"OPTIMIZING REGULATED CONTENT CREATION WITH GENERATIVE AI: A MIXED-METHODS STUDY OF AVIATIONGPT IN MRO MARKETING"

Research Paper

Elham Biglar, DBA, Swiss Business and Management (SSBM), elham.biglar22@gmail.com

”Abstract”

Aviation maintenance, repair, and overhaul (MRO) content marketing requires technical precision and regulatory compliance, often sacrificing efficiency. This study evaluates AviationGPT in revolutionizing MRO content creation. Employing a mixed-methods approach, ANOVA, Tukey's HSD, and expert evaluations, we compared AI-generated, AI-enhanced, and human-generated content across 180 samples. Results demonstrate that AI-initiated content refined by human expertise significantly outperforms fully human-driven and human-initiated, AI-refined methods in consistency, scalability, and engagement. A proposed human-AI collaboration framework integrates AI's automation with human oversight across ideation, writing, distribution, and continuous improvement, ensuring compliant, high-impact content. This research provides actionable insights for regulated industries, advancing operational efficiency while upholding ethical governance and strategic alignment, in support of the United Nations Sustainable Development Goals (SDGs).

Keywords: AviationGPT, Aviation MRO, Content Marketing, Generative AI.

1 Introduction

In an era where digital transformation affects nearly all sectors, from low risk industries suc as education, retail, hospitality, entertainment, and customer service to critical high-stakes industries, like maintenance, repair, and overhaul (MRO), healthcare, finance, aerospace, and even autonomous transportation, the advent of Generative Artificial Intelligence (GenAI) has become a pivotal factor to maintain a competitive. This technological shift aligns with the United Nations SDGs, particularly SDG 9 (Industry, Innovation, and Infrastructure), as it fosters technological advancement while maintaining compliance in highly regulated industries, SDG 8 (Decent Work and Economic Growth) and SDG 12 (Responsible Consumption and Production) by enhancing operational efficiency while minimizing resource waste, promoting sustainable industrial practices (Rane, 2023; KJ and Binu, 2024; Ziemba *et al.*, 2024).

Marketing strategies in high-stakes industries typically require extensive labor, characterized by multiple layers of legal reviews, compliance checks, and stakeholder approvals to create content pieces that adhere to stringent regulatory requirements, compliance standards, and ethical principles (Shah, 2023). While these stages are vital, they can slow down content generation, decrease agility, and increase costs.

The emergence of various general-purpose and domain-specific GenAI models has revolutionized content creation by introducing automated processes that enhance efficiency, personalization, and engagement, enabling businesses to scale marketing efforts with AI-generated content (Islam *et al.*, 2024). However, despite compelling evidence approving the utilization of these GenAI for its outstanding abilities to produce extensive volumes of high-quality content, significant challenges persist, primarily related to concerns regarding data accuracy, the risks of misinformation, and the

necessity for legal accountability (Zhiheng, Rui and Tao, 2023; Dimitrieska, 2024). Addressing these challenges is critical to advancing SDG 16 (Peace, Justice, and Strong Institutions), which emphasizes transparency, accountability, and ethical governance as foundational pillars for trustworthy AI adoption in regulated industries (Cordova and Celone, 2019).

In contrast to other industries where creative marketing content may allow for greater artistic liberty, aviation MRO, as our focused industry, requires precision-driven narratives that are supported by technical accuracy and compliance-driven messaging (Nguyet, 2024). In aviation MROs' traditional marketing approach, content creation is usually quite complex and involves subject matter experts, compliance teams, and marketing professionals to create content that is not only relevant and engaging but also compliant with industry and legal requirements.

Streamlining these processes through integrating GenAI technologies in this industry raises the question of how to guarantee that AI-generated output will possess the same level of expertise and compliance as human-written content (Tan *et al.*, 2024). It is also essential to evaluate AI-generated content for its alignment with company branding initiatives and adherence to aviation safety requirements, thus averting the dissemination of false or non-compliant information (Canaday, 2023; Shaw, 2024). Furthermore, while AI algorithms offer efficiency and scalability, overreliance on them can lead to outputs that are not always relevant or proportional to the length or complexity of the input text, raising concerns about the quality and credibility of the marketing materials. This means that there is a need for a collaborative Human-AI framework that would combine the computational power of AI, human expertise, ethical oversight, and strategic judgment. When used correctly, AI can serve as a tool to enhance the content development process instead of taking over it and making content inaccurate, inconsistent, and incompatible with the brand.

Building on these insights, this study critically assesses GenAI's effectiveness in aviation MRO content marketing by identifying its benefits, challenges, and regulatory implications. Using "AviationGPT" as a specialized GenAI model designed for the aviation industry, we conducted a direct comparison between three groups of content. The first group is human-generated content; the second group is those generated by AviationGPT as AI-generated, and the third group is the human-written content pieces enhanced with this model as AI-enhanced. This evaluation provided a structured evaluation of how AviationGPT performs in real-world aviation MRO applications and contributed to the proposal of a thorough framework for incorporating GenAI as a collaborative partner with humans in content development operations.

2 Literature Review

Generative Artificial Intelligence (GenAI), particularly large language models (LLMs), has emerged as a transformative force in content creation across various sectors, attracting considerable academic attention due to its ability to produce human-like, coherent, and engaging text. Empirical studies have demonstrated that LLMs enhance readability, clarity, and conciseness in human-authored content, making them valuable tools for academic writing and professional communication (Gmeiner and Yildirim, 2023; Moore *et al.*, 2023; Bernal, 2024). However, despite these advantages, a persistent concern is "stochastic parroting," where models mimic linguistic patterns without genuine understanding, posing significant risks in high-stakes domains such as aviation, where technical precision and regulatory compliance are paramount (Curtis, 2023). This challenge has catalyzed the development of domain-specific GenAI models, such as AviationGPT, which builds upon LLaMA-2 and Mistral architectures and is continuously fine-tuned using aviation-specific datasets to generate industry-aligned, regulation-compliant content (Wang *et al.*, 2023). Yet, research evaluating the effectiveness of such models in generating high-quality, regulation-compliant marketing content remains limited (Floridi and Chiriatti, 2020; Dale, 2021; Jandhyala, 2024).

In regulated sectors like aviation MRO, where documentation must meet strict safety and legal standards, the potential of GenAI models, both general-purpose (e.g., GPT-4) and specialized (e.g., AviationGPT), to improve content generation is increasingly recognized (Canaday, 2023; Bhatia *et al.*,

2024). These tools support tasks ranging from technical writing to stakeholder-facing marketing content, offering significant time savings. However, the risk of inaccuracies or noncompliance limits their unsupervised application, especially when communicating safety-critical information (Shah, 2023; Zhiheng, Rui, and Tao, 2023; Shaw, 2024). Consequently, the concept of human-AI collaboration has gained traction as a mitigation strategy. Grounded in cognitive load theory (Sweller, 1988). human-AI co-creation allows AI to handle routine or syntactic tasks, such as grammar correction and boilerplate drafting, while enabling human experts to focus on strategic and compliance-oriented elements (Jakesch, Hancock, and Naaman, 2023). This approach has shown promising results in content marketing, particularly in enhancing ideation efficiency and aligning outputs with brand and regulatory standards (Biglar and Palivela, 2024).

Nevertheless, while general marketing literature underscores the value of AI-human collaboration, few structured frameworks exist to support its application in regulated sectors like aviation MRO, where ethical considerations, such as algorithmic bias, lack of transparency, and the need for institutional accountability, pose additional challenges (Cordova and Celone, 2019). These ethical and procedural constraints align with the goals of Sustainable Development Goal 16 (Peace, Justice, and Strong Institutions), emphasizing the need for responsible AI governance (Tan et al., 2024; Patel & Brown, 2023). As such, there is a critical research gap in evaluating the operational viability and regulatory alignment of domain-specific GenAI models like AviationGPT. Addressing this gap is essential for developing robust human-AI collaboration models tailored to the unique needs of the aviation MRO sector.

3 Evaluating the AviationGPT in MRO Content Creation

This study employed a mixed-methods research design to evaluate the performance of AviationGPT in generating marketing content tailored to the aviation MRO sector. The evaluation compared three groups of content: AI-H (initial drafts generated by AviationGPT and subsequently refined by human experts), H-AI (initial drafts generated by humans and subsequently refined by AI tools), and H (content fully authored by humans). A total of 180 content samples (60 per group), each ranging between 600 and 900 words, were developed in response to standardized, MRO-specific marketing prompts to ensure consistency across all groups.

To evaluate content quality, the study used both automated and expert assessments. Grammarly provided composite scores for readability, clarity, and correctness, while six domain experts rated informativeness and engagement using a 5-point Likert scale. Final scores combined both sources. Data were analyzed using descriptive statistics, one-way ANOVA, and Tukey's HSD to identify significant differences among AI-generated, AI-enhanced, and human-generated content, enabling a robust comparison of AviationGPT's performance in MRO marketing.

3.1 Performance comparison results

The descriptive statistics in Table 1 reveal that AI-H content (mean = 131.35/150) significantly outperforms both H-AI (mean = 129.18) and H (mean = 118.87) in total evaluation scores. The Grammarly and expert components of the total score each reflect a similar hierarchy, indicating that integrating AI for initial drafts, refined by human experts, yields the most balanced output in terms of technical correctness, readability, and contextual engagement.

The ANOVA results are in Table 2, confirming statistically significant differences among the three content groups across all evaluation dimensions. Specifically, the F-values for the Grammarly score ($F = 58.02$, $p < 0.001$), expert evaluation score ($F = 74.41$, $p < 0.001$), and total evaluation score ($F = 96.25$, $p < 0.001$) indicate strong evidence against the null hypothesis of equal means. These findings validate the robustness of the observed differences and justify post hoc pairwise comparisons to identify where the disparities lie.

The Tukey's HSD test in Table 3 provides further granularity. The AI-H and H-AI comparison did not reach conventional statistical significance ($p = 0.066$), although the mean difference (−2.16) indicates a performance trend favoring AI-H. In contrast, both AI-H and H-AI significantly outperformed the fully human-generated content (H) with highly significant p-values ($p < 0.001$), reinforcing the added value of AI integration, whether at the ideation or refinement stage, in content development for high-stakes sectors.

Effect size analysis, as reported in Table 4, offers practical significance beyond statistical thresholds. The Cohen's d values show a medium effect size between AI-H and H-AI ($d = 0.53$), indicating that AI-first content refined by humans has a moderate advantage over human-first content polished by AI. Importantly, the effect sizes for AI-H vs. H ($d = 2.23$) and H-AI vs. H ($d = 1.74$) are both considered large, suggesting substantial improvements in quality, coherence, and compliance when AI is incorporated into the content creation pipeline.

Moreover, the eta-squared value ($\eta^2 = 0.521$) indicates that the independent variable "content generation method" explains 52.1% of the variance in the overall evaluation scores, representing a large effect size ($\eta^2 > 0.14$) according to Richardson's (2011) classification (Richardson, 2011). This further substantiates that the choice of content generation method exerts a statistically and practically significant influence on content quality, emphasizing the strategic importance of selecting the right AI-human integration model.

	Grammarly Overall Score (100)		**Expert Overall Score (50)**		**Total Evaluation Score (150)**	
Category	**mean**	**Standard Deviation**	**mean**	**Standard Deviation**	**mean**	**Standard Deviation**
AI-H	86.917	3.490	**44.428**	0.853	**131.345**	**3.582**
H-AI	**87.933**	3.940	41.248	2.057	129.182	**4.514**
H	80.300	5.060	38.568	3.982	118.868	**7.065**

Table 1. Descriptive Statistics for Overall Scores

	F-Stat	**P-Value**
Grammarly Overall Score (100)	58.01570259	4.20465E-20
Expert Overall Score (50)	74.40711921	3.53007E-24
Total Evaluation Score (150)	96.24848556	5.15313E-29

Table 2. ANOVA Test Result

Group1	**Group2**	**Meandiff**	**p-adj**	**Lower**	**Upper**	**Reject**
AI-H	H-AI	-2.1633	**0.0656**	-4.435	0.1083	**FALSE**
AI-H	H	-12.4767	**p<0.001**	-14.7483	-10.205	**TRUE**
H-AI	H	-10.3133	**p<0.001**	-12.585	-8.0417	**TRUE**

Table 3. Tukey's HSD Test Result

Comparison	**Cohen's d**
AI-H vs H-AI	**0.5309 (Medium effect)**
AI-H vs H	**2.2274 > 0.8 (Large effect)**

H-AI vs H	**1.7396 > 0.8 (Large effect)**

Table 4. Pairwise Cohen's d Results

Figure 1 depicts a radar chart for pairwise Cohen's d effect sizes of total evaluation scores, highlighting that the difference between AI-H and H-AI methods is relatively small, indicating comparable performance. In contrast, both AI-driven methods exhibit a substantial difference when compared to fully human-generated content, underscoring the superior consistency and quality of AI-influenced approaches.

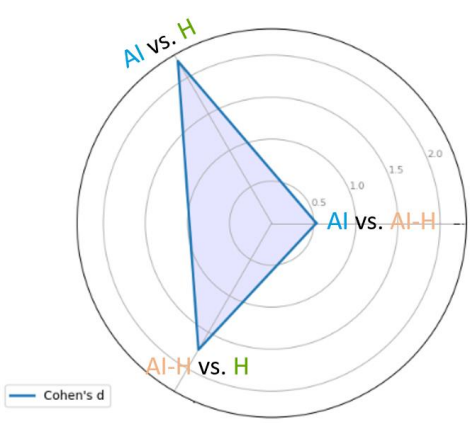

Figure 1. Radar Chart for Pairwise Cohen's d Effect Sizes of Total Evaluation Scores

3.2 Human-AI collaboration model

Building on the results of our prior investigation into the effectiveness of general-purpose generative AI models in content ideation (Biglar and Palivela, 2024), and reinforced by the empirical findings of the present study, it becomes evident that employing domain-specific GenAI as the primary content generator significantly outperforms human-centered approaches across key dimensions, including efficiency, technical accuracy, time-to-output, and consistency, even when human authors utilize AI tools in a secondary, supportive role. Integrating insights from many disciplines, such as technical terminology, compliance standards, business trends, and digital transformation initiatives, is a significant advantage of domain-specific GenAI. As a result, AI-initiated content is more comprehensive, data-driven, and in sync with market trends than human-written material, which is frequently constrained by specific knowledge and requires human research. Unlike human writers, who vary in experience, attention to detail, and writing style, AI ensures uniformity in tone and professionalism across all materials. Another major benefit is AI's built-in SEO and readability enhancements, which help businesses automatically optimize content for digital performance, integrating high-ranking keywords and improving clarity, tasks that require additional effort in human-written content.

Although human expertise is crucial for final validation, brand alignment, and creative storytelling, our research demonstrates that domain-specific GenAI significantly improves efficiency, accuracy, and consistency. By delegating drafting and structuring tasks to AI, businesses can allocate human experts to strategic enhancements, guaranteeing that content is both technically precise and captivating.

Organizations can benefit from an autonomous agent that can be built on top of this workflow to aid in content development, dissemination, and enhancement. The agent can manage the entire content life cycle, including trend analysis, content ideation and generation, platform optimization, and performance monitoring, and combine it with human validation, storytelling, and strategic oversight. This way the

collaboration can result in the creation of high-quality, most relevant and engaging content, which is further improved with the help of data and human input.

Figure 2 displays this proposed AI-driven content development workflow.

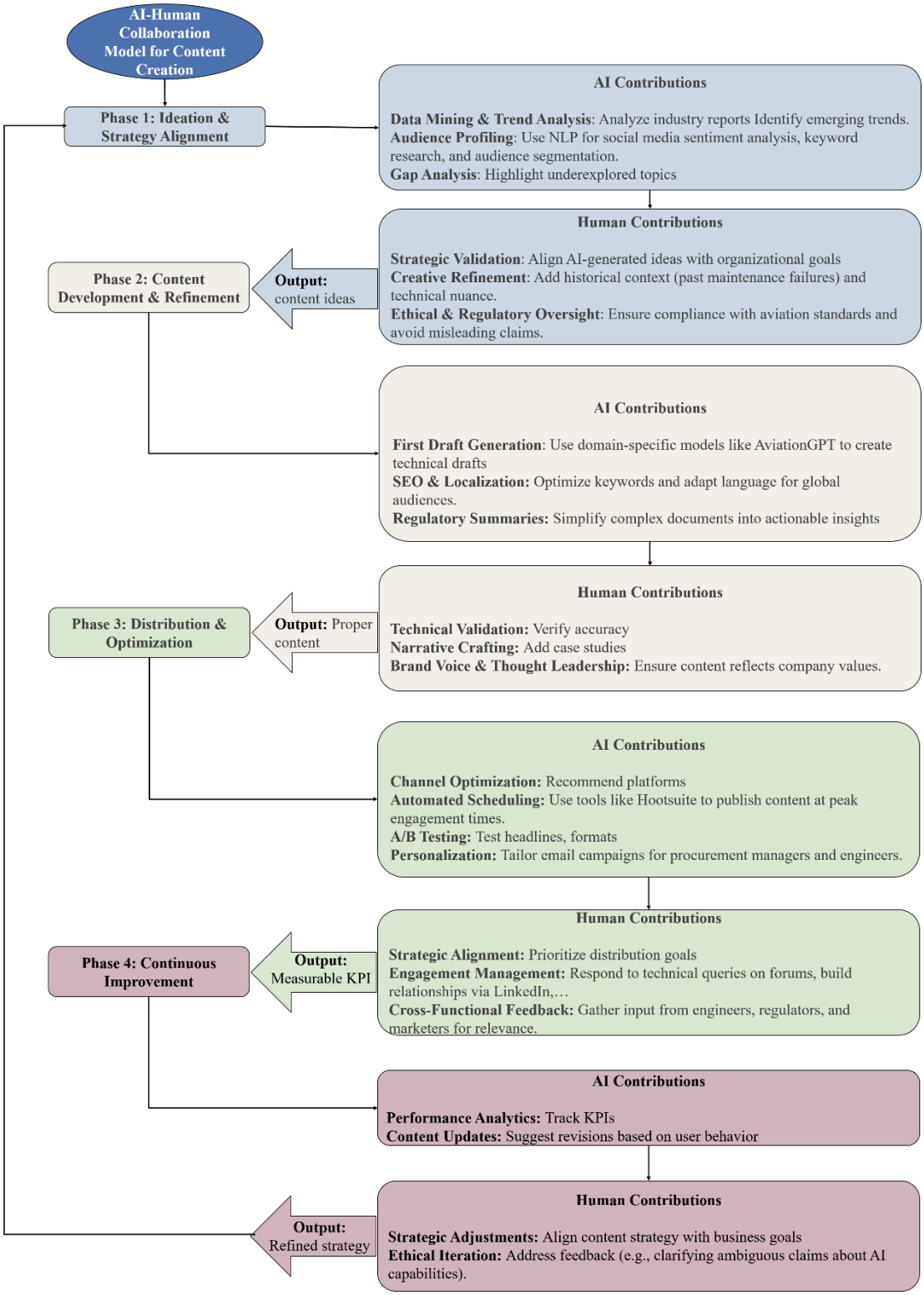

Figure 2. AI-Driven Content Development Workflow

4 SDG Alignment in AI-Human Content Creation

The AI-human collaborative framework proposed in this study aligns strategically with multiple United Nations SDGs and their respective targets. By automating high-volume, repetitive tasks in content marketing, such as ideation, SEO drafting, and performance analytics, while preserving expert oversight for compliance and strategic alignment, the model promotes inclusive industrial innovation and infrastructure (SDG 9). Specifically, it supports target 9.5 (enhance scientific research and upgrade industrial technologies) and Indicator 9.5.1 (R&D expenditure as a proportion of GDP), through AI-driven innovation.

The integration of AI into content pipelines also aligns with SDG 8 (Decent Work and Economic Growth), particularly Target 8.5, which emphasizes full and productive employment and equal pay by shifting human effort to higher-value strategic roles, enabling workforce upskilling rather than replacement.

Moreover, the framework contributes to SDG 4 (Quality Education), particularly target 4.4 (increasing the number of youth and adults with ICT skills), by necessitating continual learning for AI-literate content professionals.

It also advances SDG 12 (Responsible Consumption and Production), particularly target 12.6, which encourages companies to adopt sustainable practices, reflected in the model's reduction of redundant content workflows and its promotion of efficiency across the digital production cycle.

The model fosters transparent, ethical AI governance aligned with SDG 16, notably target 16.6 (developing accountable institutions) and target 16.10 (ensuring public access to information), by embedding AI transparency, compliance mechanisms, and human oversight to mitigate misinformation and bias.

Lastly, the framework embodies SDG 17 (Partnerships for the Goals), especially target 17.17 (promoting effective public-private-civil society partnerships), by encouraging collaborative input from compliance officers, engineers, AI developers, and marketers. This cross-functional alignment enables the co-creation of content that is not only compliant and technically accurate but also socially responsible and strategically agile. Thus, the proposed model presents a scalable blueprint for responsible, inclusive, and sustainable digital transformation in regulated sectors.

5 Conclusion

This study provides empirical evidence supporting the strategic integration of domain-specific generative AI, namely AviationGPT, within the content marketing workflows of the aviation MRO industry. By evaluating the comparative performance of AI-generated, AI-enhanced, and human-generated content across standardized metrics, the results underscore the superior effectiveness of AI-initiated, human-refined content in achieving high levels of consistency, engagement, and compliance. The proposed human-AI collaboration framework not only enhances operational efficiency but also aligns with key Sustainable Development Goals by fostering ethical innovation, workforce development, and sustainable digital transformation. As regulated industries increasingly adopt AI-driven tools, the findings of this study offer a scalable, responsible, and domain-sensitive roadmap for content creation that balances automation with human oversight. Future research should expand upon this framework by exploring longitudinal performance outcomes, industry-specific regulatory implications, and cross-sectoral generalizability to further strengthen AI's role in sustainable content ecosystems.

References

Bernal, M.E. (2024) 'Revolutionizing eLearning Assessments: The Role of GPT in Crafting Dynamic Content and Feedback', *Journal of Artificial Intelligence and Technology* [Preprint]. Available at: https://api.semanticscholar.org/CorpusID:269680963.

Bhatia, V. *et al.* (2024) 'Industry 4.0 in Aircraft Production, Maintenance, Repair, and Overhaul [MRO]: Investigating Use Cases and Associated Patent Landscape', *Recent Patents on Mechanical Engineering* [Preprint]. Available at: https://api.semanticscholar.org/CorpusID:272460434.

Biglar, E. and Palivela, H. (2024) 'Optimizing Aviation MRO Content Marketing: A Case Study Using AHP and Generative AI for Content Idea Generation', *lobal Journal of Business and Integral Security - GBIS (ISSN 2673-9690 Online)* [Preprint].

Canaday, H. (2023) 'Will Generative AI Transform The MRO Industry?', *AVIATION WEEK NETWORK, https://aviationweek.com/mro/emerging-technologies/will-generative-ai-transform-mro-industry*, 10 July.

Cordova, M.F. and Celone, A. (2019) 'SDGs and innovation in the business context literature review', *Sustainability*, 11(24), p. 7043.

Curtis, N. (2023) 'To ChatGPT or not to ChatGPT? The impact of artificial intelligence on academic publishing', *The Pediatric Infectious Disease Journal*, 42(4), p. 275.

Dale, R. (2021) 'GPT-3: What's it good for?', *Natural Language Engineering*. Cambridge University Press, pp. 113–118. Available at: https://doi.org/10.1017/S1351324920000601.

Dimitrieska, S. (2024) 'Generative Artificial Intelligence and Advertising', *Trends in Economics, Finance and Management Journal* [Preprint]. Available at: https://api.semanticscholar.org/CorpusID:270941311.

Floridi, L. and Chiriatti, M. (2020) 'GPT-3: Its nature, scope, limits, and consequences', *Minds and Machines*, 30, pp. 681–694.

Gmeiner, F. and Yildirim, N. (2023) 'Dimensions for Designing LLM-based Writing Support', in *In2Writing Workshop at CHI*.

Islam, T. *et al.* (2024) 'Transforming Digital Marketing with Generative AI', *Comput.*, 13, p. 168. Available at: https://api.semanticscholar.org/CorpusID:271082837.

Jakesch, M., Hancock, J.T. and Naaman, M. (2023) 'Human heuristics for AI-generated language are flawed', *Proceedings of the National Academy of Sciences*, 120(11), p. e2208839120.

Jandhyala, V.S.V. (2024) 'GPT-4 and Beyond : Advancements in AI Language Models', *International Journal of Scientific Research in Computer Science, Engineering and Information Technology* [Preprint]. Available at: https://api.semanticscholar.org/CorpusID:273191836.

KJ, A.M. and Binu, V.P. (2024) 'Helping Hand of AI for achieving SDG', *Journal of Applied Science, Engineering, Technology and Management*, 2(01), pp. 29–34.

Moore, S. *et al.* (2023) 'Empowering education with llms-the next-gen interface and content generation', in *International Conference on Artificial Intelligence in Education*. Springer, pp. 32–37.

Nguyet, D.T.C. (2024) 'Adoption of Generative AI in content creation: A case study from the advertising industry', *2024 IEEE Conference on Artificial Intelligence (CAI)*, pp. 111–112. Available at: https://api.semanticscholar.org/CorpusID:271573723.

Rane, N. (2023) 'Roles and challenges of ChatGPT and similar generative artificial intelligence for achieving the sustainable development goals (SDGs)', *Available at SSRN 4603244* [Preprint].

Richardson, J. (2011) 'Eta squared and partial eta squared as measures of effect size in educational research', *Educational Research Review*, 6, pp. 135–147. Available at: https://doi.org/10.1016/j.edurev.2010.12.001.

Shah, R. (2023) 'Embracing Generative AI in Pharma Regulatory Affairs - An Industry Perspective', *International Journal of Science and Research (IJSR)* [Preprint]. Available at: https://api.semanticscholar.org/CorpusID:266332207.

Shaw, B. (2024) 'Challenges & Trends in the Aircraft MRO Industry', *STS Aviation Group, https://www.stsaviationgroup.com/challenges-trends-in-the-aircraft-mro-industry/,* 24 January.

Sweller, J. (1988) 'Cognitive load during problem solving: Effects on learning', *Cognitive science*, 12(2), pp. 257–285.

Tan, Y.H. *et al.* (2024) 'Current Landscape of Generative AI: Models, Applications, Regulations and Challenges', *2024 IEEE 14th International Conference on Control System, Computing and Engineering (ICCSCE)*, pp. 168–173. Available at: https://api.semanticscholar.org/CorpusID:273048845.

Wang, L. *et al.* (2023) *AviationGPT: A Large Language Model for the Aviation Domain*.

Zhiheng, X., Rui, Z. and Tao, G. (2023) 'Safety and ethical concerns of large language models', in *Proceedings of the 22nd Chinese National Conference on Computational Linguistics (Volume 4: Tutorial Abstracts)*, pp. 9–16.

Ziemba, E.W. *et al.* (2024) 'Leveraging artificial intelligence to meet the sustainable development goals', *Journal of Economics and Management*, 46, pp. 508–583.