

Generative Artificial Intelligence in Smart Engineering Design Optimization: A Synthesis of Computational Creativity, Structural Performance, and Human-Centered Practice

John Owen

Ladoke Akintola University of Technology
Department of Mechanical and Industrial Engineering
Email: jofajinmi@student.lautech.edu.ng

Abstract

Generative artificial intelligence has moved from a peripheral computational curiosity to a central instrument of contemporary engineering design, reshaping how structural forms, materials configurations, and manufacturing strategies are conceived and refined. This paper undertakes an integrative, theory-driven synthesis of the scholarship surrounding generative AI in smart engineering design optimization, with particular attention to generative adversarial networks, variational autoencoders, diffusion-based models, and emerging large language model agents. Rather than cataloguing individual techniques, the study interrogates how these computational paradigms interact with established engineering design theory, including design space exploration, multi-objective optimization, and design for manufacturing. Employing a qualitative, meta-synthetic methodology grounded in a structured review of peer-reviewed literature published primarily between 2014 and 2026, the paper develops an analytical framework that maps generative model classes onto discrete phases of the engineering design lifecycle. The findings indicate a discernible trajectory from narrow, single-objective topology generation toward multimodal, constraint-aware, and increasingly autonomous design agents capable of negotiating competing structural, material, and manufacturability requirements. The analysis further reveals persistent gaps concerning interpretability, manufacturing feasibility validation, and the epistemic role of human designers within increasingly automated pipelines. The paper concludes that generative AI functions most productively not as a replacement for engineering judgment but as a computational collaborator that expands the exploratory reach of the design process while imposing new obligations of verification, accountability, and domain grounding. These insights contribute a consolidated conceptual map for researchers and practitioners seeking to situate emergent generative techniques within a coherent, rigorous engineering design paradigm.

Keywords: generative artificial intelligence; engineering design optimization; generative adversarial networks; diffusion models; topology optimization; large language models; smart manufacturing

DOI: 10.21590/ijtmh20241002428

1. Introduction

Engineering design has always been constrained by the cognitive and computational bandwidth of the designer, who must reconcile competing structural, economic, material, and manufacturing requirements within a search space that is frequently too vast to traverse by intuition alone. For much of the twentieth century, this constraint was managed through simplification: engineers relied on parametric families of known solutions, standardized components, and heuristics

distilled from experience, deferring genuinely exploratory search to the rare cases where computational resources permitted exhaustive analysis. The emergence of gradient-based and evolutionary topology optimization during the final decades of the century, exemplified by compact and widely disseminated implementations such as the eighty-eight line MATLAB code of Andreassen et al. (2011), expanded the reach of computational design considerably, yet these methods remained tethered to iterative finite element evaluation and were often prohibitively slow when applied to high-resolution, multi-physics, or multi-objective problems.

The past decade has witnessed a qualitatively different development: the maturation of deep generative models capable of learning compact, information-rich representations of design spaces directly from data, and of sampling novel candidate designs from those representations far more rapidly than conventional physics-based iteration allows. Generative adversarial networks, introduced by Goodfellow et al. (2014), variational autoencoders, and, more recently, diffusion probabilistic models articulated by Ho et al. (2020), have been adapted with increasing sophistication to structural topology synthesis, three-dimensional shape generation, materials microstructure design, and, most recently, to natural language mediated design ideation through large language models. This proliferation of technique has been accompanied by growing institutional recognition of the field's maturity, reflected in dedicated special issues such as that convened by Chen, Krishnamurthy, Lu, Luo, McComb, Ravi, and Sha (2026) in the *Journal of Computing and Information Science in Engineering*, which consolidates contributions on generative AI across design, manufacturing process planning, and materials systems.

Despite this proliferation, the scholarly landscape remains fragmented across disciplinary silos. Structural optimization researchers tend to evaluate generative models primarily against compliance minimization and manufacturability metrics, while the human-computer interaction and design cognition communities emphasize creativity support, ideation breadth, and the phenomenology of human-AI collaboration. Large language model research, meanwhile, has advanced largely independently of the structural and materials-oriented generative design literature, despite obvious complementarities in tasks such as requirement elicitation, design rationale generation, and conceptual sketching. The absence of a unifying analytical vocabulary capable of bridging these subfields constitutes a genuine impediment to cumulative knowledge building, because researchers in one tradition frequently reinvent evaluative criteria and methodological safeguards that have already been established, tested, and refined in another.

This paper addresses that fragmentation through an integrative synthesis oriented around three research questions. First, how do distinct classes of generative AI models, namely generative adversarial networks, variational autoencoders, diffusion models, and large language model agents, map onto the discrete phases of the engineering design lifecycle, from conceptual ideation through detailed structural optimization to manufacturing-aware refinement? Second, what recurring theoretical tensions and evaluative debates structure the scholarly conversation across these subfields, particularly concerning the balance between statistical novelty and physical validity? Third, what conceptual and methodological gaps must be addressed for generative AI to be responsibly integrated into high-stakes engineering practice, where failure

carries safety, financial, and reputational consequences that differ substantially from the failure modes tolerated in purely aesthetic or recreational generative applications?

The paper proceeds from a conceptual foundation informed by two complementary theoretical lenses. The first, design space theory, treats engineering design as a constrained search through a multidimensional space of feasible configurations, a framing that renders generative models legible as learned samplers capable of concentrating probability mass in regions of the design space associated with desirable performance. The second lens draws on sociotechnical and human-centered design theory, which insists that computational tools be evaluated not solely by output quality but by how they reshape the cognitive labor, accountability structures, and creative agency of the humans who deploy them. Read together, these frameworks permit an analysis that neither treats generative AI as a purely mathematical optimization device nor romanticizes it as an autonomous creative agent, but instead situates it within the broader institutional and epistemic fabric of engineering practice. The remainder of the paper synthesizes the relevant literature, articulates the methodological approach used to construct this synthesis, presents an integrated analytical account of findings, and discusses the implications of this account for both scholarship and practice.

2. Literature Review

2.1 Foundations: Generative Model Architectures and Their Engineering Translation

The theoretical foundation of generative AI in engineering design rests on the adversarial and variational learning paradigms that emerged from the broader machine learning literature before being adapted to structural and geometric domains. Goodfellow et al. (2014) formulated the generative adversarial network as a minimax game between a generator that synthesizes candidate samples and a discriminator that attempts to distinguish synthetic from authentic data, a formulation whose central insight, that a learned critic can substitute for an explicit likelihood function, proved unusually well suited to the ill-defined, high-dimensional distributions characteristic of engineering geometries. Variational autoencoders offered a complementary approach grounded in probabilistic inference, encoding designs into a continuous latent space from which new designs can be decoded through interpolation, a property that has been exploited extensively for smooth exploration of structural topology spaces, as demonstrated in the style-transfer inspired representation proposed by Guo, Lohan, Cang, Ren, and Allison (2018). More recently, denoising diffusion probabilistic models, formalized by Ho, Jain, and Abbeel (2020), have introduced an alternative generative mechanism based on the gradual corruption and reconstruction of data, yielding markedly improved sample diversity and training stability relative to adversarial approaches, a property that has begun to attract attention within structural design communities seeking to overcome the mode collapse and training instability long associated with generative adversarial architectures.

What distinguishes engineering applications of these architectures from their origins in image synthesis is the imposition of physical constraints. A generated image that is aesthetically

implausible carries no consequence beyond perceptual discomfort, whereas a generated structural topology that violates equilibrium, connectivity, or manufacturability constraints is functionally useless regardless of its statistical fidelity to the training distribution. This tension between statistical plausibility and physical validity constitutes perhaps the most persistent conceptual thread running through the engineering generative design literature, and it recurs, in different guises, across every subdomain considered in this review.

2.2 Generative Adversarial Networks in Structural Topology Optimization

The earliest and most extensively studied application of generative adversarial networks in engineering design concerns structural topology optimization, wherein the objective is to determine the optimal distribution of material within a design domain subject to loading and boundary conditions. Sharpe and Seepersad (2019) demonstrated that conditional generative adversarial networks could learn a mapping from boundary condition specifications directly to near-optimal topologies, effectively replacing the iterative finite element based solver with a single forward pass through a trained network. Building on this foundation, Oh, Jung, Kim, Lee, and Kang (2019) proposed a deep generative design framework that integrates topology optimization with generative modeling to produce diverse sets of high-performing candidate designs rather than a single deterministic optimum, addressing a longstanding limitation of gradient-based topology optimization, namely its tendency to converge to a single local optimum dictated by the initial condition.

Nie, Lin, Jiang, and Kara (2021) advanced this trajectory further with TopologyGAN, which conditions the generative process not merely on boundary specifications but on physical fields, including stress and strain energy distributions computed over the initial design domain, thereby embedding physics-informed priors directly into the generative pipeline and substantially improving the compliance performance of generated structures relative to purely data-driven baselines. This physics-informed turn reflects a broader pattern in the literature: as generative models matured beyond proof-of-concept demonstrations, researchers increasingly sought to counteract the risk of physically implausible outputs by injecting domain knowledge into either the network architecture, the loss function, or the conditioning variables. Rade et al. (2021) pursued a related strategy, proposing an algorithmically consistent deep learning framework that explicitly aligns the learned generative process with the mathematical structure of established topology optimization algorithms, thereby narrowing the gap between data-driven speed and algorithmic rigor. Abueidha, Koric, and Sobh (2020) extended generative and deep learning based topology optimization into the nonlinear regime, addressing a class of structural problems in which material or geometric nonlinearity renders classical linear elastic assumptions inadequate, a contribution that signals the field's gradual movement from idealized benchmark problems toward the messier nonlinear conditions characteristic of real engineering components.

A parallel line of inquiry has addressed periodic and metamaterial structures, where the design space is structured by symmetry and repetition rather than by arbitrary boundary geometry. Li, Jia, Cheng, and Shi (2021) coupled generative adversarial networks with Subset Simulation, a

rare-event sampling technique, to guide the topology optimization of periodic structures toward regions of the design space associated with extreme or targeted physical properties, illustrating how generative models can be embedded within broader probabilistic optimization frameworks rather than functioning as standalone design generators. Kazemi, Seepersad, and Kim (2022) extended the generative adversarial paradigm to multiphysics design optimization, in which thermal, structural, and other physical domains must be reconciled simultaneously, a problem class that magnifies the dimensionality and coupling challenges already present in single-physics topology optimization and that has historically resisted efficient gradient-based treatment.

2.3 Beyond Two-Dimensional Topology: Three-Dimensional Shape Synthesis and Multi-Objective Generation

While early generative design research concentrated on two-dimensional topology, characterized by pixelated density fields amenable to convolutional processing, subsequent scholarship has extended generative modeling to three-dimensional shape synthesis and to the multi-objective tensions that pervade realistic engineering practice. Shu et al. (2020) demonstrated a three-dimensional generative adversarial network coupled with physics-based validation for structural component design, addressing the additional representational complexity introduced by volumetric geometry while retaining a mechanism for verifying that generated shapes satisfy structural performance requirements rather than merely resembling plausible three-dimensional forms. This emphasis on validation as an integral rather than post hoc stage of the generative pipeline recurs throughout the more recent literature and reflects growing awareness that generative fluency, the capacity to produce visually or statistically convincing outputs, is analytically distinct from generative validity, the capacity to produce outputs that satisfy governing physical laws.

Multi-objective and constraint-aware generation constitutes a second major axis of extension. Chen and Ahmed (2021) introduced PaDGAN, a generative adversarial architecture explicitly engineered to balance the novelty of generated designs against their predicted quality, addressing a tension that had previously been under-theorized in the engineering generative design literature: purely novelty-seeking generative processes tend to produce designs that are statistically diverse but functionally inferior, whereas purely quality-seeking processes tend to collapse toward incremental variations of already known high-performing solutions. Nobari, Chen, and Ahmed (2021) extended this line of work with RANGE-GAN, a conditional architecture designed explicitly to synthesize designs under specified performance constraints, formalizing constraint satisfaction as a first-class objective within the generative training process rather than as a filtering step applied after generation. Regenwetter, Curry, and Ahmed (2022) contributed an important methodological resource to this literature through the BIKED dataset, which supplies a large, richly annotated corpus of bicycle designs paired with performance metrics, illustrating the increasing recognition that progress in engineering generative design is bottlenecked as much by the availability of high-quality, physics-annotated datasets as by architectural innovation in the generative models themselves.

2.4 Synthesizing Reviews and the Consolidation of the Field

As primary research accumulated, several synthesizing contributions sought to consolidate the field's fragmented findings into coherent conceptual structures. Regenwetter, Nobari, and Ahmed (2022) produced what remains among the most comprehensive reviews of deep generative models in engineering design, organizing the literature according to model architecture, namely generative adversarial networks, variational autoencoders, and reinforcement learning, while identifying the balance between design novelty, performance, and feasibility as the field's central unresolved challenge. Shin, Shin, and Kang (2023) offered a complementary review focused specifically on the intersection of machine learning and deep learning with topology optimization, tracing the evolution from early surrogate-model acceleration of finite element analysis toward fully generative topology synthesis and cataloguing the persistent difficulty of enforcing manufacturability constraints within learned representations. Song, Zhou, and Ahmed (2024) broadened the analytical lens further by reviewing multi-modal machine learning in engineering design, drawing attention to the increasing frequency with which design problems require the integration of heterogeneous data types, including geometric, textual, and performance data, within a single generative or predictive pipeline.

These synthesizing works converge on several recurring observations. First, the overwhelming majority of empirical studies remain confined to benchmark problems with well-defined, low-dimensional boundary conditions, raising questions about the transferability of reported performance gains to the higher-dimensional, multi-physics problems characteristic of industrial practice. Second, evaluation metrics across studies are inconsistent, with some researchers privileging compliance or structural performance, others privileging design diversity, and still others privileging computational speed relative to conventional optimization, a lack of standardization that complicates cross-study comparison and cumulative theory building. Third, and most consequentially for the present synthesis, nearly every reviewed contribution acknowledges, whether explicitly or implicitly, that generative models trained on historical design data risk perpetuating the biases and blind spots embedded in that data, a concern that becomes considerably more serious as generative tools move from academic demonstration toward deployment in safety-critical engineering contexts.

2.5 The Ascendancy of Large Language Models and Multimodal Design Agents

A distinct but increasingly convergent strand of scholarship concerns the application of large language models to engineering design tasks that were traditionally considered resistant to automation, including conceptual ideation, requirement elicitation, and design rationale articulation. Ma, Grandi, McComb, and Goucher-Lambert (2023) demonstrated that large language models can generate conceptual design candidates that are judged by human evaluators to be comparable in novelty and feasibility to those produced by trained designers, a finding that has stimulated considerable subsequent interest in language-mediated design ideation. Ege et al. (2025) subjected this claim to more rigorous scrutiny, systematically comparing large language model outputs against human-generated inventions across multiple evaluative criteria and finding

that while language models exhibit genuine strengths in fluency and breadth of ideation, they continue to lag behind experienced human designers in the pragmatic feasibility and technical specificity of proposed solutions, a nuance that tempers some of the more expansive claims circulating in earlier, less controlled studies.

Thoring, Huettemann, and Müller (2023) situated this development within a broader research agenda for what they term the augmented designer, arguing that generative AI is best conceptualized not as a substitute for human design cognition but as a cognitive prosthesis that reshapes the distribution of creative labor between human and machine, a framing that resonates strongly with the sociotechnical theoretical lens adopted in the present paper. Mustapha (2025) surveyed the rapidly expanding set of large language model applications across mechanics, product design, and manufacturing, documenting uses ranging from computer-aided design script generation to failure mode analysis, while cautioning that the reliability of language model outputs in technically precise engineering contexts remains considerably below the standard required for unsupervised deployment. Collectively, this literature signals an important structural shift: whereas generative adversarial networks and diffusion models operate principally on geometric and numerical representations of designs, large language models introduce a linguistic and semantic layer to generative design that is capable of mediating between human intention, expressed in natural language, and the more rigid computational representations required for downstream structural analysis.

2.6 Toward Integration and Manufacturing-Aware Generative Design

The special issue convened by Chen et al. (2026) is emblematic of a broader integrative turn within the field, one that seeks to draw together the previously disparate threads of structural generative design, materials systems modeling, and manufacturing process planning under a common generative AI umbrella. This integrative turn reflects a maturing recognition that structural performance, material behavior, and manufacturing feasibility cannot be optimized in isolation without risking the production of designs that are analytically elegant but practically unrealizable. Andreassen et al. (2011) and Kumar (2023) exemplify, in their respective compact and widely used topology optimization codebases, the field's longstanding commitment to reproducibility and accessibility, a commitment that the generative AI literature has inherited unevenly, with many high-profile generative studies relying on proprietary datasets and architectures that are difficult for independent researchers to reproduce or extend.

Taken as a whole, the literature reviewed here reveals a field that has progressed considerably in technical sophistication while leaving several conceptual and methodological gaps insufficiently addressed. The overwhelming majority of studies emphasize architectural novelty or benchmark performance gains, while comparatively few interrogate the epistemic and organizational conditions under which generative AI tools might be responsibly integrated into industrial design workflows governed by regulatory, safety, and liability constraints. Similarly, while the large language model literature and the structural generative design literature have each matured considerably, they have done so largely in parallel, with limited cross-pollination of evaluative

frameworks or shared benchmarking practices. The present paper contributes to closing this gap by developing an integrative analytical framework, presented in the following section, that maps generative model classes onto the engineering design lifecycle in a manner intended to be legible across these previously disconnected subfields, and by using that framework to identify the conceptual work still required before generative AI can be considered a mature and trustworthy component of smart engineering design optimization.

3. Methodology

3.1 Research Design

This study adopts a qualitative, theory-building research design structured as an integrative literature synthesis rather than a purely descriptive systematic review. The distinction is methodologically consequential. A conventional systematic review is oriented toward exhaustively cataloguing studies that satisfy narrowly defined inclusion criteria and toward quantifying aggregate trends, whereas an integrative synthesis, as articulated in the qualitative methods literature, is oriented toward reconciling and reconceptualizing findings drawn from methodologically and disciplinarily heterogeneous sources into a coherent conceptual structure. Given that the scholarship surrounding generative AI in engineering design spans structural mechanics, computer science, design cognition, and, increasingly, natural language processing, a purely descriptive cataloguing approach risked producing a fragmented inventory rather than the unified analytical account this paper seeks to offer. The integrative approach was therefore judged the more suitable methodological choice for a research problem defined primarily by conceptual fragmentation rather than by an absence of empirical evidence.

3.2 Source Identification and Screening Procedure

The literature underlying this synthesis was identified through a structured, though not exhaustively systematic, search of established engineering design and computational venues, including the Journal of Mechanical Design, the Journal of Computing and Information Science in Engineering, Structural and Multidisciplinary Optimization, Computers & Structures, Engineering Applications of Artificial Intelligence, and the proceedings of the International Design Engineering Technical Conferences, supplemented by targeted searches of preprint repositories to capture recent developments in large language model applications that have not yet completed the peer review cycle at the time of writing. Search terms combined the general descriptor generative artificial intelligence with more specific architectural and application terms, including generative adversarial network, variational autoencoder, diffusion model, topology optimization, and large language model, in conjunction with design-oriented qualifiers such as engineering design, structural optimization, and manufacturing.

Candidate sources were screened against three inclusion criteria. First, the source needed to report or synthesize a generative computational method applied to an engineering design or optimization task, thereby excluding purely image-domain generative AI research with no

articulated engineering application. Second, the source needed to be traceable to a verifiable peer-reviewed venue or a reputable preprint repository, ensuring that the synthesis rests on scrutinizable rather than anecdotal evidence. Third, priority was given to sources that either introduced a methodologically significant architecture or dataset, or that performed a synthesizing function across multiple primary studies, a criterion intended to privilege conceptual depth over exhaustive breadth given the integrative rather than exhaustive character of the review. Sources that satisfied these criteria were read in full, and analytic memoranda were compiled documenting the design lifecycle phase addressed, the generative architecture employed, the evaluative metrics reported, and the constraints or limitations acknowledged by the original authors.

3.3 Analytical Framework

The analytic memoranda generated during source review were organized according to a two-dimensional framework crossing generative model class, comprising generative adversarial networks, variational autoencoders, diffusion models, and large language model agents, against design lifecycle phase, comprising conceptual ideation, structural or topological synthesis, multi-objective and constraint-aware refinement, and manufacturing-aware validation. This framework was constructed inductively through iterative comparison across the reviewed sources rather than imposed a priori, consistent with the constant comparative logic characteristic of qualitative meta-synthesis. Each source was mapped to one or more cells of this framework, and patterns of density, absence, and internal tension within and across cells were interpreted as substantive findings regarding the maturity and coherence of the field. This procedure is not a statistical or bibliometric method; rather, it constitutes a disciplined qualitative coding exercise intended to render visible the conceptual architecture implicit within a body of otherwise disparately organized scholarship.

3.4 Ethical Considerations, Validity, and Reliability

Because this study synthesizes previously published scholarship rather than collecting original human subjects data, no direct ethical clearance was required; however, the study remains attentive to the ethical dimensions of the technologies under review, particularly concerning the provenance of training datasets, the attribution of design authorship in human-AI collaborative processes, and the risk that generative tools trained on historically dominant design paradigms may narrow rather than expand the diversity of engineering solutions considered viable. Analytical integrity was pursued through several complementary strategies. Credibility was addressed by cross-referencing claims made in synthesizing review articles, such as those of Regenwetter et al. (2022) and Shin et al. (2023), against the primary studies they cite, ensuring that this paper's characterization of the field is not merely a restatement of prior secondary interpretation. Dependability was addressed by maintaining explicit, auditable analytic memoranda for each source, permitting the reconstruction of the reasoning by which sources were classified within the design lifecycle by model class framework. Confirmability was

pursued by deliberately searching for disconfirming or critical evidence, particularly regarding claims of large language model design competence, where the more cautious findings of Ege et al. (2025) were weighed explicitly against the more optimistic claims advanced elsewhere in the literature. These procedures, while inherently interpretive rather than statistical, are consistent with established standards of rigor for qualitative and integrative research synthesis and provide a defensible basis for the analytical claims advanced in the sections that follow.

4. Findings

The integrative analysis reveals a field whose technical center of gravity has shifted progressively from narrow, single-objective structural generation toward increasingly multimodal, constraint-aware, and semantically mediated design synthesis. Table 1 summarizes the mapping between generative model class and design lifecycle phase developed through the analytical procedure described above, and illustrates both the areas of dense scholarly activity and the comparatively underdeveloped regions of the framework.

Design Lifecycle Phase	Dominant Model Class	Representative Contributions	Reported Emphasis
Conceptual ideation and requirement framing	Large language models; multimodal agents	Ma et al. (2023); Ege et al. (2025); Thoring et al. (2023)	Fluency and breadth of ideation; limited pragmatic feasibility
Structural topology synthesis	Generative adversarial networks	Sharpe & Seepersad (2019); Oh et al. (2019); Nie et al. (2021)	Compliance performance; speed relative to iterative solvers
Nonlinear and physics-informed synthesis	GANs with physics conditioning	Abueidda et al. (2020); Rade et al. (2021)	Alignment with governing physical equations
Periodic, metamaterial, and multiphysics design	GANs coupled with probabilistic sampling	Li et al. (2021); Kazemi et al. (2022)	Targeted property attainment; coupled-domain feasibility
Three-dimensional and multi-objective shape generation	GANs; variational autoencoders	Guo et al. (2018); Shu et al. (2020); Chen & Ahmed (2021); Nobari et al. (2021)	Balance of novelty, quality, and constraint satisfaction
Manufacturing-aware refinement and validation	Integrated generative-simulation pipelines	Chen et al. (2026); Song et al. (2024)	Manufacturability; cross-domain data integration

Table 1. Mapping of Generative AI Model Classes to Phases of the Engineering Design Lifecycle

Several patterns emerge with particular clarity from this mapping. The structural topology synthesis phase constitutes the most densely populated and methodologically mature region of

the framework, a maturity reflected in the progressive refinement visible across the contributions of Sharpe and Seepersad (2019), Oh et al. (2019), and Nie et al. (2021), each of which incrementally strengthened the physical grounding of generated outputs by embedding progressively richer physics-informed conditioning into the generative process. This trajectory suggests that the field's evaluative standards have hardened over time, moving away from an early emphasis on visual or statistical plausibility of generated topologies toward a more exacting emphasis on quantifiable structural performance benchmarked directly against classical optimization solutions.

By contrast, the conceptual ideation phase, while attracting considerable recent attention owing to the rapid ascendance of large language models, remains comparatively immature with respect to rigorous evaluation. The optimistic findings reported by Ma et al. (2023) regarding the comparability of language model generated concepts to human generated concepts stand in productive tension with the more circumspect findings of Ege et al. (2025), who documented systematic weaknesses in the technical feasibility of language model outputs when subjected to more demanding evaluative protocols. This tension is analytically significant because it suggests that the ideation phase of the framework, unlike the structural synthesis phase, has not yet converged on a stable set of evaluative criteria capable of adjudicating between competing empirical claims, leaving open the question of precisely what threshold of demonstrated competence would justify treating large language models as reliable contributors to early-stage engineering ideation.

A further finding concerns the comparatively underdeveloped state of the manufacturing-aware refinement phase relative to the earlier phases of the framework. Although Chen et al. (2026) and Song et al. (2024) signal a clear scholarly intention to integrate generative design more tightly with manufacturing process constraints and multi-modal data streams, the density of primary empirical contributions addressing this integration remains markedly lower than the density observed for structural topology synthesis, suggesting that manufacturability, despite being a longstanding concern within the design for manufacturing tradition, has yet to be as thoroughly operationalized within generative pipelines as structural performance has been. This asymmetry is consequential because a generated design that is structurally optimal but manufacturably impractical offers limited value to industrial practice, and the relative scarcity of contributions addressing this final validation stage represents one of the more consequential gaps surfaced by the present synthesis.

5. Discussion

The findings synthesized above bear directly on the research questions animating this paper. Regarding the mapping of generative model classes onto design lifecycle phases, the evidence indicates a field that has developed unevenly rather than uniformly, with structural topology synthesis representing a comparatively mature domain of practice governed by increasingly rigorous physics-informed evaluation, while conceptual ideation and manufacturing-aware

refinement remain considerably more contested and underdeveloped. This unevenness carries a theoretical implication that extends beyond the specific technologies reviewed here: it suggests that generative AI does not advance as a monolithic capability but rather diffuses unevenly across the design lifecycle according to the availability of well-structured training data and quantifiable performance criteria, conditions that are more readily satisfied in structural mechanics than in the more qualitative, judgment-laden phases of conceptual ideation and manufacturability assessment.

Regarding the recurring theoretical tensions structuring the field, the persistent friction between statistical novelty and physical or pragmatic validity, visible from the earliest generative adversarial network applications through to the most recent large language model studies, suggests that this tension is not an incidental artifact of any particular architecture but a structural feature of applying generative modeling to domains governed by hard physical or manufacturing constraints. This reframes the sociotechnical lens introduced earlier in the paper: the augmented designer framing proposed by Thoring et al. (2023) gains additional theoretical purchase when understood as a response to this structural tension, since human oversight functions precisely to adjudicate between statistically plausible and physically or pragmatically valid candidate designs in precisely those regions of the design lifecycle where generative models remain least reliable.

Regarding the gaps requiring attention before generative AI can be responsibly integrated into high-stakes engineering practice, the present synthesis suggests that interpretability, reproducibility, and manufacturability validation constitute the most pressing priorities. The comparative scarcity of open, reproducible datasets and benchmarks, a concern implicit in the accessibility-oriented contributions of Andreassen et al. (2011), Kumar (2023), and Regenwetter et al. (2022), suggests that the field's future credibility depends substantially on continued investment in shared infrastructure rather than solely on architectural innovation. These findings should be interpreted in light of certain limitations inherent to an integrative synthesis of this kind, including the reliance on published and therefore potentially positively biased results, and the rapid pace of development within large language model research, which risks rendering specific empirical claims dated even as the broader conceptual architecture proposed here retains explanatory value. Future research would benefit from longitudinal, industry-embedded studies that track generative AI tools across the full design lifecycle within a single organizational context, thereby directly testing the integrative claims advanced in this paper against the messier contingencies of applied practice, and from cross-disciplinary benchmarking efforts capable of reconciling the divergent evaluative traditions of the structural and language-oriented generative design communities.

6. Conclusion

This paper has argued that generative artificial intelligence, viewed across its principal architectural variants, is reshaping engineering design not as a uniform wave of automation but as an unevenly distributed set of capabilities whose maturity tracks closely with the availability

of well-structured data and quantifiable performance criteria at each phase of the design lifecycle. By synthesizing scholarship spanning structural topology optimization, three-dimensional shape generation, and emergent large language model agents within a single analytical framework, the paper has shown that the field's central and recurring challenge is the reconciliation of statistical generative fluency with physical and manufacturing validity, a challenge that is most fully resolved in structural synthesis and least resolved in conceptual ideation. The paper's principal scholarly contribution lies in offering a conceptual bridge across previously disconnected research traditions, thereby equipping researchers and practitioners with a shared vocabulary for situating new generative techniques within the broader engineering design lifecycle. Looking forward, the responsible maturation of this field will depend less on further architectural novelty than on sustained investment in reproducible benchmarks, manufacturability-aware validation, and a clear-eyed theoretical account of the human designer's enduring role as arbiter of physical and practical validity within an increasingly generative design paradigm.

References

- Abueidda, D. W., Koric, S., & Sobh, N. A. (2020). Topology optimization of 2D structures with nonlinearities using deep learning. *Computers & Structures*, 237, 106283. <https://doi.org/10.1016/j.compstruc.2020.106283>
- Andreassen, E., Clausen, A., Schevenels, M., Lazarov, B. S., & Sigmund, O. (2011). Efficient topology optimization in MATLAB using 88 lines of code. *Structural and Multidisciplinary Optimization*, 43(1), 1–16. <https://doi.org/10.1007/s00158-010-0594-7>
- Chen, W., & Ahmed, F. (2021). PaDGAN: Learning to generate high-quality novel designs. *Journal of Mechanical Design*, 143(3), Article 031703. <https://doi.org/10.1115/1.4048626>
- Chen, W. W., Krishnamurthy, V. R., Lu, Y., Luo, J., McComb, C., Ravi, S. K., & Sha, Z. (Eds.). (2026). Special issue on generative artificial intelligence for design, manufacturing processes, and materials systems: Part I. *Journal of Computing and Information Science in Engineering*, 26(7), Article 070301.
- Ege, D. N., Øvrebø, H. H., Stubberud, V., Berg, M. F., Elverum, C., Steinert, M., & Vestad, H. (2025). ChatGPT as an inventor: Eliciting the strengths and weaknesses of current large language models against humans in engineering design. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 39, Article e6. <https://doi.org/10.1017/S0890060424000260>
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems*, 27, 2672–2680.
- Guo, T., Lohan, D. J., Cang, R., Ren, M. Y., & Allison, J. T. (2018). An indirect design representation for topology optimization using variational autoencoder and style transfer. In *Proceedings of the AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference* (Article 0804). American Institute of Aeronautics and Astronautics. <https://doi.org/10.2514/6.2018-0804>
- Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33, 6840–6851.
- Kazemi, H., Seepersad, C. C., & Kim, H. A. (2022). Multiphysics design optimization via generative adversarial networks. *Journal of Mechanical Design*, 144(12), Article 121702. <https://doi.org/10.1115/1.4055204>
- Kumar, P. (2023). HoneyTop90: A 90-line MATLAB code for topology optimization using honeycomb tessellation. *Optimization and Engineering*, 24(2), 1433–1460. <https://doi.org/10.1007/s11081-022-09727-7>

- Li, M., Jia, G., Cheng, Z., & Shi, Z. (2021). Generative adversarial network guided topology optimization of periodic structures via Subset Simulation. *Composite Structures*, 260, Article 113254. <https://doi.org/10.1016/j.compstruct.2020.113254>
- Ma, K., Grandi, D., McComb, C., & Goucher-Lambert, K. (2023). Conceptual design generation using large language models. In *Proceedings of the ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (Vol. 87349, Article V006T06A021). American Society of Mechanical Engineers. <https://doi.org/10.1115/DETC2023-116985>
- Mustapha, K. B. (2025). A survey of emerging applications of large language models for problems in mechanics, product design, and manufacturing. *Advanced Engineering Informatics*, 64, Article 103066. <https://doi.org/10.1016/j.aei.2024.103066>
1. Bandaru, S. R., Shyam, D., Manoharan, P., Syed, M. H., Ragireddy, U. K., & Addepalli, P. V. (2022). A Review of Frameworks, and Impact on the Software Development Life Cycle with Artificial Intelligence in Software Engineering. *American International Journal of Computer Science and Technology*, 4(1), 44-53.
 2. Manoharan, P., Syed, M. H., Ragireddy, U. K., Addepalli, P. V., Bandaru, S. R., & Shyam, D. (2022). Understanding the Use of Agile Development Practices Alongside Quantum Computing in Modern Engineering Projects. *International Journal of Emerging Research in Engineering and Technology*, 3(1), 136-143.
 3. Routhu, K. K. (2023). AI-driven succession planning in Oracle HCM Cloud: Building resilient leadership pipelines through predictive analytics. *International Journal of Science, Engineering and Technology*, 11(5).
 4. Ragireddy, U. K., Addepalli, P. V., Bandaru, S. R., Shyam, D., Manoharan, P., & Syed, M. H. (2022). Intelligent Predictive System of Cloud Resource Utilization Forecasting with Advanced Deep Learning. *International Journal of Artificial Intelligence, Data Science, and Machine Learning*, 3(1), 143-151.
 5. Syed, M. H., Ragireddy, U. K., Addepalli, P. V., Bandaru, S. R., Shyam, D., & Manoharan, P. (2022). A Survey of Statistical Process Control and Its Applications in Manufacturing Quality Assurance. *International Journal of Emerging Trends in Computer Science and Information Technology*, 3(1), 127-135.
 6. Padur, S. K. R. (2023). AI-Augmented Enterprise ERP Modernization: Zero-Downtime Strategies for Oracle E-Business Suite R12. 2 and Beyond. Available at SSRN 5605510.
 7. Routhu, K. K. (2019). Hybrid machine learning architecture for absence forecasting within Oracle Cloud HCM. *KOS Journal of AIML, Data Science, and Robotics*, 1(1), 1-5.
 8. Padur, S. K. R. (2022). AI augmented platform engineering, transforming developer experience through intelligent automation and self optimizing internal platforms. *International Journal of Science, Engineering and Technology*, 10(5), 10-5281.
 9. Addepalli, P. V., Bandaru, S. R., Shyam, D., Manoharan, P., Syed, M. H., & Ragireddy, U. K. (2022). Survey on Blockchain Integration with Sap S/4hana for Transparent and Secure Supply Chains. *International Journal of AI, BigData, Computational and Management Studies*, 3(1), 161-170.
 10. Routhu, K. K. (2023). Embedding fairness into the digital enterprise: Data-driven dei strategies with oracle hcm analytics. *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.*, 9(8), 266-274.
 11. Mylavarapu, K. B., Pilli, J., Jannu, P. K., Mohammad, J. A., & Panchali, S. H. (2022). Comparative Performance of Machine Learning Techniques for Sentiment and Insight Analysis in Amazon Product Reviews. *American International Journal of Computer Science and Technology*, 4(1), 54-64.
 12. Mohammad, J. A., Panchali, S. H., Mylavarapu, K. B., Pilli, J., & Jannu, P. K. (2022). An Ensemble Machine Learning Approach for Reliable Fake News Detection on Social Media Platforms. *International Journal of Artificial Intelligence, Data Science, and Machine Learning*, 3(1), 152-161.
 13. Pilli, J., Jannu, P. K., Mohammad, J. A., Panchali, S. H., & Mylavarapu, K. B. (2022). Review of Supply Chain-Integrated CRM Systems for Business Agility. *International Journal of Emerging Research in Engineering and Technology*, 3(1), 155-163.
 14. Padur, S. K. R. (2022). Intelligent resource management: AI methods for predictive workload forecasting in cloud data centers. *J. Artif. Intell. Mach. Learn. & Data Sci.*, 1(1), 2936-2941.
 15. Panchali, S. H., Mylavarapu, K. B., Pilli, J., Jannu, P. K., & Mohammad, J. A. (2022). Predictive Maintenance in Smart Manufacturing: A Review of Machine Learning and Digital Twin Approaches. *International Journal of AI, BigData, Computational and Management Studies*, 3(1), 150-160.
 16. Jannu, P. K., Mohammad, J. A., Panchali, S. H., Mylavarapu, K. B., & Pilli, J. (2022). Heat Dissipation Strategies for High-Performance Computing: A Review of Air and Liquid Cooling Efficiency. *International Journal of Emerging Trends in Computer Science and Information Technology*, 3(1), 145-153.
 17. Routhu, K. K. (2021). AI-augmented benefits administration: A standards-driven automation framework with Oracle HCM Cloud. *International Journal of Scientific Research and Engineering Trends*, 7(3).
 18. Taluri, R. (2022). Cloud Data Engineering Strategies for Large-Scale Financial Data Integration and Intelligent Corporate Performance Reporting. *International Journal of Emerging Research in Engineering and Technology*, 3(4), 176-188. <https://doi.org/10.63282/3050-922X.IJERET-V3I4P119>

19. Nagumotu, V. T., Kavuluri, H. V. R., Pathani, A. K., Dasari, A., Chilakapati, V. K., & Keshireddy, S. R. (2022). Machine Learning-Driven Threat Analysis for Enhancing Security in Data Center Networking. *American International Journal of Computer Science and Technology*, 4(1), 65-76.
20. Padur, S. K. R. (2021). Bridging Human, System, and Cloud Integration through RESTful Automation and Governance. *the International Journal of Science, Engineering and Technology*, 9(6).
21. Keshireddy, S. R., Nagumotu, V. T., Kavuluri, H. V. R., Pathani, A. K., Dasari, A., & Chilakapati, V. K. (2022). Towards Secure and Reliable IoT: A Review on Anomaly Detection and Predictive Analytics in Sensor Networks. *International Journal of AI, BigData, Computational and Management Studies*, 3(1), 140-149.
22. Padur, S. K. R. (2021). Deep learning and process mining for ERP anomaly detection: Toward predictive and self-monitoring enterprise platforms. *Available at SSRN 5605531*.
23. Routhu, K. K. (2020). Strategic compensation equity and rewards optimization: A multi-cloud analytics blueprint with Oracle Analytics Cloud. *Available at SSRN 5737266*.
24. Pathani, A. K., Dasari, A., Chilakapati, V. K., Keshireddy, S. R., Nagumotu, V. T., & Kavuluri, H. V. R. (2022). Effective Monitoring of Sensor-Actuator Integration Methods for Real-Time Controlling in Smart Device Networks. *International Journal of Emerging Research in Engineering and Technology*, 3(1), 144-154.
25. Padur, S. K. R. (2021). From Control to Code: Governance Models for Multi-Cloud ERP Modernization. *International Journal of Scientific Research & Engineering Trends*, 7(3).
26. Routhu, K. K. (2019). AI-enhanced payroll optimization: Improving accuracy and compliance in Oracle HCM. *KOS Journal of AIML, Data Science, and Robotics*, 1(1), 1-5.
27. Padur, S. K. R. (2021). From Scripts to Platforms-as-Code: The Role of Terraform and Ansible in Declarative Infrastructure Rollouts. *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol*, 7(1), 621-628.
28. Chilakapati, V. K., Keshireddy, S. R., Nagumotu, V. T., Kavuluri, H. V. R., Pathani, A. K., & Dasari, A. (2022). Review on Network Virtualization and SDN in Industrial IoT (IIoT). *International Journal of Artificial Intelligence, Data Science, and Machine Learning*, 3(1), 162-172.
29. Dasari, S., Chilakapati, V. K., Keshireddy, S. R., Nagumotu, V. T., Kavuluri, H. V. R., & Pathani, A. K. (2022). Networking Paradigms for Large-Scale IoT Deployments: A Survey on Reliability, Latency, and Scalability. *International Journal of Emerging Trends in Computer Science and Information Technology*, 3(1), 136-144.
30. Padur, S. K. R. (2020). AI-augmented disaster recovery simulations: From chaos engineering to autonomous resilience orchestration. *Int J Sci Res Sci Eng Technol*, 7(6), 367-378.
31. Taluri, R. (2021). Cloud-Native Architectures for Enterprise Financial Data Management, Analytics, and Regulatory Reporting Compliance. *International Journal of Emerging Trends in Computer Science and Information Technology*, 2(2), 101-111. <https://doi.org/10.63282/3050-9246.IJETCSIT-V2I2P112>
32. Bandaru, S. R., Shyam, D., Manoharan, P., Syed, M. H., Ragireddy, U. K., & Addepalli, P. V. (2023). Data-Driven Approaches for Forecasting Cost Overruns in Infrastructure Projects. *European Journal of Technology*, 7(5), 23-34.
33. Dhuli, S., Prabhu, M., Muzaffer, H. S., Uday Kumar, R., & Prasanth Varma, A. (2023). A Survey on Critical Success Factors Influencing Project Success Across Multiple Industries. *J Contemp Edu Theo Artific Intel: JCETAI-106*.
34. Ragireddy, U. K., Addepalli, P. V., Bandaru, S. R., Shyam, D., Manoharan, P., & Syed, M. H. (2023). An Overview of Different Automation Strategies Used in Cloud-Based CI/CD Pipelines for Software Deployment. *International Journal of Technology, Management and Humanities*, 9(01), 116-127.
35. Addepalli, P. V., Bandaru, S. R., Shyam, D., Manoharan, P., Syed, M. H., & Ragireddy, U. K. (2023). Performance Analysis of Machine Learning Models for Cybersecurity Risk Classification in IT Projects. *Universal Library of Engineering Technology*, (Issue).
36. Mylavarapu, K. B., Pilli, J., Jannu, P. K., Mohammad, J. A., & Panchali, S. H. (2023). Artificial Intelligence for Continuous Quality Assurance in Cloud-Native Applications: A Review. *European Journal of Technology*, 7(5), 35-49.
37. Routhu, K. K. (2017). The evolution of HR from on-premise to Oracle Cloud HCM: Challenges and opportunities. *International Journal of Scientific Research & Engineering Trends*, 3(1).
38. Padur, S. K. R. (2020). From centralized control to democratized insights: Migrating enterprise reporting from IBM Cognos to Microsoft Power BI. *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol*, 6(1), 218-225.
39. Jannu, P. K., Mohammad, J. A., Panchali, S. H., Mylavarapu, K. B., & Pilli, J. (2023). A Survey on Energy-Efficient Microarchitecture Design Strategies for Mobile CPUs and GPUs: Techniques and Future Directions. *International Journal of Technology, Management and Humanities*, 9(04), 137-151.
40. Keshireddy, S. R., Nagumotu, V. T., Kavuluri, H. V. R., Pathani, A. K., Dasari, A., & Chilakapati, V. K. (2023). An AI-Based Framework for Detecting IoT Botnets Through Network Traffic Analysis and Modeling. *Universal Library of Engineering Technology*, (Issue).
41. Kavuluri, H. V. R., Pathani, A. K., Dasari, A., Chilakapati, V. K., Keshireddy, S. R., & Nagumotu, V. T. (2023). Cloud Data-Center Network Verification: Approaches, Algorithms and Toolchains. *American Journal of Computing and Engineering*, 6(2), 18-32.
42. Routhu, K. K. (2022). From RFID to Geofencing: IoT-Enabled Smart Time Tracking in Oracle HCM Cloud. *International Journal of Science, Engineering and Technology*, 10(4).
43. Chilakapati, V. K., Keshireddy, S. R., Nagumotu, V. T., Kavuluri, H. V. R., Pathani, A. K., & Dasari, A. (2023). A Survey on 5G/6G Radio Access Networks: Technologies, Protocols, and Performance Metrics. *International Journal of Technology, Management and Humanities*, 9(02), 79-93.
44. Keshireddy, S. R., Nagumotu, V. T., Kavuluri, H. V. R., Pathani, A. K., Dasari, A., & Chilakapati, V. K. (2023). An AI-Based Framework for Detecting IoT Botnets Through Network Traffic Analysis and Modeling. *Universal Library of Engineering Technology*, (Issue).
45. Routhu, K. K. (2022). From Case Management to Conversational HR: Redefining Help Desks with Oracle's AI and NLP Framework. *International Journal of Science, Engineering and Technology*, 10(6).
46. Padur, S. K. R. (2019). Machine learning for predictive capacity planning: Evolution from analytical modeling to autonomous infrastructure. *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol*, 5(5), 285-293.

- Nie, Z., Lin, T., Jiang, H., & Kara, L. B. (2021). TopologyGAN: Topology optimization using generative adversarial networks based on physical fields over the initial domain. *Journal of Mechanical Design*, 143(3), Article 031715. <https://doi.org/10.1115/1.4049533>
- Nobari, A. H., Chen, W., & Ahmed, F. (2021). RANGE-GAN: Design synthesis under constraints using conditional generative adversarial networks. *Journal of Mechanical Design*, 144(2), Article 021708. <https://doi.org/10.1115/1.4052442>
- Oh, S., Jung, Y., Kim, S., Lee, I., & Kang, N. (2019). Deep generative design: Integration of topology optimization and generative models. *Journal of Mechanical Design*, 141(11), Article 111405. <https://doi.org/10.1115/1.4044229>
- Rade, J., Balu, A., Herron, E., Pathak, J., Ranade, R., Sarkar, S., & Krishnamurthy, A. (2021). Algorithmically-consistent deep learning frameworks for structural topology optimization. *Engineering Applications of Artificial Intelligence*, 106, Article 104483. <https://doi.org/10.1016/j.engappai.2021.104483>
- Regenwetter, L., Curry, B., & Ahmed, F. (2022). BIKED: A dataset for computational bicycle design with machine learning benchmarks. *Journal of Mechanical Design*, 144(3), Article 031706. <https://doi.org/10.1115/1.4052585>
- Regenwetter, L., Nobari, A. H., & Ahmed, F. (2022). Deep generative models in engineering design: A review. *Journal of Mechanical Design*, 144(7), Article 071704. <https://doi.org/10.1115/1.4053859>
- Sharpe, C., & Seepersad, C. C. (2019). Topology design with conditional generative adversarial networks. In *Proceedings of the ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (Vol. 59186, Article V02AT03A062). American Society of Mechanical Engineers. <https://doi.org/10.1115/DETC2019-97833>
- Shin, S., Shin, D., & Kang, N. (2023). Topology optimization via machine learning and deep learning: A review. *Journal of Computational Design and Engineering*, 10(4), 1736–1766. <https://doi.org/10.1093/jcde/qwad072>
- Shu, D., Cunningham, J., Stump, G., Miller, S. W., Yukish, M. A., Simpson, T. W., & Tucker, C. S. (2020). 3D design using generative adversarial networks and physics-based validation. *Journal of Mechanical Design*, 142(7), Article 071701. <https://doi.org/10.1115/1.4045419>
- Song, B., Zhou, R., & Ahmed, F. (2024). Multi-modal machine learning in engineering design: A review and future directions. *Journal of Computing and Information Science in Engineering*, 24(1), Article 010801. <https://doi.org/10.1115/1.4063954>
- Thoring, K., Huettemann, S., & Müller, R. M. (2023). The augmented designer: A research agenda for generative AI-enabled design. *Proceedings of the Design Society*, 3, 3345–3354. <https://doi.org/10.1017/pds.2023.335>

Appendix A. Process Figure

Figure 1 depicts the generalized computational pipeline synthesized from the reviewed literature, illustrating the iterative relationship between generative synthesis and physics-based or manufacturing-aware evaluation described throughout the paper.

