

Networked Authenticity in AI-Mediated Architecture: A Conceptual Model and Interpretive Assessment Matrix

Article Info	Abstract
Article type: <i>Research Article</i> Article history: Received: [date] Received in revised form: [date] Accepted: [date] Available online: [date] (editorial office completes this) Keywords: Architectural authenticity, AI-mediated architecture, Networked authenticity, Distributed agency, Posthumanism	Abstract This study develops a conceptual and operational framework for rethinking authenticity in architecture mediated by artificial intelligence. Although classical accounts have established the importance of origin, authorship, agency, lived experience, meaning, value, and responsibility, their human-centered and author-centered assumptions do not fully explain design processes shaped by data, algorithms, generative models, prompts, platforms, software, and repeated human selection. To address this gap, the research adopts a qualitative, theoretical, and interpretive design based on a purposively selected interdisciplinary corpus of theoretical sources from architecture, philosophy, and technology studies. The material was examined through qualitative content analysis, conceptual analysis, comparative analysis, and theoretical inference, moving iteratively from initial codes to axial categories, comparative typologies, and relational synthesis. The findings identify five interrelated dimensions of authenticity: origin and provenance; authorship and attribution; agency and actor networks; experience, meaning, and value; and responsibility, human judgment, and process traceability. These dimensions support the formulation of networked authenticity as a relational, processual, interpretive, and traceable condition whereby human and technological contributions are differentiated but assessed within the same production network. The proposed model is further operationalized through an interpretive matrix that enables multidimensional evaluation without reducing authenticity to a single score. A constructed worked example demonstrates the application of the matrix and its capacity to distinguish among different levels of evidential adequacy across the five dimensions, although empirical validation across documented architectural cases remains for future research. The framework provides a basis for analyzing artificial intelligence-mediated architectural production while retaining human judgment, contextual interpretation, and professional responsibility. Objective: This study addresses the limited capacity of human-centered and author-centered accounts of authenticity to explain architectural production mediated by generative artificial intelligence. It aims to reconceptualize authenticity under conditions in which human judgment interacts with data, algorithms, generative models, prompts, platforms, and software. Method: A qualitative, theoretical, and interpretive research design was employed. A purposively selected interdisciplinary corpus of theoretical sources from architecture, philosophy, and technology studies was examined through qualitative content analysis, conceptual analysis, comparative analysis, and theoretical inference. The analysis moved iteratively from initial codes to axial categories, comparative typologies, and relational synthesis. Results: The study identified five analytical dimensions of authenticity: origin and provenance; authorship and attribution; agency and actor networks; experience, meaning, and value; and responsibility, human judgment, and process traceability. The findings showed that classical approaches remain conceptually important but do not adequately explain distributed production networks. Their comparison with posthumanism, actor-network theory, and technological mediation supported the formulation of a relational model of networked authenticity. Conclusions: Networked authenticity defines authenticity as a relational, processual, interpretive, and traceable condition rather than an attribute grounded solely in individual authorship or singular origin. The model was further translated into an operational interpretive matrix that enables multidimensional assessment without reducing authenticity to a single score. The matrix provides an analytical basis for examining human and technological participation while retaining human judgment and responsibility, although it has not yet been empirically tested through documented architectural cases.

Cite this article: *Last Name, Initial., Last Name, Initial., & Last Name, Initial. ([Year]). Title of the paper in sentence case. Journal of Design Thinking, [Volume]([Issue]), [pages]. [https://doi.org/\[DOI\]](https://doi.org/[DOI])*

© Author(s) retain the copyright. This is an open-access article distributed under the terms of the Creative Commons Attribution (CC-BY) License.

Publisher: University of Tehran Press. **DOI:** [https://doi.org/\[to be assigned\]](https://doi.org/[to be assigned])

JDT Manuscript

Introduction

In recent decades, the rapid development of digital technologies, particularly artificial intelligence, has altered the conditions of knowledge production and professional practice. Architecture, as a field closely connected to cultural, technological, and social change, has been directly affected by these developments. Algorithmic tools, generative models, large datasets, machine-learning systems, and digital platforms increasingly participate in the processes through which architectural form is conceived, represented, evaluated, and developed. Architectural design can therefore no longer be explained solely through a direct relation between architect, idea, and finished work. New configurations of production, representation, and decision-making have become part of the design process and increasingly shape the conditions of architectural creation (Del Campo, 2024).

This transformation has developed through successive changes in digital architectural practice. Early digital tools were used primarily for representation and modeling, while computational and parametric approaches introduced algorithmic procedures, relational systems, and the systematic exploration of design alternatives (Kolarevic, 2003; Terzidis, 2006). Generative systems have extended this trajectory by enabling the production of multiple outputs through interactions among human decisions, data, algorithms, models, prompts, and platforms (González & Garcia, 2024).

Architects working with text-to-image models, generative design tools, optimization software, and computational platforms therefore operate within design processes in which decisions are increasingly mediated by technological systems. The architect's role consequently extends beyond the direct production of form to include problem framing, system configuration, interpretation, selection, revision, and evaluation as part of a broader production network (Lee, 2025).

Much of the recent discussion of artificial intelligence in architecture has focused on technical capability, image generation, speed, variation, and software performance. Questions concerning the origin and provenance of the work, authorship and attribution, the distribution of agency, architectural meaning and value, and responsibility and traceability remain less integrated into a coherent theoretical account. Classical approaches to authenticity, developed across philosophy, aesthetics, and architectural theory, provide important accounts of human origin, authorial intention, individual creativity, lived experience, meaning, and responsibility. Their concepts, however, become difficult to apply without reformulation when architectural production involves relations among human actors, data, algorithms, generative models, prompts, platforms, contextual conditions, and successive acts of interpretation and selection (Colmenero-Acevedo, 2025).

Posthumanism, actor-network theory, and technological mediation provide theoretical resources for examining these altered relations. Rather than treating the architectural work as the product of an isolated and self-sufficient human author, these frameworks direct attention to relations among human actors, technological systems, environments, and mediating conditions. They allow authorship and agency to be reconsidered as relational and distributed while preserving distinctions among technological participation, human judgment, interpretation, and responsibility (Beerends & Aydin, 2021, 2025). The literature reviewed in this study, however, does not provide an integrated account that connects these relational approaches with classical theories of authenticity specifically in relation to AI-mediated architectural production.

The present study addresses this gap by reconsidering authenticity through a comparative analysis of classical approaches and posthuman and technology-oriented frameworks. It seeks to formulate networked authenticity as a relational, processual, interpretive, and traceable account of architectural production where the contributions of human actors, technological systems, and contextual conditions can be distinguished and examined as part of the same production network. The study further translates this conceptual model into an operational interpretive matrix for examining origin and provenance, authorship and attribution, agency and actor networks, experience, meaning and value, and responsibility, human judgment, and process traceability. The aim is therefore not to discard classical accounts of authenticity, but to reorganize and extend their conceptual components under conditions of technological mediation while retaining human interpretation and accountability.

Literature Review

Classical approaches to authenticity

Authenticity is a multilayered concept whose meanings have shifted across historical periods while remaining linked to questions of truth, origin, identity, and value (Benjamin, 1969; Taylor, 1992). Etymologically, the Persian-Arabic term derives from *asl*, meaning root, foundation, or point of origin (Ibn Manzur, 1993). The English term *authenticity* derives from the Latin *authenticus* and the Greek *authentikos*, and refers to the quality of being genuine, reliable, and faithful to an original source (Oxford University Press, 2005). In Persian usage, the concept similarly denotes rootedness, purity, and possession of an identifiable origin (Amid, 1971). In architectural discourse, these semantic associations have been extended to questions of creativity, identity, spatial experience, cultural continuity, and value. Four major theoretical orientations can be distinguished.

The romantic and individualist orientation locates authenticity in the inner life of the author, personal creativity, and the capacity for individual expression. According to this orientation, authenticity is treated primarily as an internal quality centered on the individual creator rather than on social and historical conditions (Trilling, 1972). The constructivist approach, by contrast, understands authenticity as socially and historically formed through collective habitus, institutional fields, cultural capital, conventions, and networks of cooperation. Authenticity is therefore contingent on processes of recognition, attribution, and historical formation (Bourdieu, 1993). The phenomenological orientation shifts attention to lived experience, embodiment, place, and the relation between the human subject and the world. Authenticity emerges through situated experience and through the formation of belonging, presence, and meaning in relation to place (Norberg-Schulz, 1980; Pallasmaa, 2005). The existentialist approach understands authenticity as a dynamic and formative process grounded in freedom, responsibility, existential choice, and continuing self-formation. From this perspective, authenticity is neither a fixed individual essence nor a purely social construction, but a commitment enacted under conditions of uncertainty and ambiguity (Golomb, 2012; Heidegger, 1962).

Although analytically distinguishable, these four approaches are not mutually exclusive. They overlap in architectural discourse, particularly in their treatment of creativity, identity, lived experience, meaning, choice, and responsibility. Their differences lie mainly in the position assigned to the individual, social context, embodied experience, and existential decision regarding the formation of authenticity.

Across these approaches, several recurring components can be identified: individual creativity and authorial intention; social legitimacy and historical continuity; lived experience, embodiment, and place identity; and freedom, responsibility, and self-formation. When artificial intelligence becomes involved in architectural production, these components are placed under new pressure. Questions arise concerning origin and provenance, authorship and attribution, the distribution of agency, the relation of generated outputs to architectural meaning and value, and the traceability of human and institutional responsibility.

Table 1: [Comparative overview of major theoretical approaches to authenticity.] (Source: authors)

Approach	Primary focus	Conception of authenticity	Core components
Romantic (individualist)	Individual creativity, inner expression, and authorial intention	Primarily internal and author-centered, associated with personal expression and creative individuality	Human individuality, creative genius, intrinsic creativity, authorial intention, individual expression
Constructivist	Social construction, historical formation, and collective networks	Socially and historically constituted through recognition, institutions, conventions, and collective relations	Social legitimacy, historicity, cultural capital, collective identity, social conventions
Phenomenological	Lived experience, embodiment, and relation to place	Experiential and relational, shaped through embodied encounter with place and the world	Lived experience, embodiment, spirit of place, multisensory perception, subject–world relation

Approach	Primary focus	Conception of authenticity	Core components
Existentialist	Freedom, responsibility, existential choice, and self-formation	Processual and dynamic, grounded in choice, responsibility, and continuing formation of the self	Freedom, existential responsibility, self-creation, self-formation, existential choice

Digital transformation and the rise of generative systems

The digital turn in architecture cannot be reduced to a change of instruments. It constitutes an epistemological shift that redefines how form, process, and creativity are conceived (Carpo, 2017). Early computer-aided design operated largely under a representational paradigm, whereas parametric design introduced systems of variables and relational rules through which form could emerge from defined relationships rather than from static geometric description. Computational design further expanded the role of algorithms, simulation, and systematic exploration of design space. Generative design introduced higher levels of procedural automation and the systematic production of design alternatives, with the architect defining rules, constraints, initial conditions, or conceptual parameters with which variations could be produced (Oxman, 2006). Generative artificial intelligence extends this trajectory through systems capable of producing representations, formal alternatives, and rule-conditioned design proposals.

In this environment, the architect's position shifts further away from the image of a singular author toward that of a system designer, interpreter, selector, and participant in a distributed production network. Design outcomes increasingly emerge through interactions among human decisions, computational processes, data, models, and platforms. Under these conditions, classical criteria based on singular origin, individual authorship, and self-sufficient human agency become difficult to apply without substantial reformulation (Yaneva, 2022).

Challenges to authenticity under AI mediation

Theoretical literature on AI-mediated design raises several recurring concerns about authenticity. These include the provenance of generated outputs and the entanglement of data, prompts, and generative models; the attribution of authorship across architects, systems, platforms, and data sources; the redistribution of design decisions across human and nonhuman actors; the relation of generated outputs to architectural meaning, value, context, and experience; and the transparency, traceability, documentation, and accountability of the production process (Colomero-Acevedo, 2025; Iapaolo & Lynch, 2026; Zhang et al., 2023). These concerns are closely related, but they are discussed across different bodies of literature rather than as an integrated account of authenticity. Their coexistence therefore points to a broader theoretical problem concerning how authenticity can be interpreted when architectural production is distributed across human judgment, computational systems, data, and technological mediation.

Posthuman and technology-oriented frameworks

Posthumanism, as used in this study, refers to a critical theoretical orientation that challenges the humanist conception of the autonomous, self-sufficient, and exclusively agentic human subject. Rather than eliminating the human, posthumanism understands human agency and subjectivity as relationally constituted through entanglements with bodies, environments, technologies, and nonhuman entities (Braidotti, 2013; Wolfe, 2010). This relational orientation is also informed by foundational critiques that destabilize fixed boundaries between the human and the technological, particularly Haraway's cyborg formulation and Hayles's account of embodied posthuman subjectivity in relation to information and technological systems (Haraway, 1991; Hayles, 1999).

Actor-network theory (ANT) is understood here as a relational approach that examines how action and outcomes emerge through associations among heterogeneous human and nonhuman actors rather than being attributed in advance to an autonomous human subject. It therefore provides a means of tracing how architects, design teams, data, algorithms, generative models, software, platforms, institutions, and other actants participate differently in architectural production (Latour, 2005; Yaneva, 2022). Technological mediation, as developed within postphenomenology, refers to the process through which technologies actively shape human perception, action, interpretation, and moral judgment rather than functioning as neutral instruments. Relations of embodiment, hermeneutics, alterity, and background describe different ways in which technologies mediate the designer's relation to the world and reconfigure the range of available perceptions, choices, and actions (Ihde, 1990; Verbeek, 2005, 2011).

Together, these frameworks make it possible to describe AI-mediated architectural production as a socio-technical network whereby data, models, prompts, platforms, human selection, revision, contextual interpretation, and professional judgment contribute differently to the formation of the work. They do not by themselves provide a complete theory of authenticity. Rather, they offer relational concepts through which origin, authorship, agency, meaning, and responsibility can be reconsidered beyond exclusively human-centered and author-centered accounts.

Table 2: *[Core theoretical resources for rethinking authenticity under AI mediation.]* (source: authors)

Framework	Key contribution	Relevance to authenticity in AI-mediated architecture
Posthumanism	Critique of the autonomous human subject; entanglement of human, technology, and environment	Prevents the reduction of authenticity to a solitary human origin or pure machine autonomy
Actor-network theory	Distributed agency; tracing of human and non-human actants; mediation and translation	Enables the description of design as networked rather than author-centered
Technology mediation (postphenomenology)	Technologies co-constitute perception, action, and moral evaluation	Clarifies how generative systems mediate perception, action, selection, judgment, and responsibility

Research gap and rationale

Existing studies have begun to address the relation between artificial intelligence and authenticity in architecture, but they do so from different and relatively bounded analytical perspectives. Colmenero-Acevedo (2025), for example, explicitly connects AI-generated architectural design with questions of authenticity and foregrounds its ethical implications. This work establishes an important point of departure, but it does not resolve the broader theoretical problem addressed in the present study: how the conceptual structure of architectural authenticity itself should be reformulated when origin, authorship, agency, experience, meaning, value, and responsibility are simultaneously affected by AI-mediated production. Similarly, studies of authenticity in technological environments and AI discourse provide important accounts of human-technology relations but do not integrate these architectural dimensions into a single relational framework (Beerends & Aydin, 2021, 2025). The gap addressed here therefore lies not in being the first to associate artificial intelligence with authenticity, but in the absence of an integrative framework that connects classical theories of authenticity with posthumanism, actor-network theory, and technological mediation and examines the resulting dimensions as interdependent components of the same architectural production network.

This theoretical gap becomes more evident when the components inherited from classical accounts of authenticity are considered in relation to AI-mediated architectural production. As shown in Figure 1, questions of origin, authorship, agency, meaning and value, and responsibility become interconnected once data, algorithms, generative models, prompts, platforms, and human selection enter the production process. The figure therefore synthesizes the conceptual transition from classical components of authenticity to the set of problems that motivate their reconsideration under conditions of technological mediation.

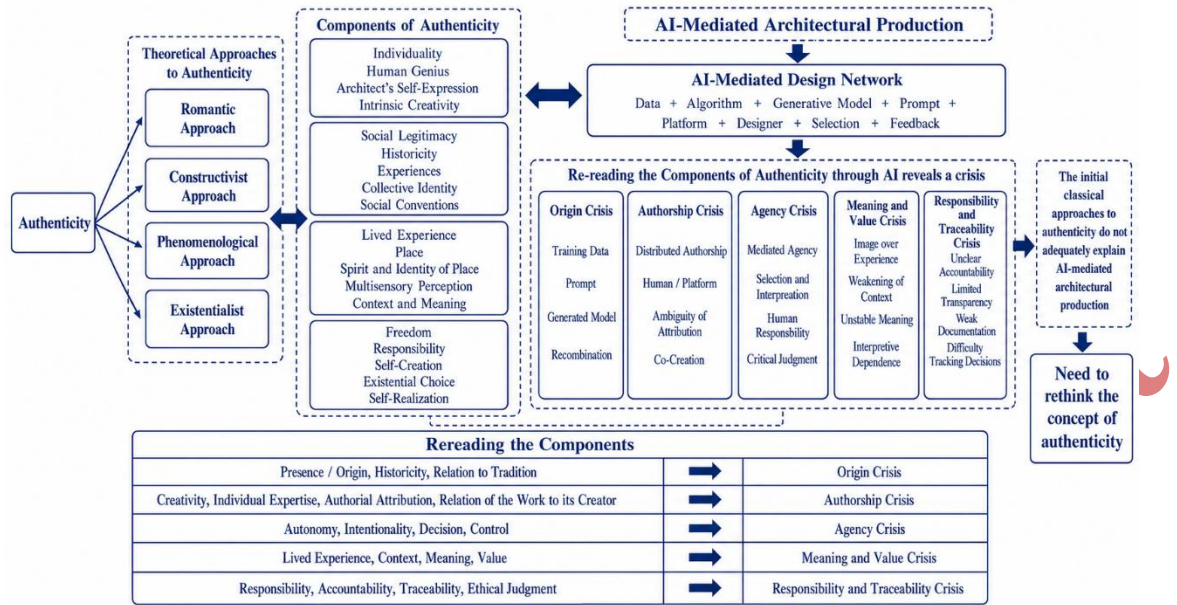


Figure 1: [Conceptual synthesis of classical authenticity components, AI-mediated design relations, and the resulting theoretical gap.] (Source: authors)

The present research therefore addresses this gap through an integrative rather than a single-lens account of authenticity. It brings classical approaches to authenticity into systematic comparison with posthumanism, actor-network theory, and technological mediation and reorganizes their implications across five interdependent dimensions: origin and provenance; authorship and attribution; agency and actor networks; experience, meaning, and value; and responsibility, human judgment, and process traceability. On this basis, the study formulates networked authenticity as a relational account in which these dimensions are examined through their mutual dependencies rather than as separate attributes. The contribution is further extended by translating this conceptual structure into an interpretive assessment matrix designed to examine the production network multidimensionally. The novelty of the study therefore lies in the combination of theoretical integration, relational reconstruction, and interpretive operationalization, rather than in introducing authenticity as a concern in AI-mediated architecture for the first time.

Methodology

This study employed a qualitative, theoretical, and interpretive research design, combining qualitative content analysis, conceptual analysis, comparative analysis, and theoretical inference to reconsider authenticity in AI-mediated architecture. The research followed a nonlinear and iterative process that moved between theoretical texts, extracted concepts, and the central research problem, integrating inductive and deductive reasoning. The analytical material consisted exclusively of theoretical texts rather than empirical observations or numerical measurements. The study aimed to identify the conceptual dimensions of authenticity, develop the model of networked authenticity, and translate its components into an operational interpretive matrix. Ethical considerations were addressed through adherence to scholarly standards of source attribution, conceptual fidelity, and analytical transparency. As the study involved no human participants, institutional ethics approval and informed consent procedures were not required.

The inquiry was structured around one main research question and three subsidiary questions rather than a testable statistical hypothesis. These questions guided source selection, coding, comparative analysis, and theoretical synthesis. Table 3 presents the research questions together with their analytical focus and expected contribution.

Table 3: [Summary of research questions and corresponding analytical levels.] (source: authors)

Research question	Analytical focus	Expected contribution
Main question: How can authenticity be reconceptualized in AI-mediated architecture beyond human-centered and author-centered accounts?	Relational reconsideration of origin, authorship, agency, experience, meaning, value, and responsibility under technological mediation	Formulation of networked authenticity and its operationalization through an interpretive assessment framework
Subsidiary question 1: What principles of classical approaches to authenticity become limited under conditions of AI-mediated architectural production?	Romantic and individualist, constructivist, phenomenological, and existential accounts of origin, authorship, agency, experience, meaning, and responsibility	Identification of the conceptual capacities and limitations of classical authenticity approaches
Subsidiary question 2: What challenges does AI mediation introduce for the interpretation of authenticity in architectural production?	Origin and provenance, authorship and attribution, distributed agency, experience, meaning and value, and responsibility and traceability	Identification and organization of the principal dimensions through which authenticity is challenged under AI mediation
Subsidiary question 3: How can posthumanism, actor-network theory, and technological mediation contribute to a relational account of authenticity?	Human and nonhuman actors, distributed and mediated agency, networks of production, technological mediation, human judgment, and responsibility	Specification of the relational requirements for networked authenticity and the conceptual basis for its interpretive assessment

The analytical corpus comprised a purposefully selected body of theoretical texts drawn from architecture, philosophy, and technology studies. Sources were selected for their conceptual relevance to authenticity, origin, authorship, agency, meaning, value, responsibility, data, algorithms, generative models, prompts, platforms, mediation, networks, and actants. The corpus included foundational texts on romantic, constructivist, phenomenological, and existential approaches to authenticity, together with contemporary studies of digital architecture, generative design, artificial intelligence in architectural processes, posthumanism, actor-network theory, and technological mediation. The sources included peer-reviewed journal articles, scholarly books, book chapters, doctoral and master's theses, and theoretical monographs. Selection prioritized conceptual relevance and explanatory capacity over geographical or institutional boundaries, producing an interdisciplinary theoretical corpus rather than a geographically bounded sample.

Text selection followed a structured, purposive, and multi-stage process designed to ensure conceptual relevance and analytical coherence. Keywords were organized along three axes: authenticity-related terms (authenticity, architectural authenticity, authorship, agency, origin); architecture and artificial intelligence terms (generative AI in architecture, AI-mediated design, computational design, algorithmic design, and generative design); and posthuman and technology-related terms (posthumanism, actor-network theory, technological mediation, postphenomenology, distributed agency, and human–technology relations). Searches were conducted in domestic databases (SID, Magiran, Noormags, and Elmnet) and international databases (Google Scholar, Web of Science, Scopus, ScienceDirect, SpringerLink, and Taylor & Francis). Titles, abstracts, and keywords were screened first, after which full texts were assessed for their theoretical contribution to the research problem. Temporal coverage was differentiated: recent publications were prioritized for artificial intelligence and digital architecture to reflect current developments, while foundational texts were retained for the core concepts of authenticity, phenomenology, existentialism, posthumanism, and actor-network theory. Conceptual saturation was operationalized as the point at which two consecutive review rounds, each comprising five newly eligible full-text sources, produced no new initial code requiring the creation of an additional axial category and no modification to the emerging five-dimensional analytical structure. Sources examined after this point contributed further examples, conceptual clarification, or refinement within existing categories rather than introducing a new category. On this basis, the search was considered conceptually saturated after ten consecutive eligible sources failed to alter the established category structure.

The research protocol unfolded through successive and interrelated stages. The first stage consisted of documentary and library-based investigation of authenticity in general and architectural authenticity in particular. Romantic, constructivist, phenomenological, and existential traditions were examined to identify their foundational concepts and establish an initial conceptual framework through qualitative content analysis, conceptual analysis, and theoretical inference. The second stage examined how artificial intelligence alters understandings of origin, authorship, agency, meaning, value, and responsibility in architectural design. Critical reading of literature on digital architecture, generative design, computational creativity, and changing design agency identified the limitations of classical human-centered and author-centered accounts under conditions of distributed agency and technological mediation. The third stage

brought the findings of the preceding stages into comparison and identified the persistence of anthropocentric and author-centered assumptions as a central theoretical tension. This stage directed the analysis toward repositioning the human designer among networks of human and nonhuman actors without removing human judgment, interpretation, and responsibility from the design process.

The fourth stage examined the capacity of posthumanism, actor-network theory, and technological mediation to support a reconceptualization of authenticity. Texts addressing human–technology relations, distributed agency, embodiment, data, algorithms, platforms, meaning, value, and responsibility were prioritized, classified, and analyzed. Sources were organized through iterative cycles of divergence and convergence, first separating materials to clarify their theoretical origins and then reconnecting them in relation to the research problem. Preliminary codes included origin, authorship, agency, experience, embodiment, data, algorithm, prompt, platform, mediation, actant, network, meaning, value, and responsibility. Through repeated reading, comparison, and refinement, these codes were organized into axial categories and then into five analytical dimensions: origin and provenance; authorship and attribution; agency and actor networks; experience, meaning, and value; and responsibility, human judgment, and process traceability. These dimensions formed the basis of the comparative typology and the intermediate relational synthesis reported in the Results. The five dimensions were not treated as independent analytical variables. Once derived through coding and category development, they were examined relationally, with particular attention to how changes in one dimension affected the interpretation of the others. Provenance, for example, conditions how authorship can be attributed; attribution shapes the distribution of agency; and these relations influence the interpretation of meaning, value, responsibility, and traceability. The relational synthesis therefore examined both the internal content of each dimension and the dependencies among them. This interdependence constituted an analytical premise of the subsequent theoretical model rather than a secondary observation added after the dimensions had been identified. The comparison also distinguished among relatively stable theoretical assumptions, procedurally modifiable components, and elements that required broader theoretical reconstruction.

Selected texts were analyzed at two levels. The first level extracted codes, concepts, and categories related to authenticity, architecture, artificial intelligence, authorship, agency, posthumanism, actor-network theory, and technological mediation. The second level examined the arguments, implications, and theoretical conclusions advanced in the sources. Overlapping concepts were organized into comparative tables that recorded both the presence of specific categories and the relative emphasis placed on them across classical traditions of authenticity and posthuman or technology-oriented frameworks. Relative emphasis was assessed qualitatively rather than through numerical frequency counts. An emphasis was classified as high when a category formed a recurring and central part of a theoretical domain's definitions, arguments, and conclusions; medium when it was present but secondary to other concerns; and limited when it appeared only occasionally or remained peripheral to the domain's main explanatory structure. These tables supported structured comparison rather than mere data aggregation.

The analytical procedure combined qualitative content analysis, conceptual analysis, comparative analysis, and theoretical inference. Propositions, definitions, arguments, concepts, and conceptual relations extracted from books, journal articles, book chapters, theses, and other theoretical texts were treated as the units of analysis. Qualitative content analysis was used to identify and organize recurring concepts and categories; conceptual analysis clarified their meanings and internal relations; comparative analysis examined similarities and differences among the theoretical traditions; and theoretical inference connected the resulting categories in the development of the relational framework and conceptual model. These strategies were adopted in accordance with the qualitative, textual, and interpretive character of the research data (Krippendorff, 2004). Rather than applying statistical procedures or seeking numerical generalizability, the analysis examined the relations among human actors, technological systems, data, algorithms, architectural works, meaning, value, and responsibility as an integrated conceptual and networked whole.

The methodological structure of the study rests on the principle that qualitative and interpretive inquiry allows a complex and multi-layered phenomenon to be examined as a coherent whole. By identifying, organizing, and recombining the ideas and concepts embedded inside it, the researcher can clarify the internal logic of the phenomenon in relation to its relevant context (Groat & Wang, 2002). Authenticity in AI-mediated architecture was therefore treated not as a fixed, independent, or predefined concept, but as a relational, mediated, and networked phenomenon whose interpretation required the simultaneous

examination of classical theories of authenticity, developments in digital architecture, and posthuman and technology-oriented frameworks.

The qualitative material was reduced, classified, and organized through successive divergent and convergent stages. In the divergent stage, concepts drawn from the three principal theoretical axes were examined separately to preserve distinctions among sources and theoretical traditions. The initial codes identified in the fourth stage were displayed in comparative tables and conceptual matrices and repeatedly examined across the theoretical corpus. In the convergent stage, related codes were compared, refined, and recombined into axial categories and subsequently organized into the five analytical dimensions. This process enabled the research to move from dispersed theoretical material to comparative analysis, the identification of different levels of conceptual change, intermediate relational synthesis, and theoretical inference.

Following coding and classification, the converged material was assembled in analytical tables and conceptual matrices and subjected to comparative content analysis. Analysis proceeded along two principal axes. The first axis examined the intersection of theoretical approaches to authenticity with the theoretical foundations of architecture and the design process. The second axis examined authenticity in relation to AI-mediated architecture through posthumanism, actor-network theory, and technological mediation. Comparison across the two axes produced the five analytical dimensions, the qualitative matrix of relative theoretical emphasis, the typology of approaches to authenticity, the distinction among stable assumptions, modifiable components, and requirements for theoretical reconstruction, and the intermediate relational synthesis presented in the Results.

The two analytical axes were compared through a combination of inductive and deductive reasoning. Origin, authorship, agency, experience, meaning, value, responsibility, and technological mediation served as the principal axes of comparison. The comparative stage produced an intermediate relational synthesis rather than the final model. This synthesis identified the conceptual relations that required further theoretical integration and provided the basis for addressing the central research question without statistical hypothesis testing. In the subsequent stage of theoretical inference, the five analytical dimensions and the intermediate relational synthesis were reconsidered through the distinct contributions of posthumanism, actor-network theory, and technological mediation. This stage produced the final model of networked authenticity by connecting the objective-technological conditions of architectural production with perceptual-interpretive conditions, human judgment, and responsibility.

Following construction of the final conceptual model, a deductive operationalization stage translated its five dimensions into an interpretive assessment matrix. For each dimension, an analytical question was formulated, and the types of information and documentation relevant to its assessment were specified. Three descriptive levels, insufficient, partial, and substantiated, were defined to indicate the adequacy with which each dimension could be examined. The numerical labels 0, 1, and 2 were used only as ordinal identifiers for these levels and were not treated as quantitative measurements of authenticity. The five assessments were designed to form a multidimensional interpretive profile and were not combined into a total score. Reconstructable provenance and identifiable human or institutional responsibility were specified as minimum interpretive conditions for an overall assessment, because the framework cannot connect production, attribution, and accountability when either relation is entirely undocumented. The matrix was developed theoretically from the categories and relations identified in the study and was subsequently illustrated through a constructed scenario; it was not empirically validated through documented architectural cases in the present research.

The overall methodological trajectory is summarized in Figure 2, which traces the movement from the theoretical corpus and initial coding to the five analytical dimensions, comparative typology, intermediate relational synthesis, final conceptual model, and operational interpretive matrix.

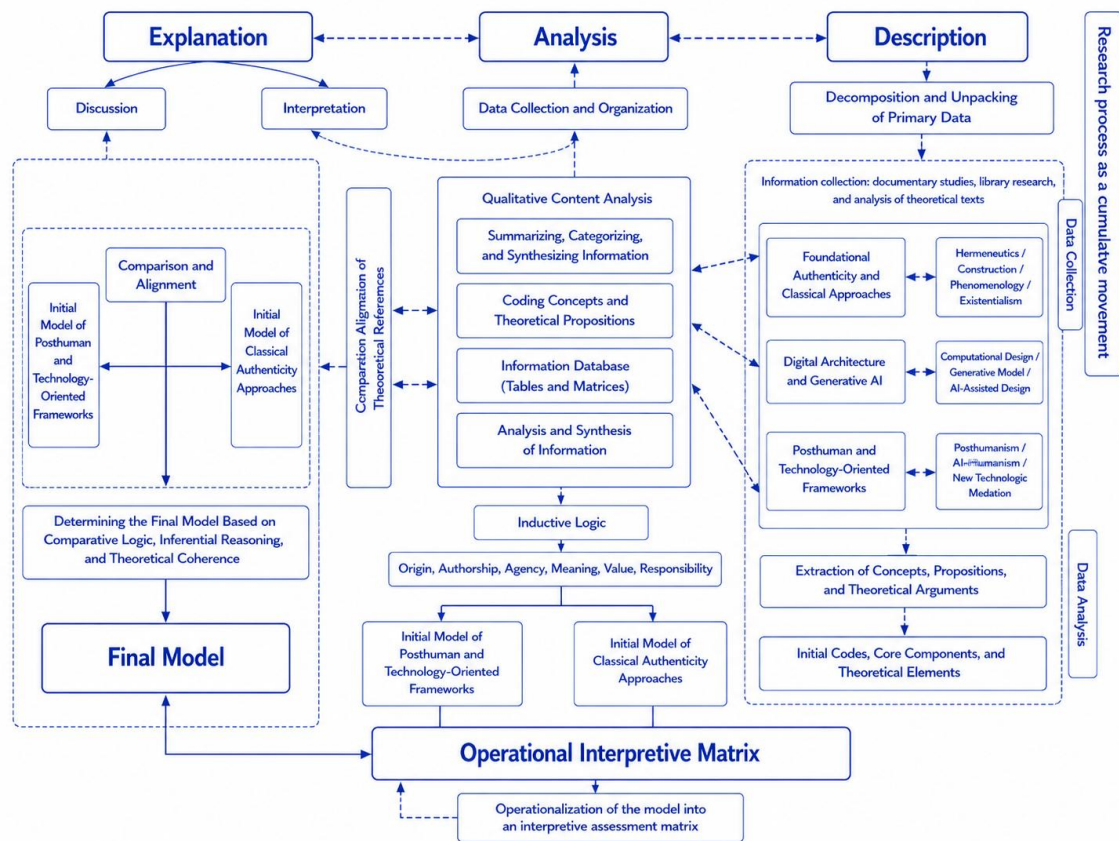


Figure 2: [Analytical methodology model of the study.] (source: authors)

Results

The analysis identified three related findings: five analytical dimensions of authenticity, a comparative typology of theoretical approaches, and an intermediate relational framework connecting the human and technological components of architectural production. The sources came from different disciplinary and theoretical traditions, yet they repeatedly addressed the same broad concerns: origin, authorship, agency, experience, meaning, value, and responsibility. They differed mainly in how they defined these concepts, the weight they assigned to each one, and whether they located them in human actors, social and historical settings, technological systems, or networks of relations. Five recurring theoretical orientations emerged from the comparison. The first concerned the movement from a single human origin to a multilayered provenance shaped by processes and networks. The second extended authorship beyond the individual creator to collective, processual, and distributed forms of production. The third treated design agency as mediated and distributed among human participants, technical systems, and other actors involved in the design process. The fourth retained experience, meaning, and value as central components of authenticity and related them to embodiment, perception, presence, and architectural context. The fifth addressed responsibility, human judgment, process transparency, and the traceability of design decisions.

The theoretical domains differed in the emphasis they gave to these orientations. Classical accounts concentrated mainly on origin, authorship, experience, and meaning, with less attention to responsibility and traceability in technologically mediated production. Literature on digital architecture and artificial intelligence gave greater attention to computational and processual origin, mediated agency, and the documentation of design processes. Posthumanist sources focused more strongly on distributed authorship and agency. Actor-network theory examined how origin, authorship, and agency arose through relations among heterogeneous actors. Technological mediation linked the role of technology in shaping action and perception with questions of human judgment and responsibility. Table 4 summarizes these qualitative differences. The categories indicate relative conceptual emphasis in the analyzed texts and do not represent statistical frequencies.

Table 4: [Relative emphasis of the analyzed theoretical domains on the five principal theoretical orientations.] (source: authors)

Theoretical orientation	Classical authenticity	Digital architecture and AI	Posthumanism	Actor-network theory	Technological mediation
Reframing origin from an individual and singular source toward a multilayered, processual, and networked origin	High	High	Medium	High	Medium
Extending authorship from a single author toward collective and distributed authorship in a production network	High	Medium	High	High	Medium
Formulating design agency as mediated, distributed, and networked	Medium	High	High	High	High
Relating experience, meaning, and value to perception, embodiment, presence, and context	High	Medium	Medium	Limited	High
Relating responsibility and human judgment to process transparency and decision traceability	Limited	High	Medium	Medium	High

The coding and category development produced five analytical dimensions: origin and provenance; authorship and attribution; agency and actor networks; experience, meaning, and value; and responsibility, human judgment, and process traceability. These dimensions brought together concepts found across the theoretical domains while keeping their differences in emphasis and meaning. These dimensions are analytically distinguishable but not conceptually independent. Their significance lies in the relations among them: the provenance of a work affects the basis on which authorship is attributed; authorship influences how agency is distributed; the configuration of agency shapes the interpretation of experience, meaning, and value; and each of these relations affects the assignment and traceability of responsibility. Conversely, limitations in responsibility or traceability can restrict how confidently provenance, authorship, and agency can be interpreted. Networked authenticity is therefore not conceived as the additive combination of five separate attributes, but as a relational configuration in which the dimensions mutually condition one another.

The first dimension, origin and provenance, included human, historical, computational, and networked accounts of how a work comes into being. Human origin was associated with authorial intention, individual creativity, and the relation between the work and its maker. Historical origin concerned the work's presence, its relation to tradition, its temporal conditions, and the possibility of reproduction. Computational origin included algorithms, parameters, systems, and procedures for generating form. Networked origin extended this account to data, generative models, prompts, platforms, software, and the participation of human and nonhuman actors.

The second dimension, authorship and attribution, included individual, historical and discursive, collective and processual, and distributed accounts of authorship. Individual authorship connected the work to the architect's intention, creativity, will, and personal expression. Historical and discursive accounts situated attribution within traditions, institutions, discourses, and systems of recognition. Collective and processual authorship included teamwork, division of labor, successive design decisions, and the participation of several human contributors. Distributed authorship included the relations among architects, design teams, data, software, generative models, prompts, and platforms.

The third dimension, agency and actor networks, included human, collective, technologically mediated, and distributed forms of agency. Human agency concerned intention, choice, design decisions, evaluation, and responsibility. Collective agency referred to coordinated action among several participants. Technologically mediated agency described how tools, interfaces, software, and models shaped available options, perceptions, and decisions. Distributed agency concerned the production of design outcomes through relations among architects, users, technical systems, data, algorithms, models, platforms, institutions, and other actors.

The fourth dimension, experience, meaning, and value, included lived experience, embodiment, spatial perception, presence, interpretive meaning, contextual value, and the perceptual qualities of the architectural work. It also included the relation between the visible outcome and the process through which the work had been produced. Historical, cultural, spatial, functional, and social contexts informed how the work was interpreted and valued. The fifth dimension, responsibility, human judgment, and process traceability, included human direction of the design process, interpretive judgment, documentation of decisions, transparency of data and production procedures, evaluation of generated outputs, ethical attribution, accountability, and trust. Classical accounts addressed this dimension less directly. It appeared more clearly in sources concerned with artificial intelligence, technological mediation, and distributed production.

The dimensions therefore function as connected analytical levels rather than as a hierarchy or a checklist of independent criteria. Origin concerns how the work came into being; authorship concerns how contributions are attributed; agency concerns how actions and decisions are distributed; experience, meaning, and value concern how the work is encountered and interpreted; and responsibility and traceability concern whether these processes and their consequences can be reconstructed and linked to accountable actors. Their analytical separation enables examination, while their relational interpretation constitutes the basis of the networked model.

The comparative analysis distinguished five broad theoretical patterns. Romantic approaches located authenticity in individual creativity, genius, personal expression, and the perceived presence of the author in the work. Origin, authorship, and agency were primarily associated with an individual human creator. Technology was generally positioned as an instrument through which the creator's intention was realized. Constructivist approaches located authenticity amid social, historical, cultural, institutional, and discursive conditions. Authorship was understood in relation to systems of recognition, traditions, and collective structures. Agency was not confined entirely to the individual, although it remained associated mainly with social and human relations rather than with active technological mediation.

Phenomenological approaches associated authenticity with lived experience, presence, embodiment, place, perception, and the disclosure of meaning through an encounter with the work. Authorship received less emphasis than the relation among the body, space, material conditions, and architectural experience. Technology entered these accounts mainly through its effects on perception and experience.

Existential approaches connected authenticity with choice, freedom, selfhood, intention, and responsibility. The human subject remained the principal bearer of agency and accountability. The authenticity of a work or action was therefore linked to the capacity of the human actor to choose, judge, and accept responsibility. Posthuman and technology-oriented frameworks produced a different arrangement of the same conceptual field. Origin was associated with relations among human actors, data, algorithms, generative models, prompts, and platforms. Authorship was formulated as processual and distributed across a production network. Agency was understood as mediated through interactions among human and nonhuman actors. Technology was consequently treated as an active mediator throughout the design process rather than only as an external instrument. This typology showed that the differences among the approaches did not arise from the complete presence or absence of origin, authorship, agency, meaning, or responsibility. They arose from the position assigned to these components and the relations established among them. Classical traditions generally organized authenticity around human origin, authorship, experience, and judgment. Later frameworks extended the analytical field to computational systems, technical mediators, data, platforms, and networks of production.

Figure 3 summarizes this typology by comparing how the five theoretical approaches located conceptual form, human agency, and nonhuman agency in relation to origin, authorship, and technological mediation

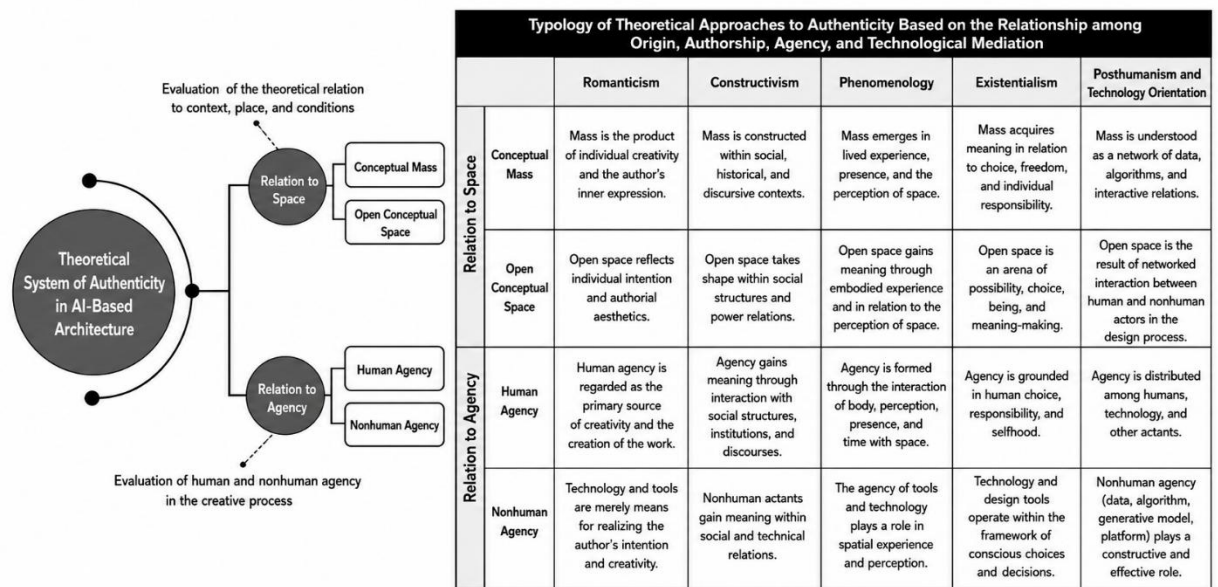


Figure 3: [Aggregation and analysis of theoretical models of authenticity based on the relationship among origin, authorship, agency, and technological mediation.] (source: authors)

The typology showed that the main difference among the approaches concerned the location of agency and the conditions through which architectural meaning was formed. Romantic and existential accounts assigned primary agency to the human subject. Constructivist and phenomenological approaches extended the analysis to social structures, embodied experience, and spatial conditions. Posthuman and technology-oriented frameworks included data, algorithms, generative models, platforms, and other nonhuman actors in the production of architectural form and meaning.

A second level of analysis distinguished among relatively stable assumptions, modifiable components, and requirements for further theoretical reconstruction. The understanding of the author as a singular creator, origin as a unique point of departure, and agency as self-sufficient human action remained closely tied to the premises of human-centered and author-centered theories. Fundamental alteration of these premises could not be achieved through minor revisions inside the same conceptual structure.

Other components were classified as open to procedural and analytical modification. These included documentation of datasets and prompts, identification of generative models and platforms, recording of successive outputs, clarification of the architect's interventions, and differentiation of the roles played by human actors and computational systems. Greater process transparency and clearer records of selection, revision, and evaluation were also included at this level.

A further group of findings required a broader theoretical reorganization. Distributed authorship, mediated agency, networked provenance, and the connection of technological operations with meaning, human judgment, and responsibility could not be accommodated through the simple addition of new technical variables to an author-centered account. These findings required the components of authenticity to be examined through their relations rather than as separate attributes.

The final stage of the findings produced an intermediate relational synthesis. Classical formulations tended to place the components of origin and authorship alongside, but partly separate from, the components of meaning and experience. This differentiated arrangement was replaced by an overlapping structure composed of a human side and a technological side.

The human side included the architect, users, design teams, social institutions, interpretive intention, experience, and judgment. The technological side included data, algorithms, generative models, prompts, platforms, and computational procedures. The intersection contained the relations through which contributions were coordinated, origin and authorship were reconsidered, and agency was redistributed.

Figure 4 presents this relational reframing. It does not constitute the final model of networked authenticity. It records the conceptual transition derived from the findings and identifies the relations that required further synthesis. At this stage, networked authenticity referred to the theoretical direction produced by the

analysis: a movement from differentiated and predominantly human-centered formulations toward an integrated account of human and technological participation in architectural production.

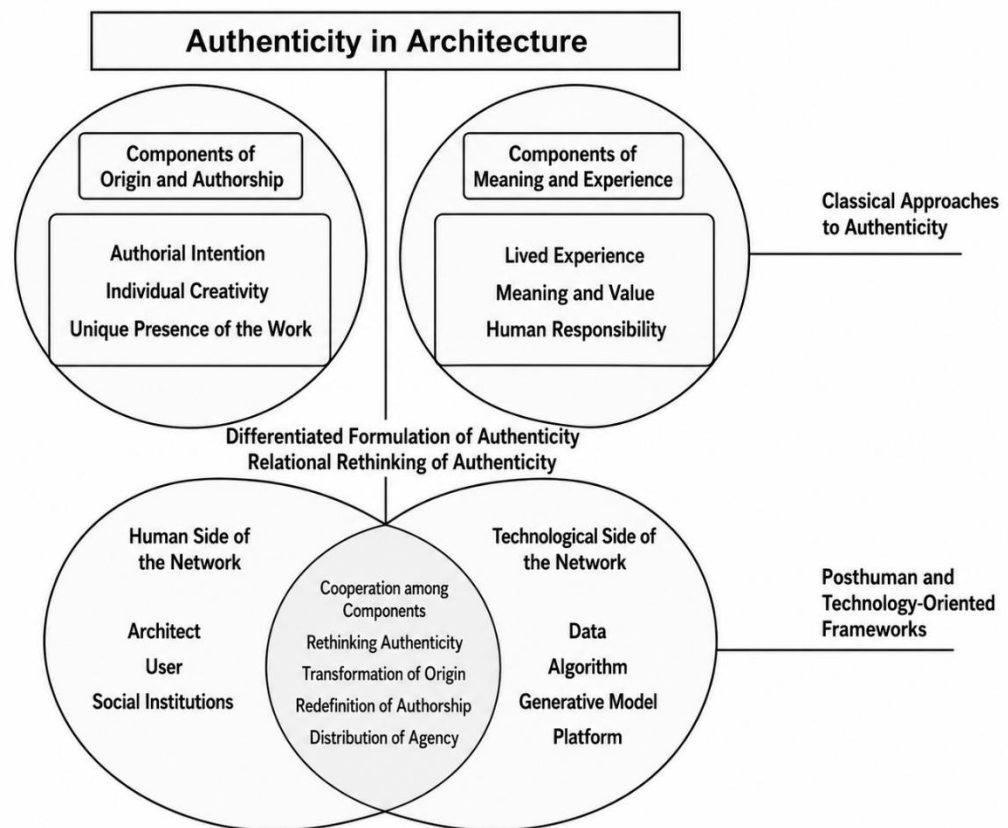


Figure 4: [Relational reframing of authenticity from differentiated classical components to an integrated human and technological configuration.] (source: authors)

Discussion

The five dimensions reported in the Results change the unit through which authenticity can be examined in AI-mediated architecture. Authenticity can no longer be assessed only through the provenance of the final object or its attribution to an individual designer.

The interdependence established analytically in the Results has important implications for how authenticity is interpreted in AI-mediated architecture. A change in provenance can alter the basis of authorship attribution; shifts in authorship can redistribute agency; and these relations can, in turn, reshape judgments of meaning, value, and responsibility. The networked model therefore differs from an additive framework in which separate dimensions are simply accumulated. Its analytical unit is the configuration of relations among the dimensions within a particular production process.

The results do not render the classical traditions irrelevant; rather, they reposition their contributions within a distributed production context. Creative intention, historical and social situatedness, lived and embodied experience, choice, and responsibility remain essential to the interpretation of authenticity. Their limitation under AI mediation lies not in these concepts themselves, but in organizing them primarily around a singular human origin, author, or self-sufficient agent.

This organization becomes difficult to maintain when architectural production includes datasets, algorithms, generative models, prompts, software environments, platforms, institutional procedures, and repeated cycles of selection and revision. In such processes, the architect does not always produce form directly. The architect may define the problem, select references, write and revise prompts, establish constraints, compare alternatives, reject outputs, modify generated proposals, and relate them to program,

site, material conditions, and cultural context. These actions remain creative, but creativity is located across a sequence of relations rather than in a single moment of individual expression.

Under AI mediation, these classical insights require extension rather than replacement. Social and institutional accounts of authorship and value must now include technical infrastructures, platforms, data conditions, and model-related constraints. Phenomenological concerns remain centered on human embodiment, situated perception, and lived experience, while acknowledging that generated representations, interfaces, and simulations can mediate how architectural possibilities are anticipated and interpreted. Likewise, distributed computational operations do not transfer existential or professional responsibility to the system: interpretation, evaluation, accountability, and the assumption of consequences remain attached to human and institutional actors.

Origin, authorship, and agency are therefore connected but should not be treated as interchangeable. In classical accounts, origin often refers to the human source from which a work proceeds, whether that source is understood as intention, genius, embodied experience, or a historically situated act of creation. In AI-mediated architecture, origin is dispersed across several stages. Human intention may initiate or direct the process, but the available outputs are also conditioned by training or reference data, model architecture, platform constraints, prompt formulation, parameters, and later acts of selection and revision. Provenance becomes historical, technical, and procedural. The analytical question is no longer limited to who first conceived the work. It also concerns how the work came into being, which sources and systems participated, and whether the production process can be reconstructed. This dispersed provenance changes the meaning of authorship. The analysis does not support treating an artificial intelligence system as a human author with consciousness, lived intention, or moral responsibility. It also does not support treating the system as a neutral instrument whose operation has no effect on the work. Generative systems organize the field of possible outputs, privilege certain formal relations, exclude or weaken other possibilities, and alter the direction of design exploration. Their contribution is operational and mediating. The architect's contribution is intentional, selective, interpretive, contextual, and responsible.

Networked authorship describes this differentiated configuration. It does not assign identical authorial status to every participant. The architect, model, dataset, prompt, platform, design team, and institution contribute in different ways. A dataset conditions the available representational field. A model structures processes of generation and transformation. A platform regulates access, interaction, and output. A prompt frames a request but remains connected to the designer who formulates, revises, and interprets it. The architect relates these operations to an architectural problem and remains responsible for the decision to accept, alter, reject, or implement an output. Agency requires the same differentiation. Actor-network theory allows the analysis to follow relations among architects, users, software, datasets, algorithms, models, platforms, institutions, and design conventions without assuming in advance that only human participants can affect the process ([Latour, 2005](#); [Yaneva, 2009](#)). Recognition of nonhuman participation does not erase distinctions among forms of action. Computational systems generate, filter, rank, predict, and redirect. Human actors formulate purposes, interpret meanings, assess consequences, and assume responsibility. Distributed agency therefore describes how effects and decisions arise through interaction. It does not imply an equal distribution of intention or ethical accountability.

Rather than functioning as interchangeable theories, the three frameworks contribute complementary analytical functions to the proposed model. As established in the Literature Review, their combined value lies in decentering singular human authorship, tracing heterogeneous human and nonhuman participation, and explaining how technological systems mediate perception and action. In the present analysis, this complementarity makes it possible to recognize distributed participation without equating computational participation with human intention, interpretation, or moral accountability.

The preceding analysis required a conceptual reorganization of authenticity. In classical formulations, origin, intention, individual creativity, lived experience, and responsibility were commonly organized around the author as the central human subject. In AI-mediated production, these components are redistributed across relations among the architect, data, algorithms, generative models, platforms, and the design process. Figure 5 presents this movement from an author-centered structure toward a relational human-technological configuration.

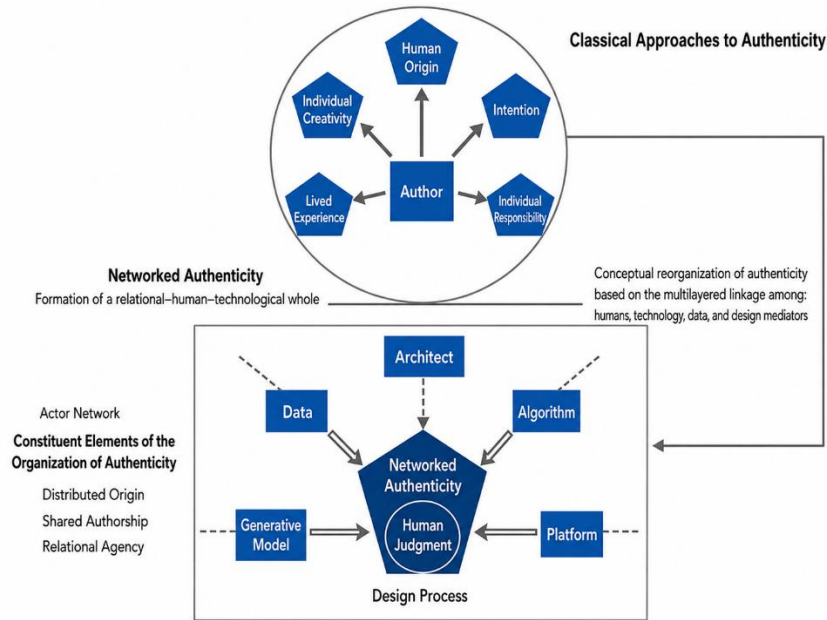


Figure 5: [Conceptual reorganization of authenticity from an author-centered classical structure to a relational human-technological configuration.] (source: authors)

As Figure 5 shows, this reorganization does not remove the architect or place human and technological actors in equivalent positions. The architect remains responsible for framing the architectural problem, interpreting generated possibilities, selecting and revising outputs, and relating them to program, site, context, and professional obligations. Data, algorithms, generative models, and platforms affect the range and organization of possible outcomes, but human judgment remains central to their architectural interpretation and evaluation. Distributed origin, shared authorship, and relational agency therefore describe differentiated forms of participation in the production network.

In this account, networked authenticity is defined as a relational, processual, interpretive, and traceable condition in which the provenance, authorship, agency, meaning, value, and responsibility of an architectural work are assessed through the organization of relations among human actors, technological systems, and contextual conditions. This definition does not make the presence of a network sufficient for authenticity. Most contemporary design processes already contain multiple people, tools, and infrastructures. Networked authenticity concerns whether their relations can be identified, differentiated, interpreted, and evaluated.

Figure 6 presents the final model. At its center is the network through which the architectural work is produced and interpreted. The perceptual and interpretive side includes experience, meaning, context, architectural judgment, and ethical and professional responsibility. The objective and technological side includes data, datasets, algorithms, generative models, prompts, platforms, production procedures, and decision records. Human and nonhuman actors connect these sides. They participate in the same production network, but they do not occupy equivalent positions. External forces also affect the network. Legal frameworks, intellectual property arrangements, platform economies, professional institutions, and social and cultural discourse influence the conditions under which a work is produced and judged. Individual interpretive conditions, including the observer's knowledge, the designer's experience, theoretical assumptions, and the context of perception, affect evaluation. Some conditions can be altered through design practice, documentation, and professional decisions. Others, including model architecture, training procedures, platform rules, and access restrictions, may remain outside the direct control of the architect.

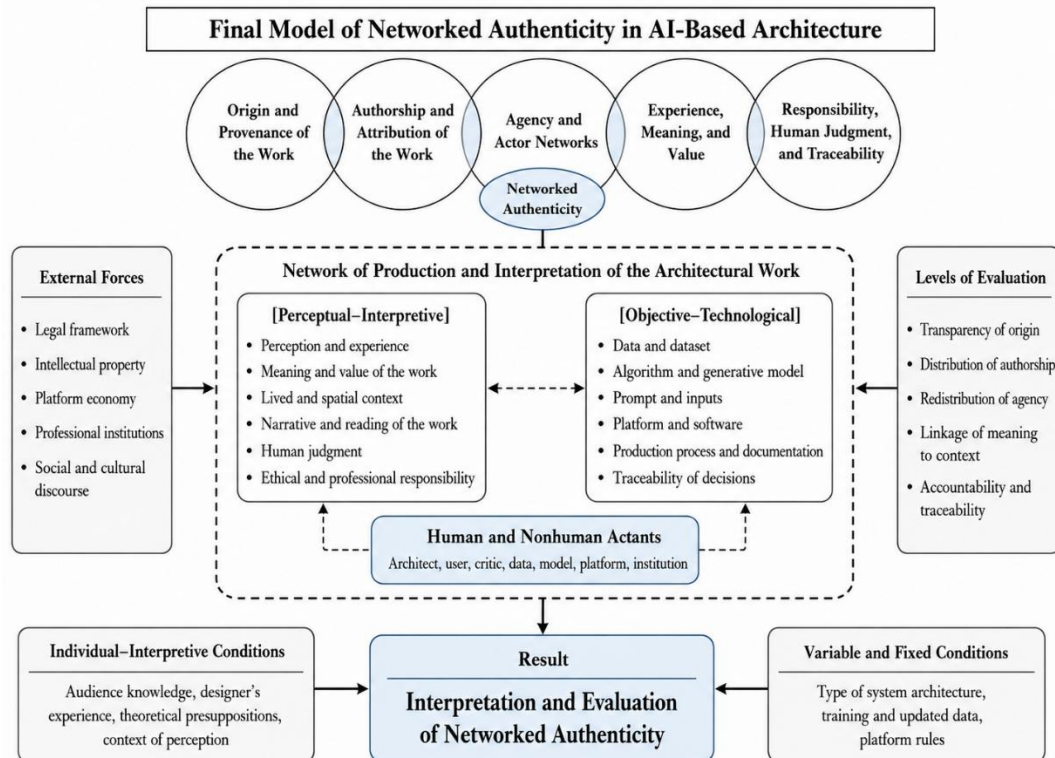


Figure 6: [Final model of networked authenticity in AI-mediated architecture.] (source: authors)

Traceability has a specific role in this model. Computational production can obscure the relation between inputs, decisions, and outputs, especially when a process contains several datasets, prompts, automated transformations, platforms, and revisions. This opacity limits informed attribution and ethical evaluation. A traceable process records enough information to reconstruct the main stages of production. Relevant evidence may include the type and version of the model, the platform used, the origin of reference materials where this can be identified, prompts and major prompt revisions, parameters or constraints, generated alternatives, human interventions, and the reasons for accepting or rejecting particular outputs.

Transparency does not establish authenticity by itself. A fully documented process may still produce an architecturally weak, contextually inappropriate, or ethically problematic result. Documentation makes evaluation possible. Without sufficient traceability, claims about origin, authorship, agency, and responsibility remain difficult to examine. Traceability therefore concerns the conditions of defensible interpretation, not proof of an intrinsic quality contained in the work.

Experience, meaning, and value add another layer. An architectural work cannot be understood only through the operations that produced it. It enters a spatial, cultural, material, and experiential context. Meaning develops through the relation among the work, its setting, its production process, and the interpreting subject. Knowledge of how an AI-mediated work was produced may affect its reception because the roles attributed to the architect, dataset, model, prompt, and platform shape judgments about intention and value. At the same time, knowledge of production does not replace direct architectural experience. Perceptible qualities, spatial use, embodiment, and contextual relations remain part of the evaluation.

Tree City and the Blur Building were selected as illustrative, pre-AI examples because each foregrounds a different form of relational architectural meaning relevant to the present argument. Tree City, the winning proposal for Toronto's Downsview Park by OMA/Rem Koolhaas and Bruce Mau, was conceived less as a fixed finished object than as an open-ended landscape strategy capable of developing over time in response to ecological, environmental, programmatic, and changing public conditions (Czerniak, 1999; Waldheim, 2001). The Blur Building by Diller + Scofidio was selected because its architectural experience depends on the interaction among bodily movement, atmospheric conditions, perception, water, and a technologically regulated mist environment, making spatial meaning contingent on relations among the visitor, site,

environment, and technical system (Hann, 2012). The two projects are therefore not presented as AI-mediated case studies or as empirical evidence for networked authenticity. Rather, they serve as illustrative precedents showing that architectural meaning can emerge relationally through interactions among systems, environment, temporality, embodiment, perception, and technological mediation. AI-mediated design extends this relational condition by introducing data, generative models, prompts, interfaces, and platforms directly into the production and interpretation of architectural work.

The type of artificial intelligence involved must also be specified. Artificial intelligence is not a uniform design category. Text-to-image systems distribute activity through language, prompt revision, interpretation, and output selection. Image-to-image systems give a larger role to reference images, transformation processes, and recombination. Generative design systems distribute decisions through constraints, parameters, performance criteria, and the evaluation of alternatives. Optimization and AI-supported BIM systems may affect prediction, coordination, or performance decisions without visibly generating the form. Co-creative systems create repeated exchanges between designer and system. These differences alter the provenance, authorship, and agency of the resulting work. Any evaluation must therefore begin by identifying the type of system and the part of the architectural process in which it intervenes.

To translate the model into an analytical procedure, the five dimensions were organized as an interpretive assessment matrix. The matrix identifies an analytical question for each dimension, the types of information and documentation relevant to its assessment, and three levels for describing the adequacy of that documentation. The levels assess the clarity and traceability of the production network. They do not measure authenticity as a material or numerical property of the work.

Table 5: [Operational interpretive matrix for assessing networked authenticity in AI-mediated architectural production.] (source: authors)

Dimension	Analytical question	Information and records to be considered	0: Insufficient	1: Partial	2: Substantiated
Origin and provenance	Can the main production lineage of the work be reconstructed?	Type and version of the AI system, platform, reference or training sources where identifiable, prompts, parameters, stages of generation, iterations, and revisions	The main sources, systems, or stages cannot be identified	Some sources and stages are documented, but important gaps remain	The main sources, systems, stages, and human interventions are documented and can be reconstructed
Authorship and attribution	Can the contributions of the participating actors be differentiated?	Roles of the architect, design team, client, users, model, data, software, platform, and relevant institutions	The work is attributed to one actor without an account of other contributions	The main participants are named, but their contributions remain partly unclear	The participants and the type of contribution made by each are explicitly differentiated
Agency and actor networks	Can the distribution of actions and design decisions be identified?	Problem framing, prompt formulation, constraints, parameters, generation, selection, rejection, revision, contextualization, and final approval	The outcome is presented without evidence of how decisions were made or who intervened	Human and technological interventions are partly identifiable, but the distribution of decisions remains incomplete	The main decisions, interventions, and relations among human and technological actors are clearly documented
Experience, meaning, and value	Is the work interpreted in relation to its architectural and contextual conditions?	Site, program, spatial experience, material conditions, cultural and historical setting, design narrative, user interpretation, and relation between form and production process	The work is evaluated mainly as an isolated image or formal output	Some spatial, contextual, or experiential relations are addressed	The work is explicitly related to its spatial, programmatic, cultural, material, and experiential conditions
Responsibility and human judgment	Can responsibility for the process and its consequences be assigned?	Human and institutional accountability, professional judgment, output evaluation, data ethics, intellectual property, process transparency, and anticipated architectural consequences	No accountable human or institutional actor can be identified	Responsibility is partly assigned, but important decisions or consequences remain ambiguous	Responsible actors, decision boundaries, and the basis of accountability are clearly identified

The matrix can be applied in four stages. First, the analyst defines the unit of analysis by identifying the architectural output, the boundaries of the design process, and the type of AI system involved. The unit may be a completed work, a competition proposal, a design phase, a sequence of generated alternatives, or another documented architectural process. The boundaries must be stated because the relevant actors and

records change according to the part of the process being examined. Second, when the matrix is applied to a documented architectural case, the analyst assembles the information and records relevant to the five dimensions. The evidence may include model and platform information, prompts, parameters, datasets or reference materials where identifiable, generated iterations, records of selection and revision, design statements, contextual documents, and records of responsibility. Absence of a particular record should be reported directly. It should not be replaced by an assumption about how the system or designer probably operated. Third, each dimension is assessed independently as insufficient, partial, or substantiated. Every rating must be linked to an identifiable record or observation. A rating of 2 does not indicate that a work is more authentic in an absolute sense. It indicates that the dimension can be examined through a sufficiently clear body of evidence. A rating of 0 identifies a gap that prevents a defensible assessment of that dimension. Fourth, the ratings are reported as a five-dimensional profile. They are not combined into a universal authenticity score because the dimensions address different relations and remain dependent on architectural context. A profile may contain strong documentation of provenance and agency but weak contextual interpretation, or a developed account of experience and meaning with limited information about data and platform conditions. Reporting the profile preserves these differences. Traceable provenance and identifiable responsibility are treated as minimum conditions for an overall evaluation. If either origin and provenance or responsibility and human judgment receives a rating of 0, the case should be reported as insufficiently assessable at the overall level. The remaining dimensions can still be examined, but the analyst should not issue a general judgment of networked authenticity. This rule does not classify the work as inauthentic. It records that the evidence needed to connect production, attribution, and accountability is absent.

Illustrative application of the interpretive matrix

To demonstrate how the interpretive matrix operates in practice, it was applied to a constructed AI-mediated architectural design scenario. The example is illustrative rather than empirical and intended to demonstrate how the three assessment levels generate differentiated interpretations when applied to a defined body of evidence. It does not constitute empirical validation of the framework.

The constructed scenario concerns the conceptual design of a small waterfront cultural pavilion developed through a text-to-image generative AI system. The architect defines the program, site conditions, design brief, and initial conceptual direction. The AI system, platform, and model version are documented, together with twelve successive prompt iterations and thirty-six generated outputs. The platform does not provide access to the model's training dataset. The architect reviews the generated alternatives, selects four outputs for further development, combines and modifies their spatial and material characteristics, and develops one proposal through subsequent architectural modeling. Records of prompt revision, output selection, rejection, and design modification are retained. The architect and design team are identified as responsible for final evaluation and approval. The design narrative addresses the site, program, spatial intention, and material concept, but the proposal remains conceptual and has not been physically constructed or evaluated through user experience.

Table 6: [Illustrative application of the interpretive matrix to a constructed AI-mediated architectural design scenario.] (Source: authors)

Rationale	Assessment	Evidence in the constructed case	Dimension
The principal design sequence is traceable, but part of the computational provenance cannot be reconstructed	1 – Partial	AI system, platform, model version, prompts, generated iterations, selections, and human revisions are documented; training data remain inaccessible	Origin and provenance
The principal participants and the nature of their contributions can be explicitly distinguished	2 – Substantiated	Architect, design team, and generative system are identified and their contributions are differentiated	Authorship and attribution
The distribution of actions and decisions between human and technological actors can be reconstructed	2 – Substantiated	Problem framing, prompting, generation, selection, rejection, modification, and final approval are documented	Agency and actor networks
Contextual and interpretive relations are	1 – Partial	Site, program, design narrative, spatial intention,	Experience, meaning, and value

Rationale	Assessment	Evidence in the constructed case	Dimension
present, but experiential evidence remains limited		and material concept are addressed, but no built experience or user evaluation exists	
Accountable human actors and decision boundaries are clearly identifiable	2 – Substantiated	The architect and design team retain final evaluation, selection, modification, and approval	Responsibility and human judgment

The resulting profile is therefore 1–2–2–1–2 rather than a cumulative score. The example demonstrates that the matrix distinguishes between dimensions that are strongly documented and those for which important evidential gaps remain. In this case, authorship, agency, and responsibility are sufficiently substantiated, whereas provenance remains partial because the underlying training data cannot be reconstructed, and experience, meaning, and value remain partial because the proposal has not entered a built or user-experienced condition. Because neither the origin-and-provenance dimension nor the responsibility and human judgment dimension receives a rating of 0, the case remains assessable at the overall level, although the two partial dimensions must remain visible in the interpretive profile. The worked example therefore illustrates the operational distinction among the three assessment levels without converting them into a single measure of authenticity.

The matrix provides a structured basis for approaching AI-mediated work in terms of attribution, design judgment, and architectural criticism. Attribution can acknowledge the effects of datasets, models, software, and platforms without treating them as human moral subjects. Design judgment can examine how outputs were framed, selected, revised, and contextualized, instead of focusing only on the novelty of the final image or form. Architectural criticism can address the visible work together with the less visible network that produced it. Responsibility remains attached to human and institutional actors capable of explanation and response, while the analysis also records how technological systems shaped their field of action.

Conclusion

This study addressed the main research question of how authenticity can be reconceptualized in AI-mediated architecture beyond human-centered and author-centered accounts. The findings indicate that authenticity can be reconceptualized as a relational, processual, interpretive, and traceable condition in which origin and provenance, authorship and attribution, agency and actor networks, experience, meaning and value, and responsibility and human judgment are examined as interdependent dimensions of the same production network. The subsidiary questions further clarified this reformulation. First, classical approaches remain conceptually important because they preserve the significance of origin, intention, creativity, lived experience, meaning, value, choice, and responsibility, but they become insufficient when these components are organized primarily around a singular human author or self-sufficient human agency. Second, AI mediation introduces challenges concerning distributed provenance, attribution, mediated and distributed agency, contextual interpretation, process transparency, and the traceability of responsibility. Third, posthumanism, actor-network theory, and technological mediation provide complementary resources for addressing these challenges by repositioning the human subject within relational networks, tracing heterogeneous human and nonhuman participation, and explaining how technologies mediate perception and action. On this basis, the study formulates networked authenticity as an account in which the provenance, authorship, agency, meaning, value, and responsibility of an architectural work are interpreted through the organized relations among human actors, technological systems, and contextual conditions. This formulation does not remove the architect from the design process or attribute human intention, lived experience, or moral responsibility to artificial intelligence. Instead, it differentiates the operational and mediating effects of computational systems from the interpretive judgment, contextual evaluation, selection, and accountability that remain human and institutional responsibilities.

The study makes three connected contributions. First, it reorganizes the classical components of authenticity as part of a relational structure suited to AI-mediated architectural production. Second, it formulates the conceptual model of networked authenticity by connecting the objective and technological

conditions of production with perceptual and interpretive conditions. Third, it operationalizes this model through a five-dimensional assessment matrix covering origin and provenance, authorship and attribution, agency and actor networks, experience, meaning and value, and responsibility, human judgment, and traceability. The matrix provides designers, critics, and researchers with a structured procedure for identifying and examining the information and documentation relevant to each dimension. It produces an interpretive profile rather than a single authenticity score, which preserves differences among the dimensions and their dependence on architectural context.

The study remains limited by its theoretical and documentary design, its reliance on selected philosophical and architectural traditions, and the absence of empirical validation across documented architectural projects. Although the constructed worked example demonstrates how the matrix can generate a differentiated interpretive profile across the five dimensions, it does not constitute empirical validation or establish inter-rater reliability. Future research should therefore test and refine the matrix through documented AI-mediated design processes, comparison among multiple analysts, and application to generative design, optimization, text-to-image, BIM-supported, and co-creative systems. Further testing should also examine data transparency, attribution, intellectual property, platform conditions, and professional accountability. Given these limits, networked authenticity provides both a theoretical account of AI-mediated architectural production and an operational framework for evaluating how technological mediation, human judgment, and responsibility are organized within the production network.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used ChatGPT to assist with translation, paraphrasing, editing, and improving the language, grammar, clarity, and readability of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

- Amid, H. (1971). *Farhang-e Amid [Amid dictionary]*. Javidan Publications.
- Beerends, S., & Aydin, C. (2021). Negotiating authenticity in technological environments. *Philosophy & Technology*, 34(4), 1665–1685. <https://doi.org/10.1007/s13347-021-00480-5>
- Beerends, S., & Aydin, C. (2025). Negotiating the authenticity of AI: how the discourse on AI rejects human indeterminacy. *AI & Society*, 40(2), 263–276. <https://doi.org/10.1007/s00146-024-01884-5>
- Benjamin, W. (1969). The work of art in the age of mechanical reproduction (H. Zohn, Trans.). In H. Arendt (Ed.), *Illuminations* (pp. 217–251). Schocken Books.
- Bourdieu, P. (1993). *The field of cultural production: Essays on art and literature*. Polity.
- Braidotti, R. (2013). *The Posthuman*. Polity.
- Carpó, M. (2017). *The second digital turn: design beyond intelligence*. MIT Press. <https://doi.org/10.7551/mitpress/9976.001.0001>
- Colmenero-Acevedo, E. (2025). Rethinking authenticity in architecture: The ethical implications of AI-generated design. In L. Martínez-Villaseñor, B. Martínez-Seis, & O. Pichardo (Eds.), *Artificial Intelligence – COMIA 2025* (pp. 313–322). Springer. https://doi.org/10.1007/978-3-031-97907-1_25
- Czerniak, J. (Ed.). (1999). *Case: Downsview Park, Toronto*. Prestel.

- Del Campo, M. (2024). Everything Can Be an Author: Rethinking Agency in the Age of Artificial Intelligence. *Architectural Design*, 94(3), 20–29. <https://doi.org/10.1002/ad.3051>
- Golomb, J. (2012). *In search of authenticity: Existentialism from Kierkegaard to Camus*. Routledge.
- González, A. F., & Garcia, M. (2024). A posthuman architectural artificial intelligence speculum? Text and images in future spaces. *Architectural Design*, 94(1), 22–29. <https://doi.org/10.1002/ad.3010>
- Groat, L., & Wang, D. (2002). *Architectural research methods*. John Wiley & Sons.
- Hann, R. (2012). Blurred Architecture: Duration and performance in the work of Diller Scofidio + Renfro. *Performance Research*, 17(5), 9-18. <https://doi.org/10.1080/13528165.2012.728434>
- Haraway, D. J. (1991). *Simians, cyborgs, and women: The reinvention of nature*. Routledge.
- Hayles, N. K. (1999). *How we became posthuman: Virtual bodies in cybernetics, literature, and informatics*. University of Chicago Press.
- Heidegger, M. (1962). *Being and time* (J. Macquarrie & E. Robinson, Trans.). Harper & Row.
- Iapaolo, F., & Lynch, C. R. (2026). Rethinking urban AI: a critical posthumanist perspective on intelligence, autonomy, and agency. *Urban Geography*, 47(1), 65–87. <https://doi.org/10.1080/02723638.2025.2510812>
- Ibn Manzur, M. I. M. (1993). *Lisan al- 'Arab* (Volume. 11). Dar Sader.
- Ihde, D. (1990). *Technology and the lifeworld: From garden to earth*. Indiana University Press.
- Kolarevic, B. (Ed.). (2003). *Architecture in the digital age: design and manufacturing*. Taylor & Francis. <https://doi.org/10.4324/9780203634561>
- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). SAGE Publications.
- Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory*. Oxford University Press. <https://doi.org/10.1093/oso/9780199256044.001.0001>
- Lee, H. J. (2025). Seeing like AI? AI urbanism reconsidered through a critical posthumanist perspective. *Urban Studies*. Advance online publication. <https://doi.org/10.1177/00420980251344914>
- Norberg-Schulz, C. (1980). *Genius Loci: Towards a Phenomenology of Architecture*. Rizzoli.
- Oxford University Press. (2005). *Oxford Dictionary of English* (2nd ed.).
- Oxman, R. (2006). Theory and design in the first digital age. *Design Studies*, 27(3), 229–265. <https://doi.org/10.1016/j.destud.2005.11.002>
- Pallasmaa, J. (2005). *The eyes of the skin: architecture and the senses* (2nd ed.). Wiley.
- Taylor, C. (1992). *The Ethics of Authenticity*. Harvard University Press. <https://doi.org/10.4159/9780674237117>
- Terzidis, K. (2006). *Algorithmic architecture*. Routledge. <https://doi.org/10.4324/9780080461298>
- Trilling, L. (1972). *Sincerity and Authenticity*. Harvard University Press. <https://doi.org/10.4159/9780674044463>

- Verbeek, P.-P. (2005). *What things do: Philosophical reflections on technology, agency, and design*. Penn State University Press. <https://doi.org/10.1515/9780271033228>
- Verbeek, P.-P. (2011). *Moralizing Technology: Understanding and Designing the Morality of Things*. University of Chicago Press. <https://doi.org/10.7208/chicago/9780226852904.001.0001>
- Waldheim, C. (2001). Park= City? The Downsview Park design competition. *Landscape Architecture*, 91(3), 80-85, 98-100.
- Wolfe, C. (2010). *What is posthumanism?* University of Minnesota Press.
- Yaneva, A. (2009). Making the social hold: Towards an actor-network theory of design. *Design and Culture*, 1(3), 273–288. <https://doi.org/10.1080/17547075.2009.11643291>
- Yaneva, A. (2022). *Latour for architects*. Routledge. <https://doi.org/10.4324/9780429328510>
- Zhang, Z., Fort, J. M., & Mateu, L. G. (2023). Exploring the potential of artificial intelligence as a tool for architectural design: A perception study using Gaudí's works. *Buildings*, 13(7), 1863. <https://doi.org/10.3390/buildings13071863>