

Review

Not peer-reviewed version

Large Language Models in Mechanical Engineering: A Scoping Review of Applications, Challenges, and Future Directions

[Christopher Baker](#)*, [Karen Rafferty](#), [Mark Price](#)

Posted Date: 27 August 2025

doi: 10.20944/preprints202508.1938.v1

Keywords: large language models; mechanical engineering; computer-aided design; multimodal artificial intelligence; generative design



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Review

Large Language Models in Mechanical Engineering: A Scoping Review of Applications, Challenges, and Future Directions

Christopher Baker *, Karen Rafferty and Mark Price

School of Electronics, Electrical Engineering and Computer Science, School of Mechanical and Aerospace Engineering, Queen's University Belfast

* Correspondence: c.baker@qub.ac.uk

Abstract

Following PRISMA-ScR guidelines, this scoping review systematically maps the landscape of Large Language Models (LLMs) in mechanical engineering. A search of four major databases (Scopus, IEEE Xplore, ACM Digital Library, Web of Science) yielded 66 studies for analysis. The findings reveal a nascent, rapidly accelerating field, with over 68% of publications from 2024, and applications concentrated on front-end design processes like conceptual design and Computer-Aided Design (CAD) generation. The technological landscape is dominated by OpenAI's GPT-4 variants. A persistent challenge identified is weak spatial and geometric reasoning, shifting the primary research bottleneck from traditional data scarcity to inherent model limitations. This, alongside reliability concerns, forms the main barrier to deeper integration into engineering workflows. A consensus on future directions points to the need for specialized datasets, multimodal inputs to ground models in engineering realities, and robust, engineering-specific benchmarks. This review concludes that LLMs are currently best positioned as powerful 'co-pilots' for engineers rather than autonomous designers, providing an evidence-based roadmap for researchers, practitioners, and educators.

Keywords: large language models; mechanical engineering; computer-aided design; multimodal artificial intelligence; generative design

1. Introduction

1.1. Background

Large Language Models (LLMs) are revolutionizing artificial intelligence with advanced capabilities in language understanding, reasoning, and knowledge synthesis. Since their widespread commercial introduction in late 2022, these models have rapidly evolved from general-purpose tools to specialised applications across various technical domains driven by approximate human level performance [1]. This rapid evolution can primarily be attributed to transformers [2] the availability of specific training data [3–6] and increased computational capacity [7–9] leading to the proliferation of LLM services for consumers worldwide. Their ability to understand and generate human-like text, combined with recent advances in handling technical content and spatial relationships, positions them as potentially powerful tools for engineering applications [10], (p. 202), [11,12].

The current landscape of computational mechanical engineering is dominated by sophisticated digital tools that have fundamentally transformed how engineers approach design and analysis. Computer-Aided Design (CAD) systems, Finite Element Analysis (FEA) software, and advanced simulation platforms form the basis of modern engineering workflows. However, these tools often present significant learning curves, require extensive manual input, and can create bottlenecks in the design process [13,14]. Due to their complexity and the specialised expertise required for effective utilisation [15] there is growing interest in utilising Large Language Models (LLM) in the design and analysis of engineering projects. LLMs and related multi-modal models offer potential solutions

through their ability to input and output natural language, facilitating integration into the natural flow of discourse that characterises engineering design processes. Their world knowledge is valuable both for interacting with humans—establishing common ground—and for reducing solution spaces to generate creative solutions. Preliminary evidence suggests that LLMs possess rich representations of the world despite being trained on simple objectives [16]. While humans are constrained by their knowledge when designing by analogy or biomimicry, LLMs can accumulate wide-ranging knowledge during training [17–20]. Models of sufficient size display strong capabilities in various reasoning tasks including step-by-step reasoning [21]. Furthermore, multi-modal LLM models (MMLM) can operate on various forms of design representations (text, tables, sketches, and 3D models), addressing the diverse representational needs throughout the product development process [22–24].

1.1.1. Multi-Modal Applications

The integration of multi-modal Large Language Models (MMLMs) has catalysed significant advancements in Computer-Aided Design (CAD) systems, which remain fundamental to industrial prototyping processes requiring feature-based modeling [25] and part editing [26]. The principal advantage of multi-modal approaches lies in their enhanced spatial understanding capabilities. Makatura et al. [10] demonstrate how these models translate natural language specifications into precise technical representations, enabling designers to develop complex components such as aerodynamic structures or functional mechanisms through intuitive textual and visual prompts. This capability minimises the iterative cycles inherent in traditional CAD workflows, substantially reducing design development timelines.

Despite their promise, contemporary MMLMs encounter considerable challenges in spatial reasoning tasks. Research on CAD-GPT [25] has identified persistent difficulties in accurately inferring three-dimensional spatial positions and orientations, resulting in geometric construction errors. These limitations manifest in practical design flaws—horizontal placement of wheels on vehicle models or incorrectly positioned table legs—underscoring the necessity for enhanced spatial reasoning mechanisms. Recent innovations such as CAD-MLLM address these challenges by offering unified frameworks for generating parametric CAD models from diverse input modalities. This system leverages command sequences from CAD models while employing sophisticated language models to align feature spaces across different data representations [18]. By creating a comprehensive translation layer between various input conditions and CAD construction sequences, these systems democratise access to complex design tools for both expert and non-expert users. The practical applications of these technological developments extend beyond basic design creation to complex engineering tasks including component integration, simulation-based optimisation, and collaborative design processes [26].

Evaluative frameworks for these systems have evolved beyond simple geometric assessment to encompass structural integrity and manufacturability constraints. Researchers have developed specialised metrics that analyse not only geometric accuracy through Chamfer Distance and F-scores but also structural integrity via topology quality measures such as Segment Error, Dangling Edge Length, and Self-Intersection Ratio [18]. These comprehensive evaluation methods ensure that generated CAD models satisfy both visual accuracy requirements and fundamental manufacturability constraints and represent an advancement over prior point cloud metrics [27,28]. Recent research demonstrates increasing robustness in multi-modal systems when handling imperfect input conditions, such as noisy or partial data. For instance, CAD-MLLM has shown remarkable resilience when processing point clouds with up to 95% of points removed or containing significant noise, vastly outperforming previous approaches. This robustness signifies the growing maturity of the field and its readiness for deployment in real-world engineering contexts where ideal input conditions rarely exist.

In summary, multi-modal applications represent a transformative paradigm in mechanical engineering design processes. By unifying diverse input modalities, enhancing spatial reasoning

capabilities, and facilitating intuitive human-machine interaction, these systems are poised to fundamentally alter traditional CAD workflows. Their continued development promises to democratise complex design processes while simultaneously increasing design efficiency and expanding the creative possibilities available to engineers across disciplines.

1.1.2. Data Challenges in LLM Enhanced Mechanical Engineering

In contrast to conventional large-scale datasets, mechanical engineering data presents distinct challenges for LLMs due to its inherent characteristics: limited sample sizes, heterogeneous sources, and intricate interdependencies [29]. This engineering data fundamentally differs from text-rich web corpora used in general LLM training, integrating specialised technical terminology, numerical parameters, dimensional specifications, and multidisciplinary knowledge spanning materials science, physics, and manufacturing processes. Consequently, several key challenges emerge for LLM application in this field. These include pronounced data scarcity, characterised by fewer documented design instances and limited exemplars of specialised components; significant multimodal complexity, necessitating the coherent integration of textual descriptions, engineering drawings, geometric CAD models, and numerical simulation outputs; and domain-specificity barriers arising from specialised lexicon and foundational mathematical principles that generalist LLMs often struggle to interpret accurately. Strategies proposed to address these limitations often involve co-learning and transfer learning methodologies, adapting models from data-abundant domains to the data-scarce engineering context [30–32]. These techniques facilitate knowledge transfer, enabling model inference even when faced with incomplete multimodal inputs characteristic of specific engineering tasks.

Issues pertaining to data quality and accessibility further compound the challenges of applying LLMs within mechanical engineering. Engineering knowledge frequently resides in tacit forms, embedded within CAD models, process sheets, and manufacturing documents that often adhere to inconsistent standards or encapsulate implicit information reflecting design rationale [33,35], typically comprehensible only to domain specialists [34]. This inherent opacity exacerbates the issue of data scarcity, as intricate engineering designs are commonly subject to stringent disclosure restrictions, preventing experts from divulging complete knowledge or data pertaining to specific project solutions, which may be archived in non-standardised or difficult-to-access formats [35]. Furthermore, real-world engineering data is characteristically siloed across diverse systems—including CAD platforms, Enterprise Resource Planning (ERP) systems, maintenance logs, and simulation software—engendering significant integration obstacles that hinder the creation of comprehensive datasets requisite for effective LLM training in this domain. These domain-specific data challenges parallel earlier difficulties encountered in scaling the training corpora for general-purpose LLMs [36], where data volume rapidly surpassed the capacity for human verification of factual accuracy. Consequently, multimodal LLMs (MMLMs) tailored for mechanical engineering grapple with a constrained supply of training data, insufficient to achieve the scale typical of generalist models. Although available datasets may be sparse, their validation is further complicated by data fragmentation across silos, often residing within proprietary systems unsuitable for open sourcing. Any feasible validation process is inherently demanding due to the requisite expert knowledge and the intricate, often bespoke nature of engineering solutions under scrutiny.

Despite these significant impediments, the transformative potential of LLMs continues to motivate intensive research efforts aimed at overcoming these data-centric barriers. Particular emphasis is now being placed on advancing multimodal large language models (MMLMs), as these architectures are inherently better suited to processing the heterogeneous data landscape of mechanical engineering. MMLMs offer the capability to integrate and reason across diverse information streams central to design and analysis, such as textual requirements, schematic drawings, geometric CAD representations, and numerical simulation results. Current research frontiers actively explore targeted mitigation strategies, including domain-specific fine-tuning on curated engineering datasets, the development of more sophisticated multimodal training paradigms

capable of capturing cross-modal relationships, the generation of high-quality synthetic data to augment limited real-world examples, and initiatives promoting standardised data formats to enhance interoperability. The subsequent section delves into practical implementations, presenting specific examples that illustrate the evolving application of these multimodal approaches in tackling complex engineering design problems, showcasing both the progress made and the ongoing development required to fully realise their potential.

1.2. Practical Implementation Examples

An early demonstration of LLM potential in engineering involved leveraging powerful, general-purpose models like GPT-4 primarily for their code-generation capabilities. In the work by Makatura et al. [10], the workflow involved users providing natural language prompts describing a desired object or component (e.g., "a sturdy bracket to hold a pipe"). The LLM then processed this prompt and generated corresponding code in a programmatic CAD language, specifically OpenJSCAD. This generated script, when executed, produced the 3D geometry of the described object. The significance lay in using an intuitive text interface to drive a CAX workflow; the resulting 3D models could then be directly exported for downstream applications like 3D printing. This approach effectively treated the LLM as an intelligent natural language-to-code translator, relying on an intermediary scripting language rather than directly manipulating native CAD features or structures, showcasing feasibility but highlighting the need for deeper domain understanding.

1.2.1. CAD-GPT: Synthesising CAD Construction Sequences with Enhanced Spatial Reasoning (Based on Wang et al. [25])

To address limitations in geometric accuracy and spatial understanding observed when directly applying general MLLMs, specialised frameworks like CAD-GPT were developed [25]. This system focuses on generating not just the final shape, but the sequence of construction steps typically used in CAD software (e.g., creating a sketch on a plane, extruding it, adding a fillet). The model takes multimodal inputs, such as textual descriptions combined with reference images or sketches. Internally, CAD-GPT incorporates architectural enhancements specifically designed to improve spatial reasoning – understanding relationships like orientation, alignment, and connectivity in 3D space, which are critical for avoiding common geometric errors (e.g., floating components, incorrect feature placement). The output is a structured sequence of CAD operations, aiming for a more accurate and potentially editable representation compared to generating raw geometry, signifying a move towards MLLMs with better intrinsic understanding of CAD procedures.

1.2.2. CAD-MLLM: Unifying Multimodality-Conditioned Parametric CAD Generation (Xu et al. [18])

Pushing towards greater flexibility and robustness, the CAD-MLLM framework [18] aims to create a unified system for generating parametric CAD models from a wide array of input types. Users can provide textual descriptions, rasterized sketches (images), or even 3D point clouds (potentially incomplete or noisy). The core mechanism involves using advanced language modeling techniques to map these diverse inputs into a common feature space. This aligned representation is then translated into sequences of commands compatible with standard CAD software kernels, crucially aiming to produce parametric models where design intent (e.g., relationships between features, driving dimensions) is preserved. A key demonstrated capability was its resilience; the system could generate plausible CAD models even when processing point clouds with significant noise or missing data. This approach represents a significant step towards MLLMs that can handle the heterogeneous and often imperfect data landscape of engineering while producing more editable and robust design outputs suitable for complex development cycles.

These examples collectively illustrate a clear trajectory: from leveraging general LLM coding abilities, to developing specialised MLLMs focusing on CAD sequence generation and spatial reasoning, further advancing to unified frameworks for robust parametric modeling from diverse

inputs, and extending towards interactive systems that couple generative techniques with evolutionary optimisation tailored to specific, user-defined simulation environments.

1.3. Rationale and Objectives

The swift proliferation and advancing capabilities of Large Language Models (LLMs) present transformative opportunities across numerous scientific and technical domains, including mechanical engineering. However, despite growing interest and preliminary applications, the specific ways in which LLMs are being utilised, the inherent limitations encountered, and the true scope of their potential within diverse mechanical engineering workflows remain largely underexplored and unsystematic in the current literature. A comprehensive understanding of this rapidly evolving landscape is currently lacking. Therefore, this scoping review is undertaken to address this knowledge gap. The primary objectives of this review are fourfold: (1) to systematically map the current landscape of LLM applications within the field of mechanical engineering; (2) to critically identify the key challenges, limitations, and opportunities associated with their practical implementation; (3) to provide insights that help guide future research and development priorities in this emergent area; and (4) ultimately, to inform the development of best practices for integrating LLMs effectively and responsibly into established engineering design, analysis, and manufacturing processes.

While prior reviews have explored the broader application of machine learning within additive manufacturing and design [12,37], these have not focused on the specific and rapidly growing impact of Large Language Models (LLMs). This review uniquely addresses that gap by systematically examining how LLMs, especially multimodal models, are being applied across the diverse landscape of mechanical engineering.

2. Methods

The methodology of this scoping review was designed to be systematic, transparent, and reproducible, adhering to the principles outlined in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) checklist [38]. The process involved a multi-stage approach encompassing a focused search strategy, a systematic screening of sources, and a detailed data charting protocol.

2.1. Research Questions

This review was guided by the following research questions:

1. What are the current applications of LLMs in mechanical engineering?
2. What are the key challenges and limitations in implementing LLMs in this domain?
3. How do LLMs complement or replace traditional computational approaches in mechanical engineering?
4. What are the emerging trends and future directions for research at the intersection of LLMs and mechanical engineering?

2.2. Search Strategy

To identify relevant literature, a systematic search was conducted in [Month, Year] across four major peer-reviewed academic databases: Scopus, IEEE Xplore, the ACM Digital Library, and Web of Science. These databases were chosen for their comprehensive coverage of engineering, computer science, and interdisciplinary research.

The search query was designed to be highly focused to precisely identify studies at the intersection of three core concepts: (A) Large Language Model technologies, (B) the mechanical engineering domain, and (C) specific engineering applications and tools. The search query, adapted for the syntax of each database, is detailed in Table 1. The search was applied to the titles, abstracts, and keywords of records published between January 2020 and the date of the search. An initial

exploratory search of the arXiv preprint server was also performed; however, due to significant technical challenges with the platform's export functionality that prevented a systematic and reproducible export of the results, a pragmatic decision was made to exclude this source from the final analysis to ensure methodological consistency.

Table 1. Overview of Search Terms, Organized by Concept.

Concept	Search Terms
A: LLM Technologies	"large language model" OR "LLM" OR "generative AI" OR "foundation model"
B: Engineering Domain	"mechanical engineering" OR "engineering design" OR "product design"
C: Specific Applications & Tools	"computer-aided design" OR "CAD" OR "generative design" OR "finite element" OR "FEA" OR "simulation" OR "CAx"

2.3. Source Selection and Screening

The source selection process followed a systematic, multi-stage screening protocol. The initial search across the four specified databases yielded 587 records. These were consolidated using Zotero (v6.0), and after the removal of duplicates, 122 unique articles remained for screening.

The initial screening of titles and abstracts was conducted using the active learning tool ASReview (v1.2.2) to ensure a systematic and efficient review. The model was seeded with a set of known relevant studies to establish prior knowledge, then proceeded iteratively, prioritizing records most likely to be relevant based on the author's classifications. A stopping rule was set to conclude the screening after 50 consecutive records were classified as 'Irrelevant'. Based on this initial review, all 122 unique studies were deemed potentially relevant and were advanced to the full-text analysis stage.

The second stage involved a full-text review of all 122 articles to determine final eligibility against a pre-defined set of criteria. Studies were included or excluded based on the following criteria:

Inclusion Criteria:

1. The study describes a direct application, framework, or analysis of an LLM within a mechanical engineering context (design, analysis, manufacturing, knowledge management).
2. The publication is a peer-reviewed journal article or conference paper.
3. The study was published between January 2020 and the date of the search.
4. The article is written in English.

Exclusion Criteria:

1. Studies where LLMs are only mentioned in passing (e.g., in future work sections).
2. Editorials, opinion pieces, or non-technical articles.
3. Review papers not containing original data or applications.
4. Studies not available in full-text.

The application of these criteria during the full-text review resulted in the final dataset for this scoping review, consisting of 66 included studies and 56 excluded studies.

2.4. Data Charting

2.4.1. Purpose and Approach

A data charting process was employed to systematically extract and organize relevant information from the final set of included sources. Following an iterative approach, the data charting form was refined as initial insights emerged from the literature, allowing for a flexible yet systematic data extraction process suitable for a scoping review.

2.4.2. Data Charting Form Categories

1. The final data charting form was organized into five key thematic areas, as detailed below:
2. Source Characteristics: Publication type (e.g., journal article, conference paper), year of publication, country of origin, and the specific engineering domain focus.
3. LLM Implementation Details: The type of LLM used (e.g., GPT-4, Llama 2), the primary application area (design, manufacturing, analysis, knowledge management), the integration approach (e.g., API-based, fine-tuned), and the specific tools or frameworks mentioned.
4. Engineering Applications: The specific engineering tasks addressed, any traditional methods being augmented or replaced, the performance metrics used for evaluation, and the reported outcomes or key findings.
5. Implementation Considerations: Any technical challenges encountered, solutions or workarounds that were developed, specific integration methods, and any discussion of safety or validation approaches.
6. Future Directions: Any identified limitations of the described approach, proposed improvements, stated research gaps, and identified needs for future development.

2.4.3. Charting Process

The charting process was designed to ensure consistency and rigor. An initial form was developed based on the review's research questions and was pilot tested on a sample of 5 sources to assess its clarity and comprehensiveness. The form was then refined based on team discussion and emerging themes from the pilot data. Throughout the extraction process, regular team meetings were held to discuss and resolve any charting challenges and ensure consistent application of the coding scheme. This iterative process of refinement continued as new patterns emerged from the literature, ensuring the final charted data was both comprehensive and systematically organized.

3. Results

The systematic search across the four specified databases initially yielded 680 records. After the removal of 12 duplicates, 668 unique articles were screened using ASReview, which resulted in the exclusion of 562 records. An additional 23 records were identified through other sources, leading to a total of 129 reports sought for retrieval. After 7 reports were not retrieved, a final pool of 122 articles underwent a comprehensive full-text review. This analysis resulted in the inclusion of 66 studies for the final synthesis. A total of 56 studies were excluded at this stage, with the most common reason being 'Review paper without original data' (n=28). A complete summary of the selection process is illustrated in the PRISMA flow diagram (Figure 1).

All 122 unique records were advanced for a comprehensive full-text review to determine final eligibility against the pre-defined inclusion and exclusion criteria. This analysis resulted in the inclusion of 66 studies. A total of 56 studies were excluded at this stage. The most common reason for exclusion was the article being a review paper without original data (n=28), followed by studies that did not describe a direct application within a mechanical engineering context (n=19).

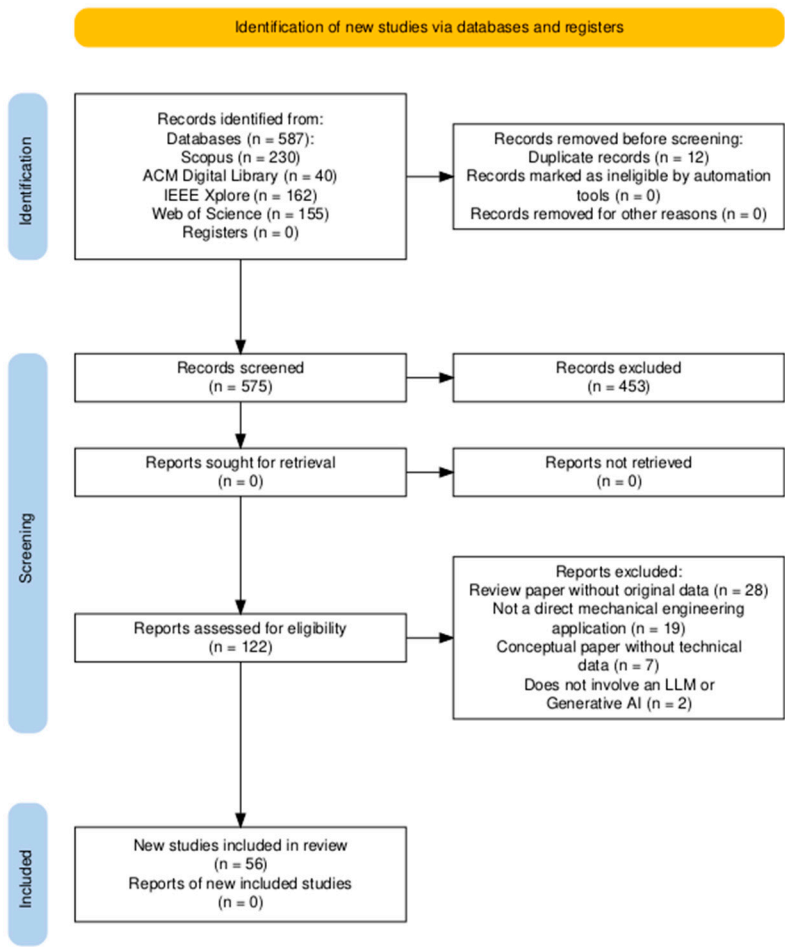


Figure 1. PRISMA Flow Diagram.

3.2. Characteristics of Evidence

The analysis of the 66 included studies reveals a rapidly emerging field of research. The temporal distribution, shown in Figure 2, indicates that research at this intersection is nascent, with the first publications appearing in 2020. However, the field has experienced a significant acceleration, with over 68% of all included studies published in 2024 alone. This highlights the recent and intense interest in applying LLMs to mechanical engineering challenges following the widespread availability of powerful models.

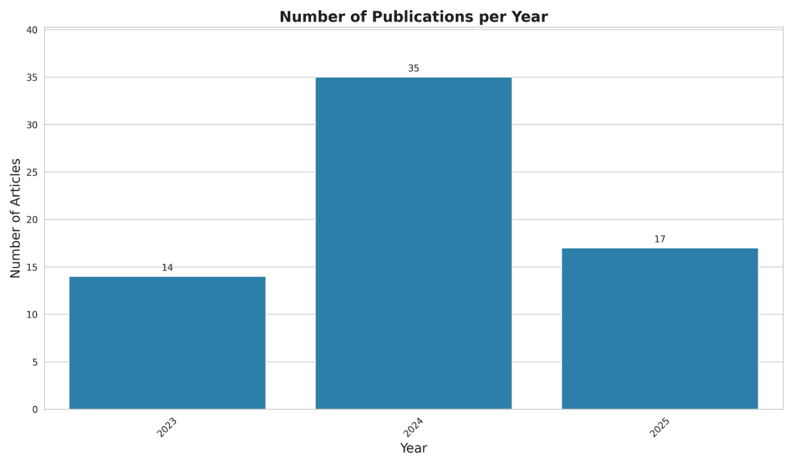


Figure 2. Publications by Year.

An analysis of the publication types (Figure 3) shows that the literature is primarily composed of preprints and conference papers, which together account for 64% of the included studies. This balance suggests a field characterized by rapid dissemination of novel ideas at conferences, with more mature research being consolidated into journal articles.

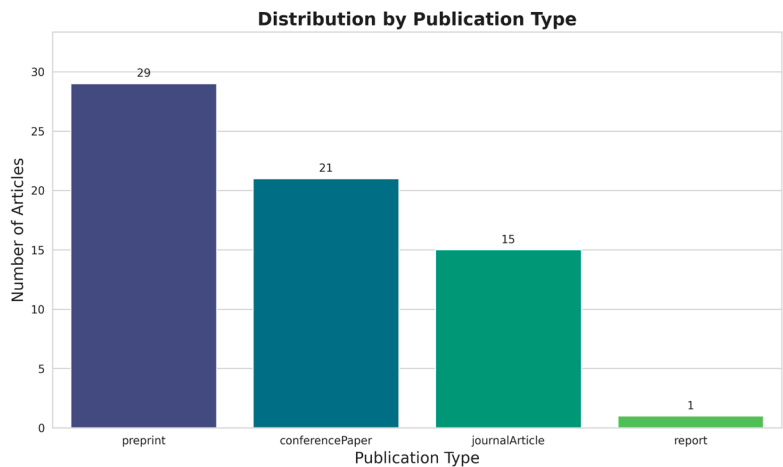


Figure 3. Publications by Type.

The geographical distribution of the research is highly concentrated, as illustrated in Figure 4. The United States is the leading contributor, accounting for nearly half of all publications, followed by China. This indicates that the current research landscape is dominated by these two countries, though a wide array of other nations contribute to a smaller but significant global effort.

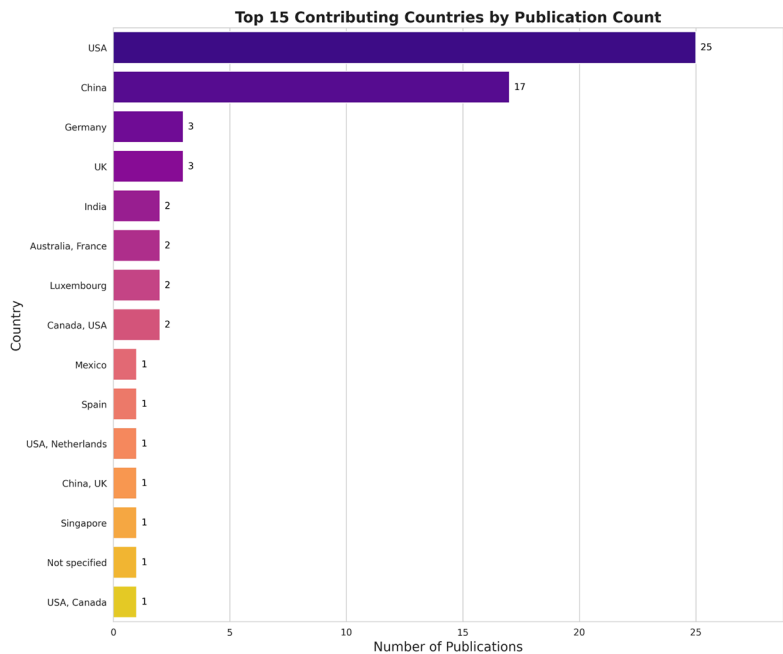


Figure 4. Publications by Geographical Distribution.

3.3. LLM Implementation Details

The analysis of implementation details reveals a strong reliance on a few dominant models and integration strategies.

As shown in Figure 5, the field is overwhelmingly dominated by OpenAI's GPT models (particularly GPT-4 and its variants), which are mentioned in over 60% of the studies that specified a

model. While other commercial and open-source models like Claude, Llama, and Gemini appear in the literature, their prevalence is significantly lower, underscoring the formative influence of the GPT series on current research.

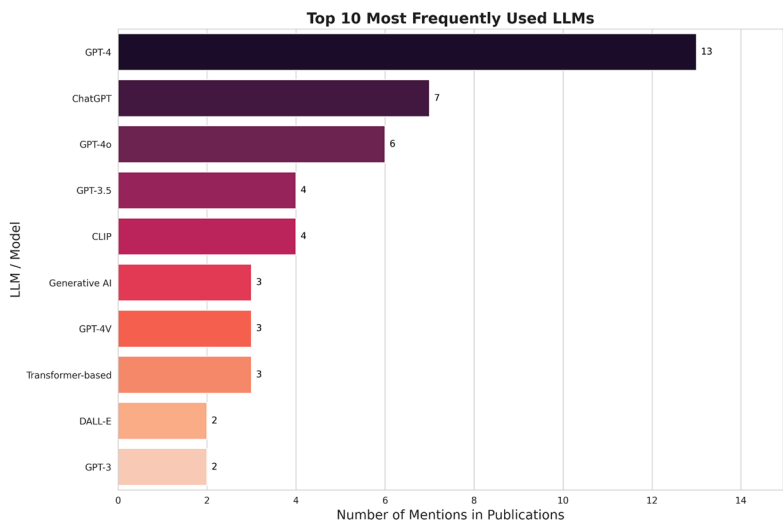


Figure 5. Publications by LLM Type.

The primary application areas for these models are heavily skewed towards the early stages of the engineering workflow (Figure 6). "Conceptual Design" and "CAD Generation from Text" are the most common applications, indicating a strong focus on leveraging LLMs to automate and enhance ideation and the creation of initial geometric models.

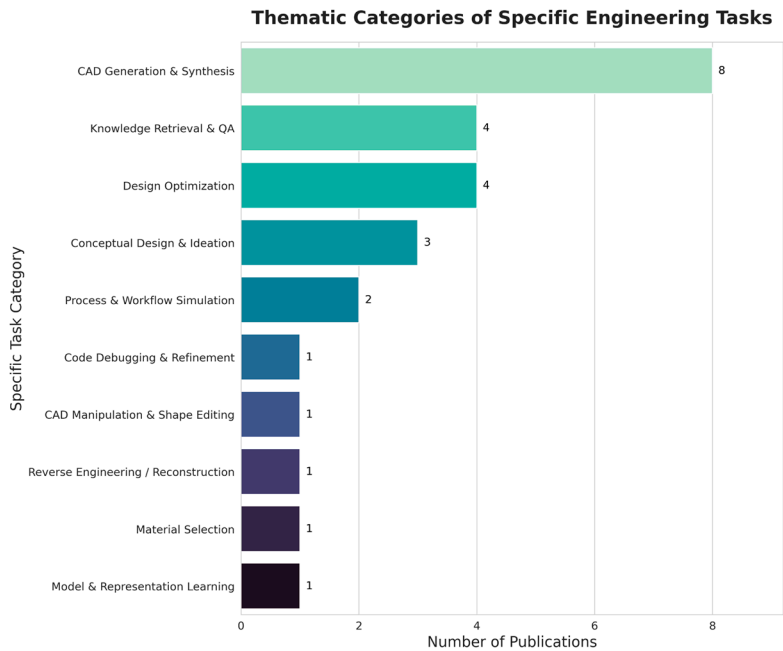


Figure 6. Task Categories.

Figure 7 illustrates the technical approaches used to integrate LLMs. The most common method is "Prompting / API Integration," where off-the-shelf models are utilized via their APIs. However, a substantial number of studies (n=31) are developing more sophisticated "Custom Frameworks or Multi-Agent Systems" to orchestrate complex tasks. "Fine-Tuning" and "Retrieval-Augmented

Generation (RAG)" represent more advanced techniques to imbue models with domain-specific knowledge, albeit used less frequently.

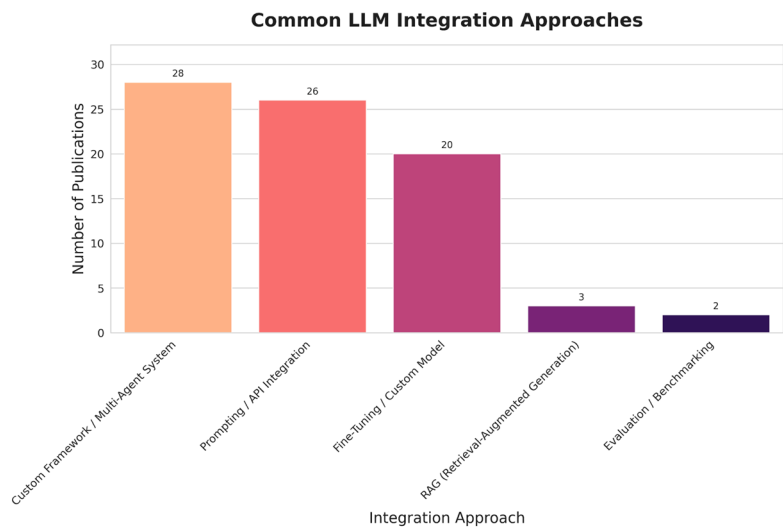


Figure 7. Publications by Integration Approaches.

3.4. Engineering Applications and Evaluation

When examining how LLMs are applied, the focus is clearly on augmenting or automating manual, knowledge-intensive tasks. As shown in Figure 8, "Manual CAD Design & Modeling" is the most frequently augmented traditional method, followed by "Conceptual Design & Ideation" and "Knowledge Management." This highlights a primary goal of using LLMs to reduce the manual effort and specialized software knowledge required in traditional CAx workflows

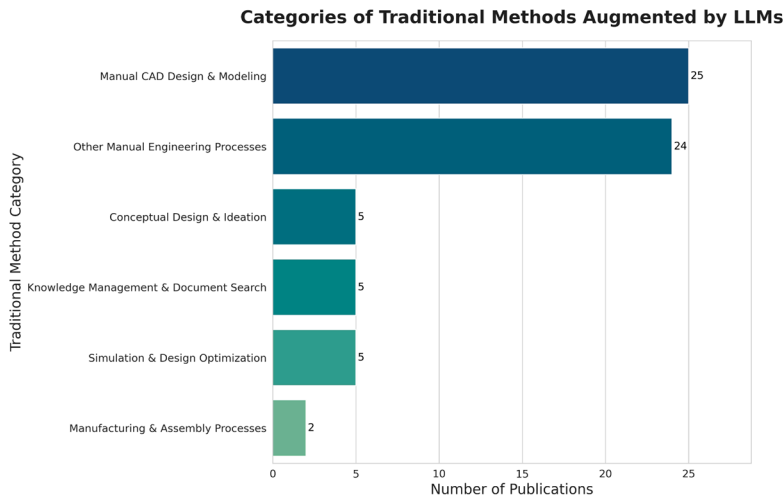


Figure 8. Manual CAD Design & Modeling.

The evaluation of these LLM-based systems employs a diverse set of metrics (Figure 9). The most common category is "Geometric Accuracy & Similarity," using metrics like Chamfer Distance and Intersection over Union (IoU) to assess the quality of generated 3D models. This is closely followed by "Task Success & NLP Metrics" (e.g., accuracy, F1-score) and "User-Based / Qualitative Evaluation," indicating that both technical performance and human-centric factors are considered critical for success.

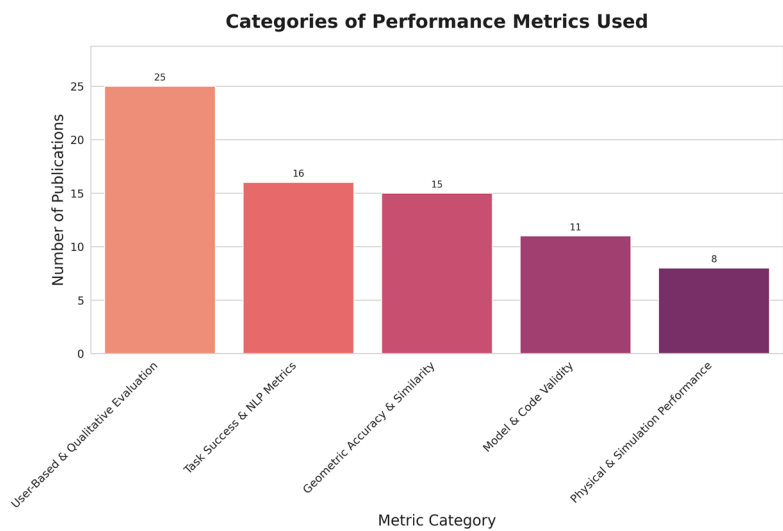


Figure 9. Performance Metrics.

3.4. Challenges, Limitations, and Future Directions

The analysis of reported challenges and limitations reveals several persistent themes. As shown in Figure 10, the most frequently cited technical challenge is the "Weak Spatial/Geometric Reasoning" of LLMs, mentioned in 19 studies. This fundamental difficulty in understanding and manipulating 3D space is a core barrier. Other significant challenges include the "Reliability & Factual Correctness" of the models (i.e., hallucination) and issues related to "Data Scarcity & Quality" for training and fine-tuning.

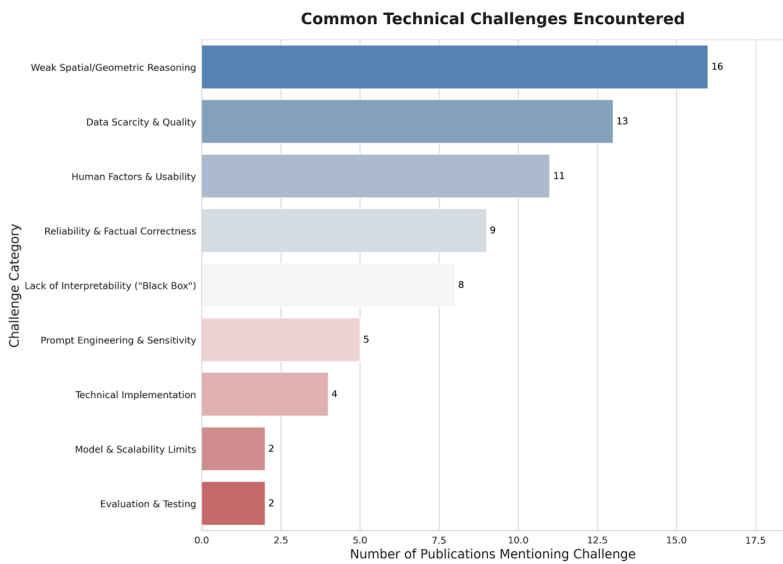


Figure 10. Common Challenges.

These challenges are directly reflected in the limitations reported in the final solutions (Figure 11). Again, "Weak Spatial/Geometric Reasoning" is the most common limitation (n=14), followed by the "Limited Scope & Task Complexity" of the proposed systems and concerns about their "Reliability & Factual Accuracy" (n=12).

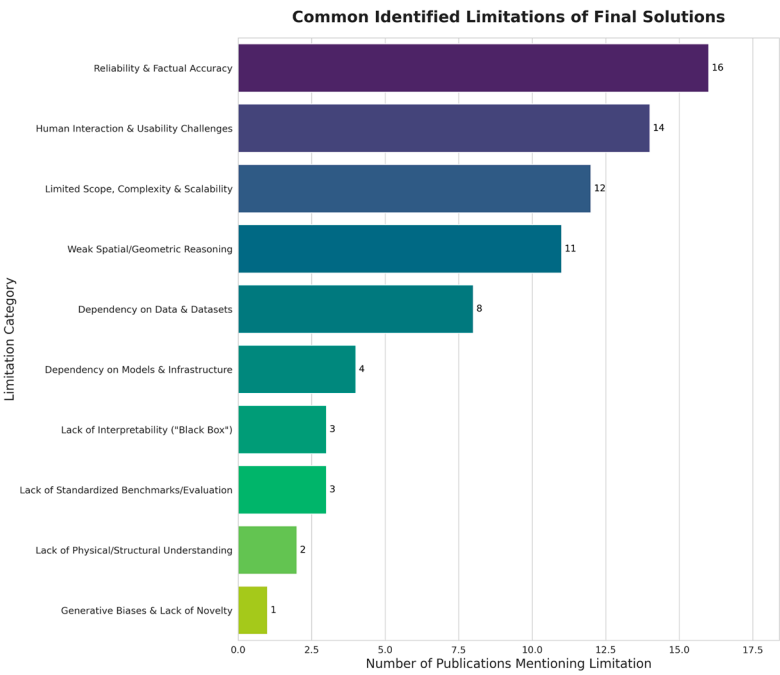


Figure 11. Common Limitations.

Finally, the proposed improvements and identified research gaps point towards a clear trajectory for the field. Figure 12 shows that the most commonly proposed improvement is to "Expand/Improve Datasets & Data Handling," followed by efforts to "Integrate Multimodal Inputs" and "Enhance Human-AI Interaction."

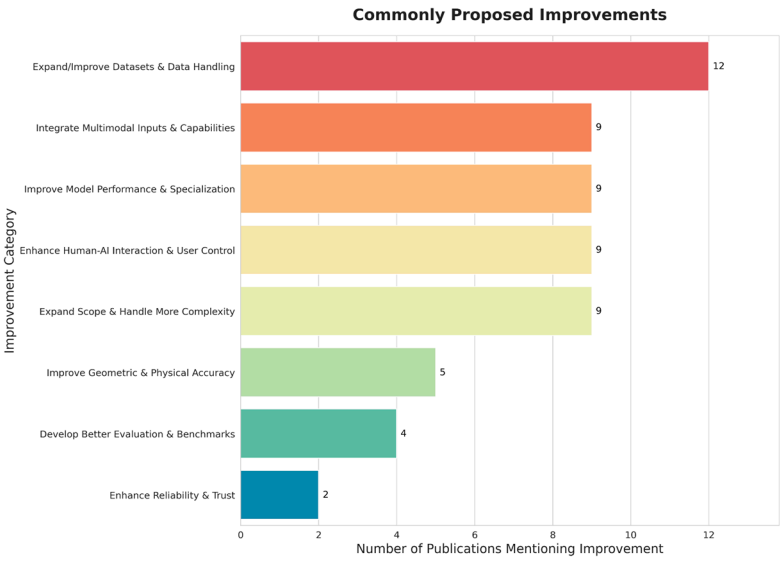


Figure 12. Proposed Improvements.

These proposed improvements directly align with the major research gaps identified in Figure 13. The most cited gap is the "Need for New Frameworks, Methods & Systems" to address complex engineering tasks, alongside a call for "Better Benchmarks & Evaluation" methods and a fundamental need to close the "Gaps in Multimodal & Spatial Capabilities."

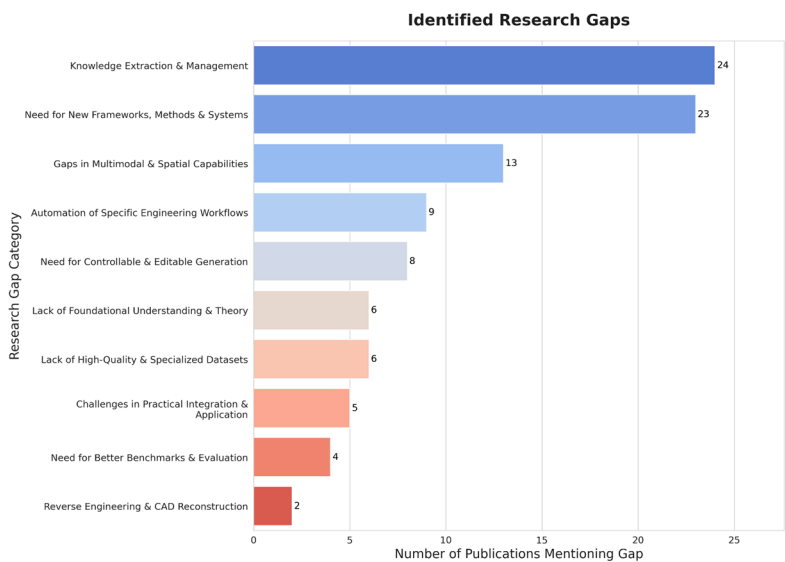


Figure 13. Research Gaps.

4. Discussion

4.1. Summary of Findings

This scoping review provides a comprehensive snapshot of a research domain that, while nascent, is experiencing explosive growth characterized by both methodological convergence and geographical concentration. The findings reveal that the application of Large Language Models in mechanical engineering, having emerged only in 2020, has since accelerated dramatically, with the vast majority of literature published in the most recent year of analysis. This rapid expansion is currently led by a concentrated cohort of researchers, primarily situated in the United States and China, who demonstrate a marked preference for fast-paced dissemination through preprints and conference papers—a publication pattern that is particularly indicative of a field navigating its formative stages while responding to rapidly evolving technological capabilities.

The technological landscape reveals a notably homogeneous foundation, characterized by significant reliance on OpenAI's GPT-4 and its variants across the surveyed literature. This pronounced dominance underscores the profound impact that a single, powerful, general-purpose model has exerted in shaping the initial trajectory and methodological approaches of the entire field. The primary application focus of these models centers heavily on the front-end of the engineering workflow, specifically augmenting and automating conceptual design processes and enabling the generation of initial CAD models from textual descriptions. This concentrated application focus suggests that researchers are strategically prioritizing the deployment of LLMs to lower barriers to entry for complex design tasks while simultaneously streamlining knowledge-intensive, manual processes that have traditionally constituted significant bottlenecks in engineering workflows.

Notably, while broader reviews of machine learning applications in engineering contexts consistently emphasize data scarcity as the principal limiting factor [37], our comprehensive synthesis reveals that for LLM applications specifically within mechanical engineering, the dominant challenge has shifted to weak spatial and geometric reasoning capabilities. This finding indicates a fundamental shift in bottleneck characteristics—transitioning from data availability constraints to inherent model architectural limitations—thereby highlighting the critical need for domain-specific design approaches in developing LLM frameworks suitable for engineering applications.

Despite the documented rapid progress across multiple application domains, a persistent and fundamental challenge continues to temper the field's broader potential: the demonstrably weak spatial and geometric reasoning capabilities exhibited by current-generation LLMs. This limitation emerged as both the most frequently cited technical challenge in the literature and the most commonly reported constraint in final implemented systems. The models' consistent inability to

reliably understand and manipulate the inherent complexities of three-dimensional spatial relationships, coupled with ongoing concerns regarding their factual accuracy and documented tendency toward hallucination, collectively represents the core barrier preventing deeper integration of these technologies into established engineering practice.

In direct response to these identified challenges, a remarkably clear consensus regarding the optimal path forward has emerged across the surveyed literature. The research community overwhelmingly points toward the urgent need for developing specialized frameworks and methodologies that fundamentally move beyond simple API-based interactions with general-purpose models. The proposed improvements and identified research gaps demonstrate notable convergence around three critical areas: first, the creation of high-quality, specialized datasets specifically designed for training and fine-tuning models on engineering-specific data and workflows; second, the strategic integration of multimodal inputs, including sketches, technical drawings, and images, to more effectively ground LLMs in the visual and spatial realities that define engineering practice; and third, the development of robust, domain-appropriate benchmarks and evaluation metrics specifically designed to assess model performance on complex, engineering-specific tasks with the rigor required for professional adoption.

4.2. Future Directions

Based on the challenges and research gaps identified in this review, several key directions for future work emerge. These can be grouped into foundational research needs, specific technical development priorities, and broader implementation challenges that must be addressed for the successful integration of LLMs into mechanical engineering.

4.2.1. Research Needs

The most pressing research need is the development of community-wide infrastructure to support rigorous and reproducible research. The literature highlights a significant lack of standardized benchmarks and evaluation metrics specifically designed for engineering tasks. Current methods often rely on general NLP metrics or geometric similarity scores that fail to capture the functional and physical validity of a design. Future work should focus on creating comprehensive benchmarks that test for manufacturability, physical plausibility, and adherence to engineering principles. Furthermore, the persistent challenge of data scarcity must be addressed. There is a clear need to develop and share high-quality, large-scale, multimodal datasets of engineering components, including their design history, associated documentation, and performance data.

4.2.2. Technical Development Priorities

The foremost technical priority is to overcome the weak spatial and geometric reasoning capabilities of current LLMs. This fundamental limitation is the primary barrier to more advanced applications. Future development should explore novel model architectures, such as graph neural networks or specialized transformers, that can better represent and manipulate the complex relationships in 3D space. Integrating multimodal inputs—such as sketches, technical drawings, and point clouds—is a critical component of this effort, as it helps to ground the models in the visual and geometric language of engineering.

Furthermore, the field must move beyond simple prompting of general-purpose APIs towards the creation of specialized, integrated frameworks. This includes developing more robust multi-agent systems that can orchestrate complex workflows, as well as fine-tuning smaller, domain-specific models on curated engineering data. The goal should be to create systems that produce controllable, editable, and parametric outputs, rather than static geometry, allowing for true integration into iterative engineering design cycles.

4.2.3. Implementation Challenges

For LLMs to be adopted in practice, their reliability and factual accuracy must be significantly improved. Future research must develop robust methods for validation, verification, and fact-checking to mitigate the risk of model hallucination in safety-critical applications. Alongside this, enhancing human-AI interaction is crucial. The development of more intuitive and collaborative interfaces is needed to move beyond the limitations of text-based prompts. These interfaces should provide engineers with greater control over the generative process and offer more interpretability into the model's "reasoning," addressing the "black box" problem and building the trust required for professional adoption.

4.3. Implications

The findings of this scoping review carry significant implications for several key stakeholder groups within the mechanical engineering community.

- **For the Research Community:** This review provides a clear, data-driven roadmap for future research priorities. The identified gaps—particularly the need for robust benchmarks, specialized datasets, and solutions to the persistent challenge of spatial reasoning—highlight the most critical areas where innovation is required. The current homogeneity in the field, with its heavy reliance on a few specific models and a concentration of research in limited geographical areas, signals a clear opportunity for diversification. Researchers can contribute by exploring alternative model architectures, developing open-source tools and datasets, and fostering broader international collaboration to enrich the field with new perspectives and approaches.
- **For Practicing Engineers and Industry:** The current state of the art, as mapped in this review, suggests that LLMs should be viewed as powerful "co-pilots" or "intelligent assistants" rather than autonomous experts. Their demonstrated strengths lie in augmenting the early stages of the design workflow, such as accelerating conceptual ideation, automating the generation of initial CAD models, and assisting in knowledge management tasks. However, the prevalent issues of reliability, factual accuracy, and weak geometric control mean that these tools are not yet suitable for detailed, safety-critical design or analysis without rigorous human oversight. The primary implication for industry is that the value of LLMs can be unlocked today by integrating them into workflows to enhance creativity and efficiency, but this must be coupled with robust validation and verification processes managed by domain experts.
- **For Engineering Education:** The rapid emergence of LLMs as tools for engineering design signals a necessary evolution in engineering curricula. The findings imply a potential shift in focus from traditional, manual software operation skills towards a new set of competencies centered on human-AI collaboration. Future engineering education will need to incorporate training on "AI literacy," including the principles of prompt engineering, understanding the inherent limitations and biases of generative models, and developing the critical thinking skills required to validate and critique AI-generated outputs. The ability to effectively leverage these tools as part of the engineering toolkit will be a critical skill for the next generation of mechanical engineers.

5. Conclusions

This scoping review has systematically mapped the rapidly expanding landscape of Large Language Model applications in mechanical engineering, revealing a research ecosystem characterized by both remarkable temporal acceleration and significant methodological challenges. The findings demonstrate a field defined not merely by nascent but exponentially accelerating growth, with research efforts demonstrating a pronounced focus on applying LLM technology to

augment the conceptual front-end of the design process, particularly in ideation workflows and the automated generation of initial CAD representations from natural language specifications. The current research landscape is dominated by a relatively narrow set of foundational models, with investigations primarily driven by coordinated efforts to automate traditionally knowledge-intensive engineering tasks while simultaneously reducing barriers to entry for sophisticated design software platforms.

Despite the considerable potential evidenced across multiple application domains, a critical and persistently recurring limitation continues to constrain broader adoption: the demonstrably weak spatial and geometric reasoning capabilities exhibited by current-generation models. This fundamental constraint, compounded by ongoing challenges in system reliability and the inherent scarcity of high-quality, domain-specific training data, collectively forms the primary barrier preventing broader deployment in autonomous applications within safety-critical engineering workflows. The path forward, as consistently indicated throughout the surveyed literature, necessitates a concerted, multi-pronged effort to develop specialized architectural frameworks, create robust and comprehensive domain-specific datasets, and establish standardized evaluation benchmarks capable of rigorously assessing performance on complex engineering tasks.

Ultimately, this comprehensive review concludes that while LLMs are strategically positioned to become powerful collaborative partners in mechanical engineering practice, their current capabilities and limitations suggest their optimal role is best characterized as intelligent assistants rather than autonomous design agents. Through providing a systematic, evidence-based synthesis of the current state of the art, its persistent challenges, and the emerging consensus on future research directions, this work establishes a foundational roadmap for researchers, industry practitioners, and engineering educators seeking to guide the effective and responsible integration of this transformative technology into the evolving landscape of engineering practice.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org. Appendix B.

Author Contributions: Conceptualization, C.B.; methodology, C.B.; software, C.B.; validation, C.B.; formal analysis, X.X.; investigation, X.X.; resources, X.X.; data curation, C.B.; writing—original draft preparation, C.B.; writing—review and editing, C.B.,K.R. and M.P.; visualization, C.B.; All authors have read and agreed to the published version of the manuscript.

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

Abbreviations

The following abbreviations are used in this manuscript:

CAD	Computer-Aided Design
API	Application Programming Interface
ASReview	Active learning for Systematic Reviews
CAX	Computer-Aided Technologies (<i>referring generally to CAD, CAM, CAE, etc.</i>)
DSL	Domain Specific Language
FEA	Finite Element Analysis
IoU	Intersection over Union
LLM	Large Language Model
MMLM	Multi-Modal Large Language Model
PRISMA-SCR	Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
RAG	Retrieval-Augmented Generation

Appendix A. Search Strategy Details

6.1.1. Scopus
TITLE-ABS-KEY ("large language model" OR "LLM" OR "generative AI" OR "foundation model") AND TITLE-ABS-KEY ("mechanical engineering" OR "engineering design" OR "product design") AND TITLE-ABS-KEY ("computer-aided design" OR "CAD" OR "generative design" OR "finite element" OR "FEA" OR "simulation" OR "CAx") AND (PUBYEAR > 2019)

6.1.2. IEEE Xplore
(((("All Metadata":"large language model") OR ("All Metadata":"LLM") OR ("All Metadata":"generative AI") OR ("All Metadata":"foundation model")) AND (("All Metadata":"mechanical engineering") OR ("All Metadata":"engineering design") OR ("All Metadata":"product design")) AND (("All Metadata":"computer-aided design") OR ("All Metadata":"CAD") OR ("All Metadata":"generative design") OR ("All Metadata":"finite element") OR ("All Metadata":"FEA") OR ("All Metadata":"simulation") OR ("All Metadata":"CAx")))))

6.1.3. ACM Digital Library
[[All: "large language model"] OR [All: "llm"] OR [All: "generative ai"] OR [All: "foundation model"]] AND [[All: "mechanical engineering"] OR [All: "engineering design"] OR [All: "product design"]] AND [[All: "computer-aided design"] OR [All: "cad"] OR [All: "generative design"] OR [All: "finite element"] OR [All: "fea"] OR [All: "simulation"] OR [All: "cax"]] AND [E-Publication Date: (01/01/2020 TO 31/12/2025)]

6.1.4. Web of Science
TS=("large language model" OR "LLM" OR "generative AI" OR "foundation model") AND TS=("mechanical engineering" OR "engineering design" OR "product design") AND TS=("computer-aided design" OR "CAD" OR "generative design" OR "finite element" OR "FEA" OR "simulation" OR "CAx")

Appendix B. Data Charting Form

The data charting form used for this scoping review is directly implemented as the structured spreadsheet containing the full dataset. The final, organized data can be found in the Excel file: LLM Mechanical Engineering Scoping Review.xlsx

This file is organized into two sheets:

Included: Contains the fully charted data for the 66 studies included in the final synthesis.

Excluded: Contains the metadata and the specific reason for exclusion for the 56 studies that did not meet the eligibility criteria.

The columns in the "Included" sheet represent the complete set of data points extracted from each study, corresponding to the five thematic areas described in the Methods section. The primary column headers are:

Source Characteristics:
Paper_ID, Paper_Title, First_Author, Year, Publication_Type, Country_of_Origin, Engineering_Domain_Focus

LLM Implementation Details:
LLM_Used, Application_Area, Integration_Approach, Tools_Frameworks_Used

Engineering Applications:
Specific_Task_Addressed, Traditional_Method_Augmented, Performance_Metrics_Used, Reported_Outcomes

Implementation Considerations:
Technical_Challenges, Proposed_Solutions, Safety_Validation_Approaches

Future Directions:
Identified_Limitations, Proposed_Improvements, Research_Gaps_Identified

Appendix C. Included Sources

1. J. Kaplan et al., "Scaling Laws for Neural Language Models," Jan. 23, 2020, arXiv: arXiv:2001.08361. doi: 10.48550/arXiv.2001.08361.
2. W. Li, G. Mac, N. G. Tsoutsos, N. Gupta, and R. Karri, "Computer aided design (CAD) model search and retrieval using frequency domain file conversion," *Additive Manufacturing*, vol. 36, p. 101554, Dec. 2020, doi: 10.1016/j.addma.2020.101554.
3. W. Tao, M. C. Leu, and Z. Yin, "Multi-modal recognition of worker activity for human-centered intelligent manufacturing," *Engineering Applications of Artificial Intelligence*, vol. 95, p. 103868, Oct. 2020, doi: 10.1016/j.engappai.2020.103868.
4. A. Chernyavskiy, D. Ilvovsky, and P. Nakov, "Transformers: 'The End of History' for Natural Language Processing?," in *Machine Learning and Knowledge Discovery in Databases. Research Track*, N. Oliver, F. Pérez-Cruz, S. Kramer, J. Read, and J. A. Lozano, Eds., Cham: Springer International Publishing, 2021, pp. 677–693. doi: 10.1007/978-3-030-86523-8_41.
5. D. Fuchs, R. Bartz, S. Kuschmitz, and T. Vietor, "Necessary advances in computer-aided design to leverage on additive manufacturing design freedom," *Int J Interact Des Manuf*, vol. 16, no. 4, pp. 1633–1651, Dec. 2022, doi: 10.1007/s12008-022-00888-z.
6. L. Mandelli and S. Berretti, "CAD 3D Model classification by Graph Neural Networks: A new approach based on STEP format," Oct. 30, 2022, arXiv: arXiv:2210.16815. doi: 10.48550/arXiv.2210.16815.
7. M. C. May, J. Neidhöfer, T. Körner, L. Schäfer, and G. Lanza, "Applying Natural Language Processing in Manufacturing," *Procedia CIRP*, vol. 115, pp. 184–189, 2022, doi: 10.1016/j.procir.2022.10.071.
8. B. Regassa Hunde and A. Debebe Woldeyohannes, "Future prospects of computer-aided design (CAD) – A review from the perspective of artificial intelligence (AI), extended reality, and 3D printing," *Results in Engineering*, vol. 14, p. 100478, June 2022, doi: 10.1016/j.rineng.2022.100478.
9. L. Regenwetter, A. H. Nobari, and F. Ahmed, "Deep Generative Models in Engineering Design: A Review," *Journal of Mechanical Design*, vol. 144, no. 7, p. 071704, July 2022, doi: 10.1115/1.4053859.
10. H. Strobelt et al., "Interactive and Visual Prompt Engineering for Ad-hoc Task Adaptation With Large Language Models," *IEEE Trans. Visual. Comput. Graphics*, pp. 1–11, 2022, doi: 10.1109/TVCG.2022.3209479.
11. S. Zhang et al., "OPT: Open Pre-trained Transformer Language Models," June 21, 2022, arXiv: arXiv:2205.01068. doi: 10.48550/arXiv.2205.01068.
12. A. A. Chien, L. Lin, H. Nguyen, V. Rao, T. Sharma, and R. Wijayawardana, "Reducing the Carbon Impact of Generative AI Inference (today and in 2035)," in *Proceedings of the 2nd Workshop on Sustainable Computer Systems*, in *HotCarbon '23*. New York, NY, USA: Association for Computing Machinery, Aug. 2023, pp. 1–7. doi: 10.1145/3604930.3605705.
13. P. Christiano, J. Leike, T. B. Brown, M. Martic, S. Legg, and D. Amodei, "Deep reinforcement learning from human preferences," Feb. 17, 2023, arXiv: arXiv:1706.03741. doi: 10.48550/arXiv.1706.03741.
14. J. A. De La Tejera, R. A. Ramirez-Mendoza, M. R. Bustamante-Bello, P. Orta-Castañón, and L. A. Arce-Saenz, "Overview of an AI-Based Methodology for Design: Case Study of a High Efficiency Electric Vehicle Chassis for the Shell Eco-Marathon," presented at the 2023 International Symposium on Electromobility, ISEM 2023, 2023. doi: 10.1109/ISEM59023.2023.10334852.

15. F. Gmeiner, H. Yang, L. Yao, K. Holstein, and N. Martelaro, "Exploring Challenges and Opportunities to Support Designers in Learning to Co-create with AI-based Manufacturing Design Tools," in Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, in CHI '23. New York, NY, USA: Association for Computing Machinery, 2023. doi: 10.1145/3544548.3580999.
16. H. Jobczyk and H. Homann, "Automatic Reverse Engineering: Creating computer-aided design (CAD) models from multi-view images," Sept. 2023, doi: 10.48550/arXiv.2309.13281.
17. J. Kaddour, J. Harris, M. Mozes, H. Bradley, R. Raileanu, and R. McHardy, "Challenges and Applications of Large Language Models," July 19, 2023, arXiv: arXiv:2307.10169. doi: 10.48550/arXiv.2307.10169.
18. O. Khattab et al., "DSPy: Compiling Declarative Language Model Calls into Self-Improving Pipelines," Oct. 05, 2023, arXiv: arXiv:2310.03714. doi: 10.48550/arXiv.2310.03714.
19. M. Kodnongbua, B. Jones, M. B. S. Ahmad, V. Kim, and A. Schulz, "ReparamCAD: Zero-shot CAD Re-Parameterization for Interactive Manipulation," in SIGGRAPH Asia 2023 Conference Papers, in SA '23. New York, NY, USA: Association for Computing Machinery, 2023. doi: 10.1145/3610548.3618219.
20. V. Liu, "Beyond Text-to-Image: Multimodal Prompts to Explore Generative AI," in Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems, in CHI EA '23. New York, NY, USA: Association for Computing Machinery, 2023. doi: 10.1145/3544549.3577043.
21. V. Liu, J. Vermeulen, G. Fitzmaurice, and J. Matejka, "3DALL-E: Integrating Text-to-Image AI in 3D Design Workflows," in Proceedings of the 2023 ACM Designing Interactive Systems Conference, in DIS '23. New York, NY, USA: Association for Computing Machinery, 2023, pp. 1955–1977. doi: 10.1145/3563657.3596098.
22. Y. Liu, A. Obukhov, J. D. Wegner, and K. Schindler, "Point2CAD: Reverse Engineering CAD Models from 3D Point Clouds," Dec. 07, 2023, arXiv: arXiv:2312.04962. doi: 10.48550/arXiv.2312.04962.
23. Y. Lou, X. Li, H. Chen, and X. Zhou, "BRep-BERT: Pre-training Boundary Representation BERT with Sub-graph Node Contrastive Learning," in Proceedings of the 32nd ACM International Conference on Information and Knowledge Management, Birmingham United Kingdom: ACM, Oct. 2023, pp. 1657–1666. doi: 10.1145/3583780.3614795.
24. L. Makatura et al., "How Can Large Language Models Help Humans in Design and Manufacturing?," July 25, 2023, arXiv: arXiv:2307.14377. doi: 10.48550/arXiv.2307.14377.
25. G. Penedo et al., "The RefinedWeb Dataset for Falcon LLM: Outperforming Curated Corpora with Web Data, and Web Data Only," June 01, 2023, arXiv: arXiv:2306.01116. doi: 10.48550/arXiv.2306.01116.
26. T. Rios, S. Menzel, and B. Sendhoff, "Large Language and Text-to-3D Models for Engineering Design Optimization," presented at the 2023 IEEE Symposium Series on Computational Intelligence, SSCI 2023, 2023, pp. 1704–1711. doi: 10.1109/SSCI52147.2023.10371898.
27. E. Ruiz, M. I. Torres, and A. Del Pozo, "Question answering models for human-machine interaction in the manufacturing industry," Computers in Industry, vol. 151, p. 103988, Oct. 2023, doi: 10.1016/j.compind.2023.103988.
28. S. Ding, X. Chen, Y. Fang, W. Liu, Y. Qiu, and C. Chai, "DesignGPT: Multi-Agent Collaboration in Design," in 2023 16th International Symposium on Computational Intelligence and Design (ISCID), Dec. 2023, pp. 204–208. doi: 10.1109/ISCID59865.2023.00056.

29. S. Samsi et al., "From Words to Watts: Benchmarking the Energy Costs of Large Language Model Inference," Oct. 04, 2023, arXiv: arXiv:2310.03003. doi: 10.48550/arXiv.2310.03003.
30. B. Song, R. Zhou, and F. Ahmed, "Multi-modal Machine Learning in Engineering Design: A Review and Future Directions," July 28, 2023, arXiv: arXiv:2302.10909. doi: 10.48550/arXiv.2302.10909.
31. D. Tas and D. Chatzinikolis, "TeamCAD -- A Multimodal Interface for Remote Computer Aided Design," Dec. 2023, doi: 10.48550/arXiv.2312.12309.
32. X. Wang, N. Anwer, Y. Dai, and A. Liu, "ChatGPT for design, manufacturing, and education," *Procedia CIRP*, vol. 119, pp. 7–14, 2023, doi: 10.1016/j.procir.2023.04.001.
33. J. Wei et al., "Chain-of-Thought Prompting Elicits Reasoning in Large Language Models," Jan. 10, 2023, arXiv: arXiv:2201.11903. doi: 10.48550/arXiv.2201.11903.
34. S. Wu et al., "BloombergGPT: A Large Language Model for Finance," Dec. 21, 2023, arXiv: arXiv:2303.17564. doi: 10.48550/arXiv.2303.17564.
35. X. Xu, P. K. Jayaraman, J. G. Lambourne, K. D. D. Willis, and Y. Furukawa, "Hierarchical Neural Coding for Controllable CAD Model Generation," June 30, 2023, arXiv: arXiv:2307.00149. doi: 10.48550/arXiv.2307.00149.
36. H. Yang, X.-Y. Liu, and C. D. Wang, "FinGPT: Open-Source Financial Large Language Models," June 09, 2023, arXiv: arXiv:2306.06031. doi: 10.48550/arXiv.2306.06031.
37. B. Zhang and H. Soh, "Large Language Models as Zero-Shot Human Models for Human-Robot Interaction," in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Oct. 2023, pp. 7961–7968. doi: 10.1109/IROS55552.2023.10341488.
38. S. Zhang, Z. Guan, H. Jiang, T. Ning, X. Wang, and P. Tan, "Brep2Seq: a dataset and hierarchical deep learning network for reconstruction and generation of computer-aided design models," *Journal of Computational Design and Engineering*, vol. 11, no. 1, pp. 110–134, Dec. 2023, doi: 10.1093/jcde/qwae005.
39. M. F. Alam et al., "From Automation to Augmentation: Redefining Engineering Design and Manufacturing in the Age of NextGen-AI," *An MIT Exploration of Generative AI*, Mar. 2024, doi: 10.21428/e4baedd9.e39b392d.
40. K. Alrashedy, P. Tambwekar, Z. Zaidi, M. Langwasser, W. Xu, and M. Gombolay, "Generating CAD Code with Vision-Language Models for 3D Designs," Oct. 07, 2024, arXiv: arXiv:2410.05340. doi: 10.48550/arXiv.2410.05340.
41. A. Badagabettu, S. S. Yarlagadda, and A. B. Farimani, "Query2CAD: Generating CAD models using natural language queries," May 31, 2024, arXiv: arXiv:2406.00144. doi: 10.48550/arXiv.2406.00144.
42. G. Bai et al., "Beyond Efficiency: A Systematic Survey of Resource-Efficient Large Language Models," Dec. 29, 2024, arXiv: arXiv:2401.00625. doi: 10.48550/arXiv.2401.00625.
43. Y. Cao, A. Taghvaie Nakhjiri, and M. Ghadiri, "Different applications of machine learning approaches in materials science and engineering: Comprehensive review," *Engineering Applications of Artificial Intelligence*, vol. 135, p. 108783, Sept. 2024, doi: 10.1016/j.engappai.2024.108783.
44. L. Chen et al., "Toward Controllable Generative Design: A Conceptual Design Generation Approach Leveraging the Function–Behavior–Structure Ontology and Large Language Models," *Journal of Mechanical Design*, vol. 146, no. 12, p. 121401, Dec. 2024, doi: 10.1115/1.4065562.

45. S. Chen, J. Ding, Z. Shao, Z. Shi, and J. Lin, "Neural surrogate-driven modelling, optimisation, and generation of engineering designs: A concise review," presented at the Materials Research Proceedings, 2024, pp. 493–502. doi: 10.21741/9781644903254-53.
46. V. Chheang et al., "A Virtual Environment for Collaborative Inspection in Additive Manufacturing," in Extended Abstracts of the CHI Conference on Human Factors in Computing Systems, in CHI EA '24. New York, NY, USA: Association for Computing Machinery, 2024. doi: 10.1145/3613905.3650730.
47. L. Chong, J. Rayan, S. Dow, I. Lykourantzou, and F. Ahmed, "CAD-PROMPTED GENERATIVE MODELS: A PATHWAY TO FEASIBLE AND NOVEL ENGINEERING DESIGNS," presented at the Proceedings of the ASME Design Engineering Technical Conference, 2024. doi: 10.1115/DETC2024-146325.
48. S. Colabianchi, F. Costantino, and N. Sabetta, "Assessment of a large language model based digital intelligent assistant in assembly manufacturing," Computers in Industry, vol. 162, p. 104129, Nov. 2024, doi: 10.1016/j.compind.2024.104129.
49. A. C. Doris, D. Grandi, R. Tomich, M. F. Alam, H. Cheong, and F. Ahmed, "DesignQA: BENCHMARKING MULTIMODAL LARGE LANGUAGE MODELS ON QUESTIONS GROUNDED IN ENGINEERING DOCUMENTATION," presented at the Proceedings of the ASME Design Engineering Technical Conference, 2024. doi: 10.1115/DETC2024-139024.
50. Y. Du et al., "BlenderLLM: Training Large Language Models for Computer-Aided Design with Self-improvement," Dec. 2024, doi: 10.48550/arXiv.2412.14203.
51. A. Ganeshan, R. Huang, X. Xu, R. K. Jones, and D. Ritchie, "ParSEL: Parameterized Shape Editing with Language," ACM Trans. Graph., vol. 43, no. 6, pp. 1–14, Dec. 2024, doi: 10.1145/3687922.
52. J. Göpfert, J. M. Weinand, P. Kuckertz, and D. Stolten, "Opportunities for large language models and discourse in engineering design," Energy and AI, vol. 17, p. 100383, Sept. 2024, doi: 10.1016/j.egyai.2024.100383.
53. D. Grandi, Y. P. Jain, A. Groom, B. Cramer, and C. McComb, "Evaluating Large Language Models for Material Selection," Apr. 23, 2024, arXiv: arXiv:2405.03695. doi: 10.48550/arXiv.2405.03695.
54. M. U. Hadi et al., "Large Language Models: A Comprehensive Survey of its Applications, Challenges, Limitations, and Future Prospects," Aug. 12, 2024, Preprints. doi: 10.36227/techrxiv.23589741.v6.
55. Y. Huang et al., "TrustLLM: Trustworthiness in Large Language Models," Sept. 30, 2024, arXiv: arXiv:2401.05561. doi: 10.48550/arXiv.2401.05561.
56. Y. Jang and K. H. Hyun, "Advancing 3D CAD with Workflow Graph-Driven Bayesian Command Inferences," in Extended Abstracts of the CHI Conference on Human Factors in Computing Systems, in CHI EA '24. New York, NY, USA: Association for Computing Machinery, 2024. doi: 10.1145/3613905.3650895.
57. A. Jignasu, K. Marshall, B. Ganapathysubramanian, A. Balu, C. Hegde, and A. Krishnamurthy, "Evaluating Large Language Models for G-Code Debugging, Manipulation, and Comprehension," in 2024 IEEE LLM Aided Design Workshop (LAD), San Jose, CA, USA: IEEE, June 2024, pp. 1–5. doi: 10.1109/LAD62341.2024.10691700.
58. M. Jung, M. Kim, and J. Kim, "ContrastCAD: Contrastive Learning-based Representation Learning for Computer-Aided Design Models," Apr. 2024, doi: 10.48550/arXiv.2404.01645.

59. T. Kapsalis, "CADgpt: Harnessing Natural Language Processing for 3D Modelling to Enhance Computer-Aided Design Workflows," Jan. 2024, doi: 10.48550/arXiv.2401.05476.
60. R. Kasar and T. Kumar, "Digital Twin and Generative AI for Product Development," presented at the Procedia CIRP, 2024, pp. 905–910. doi: 10.1016/j.procir.2024.06.043.
61. M. S. Khan, S. Sinha, T. U. Sheikh, D. Stricker, S. A. Ali, and M. Z. Afzal, "Text2CAD: Generating Sequential CAD Models from Beginner-to-Expert Level Text Prompts," Sept. 25, 2024, arXiv: arXiv:2409.17106. doi: 10.48550/arXiv.2409.17106.
62. P. Krus, "LARGE LANGUAGE MODEL IN AIRCRAFT SYSTEM DESIGN," presented at the ICAS Proceedings, 2024. [Online]. Available: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85208784130&partnerID=40&md5=fe0b7c9e9020fd5bd4be3770e62aef59>
63. S. Lambert, C. Mathews, and A. Jaddoa, "CONCEPT TO PRODUCTION WITH A GEN AI DESIGN ASSISTANT: AIDA," presented at the Proceedings of the 26th International Conference on Engineering and Product Design Education: Rise of the Machines: Design Education in the Generative AI Era, E and PDE 2024, 2024, pp. 235–240. doi: 10.35199/epde.2024.40.
64. K. Li, A. K. Hopkins, D. Bau, F. Viégas, H. Pfister, and M. Wattenberg, "Emergent World Representations: Exploring a Sequence Model Trained on a Synthetic Task," June 26, 2024, arXiv: arXiv:2210.13382. doi: 10.48550/arXiv.2210.13382.
65. X. Li, Y. Sun, and Z. Sha, "LLM4CAD: Multi-Modal Large Language Models for 3D Computer-Aided Design Generation," in Volume 6: 36th International Conference on Design Theory and Methodology (DTM), Washington, DC, USA: American Society of Mechanical Engineers, Aug. 2024, p. V006T06A015. doi: 10.1115/DETC2024-143740.
66. Y. Li et al., "Large Language Models for Manufacturing," Oct. 28, 2024, arXiv: arXiv:2410.21418. doi: 10.48550/arXiv.2410.21418.
67. Y. Liu et al., "Understanding LLMs: A Comprehensive Overview from Training to Inference," Jan. 06, 2024, arXiv: arXiv:2401.02038. doi: 10.48550/arXiv.2401.02038.
68. Y. Liu, J. Chen, S. Pan, D. Cohen-Or, H. Zhang, and H. Huang, "Split-and-Fit: Learning B-Reps via Structure-Aware Voronoi Partitioning," June 07, 2024, arXiv: arXiv:2406.05261. doi: 10.48550/arXiv.2406.05261.
69. L. Makatura et al., "Large Language Models for Design and Manufacturing," An MIT Exploration of Generative AI, Mar. 2024, doi: 10.21428/e4baedd9.745b62fa.
70. D. Mallis et al., "CAD-Assistant: Tool-Augmented VLLMs as Generic CAD Task Solvers?," Dec. 18, 2024, arXiv: arXiv:2412.13810. doi: 10.48550/arXiv.2412.13810.
71. B. Murugadoss et al., Evaluating the Evaluator: Measuring LLMs' Adherence to Task Evaluation Instructions. 2024. doi: 10.48550/arXiv.2408.08781.
72. N. Widulle, F. Meyer, and O. Niggemann, "Generating Assembly Instructions Using Reinforcement Learning in Combination with Large Language Models," in 2024 IEEE 22nd International Conference on Industrial Informatics (INDIN), Aug. 2024, pp. 1–7. doi: 10.1109/INDIN58382.2024.10774545.
73. D. Nath, Ankit, D. R. Neog, and S. S. Gautam, "Application of Machine Learning and Deep Learning in Finite Element Analysis: A Comprehensive Review," Arch Computat Methods Eng, vol. 31, no. 5, pp. 2945–2984, July 2024, doi: 10.1007/s11831-024-10063-0.
74. H. Naveed et al., "A Comprehensive Overview of Large Language Models," Oct. 17, 2024, arXiv: arXiv:2307.06435. doi: 10.48550/arXiv.2307.06435.

75. K. C. Pierson and M. J. Ha, "Usage of ChatGPT for Engineering Design and Analysis Tool Development," presented at the AIAA SciTech Forum and Exposition, 2024, 2024. doi: 10.2514/6.2024-0914.
76. A. Ray, "Smart Design Evolution with GenAI and 3D Printing," presented at the 2024 IEEE Integrated STEM Education Conference, ISEC 2024, 2024. doi: 10.1109/ISEC61299.2024.10665095.
77. D. Rukhovich, E. Dupont, D. Mallis, K. Cherenkova, A. Kacem, and D. Aouada, "CAD-Recode: Reverse Engineering CAD Code from Point Clouds," Dec. 18, 2024, arXiv: arXiv:2412.14042. doi: 10.48550/arXiv.2412.14042.
78. S. Wang et al., "CAD-GPT: Synthesising CAD Construction Sequence with Spatial Reasoning-Enhanced Multimodal LLMs," Dec. 27, 2024, arXiv: arXiv:2412.19663. doi: 10.48550/arXiv.2412.19663.
79. X. Wang, J. Zheng, Y. Hu, H. Zhu, Q. Yu, and Z. Zhou, "From 2D CAD Drawings to 3D Parametric Models: A Vision-Language Approach," Dec. 17, 2024, arXiv: arXiv:2412.11892. doi: 10.48550/arXiv.2412.11892.
80. Z. Wang et al., "LLaMA-Mesh: Unifying 3D Mesh Generation with Language Models," Nov. 14, 2024, arXiv: arXiv:2411.09595. doi: 10.48550/arXiv.2411.09595.
81. M. Wong, T. Rios, S. Menzel, and Y. S. Ong, "Prompt Evolutionary Design Optimization with Generative Shape and Vision-Language models," presented at the 2024 IEEE Congress on Evolutionary Computation, CEC 2024 - Proceedings, 2024. doi: 10.1109/CEC60901.2024.10611898.
82. S. Wu et al., "CadVLM: Bridging Language and Vision in the Generation of Parametric CAD Sketches," Sept. 26, 2024, arXiv: arXiv:2409.17457. doi: 10.48550/arXiv.2409.17457.
83. L. Xia, C. Li, C. Zhang, S. Liu, and P. Zheng, "Leveraging error-assisted fine-tuning large language models for manufacturing excellence," *Robotics and Computer-Integrated Manufacturing*, vol. 88, p. 102728, Aug. 2024, doi: 10.1016/j.rcim.2024.102728.
84. J. Xu, C. Wang, Z. Zhao, W. Liu, Y. Ma, and S. Gao, "CAD-MLLM: Unifying Multimodality-Conditioned CAD Generation With MLLM," Nov. 07, 2024, arXiv: arXiv:2411.04954. doi: 10.48550/arXiv.2411.04954.
85. X. Xu, J. G. Lambourne, P. K. Jayaraman, Z. Wang, K. D. D. Willis, and Y. Furukawa, "BrepGen: A B-rep Generative Diffusion Model with Structured Latent Geometry," Nov. 03, 2024, arXiv: arXiv:2401.15563. doi: 10.48550/arXiv.2401.15563.
86. Z. Yuan, J. Shi, and Y. Huang, "OpenECAD: An efficient visual language model for editable 3D-CAD design," *Computers & Graphics*, vol. 124, p. 104048, Nov. 2024, doi: 10.1016/j.cag.2024.104048.
87. S. Zarghami, H. Kouchaki, L. Yangand, and P. M. Rodriguez, "EXPLAINABLE ARTIFICIAL INTELLIGENCE IN GENERATIVE DESIGN FOR CONSTRUCTION," presented at the Proceedings of the European Conference on Computing in Construction, 2024, pp. 556–563. doi: 10.35490/EC3.2024.277.
88. H. Zhang, P. Chen, X. Xie, Z. Jiang, Z. Zhou, and L. Sun, "A Hybrid Prototype Method Combining Physical Models and Generative Artificial Intelligence to Support Creativity in Conceptual Design," *ACM Trans. Comput.-Hum. Interact.*, vol. 31, no. 5, pp. 1–34, Oct. 2024, doi: 10.1145/3689433.
89. Z. Zhang, S. Sun, W. Wang, D. Cai, and J. Bian, "FlexCAD: Unified and Versatile Controllable CAD Generation with Fine-tuned Large Language Models," Nov. 05, 2024, arXiv: arXiv:2411.05823. doi: 10.48550/arXiv.2411.05823.

90. Z. Zhao et al., "ChatCAD+: Towards a Universal and Reliable Interactive CAD using LLMs," *IEEE Trans. Med. Imaging*, vol. 43, no. 11, pp. 3755–3766, Nov. 2024, doi: 10.1109/TMI.2024.3398350.
91. Y. Zhu, G. Zhao, H. Liu, D. Sun, and Y. He, "Refining Bridge Engineering-based Construction Scheme Compliance Review with Advanced Large Language Model Integration," in *Proceedings of the 2024 8th International Conference on Big Data and Internet of Things*, in BDIOT '24. New York, NY, USA: Association for Computing Machinery, 2024, pp. 297–305. doi: 10.1145/3697355.3697404.
92. Q. Zou, Y. Wu, Z. Liu, W. Xu, and S. Gao, "Intelligent CAD 2.0," *Visual Informatics*, vol. 8, no. 4, pp. 1–12, 2024, doi: 10.1016/j.visinf.2024.10.001.
93. M. F. Alam and F. Ahmed, "GenCAD: Image-Conditioned Computer-Aided Design Generation with Transformer-Based Contrastive Representation and Diffusion Priors," Apr. 2025, doi: 10.48550/arXiv.2409.16294.
94. Y. Ao, S. Li, and H. Duan, "Artificial Intelligence-Aided Design (AIAD) for Structures and Engineering: A State-of-the-Art Review and Future Perspectives," *Archives of Computational Methods in Engineering*, 2025, doi: 10.1007/s11831-025-10264-1.
95. O. Bleisinger and M. Eigner, "AI Applications in Engineering New Technologies, New Opportunities?," *ZWF Zeitschrift fuer Wirtschaftlichen Fabrikbetrieb*, vol. 120, no. s1, pp. 39–43, 2025, doi: 10.1515/zwf-2024-0173.
96. D. Byrne, V. Hargaden, and N. Papakostas, "Application of generative AI technologies to engineering design," presented at the *Procedia Computer Science*, 2025, pp. 147–152. doi: 10.1016/j.procir.2025.01.025.
97. R. P. Cardoso Coelho, A. F. Carvalho Alves, T. M. Nogueira Pires, and F. M. Andrade Pires, "A composite Bayesian optimisation framework for material and structural design," *Computer Methods in Applied Mechanics and Engineering*, vol. 434, p. 117516, Feb. 2025, doi: 10.1016/j.cma.2024.117516.
98. C. Chen et al., "CADCrafter: Generating Computer-Aided Design Models from Unconstrained Images," Apr. 2025, doi: 10.48550/arXiv.2504.04753.
99. A. Daareyni, A. Martikkala, H. Mokhtarian, and I. F. Ituarte, "Generative AI meets CAD: enhancing engineering design to manufacturing processes with large language models," *International Journal of Advanced Manufacturing Technology*, 2025, doi: 10.1007/s00170-025-15830-2.
100. A. C. Doris et al., "DesignQA: A Multimodal Benchmark for Evaluating Large Language Models' Understanding of Engineering Documentation," *Journal of Computing and Information Science in Engineering*, vol. 25, no. 2, 2025, doi: 10.1115/1.4067333.
101. A. C. Doris, M. F. Alam, A. H. Nobari, and F. Ahmed, "CAD-Coder: An Open-Source Vision-Language Model for Computer-Aided Design Code Generation," May 2025, doi: 10.48550/arXiv.2505.14646.
102. Y. Etesam, H. Cheong, M. Ataei, and P. K. Jayaraman, "Deep Generative Model for Mechanical System Configuration Design," presented at the *Proceedings of the AAAI Conference on Artificial Intelligence*, 2025, pp. 16496–16504. doi: 10.1609/aaai.v39i16.33812.
103. N. Heidari and A. Iosifidis, "Geometric Deep Learning for Computer-Aided Design: A Survey," July 2025, doi: 10.48550/arXiv.2402.17695.

104. B. T. Jones, F. Hähnlein, Z. Zhang, M. Ahmad, V. Kim, and A. Schulz, "A Solver-Aided Hierarchical Language for LLM-Driven CAD Design," Feb. 2025, doi: 10.48550/arXiv.2502.09819.
105. J. Kim, "On Transdisciplinary Research through Data Science and Engineering Education," in Proceedings of the 2024 16th International Conference on Education Technology and Computers, in ICETC '24. New York, NY, USA: Association for Computing Machinery, 2025, pp. 523–528. doi: 10.1145/3702163.3702465.
106. J. Li, W. Ma, X. Li, Y. Lou, G. Zhou, and X. Zhou, "CAD-Llama: Leveraging Large Language Models for Computer-Aided Design Parametric 3D Model Generation," June 2025, doi: 10.48550/arXiv.2505.04481.
107. K.-Y. Li, C.-K. Huang, Q.-W. Chen, H.-C. Zhang, and T.-T. Tang, "Generative AI and CAD automation for diverse and novel mechanical component designs under data constraints," Discover Applied Sciences, vol. 7, no. 4, 2025, doi: 10.1007/s42452-025-06833-5.
108. X. Li, Y. Sun, and Z. Sha, "LLM4CAD: Multimodal Large Language Models for Three-Dimensional Computer-Aided Design Generation," Journal of Computing and Information Science in Engineering, vol. 25, no. 2, 2025, doi: 10.1115/1.4067085.
109. X. Li and Z. Sha, "Image2CADSeq: Computer-Aided Design Sequence and Knowledge Inference from Product Images," Jan. 2025, doi: 10.48550/arXiv.2501.04928.
110. X. Liu, J. A. Erkoyuncu, J. Y. H. Fuh, W. F. Lu, and B. Li, "Knowledge extraction for additive manufacturing process via named entity recognition with LLMs," Robotics and Computer-Integrated Manufacturing, vol. 93, p. 102900, June 2025, doi: 10.1016/j.rcim.2024.102900.
111. Z. Liu, Y. Chai, and J. Li, "Toward Automated Simulation Research Workflow through LLM Prompt Engineering Design," Journal of Chemical Information and Modeling, vol. 65, no. 1, pp. 114–124, 2025, doi: 10.1021/acs.jcim.4c01653.
112. T. Mao, S. Yang, and B. Fu, "A Multi-Agent Framework for Multi-Source Manufacturing Knowledge Integration and Question Answering," in Companion Proceedings of the ACM on Web Conference 2025, in WWW '25. New York, NY, USA: Association for Computing Machinery, 2025, pp. 1687–1695. doi: 10.1145/3701716.3716884.
113. W. P. McCarthy et al., "mrCAD: Multimodal Refinement of Computer-aided Designs," Apr. 2025, doi: 10.48550/arXiv.2504.20294.
114. A. Murphy et al., "An Analysis of Decoding Methods for LLM-based Agents for Faithful Multi-Hop Question Answering," Mar. 2025, doi: 10.48550/arXiv.2503.23415.
115. K. J. Offor, "Leveraging Generative AI to Simulate Stakeholder Involvement in the Engineering Design Process: A Case Study of MSc Team-Based Projects," presented at the IEEE Global Engineering Education Conference, EDUCON, 2025. doi: 10.1109/EDUCON62633.2025.11016557.
116. C. Picard et al., "From concept to manufacturing: evaluating vision-language models for engineering design," Artificial Intelligence Review, vol. 58, no. 9, 2025, doi: 10.1007/s10462-025-11290-y.
117. U. L. Roncoroni, V. Crousse de Vallongue, and O. Centurion Bolaños, "Computational creativity issues in generative design and digital fabrication of complex 3D meshes," International Journal of Architectural Computing, vol. 23, no. 2, pp. 582–600, 2025, doi: 10.1177/14780771241260850.
118. S. Wang et al., "CAD-GPT: Synthesising CAD Construction Sequence with Spatial Reasoning-Enhanced Multimodal LLMs," AAAI, vol. 39, no. 8, pp. 7880–7888, Apr. 2025, doi: 10.1609/aaai.v39i8.32849.

119. B. Yang, J. J. Dudley, and P. O. Kristensson, "Design Activity Simulation: Opportunities and Challenges in Using Multiple Communicative AI Agents to Tackle Design Problems," in *Proceedings of the 7th ACM Conference on Conversational User Interfaces*, in CUI '25. New York, NY, USA: Association for Computing Machinery, 2025. doi: 10.1145/3719160.3736609.
120. C. Zhang et al., "A survey on potentials, pathways and challenges of large language models in new-generation intelligent manufacturing," *Robotics and Computer-Integrated Manufacturing*, vol. 92, p. 102883, Apr. 2025, doi: 10.1016/j.rcim.2024.102883.
121. L. Zhang, B. Le, N. Akhtar, S.-K. Lam, and T. Ngo, "Large Language Models for Computer-Aided Design: A Survey," May 2025, doi: 10.48550/arXiv.2505.08137.
122. J. Zhou and J. D. Camba, "The status, evolution, and future challenges of multimodal large language models (LLMs) in parametric CAD," *Expert Systems with Applications*, vol. 282, 2025, doi: 10.1016/j.eswa.2025.127520.

References

1. A. Wang et al., 'SuperGLUE: A Stickier Benchmark for General-Purpose Language Understanding Systems'.
2. A. Chernyavskiy, D. Ilvovsky, and P. Nakov, 'Transformers: "The End of History" for Natural Language Processing?', in *Machine Learning and Knowledge Discovery in Databases. Research Track*, N. Oliver, F. Pérez-Cruz, S. Kramer, J. Read, and J. A. Lozano, Eds, Cham: Springer International Publishing, 2021, pp. 677–693. doi: 10.1007/978-3-030-86523-8_41.
3. G. Penedo et al., 'The RefinedWeb Dataset for Falcon LLM: Outperforming Curated Corpora with Web Data, and Web Data Only', June 01, 2023, arXiv: arXiv:2306.01116. doi: 10.48550/arXiv.2306.01116.
4. B. Zhang and H. Soh, 'Large Language Models as Zero-Shot Human Models for Human-Robot Interaction', in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Oct. 2023, pp. 7961–7968. doi: 10.1109/IROS55552.2023.10341488.
5. H. Yang, X.-Y. Liu, and C. D. Wang, 'FinGPT: Open-Source Financial Large Language Models', June 09, 2023, arXiv: arXiv:2306.06031. doi: 10.48550/arXiv.2306.06031.
6. S. Wu et al., 'BloombergGPT: A Large Language Model for Finance', Dec. 21, 2023, arXiv: arXiv:2303.17564. doi: 10.48550/arXiv.2303.17564.
7. S. Samsi et al., 'From Words to Watts: Benchmarking the Energy Costs of Large Language Model Inference', Oct. 04, 2023, arXiv: arXiv:2310.03003. doi: 10.48550/arXiv.2310.03003.
8. A. A. Chien, L. Lin, H. Nguyen, V. Rao, T. Sharma, and R. Wijayawardana, 'Reducing the Carbon Impact of Generative AI Inference (today and in 2035)', in *Proceedings of the 2nd Workshop on Sustainable Computer Systems*, in HotCarbon '23. New York, NY, USA: Association for Computing Machinery, Aug. 2023, pp. 1–7. doi: 10.1145/3604930.3605705.
9. J. Kaddour, J. Harris, M. Mozes, H. Bradley, R. Raileanu, and R. McHardy, 'Challenges and Applications of Large Language Models', July 19, 2023, arXiv: arXiv:2307.10169. doi: 10.48550/arXiv.2307.10169.
10. L. Makatura et al., 'Large Language Models for Design and Manufacturing', *MIT Explor. Gener. AI*, Mar. 2024, doi: 10.21428/e4baedd9.745b62fa.
11. Y. Li et al., 'Large Language Models for Manufacturing', Oct. 28, 2024, arXiv: arXiv:2410.21418. doi: 10.48550/arXiv.2410.21418.
12. C. Zhang et al., 'A survey on potentials, pathways and challenges of large language models in new-generation intelligent manufacturing', *Robot. Comput.-Integr. Manuf.*, vol. 92, p. 102883, Apr. 2025, doi: 10.1016/j.rcim.2024.102883.

13. D. J. Kasik, W. Buxton, and D. R. Ferguson, 'Ten CAD challenges', *IEEE Comput. Graph. Appl.*, vol. 25, no. 2, pp. 81–92, Mar. 2005, doi: 10.1109/MCG.2005.48.
14. T. Heikkinen, J. Johansson, and F. Elgh, 'Review of CAD-model capabilities and restrictions for multidisciplinary use', *Comput.-Aided Des. Appl.*, vol. 15, no. 4, pp. 509–519, July 2018, doi: 10.1080/16864360.2017.1419639.
15. A. Abdul-Aziz, G. Y. Baaklini, D. Zagidulin, and R. W. Rauser, 'Challenges in integrating nondestructive evaluation and finite-element methods for realistic structural analysis', in *Nondestructive Evaluation of Aging Materials and Composites IV*, SPIE, May 2000, pp. 35–46. doi: 10.1117/12.385504.
16. K. Li, A. K. Hopkins, D. Bau, F. Viégas, H. Pfister, and M. Wattenberg, 'Emergent World Representations: Exploring a Sequence Model Trained on a Synthetic Task', June 26, 2024, arXiv: arXiv:2210.13382. doi: 10.48550/arXiv.2210.13382.
17. L. Xia, C. Li, C. Zhang, S. Liu, and P. Zheng, 'Leveraging error-assisted fine-tuning large language models for manufacturing excellence', *Robot. Comput.-Integr. Manuf.*, vol. 88, p. 102728, Aug. 2024, doi: 10.1016/j.rcim.2024.102728.
18. J. Xu, C. Wang, Z. Zhao, W. Liu, Y. Ma, and S. Gao, 'CAD-MLLM: Unifying Multimodality-Conditioned CAD Generation With MLLM', Nov. 07, 2024, arXiv: arXiv:2411.04954. doi: 10.48550/arXiv.2411.04954.
19. L. Chen et al., 'Toward Controllable Generative Design: A Conceptual Design Generation Approach Leveraging the Function–Behavior–Structure Ontology and Large Language Models', *J. Mech. Des.*, vol. 146, no. 12, p. 121401, Dec. 2024, doi: 10.1115/1.4065562.
20. A. Ganeshan, R. Huang, X. Xu, R. K. Jones, and D. Ritchie, 'ParSEL: Parameterized Shape Editing with Language', *ACM Trans. Graph.*, vol. 43, no. 6, pp. 1–14, Dec. 2024, doi: 10.1145/3687922.
21. J. Wei et al., 'Chain-of-Thought Prompting Elicits Reasoning in Large Language Models', Jan. 10, 2023, arXiv: arXiv:2201.11903. doi: 10.48550/arXiv.2201.11903.
22. T. Rios, F. Lanfermann, and S. Menzel, 'Large Language Model-assisted Surrogate Modelling for Engineering Optimization'.
23. K. Alrashedy, P. Tambwekar, Z. Zaidi, M. Langwasser, W. Xu, and M. Gombolay, 'Generating CAD Code with Vision-Language Models for 3D Designs', Oct. 07, 2024, arXiv: arXiv:2410.05340. doi: 10.48550/arXiv.2410.05340.
24. X. Li, Y. Sun, and Z. Sha, 'LLM4CAD: Multi-Modal Large Language Models for 3D Computer-Aided Design Generation', in *Volume 6: 36th International Conference on Design Theory and Methodology (DTM)*, Washington, DC, USA: American Society of Mechanical Engineers, Aug. 2024, p. V006T06A015. doi: 10.1115/DETC2024-143740.
25. S. Wang et al., 'CAD-GPT: Synthesising CAD Construction Sequence with Spatial Reasoning-Enhanced Multimodal LLMs', Dec. 27, 2024, arXiv: arXiv:2412.19663. doi: 10.48550/arXiv.2412.19663.
26. B. Song, R. Zhou, and F. Ahmed, 'Multi-modal Machine Learning in Engineering Design: A Review and Future Directions', July 28, 2023, arXiv: arXiv:2302.10909. doi: 10.48550/arXiv.2302.10909.
27. Y. Liu, A. Obukhov, J. D. Wegner, and K. Schindler, 'Point2CAD: Reverse Engineering CAD Models from 3D Point Clouds', Dec. 07, 2023, arXiv: arXiv:2312.04962. doi: 10.48550/arXiv.2312.04962.
28. Y. Liu, J. Chen, S. Pan, D. Cohen-Or, H. Zhang, and H. Huang, 'Split-and-Fit: Learning B-Reps via Structure-Aware Voronoi Partitioning', June 07, 2024, arXiv: arXiv:2406.05261. doi: 10.48550/arXiv.2406.05261.
29. Y. Cao, A. Taghvaie Nakhjiri, and M. Ghadiri, 'Different applications of machine learning approaches in materials science and engineering: Comprehensive review', *Eng. Appl. Artif. Intell.*, vol. 135, p. 108783, Sept. 2024, doi: 10.1016/j.engappai.2024.108783.
30. A. Frome et al., 'DeViSE: A Deep Visual-Semantic Embedding Model'.

31. Y.-H. H. Tsai, P. P. Liang, A. Zadeh, L.-P. Morency, and R. Salakhutdinov, 'Learning Factorized Multimodal Representations', May 14, 2019, arXiv: arXiv:1806.06176. doi: 10.48550/arXiv.1806.06176.
32. S. Reed, Z. Akata, B. Schiele, and H. Lee, 'Learning Deep Representations of Fine-grained Visual Descriptions', May 17, 2016, arXiv: arXiv:1605.05395. doi: 10.48550/arXiv.1605.05395.
33. R. Bracewell, K. Wallace, M. Moss, and D. Knott, 'Capturing design rationale', *Comput.-Aided Des.*, vol. 41, no. 3, pp. 173–186, Mar. 2009, doi: 10.1016/j.cad.2008.10.005.
34. K. Jokinen, M. Hajda, and J. Borgman, 'CHALLENGES WITH PRODUCT DATA EXCHANGE IN PRODUCT DEVELOPMENT NETWORKS'.
35. W. Li, G. Mac, N. G. Tsoutsos, N. Gupta, and R. Karri, 'Computer aided design (CAD) model search and retrieval using frequency domain file conversion', *Addit. Manuf.*, vol. 36, p. 101554, Dec. 2020, doi: 10.1016/j.addma.2020.101554.
36. J. Kaplan et al., 'Scaling Laws for Neural Language Models', Jan. 23, 2020, arXiv: arXiv:2001.08361. doi: 10.48550/arXiv.2001.08361.
37. S. Inayathullah and R. Buddala, 'Review of machine learning applications in additive manufacturing', *Results Eng.*, vol. 25, p. 103676, Mar. 2025, doi: 10.1016/j.rineng.2024.103676.
38. A. C. Tricco et al., 'PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation', *Ann. Intern. Med.*, vol. 169, no. 7, pp. 467–473, Oct. 2018, doi: 10.7326/M18-0850.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.