www.itcon.org - Journal of Information Technology in Construction - ISSN 1874-4753

FROM AUTOMATION TO AUTHORSHIP: COMPUTATIONAL REASONING AND DESIGN AGENCY IN ARCHITECTURAL AI TOOLS

SUBMITTED: September 2025
PUBLISHED: November 2026
EDITOR: Frédéric Bosché
DOI: 10.36680/j.itcon.2026.055

Seeja Sudhakaran
University of West London
https://orcid.org/0009-0002-8575-3505
seeja.sudhakaran@uwl.ac.uk

Changfeng Fu, Professor
University of West London
https://orcid.org/0000-0002-2019-5445
Charlie.Fu@uwl.ac.uk

Efcharis Balodimou, Dr
University of West London
https://orcid.org/0000-0003-1249-3941
Efcharis.Balodimou@uwl.ac.uk

Mukhtar Maigari, Dr
University of West London
mukhtarmaigari@gmail.com

Prapooja KC
University of West London
prapooja.kc@research.uwl.ac.uk

Paul Coates, Dr
University of Salford
https://orcid.org/0000-0002-5437-1777
S.P.Coates@salford.ac.uk

SUMMARY: *The increasing presence of AI in architectural practice is renegotiating who, or what, holds creative authority over a project. This paper introduces a novel two-construct framework, AI Logic (AL) and Architectural Value (AV), to operationalise this shift. AL captures a tool's computational reasoning across automation, optimisation, and predictive capability, while AV captures design agency support, contextual sensitivity, and exploratory capacity. Positioning fifty AI-driven tools within this AL-AV construct via tension map analysis identifies stage-specific patterns within the selected dataset. It shows that design agency is redistributed unevenly across the project lifecycle, expanding in early design stages where tools act as creative collaborators, and contracting under efficiency-driven logics in later stages where automation prioritises monitoring and compliance over qualitative judgement. Cross-referencing against secondary parameters such as building typology, user group, tool maturity, and regulatory context surfaces fundamental gaps in current AI provision. The selected tools predominantly support residential and commercial applications. Documented regional orientations include North American and European standards, while coverage of other regions and typologies is less evident within the sample. To assess scoring consistency, ten AEC professionals and the author independently evaluated a subset of five tools, with one tool representing each of RIBA Stages 1–5. Intraclass correlation analysis demonstrated strong panel-level inter-rater reliability for both constructs, while comparison of the author's scores with the expert consensus indicated close alignment. Following this calibration, the author independently applied the same evaluation framework to the remaining forty-five tools. Beyond its analytical application, the framework offers a practical means of informing future AI development and adoption. It gives toolmakers a vocabulary for embedding stakeholder values in development decisions, and architects a means of recognising embedded biases before adoption, so that increasing automation extends rather than erodes the disciplinary and cultural values of architecture as a whole.*

KEYWORDS: *Artificial Intelligence, architectural design, design agency, computational design, project lifecycle.*

REFERENCE: *Sudhakaran, S., Fu, C., Balodimou, E., Maigari, M., KC, P., & Coates, P. (2026). From automation to authorship: Computational reasoning and design agency in architectural AI tools. Journal of Information Technology in Construction (ITcon), 31, 1336-1356. https://doi.org/10.36680/j.itcon.2026.055*

COPYRIGHT: *© 2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.*

1. INTRODUCTION

The Architecture, Engineering, and Construction (AEC) industry is undergoing a rapid transformation driven by advances in Artificial Intelligence (AI). Traditionally reliant on manual coordination and rule-based automation, the industry is now incorporating AI-powered tools that enhance design optimisation, streamline construction workflows, and improve sustainability outcomes (Caetano et al., 2020; Yang, 2025). These tools which range from generative design platforms and machine learning based simulations to robotic systems and AI-enhanced building information modelling (BIM), are reshaping how buildings are conceived, planned, and executed (J. Li et al., 2024).

The emergence of AI in architecture does not arrive in isolation, but rather as part of a broader historical trajectory in computational design. The evolution of design methodology over the last several decades can be described in phases; from modularity, to computational design, to parametricism, and now to artificial intelligence (Chaillou, 2019). Each of these paradigms has responded to the limitations of the previous one. Modularity offered efficient construction but limited variation, parametricism enabled fluid formal control but often yielded stylistic repetition. In contrast, AI is increasingly positioned as a paradigm that transcends prior constraints, offering not only technical power but also new ways of generating, evaluating, and adapting architectural form and function.

Unlike earlier tools, AI systems are not strictly deterministic. Many function as probabilistic models that learn from data and generate design options without explicit user direction. This shift from assistive to generative computation introduces both novel opportunities and significant epistemological tensions. AI systems embed their own internal logics, favouring efficiency, performance, and prediction (Albukhari, 2025). This may or may not align with traditional architectural values such as contextual sensitivity, ambiguity, material richness, and authorship.

The increasing integration of AI tools within architectural workflows raises questions about their influence on design thinking, decision-making, and authorship. Central to this inquiry is the concept of design agency, understood as the architect's capacity to define problems, assert creative intent, and shape spatial and formal outcomes (Awan et al., 2011). Drawing from architectural theory and socio-technical design studies, design agency is approached not as an abstract ideal, but as a dynamic and distributed condition shaped by tools, systems, and institutional structures (Newton, 2021).

Rather than offering a catalogue of technologies or a speculative vision of AI-led futures, this paper provides a critical framework for understanding how AI reconfigures design agency through the logics it prioritises. These include computational modes of optimisation, automation, and prediction, which may reinforce or contradict architectural traditions depending on their application. The paper introduces a conceptual lens that maps the tensions between AI logics and architectural values, and applies this framework to key AI applications used across the architectural lifecycle.

Building on this conceptual foundation, the study seeks to examine the relationship between computational reasoning and architectural design agency through a structured evaluation of current AI tools across the project lifecycle. Three research questions guide this investigation: (RQ1) How can AI Logic and Architectural Values be systematically evaluated to examine the relationship between computational reasoning and design agency in AI-driven architectural tools? (RQ2) How do these tools prioritise AI Logic and Architectural Values across different stages of the architectural project lifecycle? (RQ3) What patterns and gaps emerge from examining AI tools across multiple architectural and technological parameters, including building typologies and geographic reach?

2. CONCEPTUAL FRAMEWORK: AI LOGICS, ARCHITECTURAL VALUES, AND DESIGN AGENCY

A conceptual framework provides the theoretical foundation for critically examining how AI reshapes architectural practice (Miles et al., 2014). It explicitly defines core concepts like AI logics, architectural values, and design agency while establishing the analytical lens through which their interactions are evaluated. By positioning AI not merely as a tool, but as a mode of reasoning embedded with particular assumptions and values, the framework supports a structured analysis of AI's impact on architectural practice beyond technical functionality alone.

CC BY

2.1 AI logics

AI Logics (AL) refer to computational principles and assumptions underlying contemporary AI systems used in architecture (Chaillou, 2022). Unlike earlier deterministic design tools, contemporary AI systems predominantly operate through probabilistic approaches, utilising machine learning models (Tamke et al., 2018) and generative algorithms trained on extensive datasets (As et al., 2018; Nagy et al., 2018). These approaches underpin three principal computational logics: optimisation, prediction, and automation.

Optimisation logic emphasises measurable performance criteria, such as structural efficiency, energy usage, cost reduction, or spatial utilisation (Carta, 2019). Prediction logic, facilitated by machine learning, leverages historical and real-time data to forecast building performance, occupancy patterns, environmental responsiveness, and future design scenarios (J. Li et al., 2024). Automation logic reduces manual intervention by delegating routine or complex tasks to algorithmically driven processes (Caetano et al., 2020). These tasks include varied processes like regulatory compliance verification, construction oversight, or iterative form generation.

AI tools often emphasise on performance optimisation, cost efficiency, and predictability. Thus, redefining standards of architectural success in ways that may conflict with the traditional architectural values.

2.2 Architectural values

Architectural Values (AV) represent qualitative dimensions historically significant to the discipline. These values include contextual sensitivity, ambiguity, experiential richness, and aesthetic authorship, all emphasising architecture's cultural, sensory, and contextual qualities (Leatherbarrow, 2009; Pérez-Gómez, 2016).

Contextual sensitivity values responsiveness to specific cultural, historical, social, and environmental settings, advocating site-specific rather than universal or standardised solutions (Pallasmaa, 2012). Ambiguity highlights openness to interpretation and adaptability, enabling spaces to support diverse uses and interpretations over time (Baumberger, 2009). Architectural authorship values creative intention, originality, and subjective judgments, foregrounding the human architect as central to the design process. These qualitative values frequently exist in productive tension with the objective and quantifiable nature of computational logics, generating critical questions about how AI systems reshape the discipline's value structures (del Campo, 2024; Schneider et al., 2025).

2.3 Intersection of AI logics and architectural values: Design agency

At the intersection of AI logics and architectural values lies design agency, the concept previously defined as the capacity of architects to frame design problems, assert creative intent, and meaningfully shape spatial and formal outcomes (Awan et al., 2011; del Campo, 2024). The interaction between computational logic and architectural values significantly reconfigures this agency, redistributing decision-making authority between human designers and computational tools (Schneider et al., 2025).

Evaluating this redistribution requires examining how architectural decision-making authority is shifted, reframed, or constrained by the assumptions embedded in AI systems. Design agency thus serves as a critical analytical term linking computational and disciplinary values, highlighting not only who or what decides, but how the conditions of architectural decisions are being transformed (Pasquinelli & Joler, 2021).

Previous reviews have examined the evolution of AI technologies in architecture, their applications across design stages, and approaches to evaluating computational and architectural outcomes (Albukhari, 2025; Cocho-Bermejo, 2025; Jang et al., 2025; Khan et al., 2025; C. Li et al., 2025; Zhuang et al., 2025). As summarised in Table 1, their methods range from bibliometric and systematic reviews to criteria-based model evaluations. Building on their consideration of designer input, subjective judgement and human-centred evaluation, the present study assesses AI Logic and Architectural Value across fifty contemporary AI tools mapped to RIBA Stages 1–5. Expert assessment of five tools and author calibration support subsequent author-led evaluation of the remaining forty-five, providing a structured basis for examining tools' implications for architectural design agency.

3. METHODOLOGY

This study employs a mixed-method approach combining a scoping review with structured mapping to examine the relationship between artificial intelligence and design agency in architectural practice. The objective is to evaluate how AI tools influence, reconfigure, or displace human authorship and decision-making across the

CC BY

architectural design lifecycle, with design agency serving as the principal analytical lens. The approach is intentionally evaluative rather than enumerative: rather than cataloguing tools as technical artefacts, the methodology is designed to surface the embedded assumptions and value priorities that distinguish one application from another within a shared disciplinary context.

Table 1: Comparison of related reviews and the present study.

Study	Scope and Method	Design Agency Perspective	Lifecycle Coverage	Validation Approach
Albukhari (2025)	Review of AI technologies & architectural applications using PRISMA-labelled selection, thematic synthesis and illustrative comparisons.	Discusses creativity, human authorship, ethical concerns and architectural identity.	Broad architectural and urban applications; no systematic lifecycle-stage mapping.	Reports literature-quality screening; no independent expert-rating or reliability assessment reported.
Cocho-Bermejo (2025)	Bibliometric review of genetic algorithms, neural networks and agent-based systems across Scopus, Web of Science and CumInCAD.	Examines computational evolution rather than directly assessing design agency.	Historical development, stated as 2000–2022; not organised by project stages.	Cross-database comparison and literature triangulation.
Jang et al. (2025)	PRISMA-guided review of 161 papers, classifying AI models, design tasks, data and evaluation methods.	Addresses designer intent, subjective judgement and human-in-the-loop control.	Pre-design through construction documentation; 68.94% of studies concern schematic design.	Authors checked screening outputs; underlying studies' evaluation methods are analysed rather than independently validating the review's framework.
Li et al. (2025)	Technical and application-based review of generative AI models, principally covering architectural research from 2020 onward.	Discusses designer-guided generation, customisation and interactive editing.	Six design tasks, from concept images to structural systems and sections.	Literature synthesis and illustrative examples; no independent expert-rating or reliability study reported.
Zhuang et al. (2025)	Review of ML algorithms, data representations and applications in generative architectural design.	Addresses controllability, authorship, interpretability and human oversight.	Primarily early-stage design generation and evaluation across multiple spatial scales.	Discusses assessments in prior studies; no separate expert validation of the review's synthesis reported.
Khan et al. (2025)	Review of six generative-model families using literature-derived scoring, reported AHP weighting and qualitative comparison.	Explicitly addresses designer input, co-creation and architects' evolving roles.	Design automation and CAD/BIM integration; no systematic lifecycle-stage mapping.	Reports an AHP consistency check; no independent expert panel or inter-rater reliability assessment reported.
Present study	Structured AL–AV assessment and tension mapping of fifty contemporary AI tools using six scoring criteria.	Evaluates computational reasoning alongside design-agency support, contextual sensitivity and exploratory capacity.	RIBA Stages 1–5; tools classified by primary application stage.	Ten AEC experts independently rated five tools, demonstrating strong panel-level reliability. Author ratings showed strong correlation with expert consensus, followed by author-led evaluation of the remaining forty-five tools.

3.1 Research design

The research is grounded in a qualitative, interpretivist paradigm that recognises the contextual and socio-technical nature of AI adoption in architecture (Creswell & Poth, 2018). Whilst the tool mapping process follows a systematic and replicable structure, the analysis foregrounds interpretive frameworks relating to design agency,

CC BY

disciplinary values, and technological mediation rather than seeking purely quantitative outcomes. This position acknowledges that the relationship between computational tools and architectural authorship cannot be fully captured through technical specification alone, and that evaluative judgements about design agency require a framework that is sensitive to professional and cultural context.

The study proceeds in two phases. The first involves the selection and categorisation of fifty AI-based tools currently applied in architectural workflows, organised according to the RIBA Plan of Work stages 1 to 5 (*RIBA Plan of Work*, n.d.). In instances where tools spanned multiple lifecycle stages, categorisation was based on primary use case and dominant functionality. The second phase maps design agency against AI capabilities at each lifecycle stage, examining how architects' traditional roles align with or diverge from the functions performed by AI tools, and identifying patterns of augmentation, displacement, or redistribution of authorship. To support this primary analysis, each tool is additionally cross-referenced against five secondary parameters: building typology, user type and intended audience, geographic or regulatory context, tool maturity and development stage, and design logic or functionality. These dimensions enable the identification of structural trends, systemic gaps, and underexplored areas across the broader tool landscape. The full dataset of fifty tools, organised by lifecycle stage, is listed in Figure 1.

Stage 1 Preparation and Briefing	Stage 2 Concept Design	Stage 3 Spatial Coordination	Stage 4 Technical Design	Stage 5 Manufacturing & Construction
AECOM PlanEngage	Architechtures	Augmenta	Alice Technologies	AI-SpaceFactory
ARKdesign.ai	ArkAI	BricsCAD BIM AI	CityEngine	AirWorks
Archistar	TestFit	NeuralConcept	Ecopilot	Avvir
Delve	Chaos Veras	Preoptima	Generative AIBIM	Buildots
Orkestra	Cove.tool	QBiq AI	Tally	Deep Space
UrbanForm	Finch3D	Revit (AI plugins)		Disperse
Urbanist AI	GroundUp	Rhino (AI plugins)		Doxel
iPLAN	Hypar	Swapp		Imerso
	LookX	UpCodes AI		OpenSpace.ai
	Maket.ai			SiteAware
	Midjourney			Verisk AI
	MNML			
	Rerender			
	Spacemaker			
	Spacio			
	Stable Diffusion			
	Adobe Firefly			

Figure 1: Distribution of the fifty AI tools examined in this study across the five project lifecycle stages, from early design to construction.

3.2 Selection criteria for AI tools

An initial scoping phase surveyed AI tools currently active in architectural practice, drawing on academic databases (Scopus, Google Scholar), industry publications, architectural technology blogs, startup directories, and AI-focused software aggregators. Tools were considered eligible if they employed core AI technologies such as machine learning, computer vision, or generative algorithms; were in active professional, academic, or open-

CC BY

source use; and addressed one or more stages of the architectural design lifecycle. Deployable experimental tools were eligible where sufficient documentation supported assessment. Non-functional concepts, speculative proposals and projects no longer in active development were excluded.

Specific inclusion criteria further required that each tool have sufficient publicly available documentation to permit structured analysis of its features and capabilities, and that it exist in a deployable, non-conceptual state. The combination of these criteria was intended to produce a dataset that is both diverse and representative, supporting meaningful comparison across technical, architectural, and contextual dimensions. From an initial pool of over 150 tools, a final sample of fifty was selected through iterative refinement against these criteria. Evaluation Framework and Scoring Rubric

The structure of the evaluation framework is such that each tool is assessed against two parallel and analytically independent constructs (Figure 2). AI Logic and Architectural Value are each operationalised through three sub-criteria, rated on a 0–10 scale with interpretive anchors at the low and high ends to support consistent scoring across raters. The two constructs are designed to vary independently of one another: a tool may score high on both, low on both, or inversely across the pair, and it is precisely this variation that the framework is designed to capture. Composite scores for each construct are derived by averaging the three sub-criterion ratings, and all assessments were grounded in a standardised Tool Documentation Package compiled for each application to minimise experiential bias.

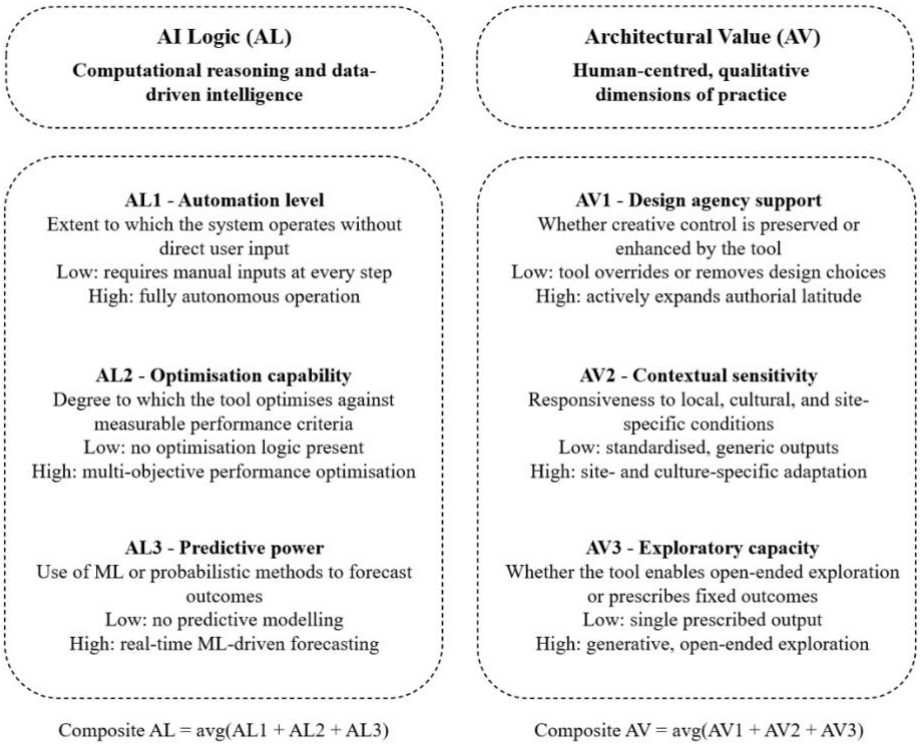


Figure 2: The evaluation framework showing AI Logic (AL) and Architectural Value (AV) constructs, each comprising three equally weighted sub-criteria with composite scoring.

3.3 Expert validation and inter-rater reliability

To validate the framework prior to its broader application, ten AEC professionals were recruited as independent raters. All participants possessed relevant domain knowledge pertaining to digital design tools and emerging AI applications within architectural workflows. Five tools were selected, one representing each of RIBA Stages 1–5, to examine the framework's application across different project stages. Each tool required approximately fifteen minutes to assess, resulting in an estimated seventy-five-minute assessment per practitioner. This selection balanced stage coverage with a manageable assessment workload. Following author calibration against the expert

CC BY

ratings, the remaining forty-five tools were assessed solely by the author. Figure 3 shows one individual rater's scores across all 5 tools while Figure 4 shows all ten raters' scores for a single tool.

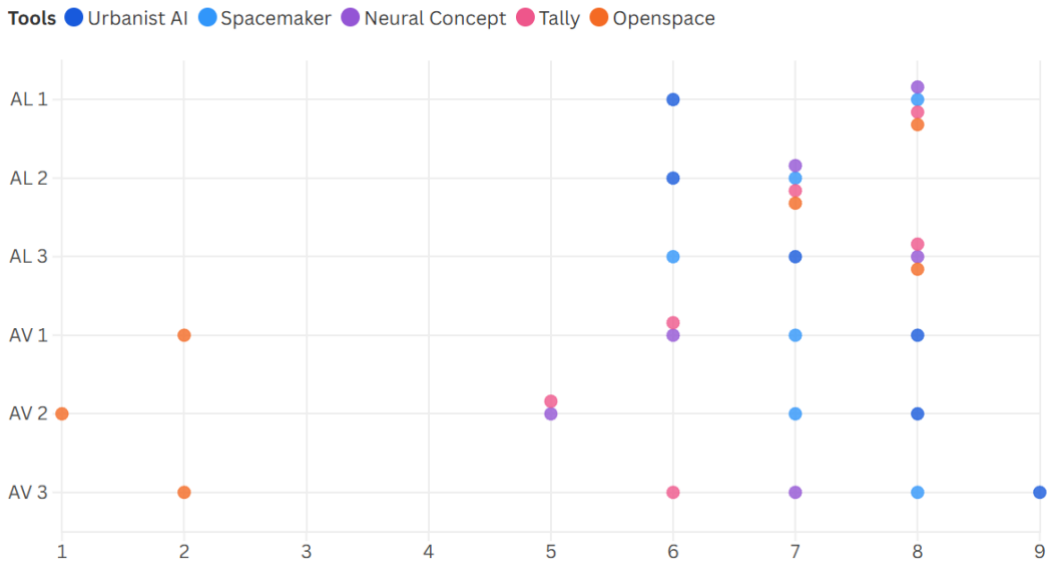


Figure 3: Facet-level scores (AL1–AL3, AV1–AV3) assigned by a single expert rater across the five validation tools. While ten experts completed the evaluation, this figure shows one rater's responses as a representative example of the scoring structure.

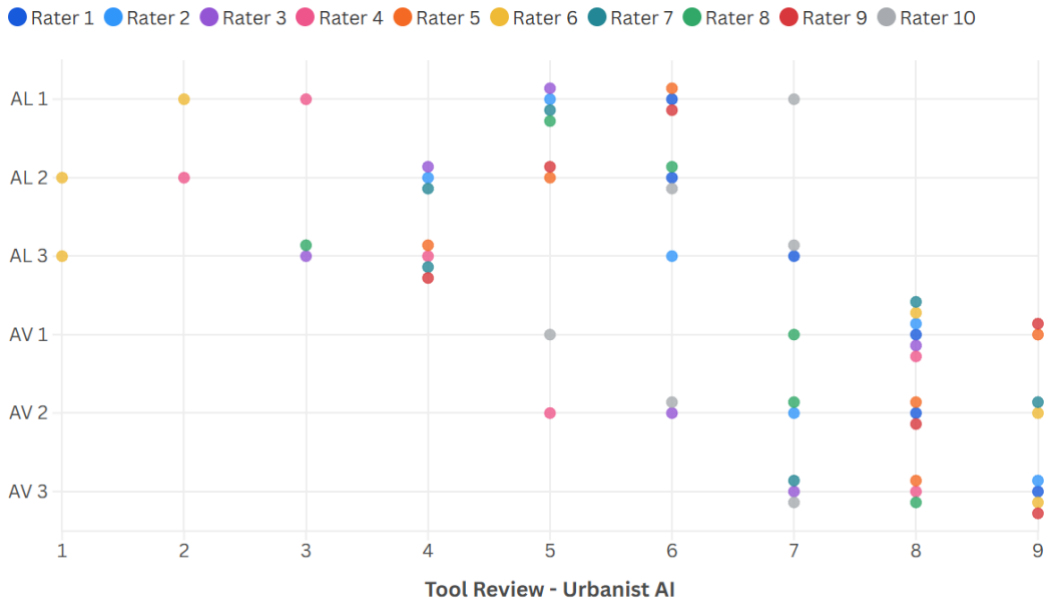


Figure 4: Facet-level scores (AL1–AL3, AV1–AV3) assigned by all ten expert raters for the tool Urbanist AI, illustrating the distribution of ratings across evaluators prior to aggregation.

Each rater was provided with the scoring rubric and the standardised Tool Documentation Package for five selected tools chosen to represent a range of lifecycle stages, AI capabilities, and architectural orientations: UrbanistAI, Spacemaker, NeuralConcept, Tally, and OpenSpace.ai. Raters were explicitly instructed to base their assessments solely on the provided documentation, rather than on prior hands-on experience or subjective familiarity with the

CC BY

tools. This constraint ensured that all evaluations were grounded in uniform evidentiary information, thereby reducing experiential bias and strengthening cross-rater comparability. All responses were collected anonymously, and only anonymised data were used in subsequent statistical analyses.

Statistical analysis was conducted in IBM SPSS Statistics. Inter-rater reliability was assessed using the Intraclass Correlation Coefficient (ICC) under a two-way mixed effects model with absolute agreement, following the recommendations for measurement instruments with a fixed rater panel evaluating multiple targets (Koo & Li, 2016). Both single-measure (ICC(3,1)) and average-measure (ICC(3,k)) coefficients were calculated; the latter constitutes the primary index of reliability, as the framework relies on aggregated panel scores rather than individual judgements.

The results, presented in Table 2, confirm strong panel-level reliability for both constructs. AI Logic achieved ICC(3,k) = .919 (95% CI [.749, .990]) and Architectural Value achieved ICC(3,k) = .990 (95% CI [.969, .999]), indicating that expert ratings are highly consistent when averaged and that the framework produces stable, reproducible outcomes across independent raters. At the individual level, Architectural Value also demonstrated high single-rater reliability (ICC(3,1) = .909), whilst AI Logic produced a moderate coefficient (ICC(3,1) = .532), indicating that assessments of computational sophistication are more susceptible to individual interpretive variability and benefit more from panel averaging. This pattern likely reflects the more intuitive and professionally shared nature of architectural judgement compared with the technical specificity required to assess AI mechanisms such as predictive modelling and optimisation logic. The significant F-statistics for both constructs confirm that experts meaningfully differentiated amongst the five tools rather than assigning undifferentiated or arbitrary scores (AL: $F(4,36) = 12.039$, $p < .001$; AV: $F(4,36) = 138.282$, $p < .001$).

Table 2: Intraclass Correlation Coefficients for AI Logic and Architectural Value.

Construct	ICC(3,1)	ICC(3,k)	95% CI [Average]	F(df1, df2)	p
AI Logic (AL)	0.532	0.919	[.749, .990]	12.039 (4,36)	< .001
Architectural Value (AV)	0.909	0.990	[.969, .999]	138.282 (4,36)	< .001

Table 3: Descriptive Statistics for Expert Ratings Across Five Validation Tools.

Tool	Mean AL	SD AL	Mean AV	SD AV
UrbanistAI	4.53	1.54	7.73	0.87
Spacemaker	7.40	0.72	7.23	0.69
NeuralConcept	7.43	0.75	5.63	0.91
Tally	5.67	1.71	4.73	0.83
OpenSpace.ai	7.50	0.96	1.60	0.44

Descriptive statistics for the five validation tools are presented in Table 3. For AI Logic, OpenSpace.ai (M = 7.50), NeuralConcept (M = 7.43), and Spacemaker (M = 7.40) received the highest scores, reflecting strong perceived AI sophistication, whilst UrbanistAI scored lowest (M = 4.53), consistent with its orientation towards design exploration over high-automation processing. Tally showed the greatest inter-rater variability (SD = 1.71), suggesting divergent expert interpretations of its underlying computational logic. For Architectural Value, UrbanistAI ranked highest (M = 7.73) and OpenSpace.ai scored lowest (M = 1.60), with generally smaller standard

CC BY

deviations across all tools than those observed for AI Logic, reinforcing the higher individual-level reliability found in the ICC results.

The inverse scoring profiles of UrbanistAI and OpenSpace.ai across the two constructs are analytically significant: one scores low on AI Logic but highest on Architectural Value, whilst the other shows the precise reverse pattern. This divergence provides empirical confirmation that computational sophistication and architectural contribution are independent dimensions rather than proxies for a single underlying quality, and supports the construct validity of the evaluation model.

3.4 Author calibration

To assess the consistency of the framework when applied by a single independent evaluator, the author rated the same five validation tools using the identical scoring protocol and documentation, with no access to the expert panel's scores during the rating process. Pearson correlation analysis was subsequently conducted to examine ranking consistency between the author and the expert panel. For Architectural Value, the correlation was perfect ($r = 1.000$, $p < .001$), indicating identical rank ordering across all five tools. For AI Logic, the correlation was strong ($r = .778$), though statistical significance was not reached given the small sample size ($N = 5$).

Bias assessment through mean difference scores (Author minus Expert) returned values of 0.094 for AI Logic and 0.018 for Architectural Value, both near zero, indicating little average directional difference between the author's ratings and expert consensus within the five-tool subset. Mean Absolute Error (MAE) was 0.694 for AI Logic and 0.066 for Architectural (Table 4). On a ten-point scale, both values are modest and fall comfortably within the bounds of natural inter-rater variability documented during the expert validation phase, suggesting that the author's independent scoring may remain aligned with expert consensus. However, this expectation is based on the five-tool calibration subset.

Table 4: Bias Assessment: Mean Difference and Mean Absolute Error (Author vs Expert Mean).

Construct	Mean Difference (Author - Expert)	SD	MAE	SD
AI Logic (AL)	0.094	0.865	0.694	0.397
Architectural Value (AV)	0.018	0.074	0.066	0.022

Graphical comparison of expert mean scores and author ratings provides further corroboration (Figure 5). For Architectural Value, near-horizontal slope lines across all five tools reflect close numerical agreement and an identical ranking structure. For AI Logic, minor crossovers are visible, consistent with the less-than-perfect correlation coefficient, yet deviations remain within acceptable bounds relative to expert variability. Taken together, the correlation analysis, mean differences and graphical comparison indicate alignment between the author's ratings and expert consensus for the five selected tools, supporting subsequent author-led application of the rubric. However, calibration does not eliminate interpretive subjectivity or establish equivalent agreement for the remaining forty-five tools.

3.5 Application to the full dataset

Following validation and calibration, the evaluation protocol was extended to the remaining forty-five tools using the identical rubric and documentation procedure. A standardised Tool Documentation Package was compiled for each application, drawing on publicly available technical literature and product documentation, and assessments were based exclusively on this material to maintain methodological consistency with the validation phase. Each tool was scored across all six facets on the 0 to 10 anchored scale, and composite AL and AV scores were computed by averaging the respective sub-criterion scores and rounding to the nearest 0.25 increment.

Because the rubric captures both computational sophistication and the preservation of architectural agency, the resulting dataset supports multidimensional analysis rather than a purely technical ranking. Application of the same rubric and documentation procedure maintains methodological continuity between the five-tool calibration subset and the remaining forty-five tools. The full dataset of fifty tools thereby forms the basis for the lifecycle

CC BY

mapping, tension map analysis, and cross-referencing with secondary parameters presented in the sections that follow.

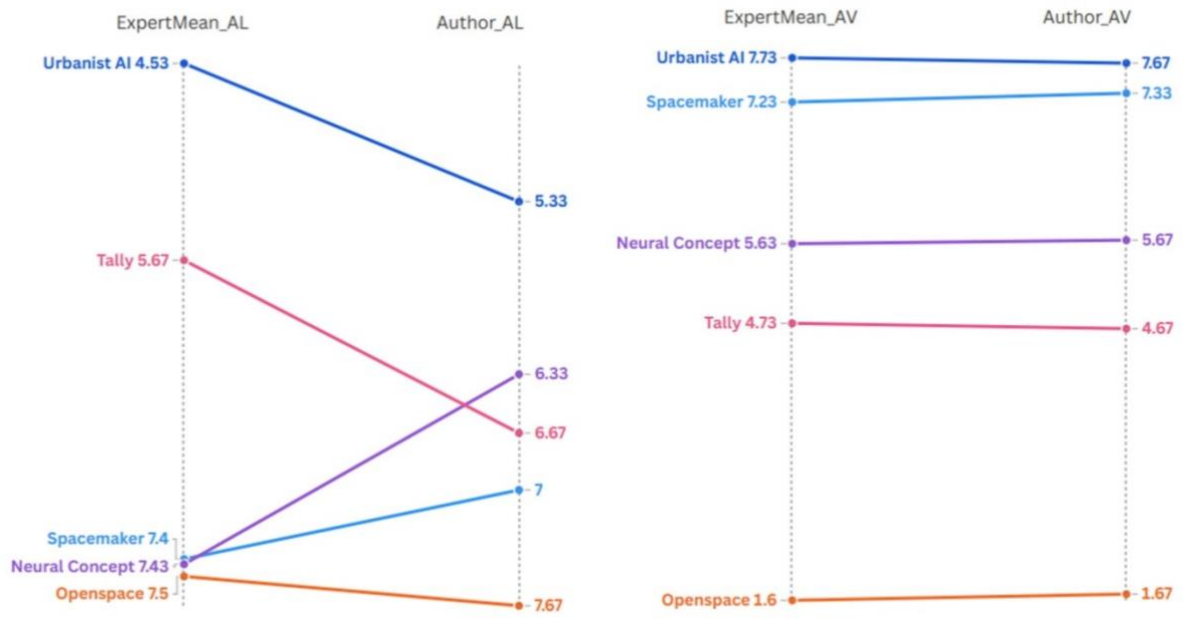


Figure 5: Comparison between expert mean scores and author-assigned scores for AI Logic (AL) and Architectural Value (AV) across the five validation tools.

3.6 Tension map analysis

The tension map (Figure 6) reveals clear clustering patterns that align closely with RIBA lifecycle stages. The upper-right quadrant (high AI Logic, high Architectural Values) is dominated by early-stage design tools from Stages 1–2, such as *Spacemaker* (Harrouk, 2020; Landes, 2022), *ARKdesign.ai* (*AECplusTech ARKDesign*, n.d.; *ARKDesign AI*, n.d.), and *Finch3D* (*Finch - AI for How the World Builds*, n.d.). These combine computational optimisation with strong design agency, enabling contextual, aesthetic, and exploratory decision-making.

The lower-right quadrant (high AI Logic, low Architectural Values) is primarily occupied by Stages 4–5 tools such as *Alice Technologies*, *OpenSpace.ai* (*OpenSpace: The Visual Intelligence Platform for Builders*, n.d.), and *Buildots* (*Construction Intelligence Platform | Buildots*, n.d.). These platforms focus on construction sequencing, monitoring, and verification. These are typically tasks where efficiency and precision are paramount, but opportunities for qualitative design input are minimal.

Stage 3 - Spatial Coordination tools tend to occupy a middle ground, balancing technical optimisation with some retention of design intent. Examples such as *NeuralConcept* and *Revit (AI plugins)* apply automation to modelling and simulation while still relying on user input to shape spatial outcomes.

The upper-left quadrant (low AI Logic, high Architectural Values) is sparsely populated, indicating a lack of tools that foreground designer intuition while employing only moderate automation. The lower-left quadrant (low AI Logic, low Architectural Values) contains almost no entries, reflecting limited adoption of low-automation, low-impact AI solutions in current architectural practice.

The distribution highlights a consistent tension: as AI Logic increases, Architectural Values often decrease unless tools are explicitly designed to keep the human in the loop. In early stages, AI tends to act as a collaborator, augmenting creative exploration. By later stages, it shifts toward directive automation, replacing human decision-

CC BY

making with computational certainty. This reflects broader debates on design agency, raising questions about how to preserve qualitative, context-driven decision-making in an increasingly optimised workflow.

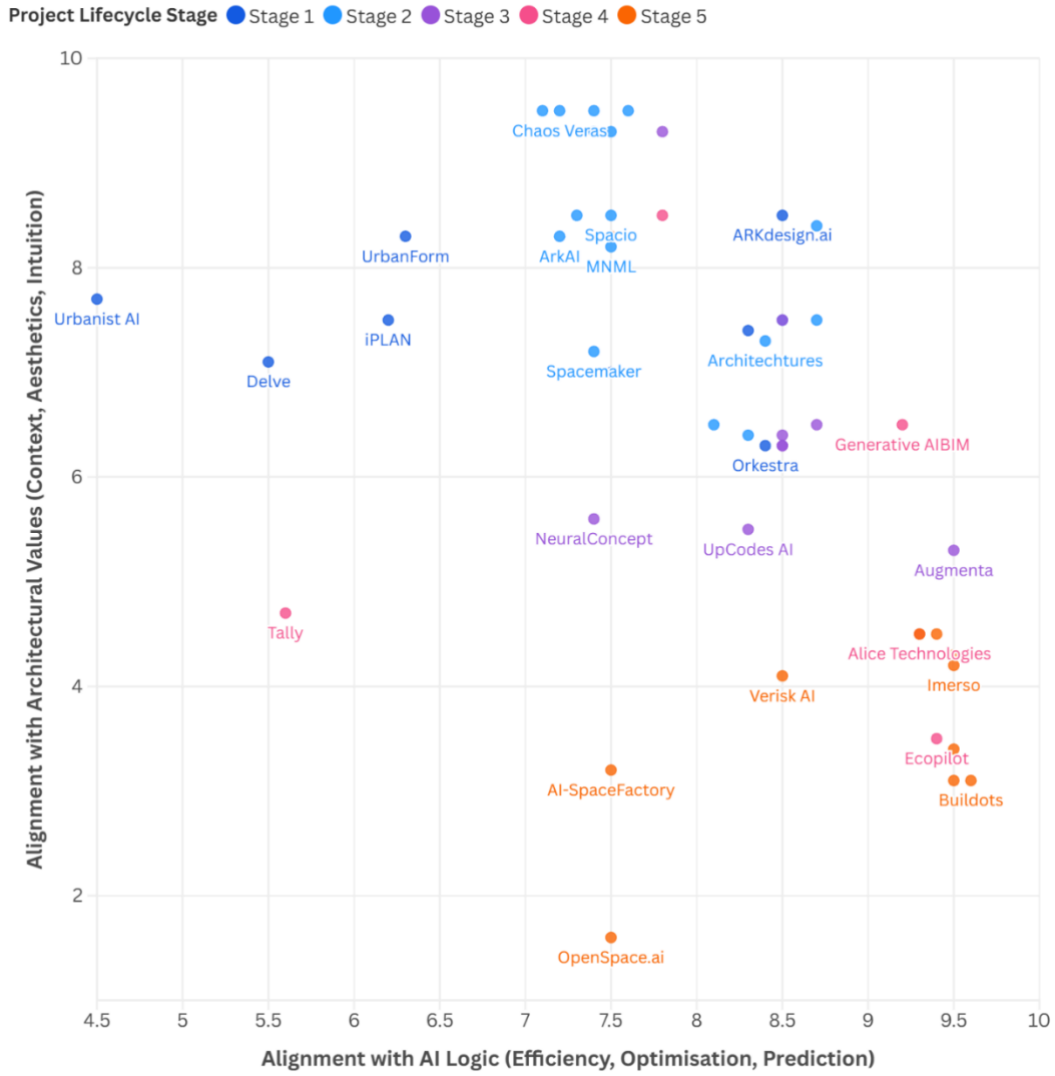


Figure 6: Distribution of the fifty AI tools across AI Logic and Architectural Values, colour-coded by project lifecycle stage (Stages 1–5).

3.7 Gaps and opportunities

Two major gaps emerge from the analysis. First, the upper-left quadrant remains underdeveloped, representing potential for tools that are high in Architectural Values yet rely on lighter, more interpretable AI automation. This enables designers to retain control while benefiting from selective computational assistance. Second, late-stage AI tools are overwhelmingly efficiency-oriented, suggesting opportunities to embed design-awareness in construction, monitoring, and manufacturing systems so that architectural intent is preserved through to project delivery. The diagram (Figure 7) demonstrates this with the average position of the tools across project lifecycle stages.

CC BY

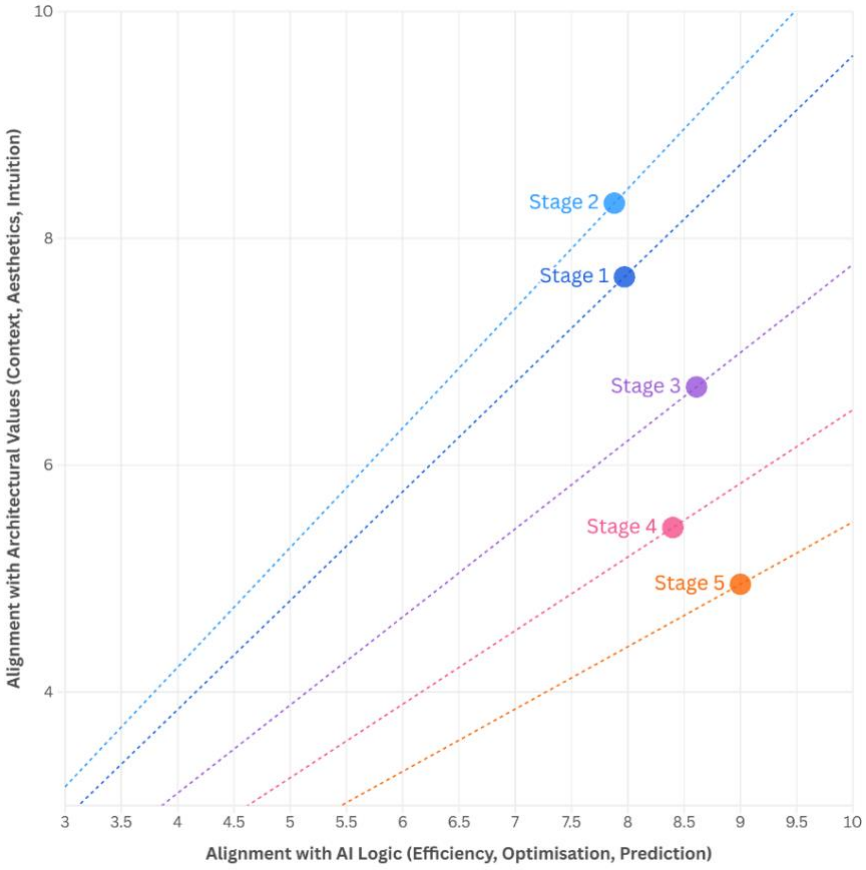


Figure 7: Average positions of AI tools by lifecycle stage, showing higher Architectural Values in early stages and stronger AI Logic in later stages.

4. CROSS-REFERENCING WITH OTHER PARAMETERS

4.1 Building typologies

The previously selected 50 applications were tested against a range of building typologies to understand their distribution and applicability across different sectors of the built environment. This assessment aims to reveal patterns in how these tools are being used, highlight areas of concentrated activity, and identify typologies or project stages that may be underserved. To visualise these relationships, a diagram (Figure 8) was created, mapping the connection between project lifecycle stages; Preparation and Briefing, Concept Design, Spatial Coordination, Technical Design, and Manufacturing & Construction and building typologies which are Residential, Commercial, Mixed Use, Urban/Master planning, Industrial, Infrastructure, and Institutional.

The diagram (Figure 8) serves as both a diagnostic and comparative tool. By tracing the flow of connections, it becomes possible to see where the majority of AI-enabled processes are concentrated and where coverage thins out. Wider bands indicate typology–stage combinations with greater tool support, while narrower flows expose underrepresented areas that may present opportunities for innovation or targeted development. This visual mapping not only clarifies the strengths of current digital tool deployment but also provides a foundation for strategic decision-making in technology adoption, investment, and research focus.

From the visualisation, Residential and Commercial typologies emerge as the most heavily connected across nearly all stages, suggesting that these sectors have the broadest tool adoption and applicability in practice. Concept Design shows strong flows to Residential and Mixed Use, highlighting these as early-stage priorities in many projects. Spatial Coordination appears to have the most balanced connections across multiple typologies, including Commercial and Urban/Masterplanning, indicating its central role in ensuring integration across diverse project

CC BY

types. Manufacturing & Construction connects widely with Industrial and Infrastructure, revealing the sector-specific demand for construction-phase tools in these areas.

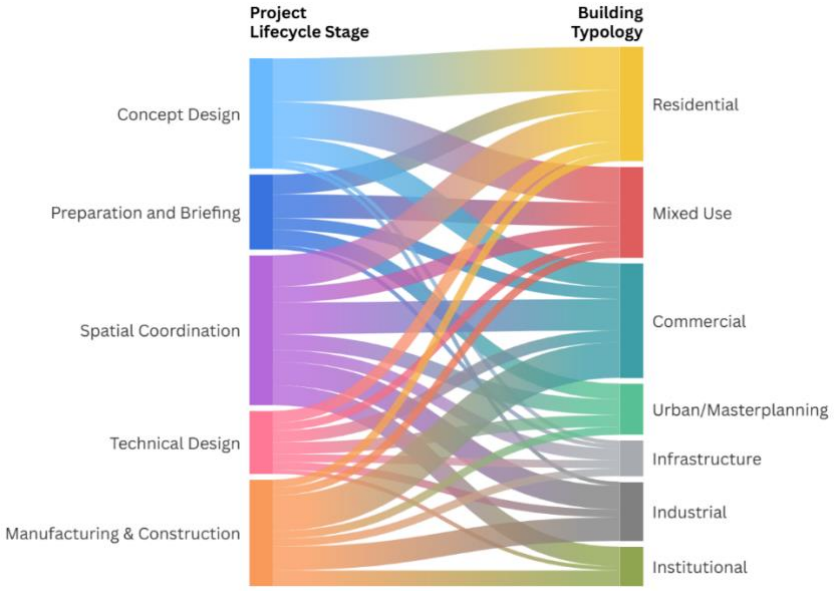


Figure 8: Relationship between project lifecycle stages and building typologies, showing concentrations of AI tool adoption across residential, commercial, and mixed-use projects.

The diagram also reveals gaps and opportunities. For example, Institutional and Infrastructure typologies have fewer strong connections from earlier stages like Concept Design and Preparation and Briefing, implying potential underrepresentation in early-stage design toolsets. This suggests that future development of AI or digital workflows could focus on expanding coverage for these typologies earlier in the project lifecycle. Overall, the Sankey chart provides a clear, high-level map of how different project phases align with building sectors, highlighting both strengths in coverage and potential areas for innovation.

4.2 Secondary parameters – user type, tool maturity, tool design logic and regulatory constraints

This stage of the study organises the previously selected 50 AI-powered AEC tools into a multi-criteria framework to better understand their applicability, market readiness, and technical characteristics. Each tool is tested against four key parameters:

- User Type – categorising the primary professional audience for each tool (e.g., architect, urban planner, engineer, contractor).
- Regulatory Context – combining geographic focus and compliance alignment, including factors such as US zoning/building codes, EU energy standards, or APAC building regulation standards.
- Tool Maturity / Development Stage – classifying tools as Prototype / Experimental, Early-Stage, or Commercial Product based on market readiness and deployment.
- Design Logic / Functionality – identifying the underlying computational or AI approach, such as Generative Design, Machine Learning (ML), Generative Adversarial Networks (GAN), Analytical / Simulation, or Monitoring / Quality Assurance (QA).

By consolidating these dimensions, the analysis provides a holistic view of not only where a tool fits in the design and construction process but also the contexts in which it is most likely to be effective and compliant. This classification allows for deeper cross-referencing between tool capabilities, intended users, and the environments in which they operate. For example, a generative design tool classified as “Commercial Product” with global BIM

CC BY

compliance offers a very different adoption potential compared to an early-stage GAN-based concept tool focused on visualisation in APAC markets.

The inclusion of regulatory constraints ensures that market suitability is understood not just from a technical perspective but also through the lens of local codes, standards, and policy environments. This dataset can be used to identify market gaps, highlight areas where innovation is concentrated, and guide strategic decisions on technology adoption, development priorities, or regional market entry.

User Type

Architects are the primary users across early to mid-stages, while contractors dominate Stage 5, reflecting the construction-focused nature of late-stage AI applications (Figure 9). Engineers are more evenly distributed across stages, and sustainability specialists are underrepresented throughout, highlighting a gap in environmentally driven AI solutions.

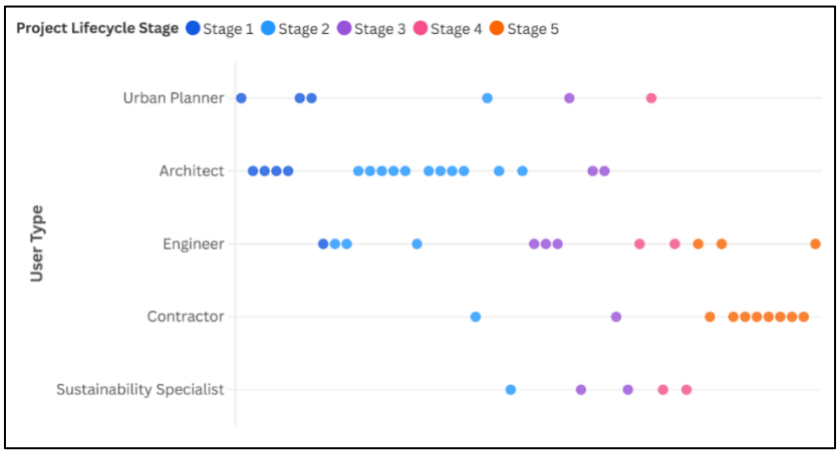


Figure 9: Distribution of AI tools by user type across project lifecycle stages.

Geographic / Regulatory Context

The majority of tools position themselves as globally applicable, with smaller clusters targeting North America/US and Europe (Figure 10). APAC-focused tools are notably scarce, indicating regional adoption gaps, while a significant share of tools have unspecified geographic/regulatory contexts, reducing clarity on compliance alignment.

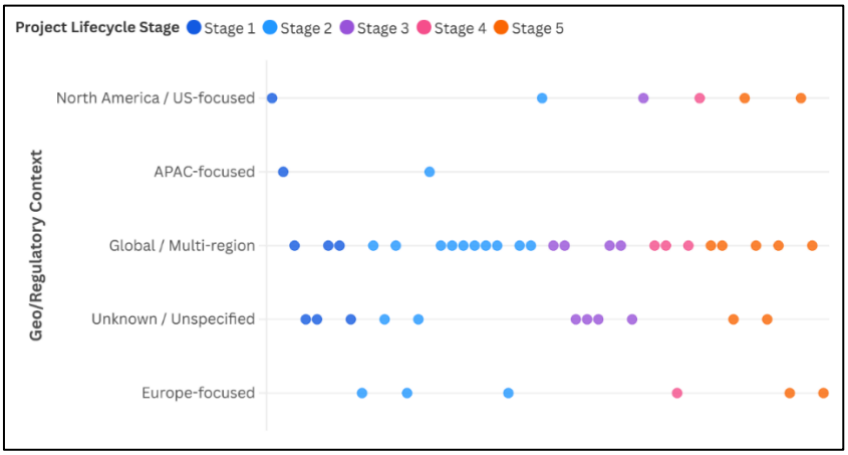


Figure 10: Distribution of AI tools by Geographic and Regulatory Context.

CC BY

Tool Maturity / Development Stage

Most tools are commercially available and span all lifecycle stages (Figure 11), with early-stage and prototype/experimental tools concentrated in Stages 2–3, where design innovation is most active.

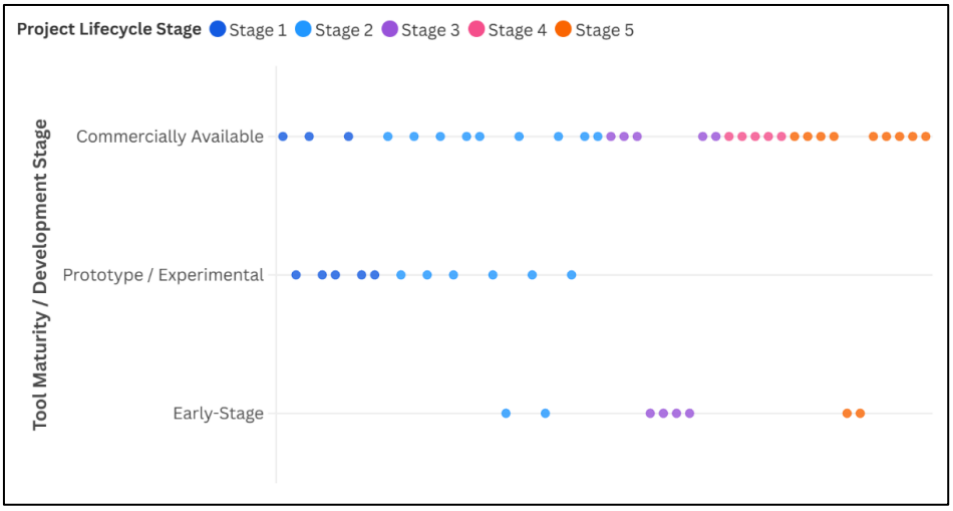


Figure 11: Distribution of AI tools by Tool Maturity Stage across project lifecycle stages.

Design Logic / Functionality

Analytical/Simulation and Generative Design dominate the early to mid-project lifecycle stages (1–3), supporting conceptual and iterative design workflows (Figure 12). Machine Learning (ML) and Monitoring/QA tools are concentrated in later stages (4–5), aligning with delivery, optimisation, and verification. GAN-based tools are rare and mainly appear in mid-stages, reflecting niche visualisation and generative applications.

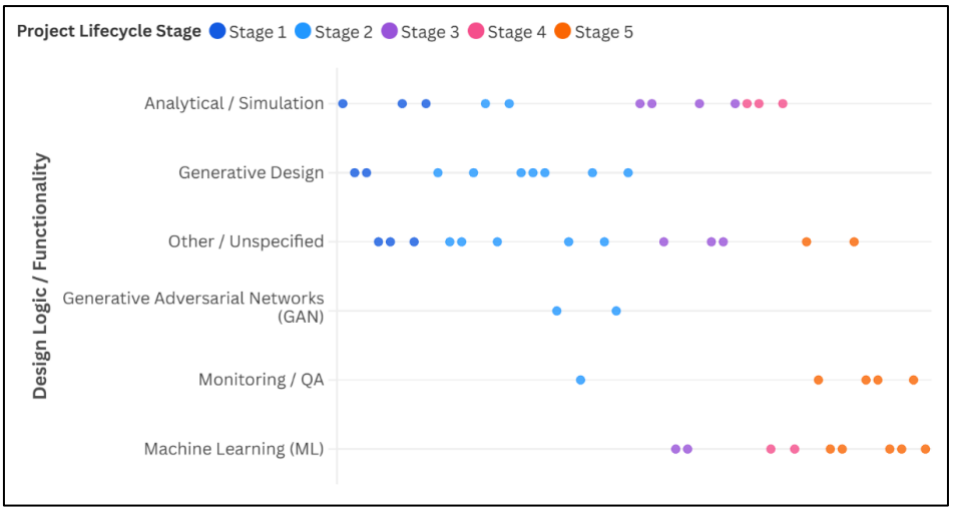


Figure 12: Distribution of AI tools by its Design Logic across project lifecycle stages.

5. DISCUSSIONS

This study set out to examine how AI tools are reshaping design agency in architecture by evaluating fifty applications across the design lifecycle. By combining a scoring rubric for AI Logic (AL) and Architectural Values (AV) with secondary parameters such as typology, user groups, tool maturity, design logic, and regulatory context,

CC BY

the analysis provides a multi-scalar picture of how automation, optimisation, and prediction are reconfiguring the relationship between technology and authorship.

The scoring of tools against AL and AV revealed a consistent gradient across the project lifecycle. Early-stage platforms such as ARKdesign.ai and Spacemaker combine high levels of computational reasoning with strong support for architectural values, positioning AI as a collaborator that expands design exploration and contextual responsiveness. By contrast, late-stage tools such as Alice Technologies and OpenSpace.ai are dominated by efficiency-driven logics of optimisation and monitoring, offering little support for creative authorship or qualitative design dimensions. Tools in the middle stages, such as NeuralConcept, straddle these conditions, offering technical performance insights while still requiring human input for spatial judgement. As shown in Figure 13, this distribution underscores a broader tendency; as AI Logic increases, Architectural Values often decline unless systems are deliberately designed to keep designers in the loop.

When these scores are cross-referenced with building typologies, the unevenness of AI adoption becomes visible. Residential, commercial, and mixed-use projects dominate the dataset, especially in early design stages where generative and simulation tools are prevalent. These sectors benefit from high AL and relatively high AV scores, suggesting that commercial markets are best served by tools that combine automation with exploratory potential. By contrast, institutional and infrastructure typologies appear underrepresented, particularly in early stages, and skew toward tools that score high in AL but low in AV. This points to a disciplinary imbalance: AI is consolidating its role in profit-driven contexts while underserving civic or public typologies where cultural and contextual values are critical.

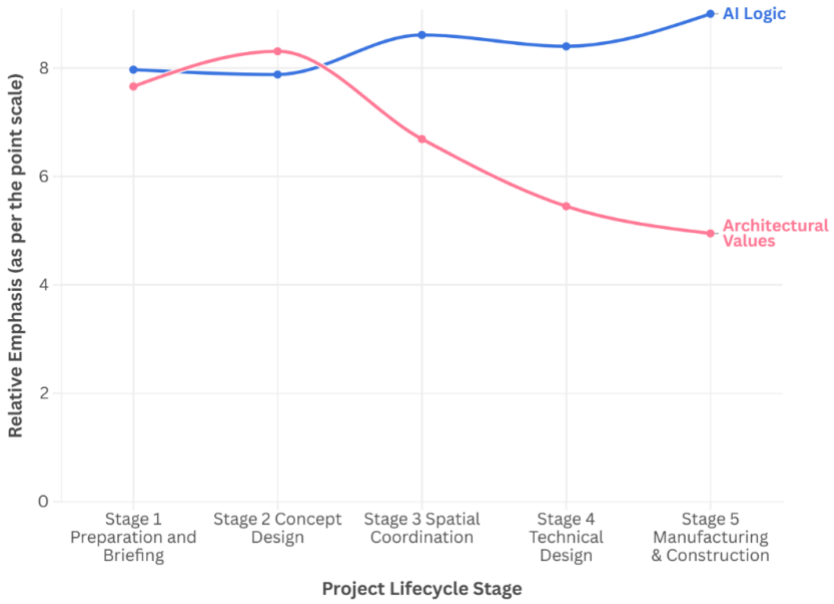


Figure 13: Redistribution of design agency across the lifecycle, with AI Logic increasing and Architectural Values declining from early design to construction.

Cross-referencing user groups with AL and AV scores also highlights the redistribution of agency across professions. Architects primarily engage with early-stage tools that support exploration, often scoring higher on Architectural Values. Contractors, by contrast, dominate the use of late-stage monitoring and quality assurance tools, where AI Logic is high but design values are minimal. Engineers occupy a more evenly spread role, particularly in simulation and optimisation tasks, while very few tools are targeted towards sustainability as intended users in the dataset. This suggests that sustainability, though central to architectural discourse, remains framed largely as a technical optimisation problem rather than a value-laden design dimension, raising questions about the adequacy of current AI approaches to environmental challenges.

As shown in Figure 14, patterns also emerge when tool maturity is considered alongside design logics. Early-stage generative tools are often experimental or in beta stages, reflecting the ongoing innovation in conceptual design workflows. By contrast, late-stage construction tools are overwhelmingly mature, commercial products, aligning

CC BY

with industry imperatives of cost reduction and risk management. When mapped against AV scores, a clear paradox emerges: the commercially mature tools in the sample are also those least engaged with architectural values, while the most experimental tools are those that most actively support design authorship and contextual sensitivity. This suggests that the consolidation of AI within the profession may be reinforcing optimisation logics at the expense of creative ones.

Regulatory and geographic contexts add another layer of complexity. The majority of tools position themselves as globally applicable, yet in practice they are implicitly oriented toward North American and European regulatory environments, particularly US zoning and EU energy standards and very few address APAC or Global South contexts. Cross-referenced with typology and maturity, this reveals a form of embedded bias, globally marketed tools reinforce Western standards while marginalising alternative regulatory frameworks and localised design priorities. As a result, AI adoption risks reproducing dominant value systems rather than supporting pluralistic, context-sensitive practice.

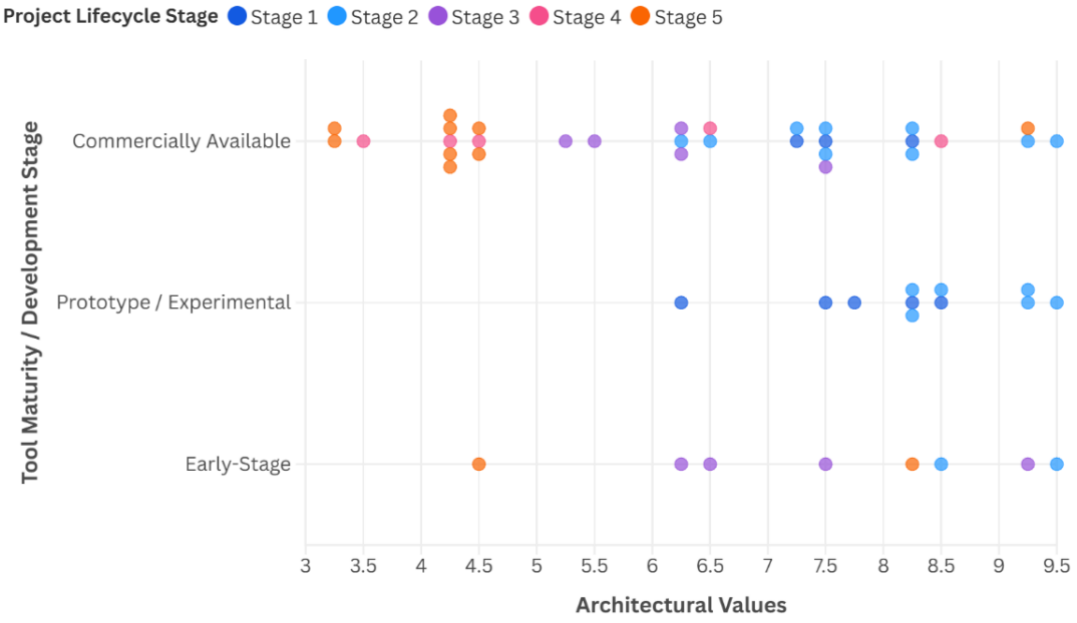


Figure 14: Distribution of AI tools across architectural values and tool maturity, color-coded by project lifecycle stage illustrating the paradox between creative potential and market maturity.

Taken together, these overlapping patterns suggest that design agency in the age of AI is not a singular phenomenon but a distributed condition shaped by multiple, intersecting parameters (Cocho-Bermejo, 2025). The lifecycle stage largely determines whether AI operates as a collaborator or a controller; and typologies dictate where disciplinary values are supported or constrained. Tool maturity and market adoption skew toward efficiency over creativity and regulatory contexts privilege certain standards over others. While AI clearly extends the architect's capacity in some domains, particularly in early-stage exploration, it also narrows the scope of agency in others, particularly in construction and delivery where algorithmic certainty takes precedence.

Ten experts assessed five tools representing RIBA Stages 1–5, while the remaining forty-five tools were assessed solely by the author following calibration. The reliability findings therefore apply to the five-tool subset. Although calibration supported consistency with expert consensus, single-rater interpretation remains a limitation of the wider assessment. Further independent ratings would help establish whether similar agreement extends across the full dataset. However, by making the criteria explicit, the rubric opens the possibility for further inter-rater testing, comparative studies, and more robust validation across different evaluators and professional contexts. This methodological transparency helps characterise the selected AI tools and establishes a baseline against which future tools and their support for architectural design agency can be systematically assessed.

CC BY

The scope of the present study is focused on design and planning, and it leaves open opportunities for extension into domains such as structural engineering, mechanical, electrical and plumbing (MEP) coordination, and construction logistics, where AI tools are already proliferating. These domains are central to the redistribution of agency in later project stages, as automation is particularly pronounced in clash detection, systems integration, scheduling, and predictive maintenance. A future iteration of this framework could therefore broaden the lens, applying the same AI Logic and Architectural Values rubric to evaluate tools across these technically intensive subfields. Such an extension would reveal whether the same trade-offs between automation and architectural values hold true, or whether new forms of agency emerge when AI is embedded in highly coordinated, systems-driven aspects of the building lifecycle.

6. CONCLUSION

This study has examined fifty AI tools across the architectural lifecycle, evaluating each against two analytically distinct constructs: AI Logic, capturing computational sophistication through automation, optimisation, and predictive capability, and Architectural Value, capturing the degree to which tools preserve design agency, contextual sensitivity, and exploratory capacity. The findings demonstrate that AI does not eliminate design agency but redistributes it, expanding creative latitude in early design stages whilst progressively constraining qualitative judgement as projects advance towards construction and delivery.

The principal focus of this work is the development and expert validation of a structured evaluation framework capable of assessing AI tools beyond purely technical benchmarking. Ten AEC professionals independently rated five representative tools against the scoring rubric, yielding strong panel-level inter-rater reliability for both constructs (ICC(3,k) = .919 for AI Logic; ICC(3,k) = .990 for Architectural Value) and confirming that the two dimensions are analytically independent. Author calibration against the expert consensus returned near-zero directional bias and Mean Absolute Error values well within natural inter-rater variability, establishing the framework as a replicable instrument that can be applied consistently as the AI landscape continues to evolve. By making the evaluative criteria explicit, the rubric also opens the possibility for comparative and longitudinal studies across different evaluators and professional contexts.

Cross-referencing with secondary parameters revealed systemic imbalances: AI adoption is concentrated in residential and commercial typologies, relatively few tools in the selected sample explicitly identify sustainability specialists as intended users, and most commercially mature tools are implicitly oriented towards North American and European regulatory environments. These patterns suggest that current AI development in architecture risks reinforcing existing disciplinary and geographic biases rather than expanding the conditions of practice.

The framework provides a structured basis for assessing how AI tools support architectural design agency and for identifying opportunities to strengthen contextual sensitivity, equity and design awareness in future tool development. Future work should extend the rubric to structural engineering and construction logistics domains, and broaden the expert panel to encompass more diverse geographic and professional backgrounds, further strengthening both the reliability and representativeness of the instrument.

The findings within the selected dataset suggest that AI tools' support for architectural values varies across intended functions and project stages, alongside their technical capabilities. This pattern raises concerns that an emphasis on efficiency may leave contextual and qualitative considerations underrepresented. Incorporating sustainability and contextual responsiveness into future tool development could strengthen support for these values. The framework provides a structured basis for examining these priorities and informing AI tool selection and development within the built environment.

REFERENCES

AECplusTech ARKDesign. (n.d.). Retrieved 31 May 2026, from https://www.aecplustech.com/tools/arkdesign-ai

Albukhari, I. N. (2025). The role of artificial intelligence (AI) in architectural design: A systematic review of emerging technologies and applications. *Journal of Umm Al-Qura University for Engineering and Architecture*, *16*(4), 1457–1476. https://doi.org/10.1007/s43995-025-00186-1

ARKDesign AI. (n.d.). Retrieved 31 January 2026, from https://arkdesign.ai/

As, I., Pal, S., & Basu, P. (2018). Artificial intelligence in architecture: Generating conceptual design via deep learning. *International Journal of Architectural Computing*, *16*(4), 306–327. https://doi.org/10.1177/1478077118800982

Awan, N., Schneider, T., & Till, J. (2011). *Spatial Agency: Other Ways of Doing Architecture*. Routledge.

Baumberger, C. (2009). From Logic to Art. In G. Ernst, J. Steinbrenner, & O. R. Scholz (Eds), *Themes from Nelson Goodman* (pp. 293–320). De Gruyter. https://doi.org/10.1515/9783110327199.293

Caetano, I., Santos, L., & Leitão, A. (2020). Computational design in architecture: Defining parametric, generative, and algorithmic design. *Frontiers of Architectural Research*, *9*(2), 287–300. https://doi.org/10.1016/j.foar.2019.12.008

Carta, S. (2019). *Book Big Data, Code and the Discrete City: Shaping Public Realms* (1st Edition). Routledge. https://doi.org/https://doi.org/10.4324/9781351007405

Chaillou, S. (2019). *AI + Architecture | Towards a New Approach*. Harvard University.

Chaillou, S. (2022). *Artificial Intelligence and Architecture: From Research to Practice*. Birkhäuser.

Cocho-Bermejo, A. (2025). Artificial Intelligence and Architectural Design Before Generative AI: Artificial Intelligence Algorithmics Approaches 2000–2022 in Review. *Engineering Reports*, *7*(4), e70114. https://doi.org/10.1002/eng2.70114

Construction Intelligence Platform | Buildots. (n.d.). Retrieved 25 January 2026, from https://buildots.com/

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th Edition). Sage.

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

Finch—AI for how the world builds. (n.d.). Retrieved 24 February 2026, from https://www.finch3d.com/

Harrouk, C. (2020). Spacemaker Proposes AI-Powered Generative Design to Create More Sustainable Spaces and Cities | ArchDaily. *ArchDaily*. https://www.archdaily.com/952850/spacemaker-proposes-ai-powered-generative-design-to-create-more-sustainable-spaces-and-cities

Jang, S., Roh, H., & Lee, G. (2025). Generative AI in architectural design: Application, data, and evaluation methods. *Automation in Construction*, *174*, 106174. https://doi.org/10.1016/j.autcon.2025.106174

Khan, A., Chang, S., & Chang, H. (2025). Generative AI approaches for architectural design automation. *Automation in Construction*, *180*, 106506. https://doi.org/10.1016/j.autcon.2025.106506

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Landes, J. (2022). Spacemaker.Ai: Using AI in developing urban block variations. In I. As, P. Basu, & P. Talwar (Eds), *Artificial Intelligence in Urban Planning and Design* (pp. 263–291). Elsevier. https://doi.org/10.1016/B978-0-12-823941-4.00010-X

Leatherbarrow, D. (2009). *Architecture oriented otherwise*. Princeton Architectural Press.

Li, C., Zhang, T., Du, X., Zhang, Y., & Xie, H. (2025). Generative AI models for different steps in architectural design: A literature review. *Frontiers of Architectural Research*, *14*(3), 759–783. https://doi.org/10.1016/j.foar.2024.10.001

Li, J., Liu, Z., Han, G., Demian, P., & Osmani, M. (2024). The Relationship Between Artificial Intelligence (AI) and Building Information Modeling (BIM) Technologies for Sustainable Building in the Context of Smart Cities. *Sustainability*, *16*(24), 10848. https://doi.org/10.3390/su162410848

Miles, M., Huberman, M., & Saldana, J. (2014). *Qualitative Data Analysis: A Method Sourcebook* (3rd Edition). Sage.

CC BY

Nagy, D., Villaggi, L., & Benjamin, D. (2018). *Generative Urban Design: Integrating Financial and Energy Goals for Automated Neighborhood Layout*. https://doi.org/10.22360/simaud.2018.simaud.025

Newton, D. (2021). Toward dynamic and explorative optimization for architectural design. In *The Routledge Companion to Artificial Intelligence in Architecture* (1st Edition, p. 21). Routledge.

OpenSpace: The Visual Intelligence Platform for Builders. (n.d.). Retrieved 31 January 2026, from https://www.openspace.ai/

Pallasmaa, J. (2012). Newness, Tradition and Identity: Existential Content and Meaning in Architecture. *Architectural Design*, *82*(6), 14–21. https://doi.org/10.1002/ad.1486

Pasquinelli, M., & Joler, V. (2021). The Nooscope manifested: AI as instrument of knowledge extractivism. *AI & SOCIETY*, *36*(4), 1263–1280. https://doi.org/10.1007/s00146-020-01097-6

Pérez-Gómez, A. (2016). *Attunement: Architectural Meaning after the Crisis of Modern Science*. The MIT Press. https://doi.org/https://doi.org/10.7551/mitpress/10703.001.0001

RIBA Plan of Work. (n.d.). Retrieved 31 May 2026, from https://www.riba.org/work/insights-and-resources/riba-plan-of-work/

Schneider, J., Sinem, K., & Stockhammer, D. (2025). *Empowering Clients—Transformation of Design Processes Due to Generative AI* (arXiv:2411.15061). arXiv. https://doi.org/10.48550/arXiv.2411.15061

Tamke, M., Nicholas, P., & Zwierzycki, M. (2018). Machine learning for architectural design: Practices and infrastructure. *International Journal of Architectural Computing*, *16*(2), 123–143. https://doi.org/10.1177/1478077118778580

Yang, L. (2025). Generative AI for Sustainable Architectural Design Optimization. *Scientific Journal of Technology*, *7*(5), 96–106. https://doi.org/10.54691/czaf4b21

Zhuang, X., Zhu, P., Yang, A., & Caldas, L. (2025). Machine learning for generative architectural design: Advancements, opportunities, and challenges. *Automation in Construction*, *174*, 106129. https://doi.org/10.1016/j.autcon.2025.106129