- From interface to embodiment. XR is treated not as a neutral interface choice but as a design feature that shapes presence, attribution, and perceived authority. This makes disclosure and scope control essential rather than optional.
- From ethical principles to governance architectures. The framework links identity infrastructures (DIDs, verifiable credentials) with AI risk governance, suggesting that trustworthy leader twins require concrete mechanisms of verification, documentation, and lifecycle control.

4.3. *Theoretical implications*

The first contribution of the framework is conceptual clarification. Current discussion of leader digital twins often oscillates between enthusiasm for AI augmentation and anxiety about replacement, but it frequently lacks a precise vocabulary for distinguishing simulation quality from identity legitimacy. QELDT helps address that gap by showing that the relevant question is not only whether an AI can imitate a leader convincingly, but whether it can be governed as a legitimate extension of leadership without eroding accountability or authenticity.

The second contribution is to reposition the leader digital twin as a representation system rather than a mere productivity tool. This shift matters because leadership is not only informational but also symbolic and relational. A system that stores knowledge, suggests options, and expands scenario analysis may be relatively straightforward to justify. A system that appears to speak as the leader requires a much stronger

governance architecture because the risks shift from technical inaccuracy to legitimacy distortion. In this respect, legitimacy theory is useful because it highlights how organizational actors evaluate appropriateness, credibility, and social acceptability at multiple levels (Bitektine & Haack, 2015).

The third contribution is to make visible the role of XR in that distortion or legitimization process. XR increases the twin's practical usability, but it also changes how people interpret the twin. Embodied systems are more likely to evoke social presence, emotional attribution, and symbolic authority. That makes disclosure, role clarity, and controlled scope essential design requirements rather than optional ethical add-ons.

Finally, the framework links identity infrastructures with AI risk governance. It suggests that trustworthy leader digital twins require not only ethical principles, but also concrete architectures of verification, documentation, and lifecycle control. This provides a more robust basis for future empirical inquiry than framing leader twins simply as “personalized AI.”

4.4. Managerial implications

For management practice, the framework suggests that organizations should not deploy leader digital twins as undifferentiated AI clones. Instead, they should design a graded delegation architecture. Low-risk analytical and monitoring tasks may be appropriate for substantial augmentation. Scenario exploration and interpretive support may be suitable under active human review. Identity-intensive and ethically charged tasks should remain predominantly human-led.

Organizations should also govern leader twins as representation systems. This implies separating at least four layers of control: identity anchoring, memory governance, delegation scope, and disclosure protocol. These layers can be operationalized through the following checklist:

Table 2. Deployment checklist for leader digital twins

Governance layer	Practical question
Identity anchoring	Who authorized this twin to represent the leader?
Memory governance	Which personal and organizational data may it retain and update?
Delegation scope	Which tasks may it perform, recommend, or never execute?
Disclosure protocol	How will users know when they are interacting with the twin?

Identity anchoring establishes who the twin is authorized to represent and under what conditions. Memory governance determines what personal and organizational data may be used, retained, or updated. Delegation scope defines which tasks the twin may perform, advise on, or never execute autonomously. Disclosure protocol determines how the system's status is communicated to employees, customers, and other stakeholders.

A further implication concerns transparency. Disclosure is ethically necessary, especially where AI-mediated leadership communication may otherwise be mistaken for direct human communication. However, transparency should not be treated as a simple reputational safeguard. If disclosure reduces legitimacy perceptions, organizations must complement it with clear safeguards, defined scope, and visible human accountability.

4.5. Limitations and future research

The present paper is conceptual, and the empirical evidence discussed is exploratory. The survey sample is limited and cannot sustain strong generalization claims. Moreover, the use of a quantum metaphor, even

in bounded form, carries the risk of interpretive overextension if read literally. Its value lies in clarifying coupling and coherence, not in importing physical ontology into organizational life. The term “entanglement” refers only to structured, governable coupling mediated by data flows and feedback loops; it does not imply instantaneous transmission of intention, agency, or influence. Similarly, “superposition” is used solely as a metaphor for the temporary coexistence of alternative analytical scenarios; it does not suggest that organizational decisions behave like quantum states.

Future research should therefore proceed in several directions. First, it should develop measurement instruments that distinguish perceived usefulness from perceived legitimacy, authenticity, identity continuity, and acceptable delegation scope. Second, it should move beyond hypothetical scenarios toward field studies, longitudinal research, and pilot implementations. Third, it should test interface effects directly: the same underlying twin may be evaluated very differently when encountered as text, voice, avatar, or immersive XR presence. Fourth, it should compare governance architectures to determine whether stronger verifiable identity mechanisms increase trust, constrain use, or alter perceptions of authority.

5. Conclusions

AI-powered leader digital twins are becoming thinkable before they are fully theorized. This paper has proposed the QELDT framework as a more academically grounded way to conceptualize them. By integrating human digital twin scholarship, digital twin typologies, digital identity infrastructures, trust research, XR-mediated embodiment, and a bounded entanglement metaphor, the framework explains how a leader twin may function as an extension of leadership capacity rather than as a substitute for leadership itself.

The central claim is straightforward: the viability of a leader digital twin depends less on mimicry alone than on identity anchoring, continuous coupling, bounded delegation, transparent embodiment, and persistent human accountability. XR expands the twin’s organizational usefulness, but also its ethical exposure. The exploratory survey evidence is compatible with this position, suggesting that leaders are more willing to delegate structured and analytical work than tasks deeply tied to authenticity, symbolic authority, and moral judgment.

A publishable research agenda follows from this framework. Future studies should test how identity infrastructures, disclosure designs, and interface choices affect trust in leader digital twins across different organizational contexts. They should also examine whether bounded augmentation can preserve continuity without dissolving the human distinctiveness that gives leadership its legitimacy.

References

- Autio, C., R. Schwartz, J. Dunietz, S. Jain, M. Stanley, E. Tabassi, P. Hall and K. Roberts (2024). Artificial intelligence risk management framework: Generative artificial intelligence profile. Technical Report NIST AI 600-1. National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.AI.600-1>.
- Bitektine, A. and P. Haack (2015). The macro and the micro of legitimacy: Toward a multilevel theory of the legitimacy process. *Academy of Management Review* **40**(1), 49–75. <https://doi.org/10.5465/amr.2013.0318>.
- Clark, A. and D. Chalmers (1998). The extended mind. *Analysis* **58**(1), 7–19. <https://doi.org/10.1093/analysis/58.1.7>.
- De Angelis, T., E. Nagali, F. Sciarrino and F. De Martini (2007). Experimental test of the no-signaling theorem. *Physical Review Letters* **99**(19), 193601. <https://doi.org/10.1103/PhysRevLett.99.193601>.
- European Union (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). EUR-Lex. <http://data.europa.eu/eli/reg/2024/1689/oj>.

- Gerlich, M. (2024). Exploring motivators for trust in the dichotomy of human–AI trust dynamics. *Social Sciences* **13**(5), 251. <https://doi.org/10.3390/socsci13050251>.
- Horodecki, R., P. Horodecki, M. Horodecki and K. Horodecki (2009). Quantum entanglement. *Reviews of Modern Physics* **81**(2), 865–942. <https://doi.org/10.1103/RevModPhys.81.865>.
- Jaakkola, E. (2020). Designing conceptual articles: Four approaches. *AMS Review* **10**(1–2), 18–26. <https://doi.org/10.1007/s13162-020-00161-0>.
- Kabalska, A. and R. Wagner (2025). The better ones: The rise of human digital twins—our future or our demise?. *Future Business Journal* **11**, 155. <https://doi.org/10.1186/s43093-025-00582-y>.
- Kritzinger, W., M. Karner, G. Traar, J. Henjes and W. Sihn (2018). Digital twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine* **51**(11), 1016–1022. <https://doi.org/10.1016/j.ifacol.2018.08.474>.
- Logg, J. M., J. A. Minson and D. A. Moore (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes* **151**, 90–103. <https://doi.org/10.1016/j.obhdp.2018.12.005>.
- Mayer, A., L. Greif, T. M. Häußermann, S. Otto, K. Kastner, S. El Bobbou, J.-R. Chardonnet, J. Reichwald, J. Fleischer and J. Ovtcharova (2025). Digital twins, extended reality, and artificial intelligence in manufacturing reconfiguration: A systematic literature review. *Sustainability* **17**(5), 2318. <https://doi.org/10.3390/su17052318>.
- Schardong, F. and R. Custódio (2022). Self-sovereign identity: A systematic review, mapping and taxonomy. *Sensors* **22**(15), 5641. <https://doi.org/10.3390/s22155641>.
- Schilke, O. and M. Reimann (2025). The transparency dilemma: How AI disclosure erodes trust. *Organizational Behavior and Human Decision Processes* **188**, 104405. <https://doi.org/10.1016/j.obhdp.2025.104405>.
- Stacchio, L., A. Angeli and G. Marfia (2022). Empowering digital twins with extended reality collaborations. *Virtual Reality & Intelligent Hardware* **4**(6), 487–505. <https://doi.org/10.1016/j.vrih.2022.06.004>.
- Tabassi, E. (2023). Artificial intelligence risk management framework (AI RMF 1.0). Technical Report NIST AI 100-1. National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.AI.100-1>.
- Tomlinson, E. C. and A. Schnackenberg (2022). The effects of transparency perceptions on trustworthiness perceptions and trust. *Journal of Trust Research* **12**(1), 1–23. <https://doi.org/10.1080/21515581.2022.2060245>.
- van der Valk, H., H. Haße, F. Möller and B. Otto (2022). Archetypes of digital twins. *Business & Information Systems Engineering* **64**(3), 375–391. <https://doi.org/10.1007/s12599-021-00727-7>.
- World Wide Web Consortium (2022). Decentralized identifiers (DIDs) v1.0. W3C Recommendation. <https://www.w3.org/TR/did-core/>.
- World Wide Web Consortium (2025). Verifiable credentials data model v2.0. W3C Recommendation. <https://www.w3.org/TR/vc-data-model-2.0/>.
- Zhu, D. and W. Sulaiman Khail (2025). What makes a digital human twin more than a simulation? a computational-ecological stance. *AI & Society*. Advance online publication. <https://doi.org/10.1007/s00146-025-02796-8>.
- Zhu, E. and S. Yang (2025). Towards human digital twin: Reviewing human modelling and simulation. *Journal of Industrial Information Integration* **48**, 100975. <https://doi.org/10.1016/j.jii.2025.100975>.