LLM-based System for Technical Writing Real-time Review in Urban Construction and Technology

Laura Cruz-Castro, Ph.D., Gabriel Castelblanco, Ph.D., and Pavlo Antonenko, Ph.D.
University of Florida
Gainesville, Florida

This pilot study addresses the critical challenge of providing timely and comprehensive feedback on technical reports in Architectural, Engineering, and Construction (AEC) education. The significance of effective writing skills in the AEC industry is paramount for technical reports. Traditional methods of evaluation and feedback are time-consuming, leading to a deficiency in students' exposure to writing practice and hindering their holistic skill development. This study introduces a Large Language Model (LLM)-based System for Real-time Technical Writing Review in AEC, aiming to assess the reliability of LLMs in offering constructive feedback and grading for technical reports, focusing on Construction Capstone Projects. The proposed system aligns with pedagogical frameworks, such as Writing Across the Curriculum and AI Across the Curriculum, generative learning theory, and the Feedback Model. This model is applied Construction Capstone Project at the University of Florida, focusing on targeted writing assignments. Preliminary results indicate the model's ability to evaluate sustainability aspects of projects, providing detailed criteria-based feedback. This pilot study aims to lay the groundwork for an AI-assisted system tailored for AEC education, offering real-time, personalized feedback to enhance students' writing skills. The findings hold implications for researchers, students, and educators seeking innovative solutions to address the challenges in technical writing education within AEC.

Key Words: Large Language Model, Writing Across the Curriculum, AI Across the Curriculum, generative learning theory, feedback model

Introduction

In environments dedicated to fostering the growth of engineering formation in AEC (Architectural, Engineering, Construction) domains, it is expected that professionals may be able to generate high-quality technical reports that serve as critical documents that encapsulate detailed information about various aspects of a construction project (Castelblanco & Guevara, 2024; Castelblanco et al., 2024). These reports play a pivotal role in ensuring effective communication, facilitating informed decision-making, and documenting the comprehensive lifecycle of projects providing and systematic overview of the various stages and components of a construction project, enabling stakeholders to gain a holistic understanding of the project's progress, challenges, and outcomes (Khan, 2019). These reports serve as a repository of essential information related to the project's design, planning, execution, and completion phases, documenting the progress of the project (Hering, 2019). Technical reports also serve as a crucial communication tool, facilitating effective interaction and information exchange among various stakeholders, including project managers, engineers, clients, contractors, and regulatory authorities.
Despite the utmost relevance of writing adequate technical reports, instructors of AEC courses are not able to incorporate this within evaluation rubrics nor provide relevant feedback on writing skills to students because of the excessive time required to analyze these documents through traditional analytical techniques (Ely & Chen, 2021). Moreover, in the few cases where the instructor is willing to provide some feedback regarding writing, the feedback is not timely due to the lengthy process of reviewing and elaborating detailed feedback. Further challenges to providing relevant feedback to students are related to scalability issues, time consumption, bias, interrater reliability, and intrarater reliability (Stevens & Levi, 2023).

Lacking proper feedback introduces deficiencies in student exposure to attain deeper and stronger learning experiences related to writing technical reports within the context of AEC education. There is a significant scarcity of learning opportunities to practice writing skills in technical contexts, resulting in restricting students’ knowledge understanding, information retention, creativity, and critical thinking (Choudhury et al., 2003). The challenge, therefore, is about how students can receive comprehensive and timely feedback to develop such skills while instructors can enhance their rubrics to assess these skills.

In the initial phase of developing a Large Language Model (LLM)-based System for Real-time Technical Writing Review in AEC, this pilot study seeks to assess the reliability of LLMs in providing feedback and evaluating technical reports within the context of Construction Capstone Projects.

The primary objective of this pilot study is to explore the potential of LLMs in offering constructive feedback and grading for technical reports in the field of AEC. The outcomes of this investigation hold significant implications for researchers, students, and educators. Furthermore, the findings will lay the groundwork for the future development of a comprehensive AI-assisted system aimed at enhancing writing skills specifically tailored for technical reports within AEC education.

**Theoretical Background**

Teaching writing skills to Architectural, Engineering, and Construction (AEC) students is of paramount importance, especially from the engineering formation perspective when considering their future roles in the AEC field (Conrad, 2017). Effective writing in the AEC industry plays a crucial role in ensuring clarity, precision, and coherence in the communication of complex technical information (Yoritomo et al., 2018). Clear written documentation is vital in professional practice, from project proposals to construction specifications, as it reduces the risk of misunderstandings and errors (American Institute of Architects, 2020). Engineers also rely on written reports and documentation to convey critical engineering concepts, plans, and project updates, which underscores the necessity of proficient writing skills in this field. Additionally, construction professionals heavily rely on written specifications to ensure that construction projects meet regulatory requirements and quality standards. Teaching AEC students writing skills not only enhances their ability to communicate effectively but also enables them to excel in an industry where precise and comprehensive documentation is crucial for project success and regulatory compliance (Conrad et al. 2012).

Despite the importance of writing skills in the professional development of students of AEC programs, the focus on teaching these essential skills remains relatively inadequate. In instances where writing courses are offered (if offered), they are often concentrated within the initial years of the curriculum, typically when students are still in the process of acquiring foundational disciplinary and technical knowledge (Choudhury et al., 2003). However, as students progress through the academic journey and gain in-depth expertise in their respective fields, the provision of comprehensive, targeted, and timely feedback on their writing tends to diminish significantly, leading to a gap in their holistic skill development (Yoritomo et al., 2018). Furthermore, furnishing proper feedback on writing technical reports by instructors poses a considerable challenge, primarily due to the imperative of delivering feedback in a timely and effective manner (Dawson et al., 2019). Timely and effective feedback stands as a cornerstone in the pedagogical process, playing a pivotal role in students’ writing abilities by helping them discern errors, offering clear paths for correction, and...
fostering continuous development. The significance of timely feedback lies in its ability to guide students toward a deeper understanding of their writing, preventing them from veering off course and ensuring a meaningful connection between the feedback and their ongoing writing processes. Delays in feedback run the risk of students losing focus on their writing objectives, hindering the assimilation of valuable insights and impeding the overall learning. Insufficient feedback may fall short of providing essential guidance, creating challenges for students striving to master the nuances of effectively writing technical reports. Consequently, integrating timely and effective feedback into writing technical reports instruction becomes imperative for AEC students, serving as a vital link between instructional efforts and the acquisition of proficient writing skills (Graham & Perin, 2007; Hattie & Timperley, 2007).

In this pilot study, the construction domain has been selected as the focus and test-bed of the project considering the inherent complexity associated with writing technical reports within the AEC field. Creating these reports demands a unique blend of generic writing skills and an in-depth understanding of the technical intricacies related to the conceptualization, feasibility, design, and construction of buildings and infrastructure. AEC students are required to integrate knowledge from diverse disciplines, including architecture, construction management, civil engineering, structural engineering, and mechanical engineering, into a coherent and comprehensive report. The challenge intensifies as students must not only grasp the technical aspects but also effectively articulate their ideas in writing, a skill particularly daunting for students who still developing their expertise. Compounding this, the difficulties in teaching technical report writing are prevalent, given that instructors in AEC fields typically possess backgrounds in architecture, engineering, or construction management. While their technical expertise is robust, their lack of formal training in writing or composition poses a hurdle in imparting nuances such as structure, style, and clarity to students. AEC programs, emphasizing technical knowledge and skills, often lead instructors to prioritize teaching the technical aspects of the field, inadvertently overlooking the critical importance of effective communication through technical writing. Furthermore, the evaluation of technical writing using traditional methods is subjective and time-consuming, leaving instructors grappling with providing timely and comprehensive feedback on a substantial volume of technical reports. This hinders the effectiveness of the learning process within engineering formation. The coexistence of students’ challenges and instructors’ limitations in providing timely and comprehensive feedback frequently results in the exclusion of writing assessments from rubrics. This, in turn, discourages students from actively improving their writing skills. These barriers compound existing disparities among students, potentially disadvantaging minorities and underrepresented groups who may have received less instruction on writing before commencing their undergraduate programs. This pilot study aims to serve as a foundation for supporting AEC instructors (and eventually all STEM educators) to fill the existing teaching and learning gap in writing technical reports.

Pedagogically, an AI-based feedback agent may be conceptualized as a computer-mediated communication (CMC) and a computer-supported collaborative learning (CSCL) environment based on the premise that this agent serves as a medium through which users receive information and engage in a communicative exchange and provides feedback on collaborative projects, fostering discussions. CMC is defined as the use of computers and digital technologies as mediums for communication between individuals or groups (Yao & Ling, 2020). CSCL is an educational approach that leverages computer technology and digital tools to facilitate collaborative learning experiences among students (Jeong & Hmelo-Silver, 2016). We envision that the pilot project presented in this paper may be a basis for creating a novel CSCL learning environment, fostering thinking by providing students with extended periods of scaffolded, self-controlled reflection on feedback for writing technical reports and encouraging active cognitive engagement (Moore, 2002; Tanner, 2012). The learning strategy in this pilot is underpinned by the social agency principle, emphasizing the role of social cues provided via personalization. Through continuous training based on past students’ technical reports, LLM models may help students achieve deeper cognition and improved learner performance (Lachner et al., 2017).
These cues can potentially enhance emotional engagement with feedback and peer collaboration, resulting in motivationally and cognitively rich problem-solving interactions (Camfield, 2016).

Furthermore, strengthening writing skills in AEC students aligns with the Writing Across the Curriculum (WAC) framework, an interdisciplinary pedagogical approach integrating writing into diverse academic disciplines. This framework underscores faculty collaboration and the creation of writing-intensive courses across the curriculum (Yancey, 1999). Complementary, AI Across the Curriculum encompasses all AI applications and aligns with the interdisciplinary character of WAC for imparting AI knowledge and skills to the upcoming generation (Southworth et al., 2023). Recognizing the significance of training in implementing WAC and AI Across the Curriculum initiatives, this pilot study aims to enhance engineering formation (McLeod & Soven, 1992). Ultimately, LLM models can be used to elevate writing as a tool for learning, critical thinking, and communication across diverse AEC fields, ultimately enhancing students' writing proficiency for engineering formation.

This pilot study aims to integrate specific well-established educational frameworks: the generative learning theory, the WAC framework, the AI Across the Curriculum, and the Feedback Model (Graham & Perin, 2007; Hattie & Timperley, 2007; Hand et al., 1992). By combining these frameworks, an LLM-based model for feedback in writing technical reports may create a dynamic educational tool that enhances students' writing skills. The generative learning theory is incorporated to highlight that learners can generate semantic and distinctive idiosyncratic associations between stimuli and stored information allowing learners to self-generate inferences and mental models (Hand et al., 1992). Complementary, the WAC and AI Across the Curriculum frameworks aim to integrate writing and AI across academic disciplines, allowing students to practice writing within their curriculum (Graham & Perin, 2007) and Feedback Models emphasize the significance of timely, specific, and actionable feedback for learning and improvement (Hattie & Timperley, 2007). Leveraging AI may provide students with opportunities to write within the curriculum, delivering feedback aligned with the principles of both frameworks. LLM-based models may equip students with valuable writing skills while fostering their ability to receive adequate and immediate feedback, ultimately enhancing their overall learning experience. Through this pilot study, we aspire to offer a practical solution that encourages writing proficiency and empowers students in their educational journey.

Specifically, LLM-based models present a unique opportunity to revolutionize technical report writing education in AEC. This pilot study aims to offer the affordances that provide an in-depth learning experience leveraged by artificial intelligence. By receiving real-time feedback, students can navigate an immersive data-rich environment of various technical reports, observe strengths and weaknesses, and receive specific feedback on their own reports. LLM-based models may cover grammar, structural weaknesses, and suggestions for clarity, all while considering the articulation of technical information. Tailored to individual student needs, AI is useful for leveraging timely, personalized feedback, relieving instructors of some of the time-consuming aspects of grading, and offering students opportunities for prompt iteration and improvement.

The goal of this pilot study aligns with the critical components of effective feedback in education. It strives for timeliness, specificity, clarity, and actionable guidance—essential elements for proper feedback (Hattie & Timperley, 2007). This study aims to offer a path to address the necessity of prompt feedback, enabling learners to make timely adjustments to their understanding and performance. It also emphasizes specific and precise feedback, helping students comprehend their strengths and areas for improvement, and provides actionable advice for tangible steps for improvement (Hattie & Timperley, 2007). Moreover, the LLM will be specifically crafted for AEC, incorporating precise terminology, concepts, and communication standards. This tailored focus ensures that feedback and guidance align with the technical intricacies of architectural design, engineering principles, and construction processes.
Methods

Context and participants
This ongoing study takes place in a Construction Capstone Project (BCN 4787C) at the University of Florida during Fall 2023. Consequently, the experiments conducted will be completed by the first week of December of 2023 and only preliminary results are presented in this draft up to November 2023. During the Fall 2023 semester, the course has a total of 37 students. The course learning objectives are related to simulating a construction project with each student being responsible for designing, developing, estimating, scheduling, contracting, and administering the work for the completion of a small commercial, residential, civil, or light industrial project. In addition, the course has three major deliveries, and this study focuses on one of those deliveries.

Targeted Writing Assignments
Within the curriculum, three individual deliveries are required, each of them is based on a technical report written purposely to explain in detail the decisions made, the justification of these decisions, the assumptions made, the technical considerations taken into account, the sustainability implications, and answer specific questions from the project owner in a complete, argumentative, creative, and thoughtful manner. This assignment accounts for 32% percent of the total grade and is meant to be an authentic assignment. In this assignment, students are required to provide consulting services for a new building project in Gainesville, FL. To start with, the student may select a site for their project among three potential properties provided. Students may support the right balance between land cost and potential for rental income to recommend a specific site and justify its appropriateness for the investor, taking into account environmental considerations. Students are also required to achieve a minimum LEED Certified certification for the proposed project. This often requires special consideration during design to ensure that daylighting, energy, water use, and indoor air quality requirements are all being met. Students are tasked with making a recommendation on how to achieve this goal (i.e., which credits to pursue), as well as demonstrating an understanding of the LEED certification process.

Human Assessment of the Final Project
The final project in this course is usually assessed by the class instructor at the end of the semester. The final project has a total of 270 points, and the criteria used to evaluate this assignment follow Site Selection, Conceptual Estimate, Sustainability, and Financial Feasibility with the descriptions.

Large Language Models
Currently, there is a vast offering of LLM alternatives. For instance, OperAI’s GPT series (GPT-3 and GPT-4), BERT, T5, XLNET, Transformer-XL, and ELECTRA to name a few. All of these tools are based on deep neuronal networks and rely on regenerative architectures, however, depending on the task that they specialize in they will be more appropriate for certain tasks. For the task at hand, GTP-4 has been selected, due to the following reasons: (1) GPT-4 conversational agent features that permit easy retraining, (2) the availability of the tool makes it suitable for replicability (3) It permits the use of data formatted in different ways including documents in pdf.

Experimental design
From the total number of assignments, nine assignments were selected. These nine assignments represent submissions that were selected because they represent different qualities of students’ work. An instructor of the class evaluated the submissions and using the assessment classified these nine submissions as Above-Average Submission (AAS), Average Submission (AS), and Below-Average Submission (BAS). A table representing all the submissions and the assigned code is presented in Table
AAS: Above Average Submission. AS: Average Submission. BAS Bellow Average Submission.

We use these submissions to understand the capability of GPT-4 to assess an aspect of the project using its trained data. For this experiment, the focus was to understand the aspects that the LLM model assessed, as well as the type of feedback that was produced from it. For this purpose, we designed an experiment in which in every trial three projects that were assigned to each category were provided to the model to be ranked (see Table 2). With this information, we asked the model the following prompt: "Read over these projects and evaluate them according to sustainability. Rank the 3 projects at the end."

<table>
<thead>
<tr>
<th>Trial 1</th>
<th>Trial 2</th>
<th>Trial 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>BAS</td>
<td>Project 1</td>
<td>Project 2</td>
</tr>
<tr>
<td>AS</td>
<td>Project 4</td>
<td>Project 5</td>
</tr>
<tr>
<td>AAS</td>
<td>Project 7</td>
<td>Project 8</td>
</tr>
</tbody>
</table>

AAS: Above Average Submission. AS: Average Submission. BAS Bellow Average Submission.

**Results**

For Trial 1, Project 1, Project 4, and Project 7 were provided to GPT-4 to be evaluated in terms of sustainability (see Table 2). Without any other parameter, GPT-4 selects five criteria that differ depending on the project (see Table 3).

<table>
<thead>
<tr>
<th>Project 1</th>
<th>Project 4</th>
<th>Project 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEED Certification</td>
<td>LEED Certification</td>
<td>LEED Certification</td>
</tr>
<tr>
<td>Site Selection</td>
<td>Site Selection</td>
<td>Site Selection</td>
</tr>
<tr>
<td>Sustainable Building Practices</td>
<td>Materials and Resources</td>
<td>Environmental Protection and Waste Management</td>
</tr>
<tr>
<td>Water Efficiency</td>
<td>Energy and Atmosphere</td>
<td>Water and Energy Efficiency</td>
</tr>
<tr>
<td>Energy Sufficiency</td>
<td>Indoor Environmental Quality</td>
<td>Community and Patient Wellbeing</td>
</tr>
</tbody>
</table>
The results for Trial 1, Trial 2, and Trial 3 are summarized in Table 4. Similar comments were made about the project evaluated in the different trials. However, in Trial 2 and Trial 3, there were no categories provided for the ranking.

| Table 4 |
|---|---|---|---|
| Results of the ranking in the different trials | | | |
| | Trial 1 | Trial 2 | Trial 3 |
| Expected | Obtained | Expected | Obtained | Expected | Obtained |
| Project 7 | Project 1 | Project 5 | Project 5 | Project 6 | Project 6 |
| Project 1 | Project 7 | Project 8 | Project 2 | Project 9 | Project 3 |
| Project 4 | Project 4 | Project 2 | Project 8* | Project 3 | Project 9 |

*Project 8 was rejected due to formatting, and it was not ranked provided by GPT-4. After exploring the reason, it was concluded that Project 8 contained too many tables and figures.

In each ranking provided, some summary comments were given on the project. These summary comments spoke to the strengths and weaknesses of each of the projects.

In Trial 1, GPT-4 exhibited a reversed ranking for the above-average (Project 7) and average (Project 1) projects compared to manual grading. While manual grading deemed the explanation of achieving LEED points complete for Project 7 and incomplete for Project 1, GPT-4 emphasized a higher number of sustainability measures in Project 1 (e.g., energy-efficient systems, low-VOC materials, water-efficient plumbing, a green roof, and recycled materials) compared to Project 7 (e.g., stormwater management, local material usage, and energy efficiency). Notably, the rubric focuses on supporting the necessary measures for LEED certification, while GPT-4 appears to prioritize the quantity of enunciated measures. Despite this, there was agreement between manual grading and GPT-4 regarding the below-average submission.

In Trial 2, GPT-4 faced challenges evaluating Project 8 due to the document format, which indicated the need for a more precise sustainability assessment, especially for Project 8, suggesting additional descriptive information or context. Upon reviewing Project 8, it became apparent that the sustainability information was presented in a segmented manner, with each element subdivided into subtitles, each containing only two to three sentences. It seems that GPT-4’s capabilities may be hindered by these segmented paragraphs organized into two levels of aggregation in titles. Upon excluding Project 8 from evaluation, there was a concurrence between manual grading and GPT-4 in ranking Project 5 as superior to Project 2. Specifically, GPT-4 noted that Project 2, while including various sustainable features, lacked the depth and breadth of sustainability efforts compared to Project 5.

In the final Trial 3, GPT-4 and manual grading reached a consensus regarding the above-average project (Project 6). However, divergence emerged in the assessment of the average and below-average projects. Manual grading deemed the LEED narrative in Project 9 more comprehensive than in Project 3, whereas GPT-4 held the opposite view. Upon manual review of both projects, it became apparent that Project 9 provided a more thorough explanation of fewer elements, sufficient to support the LEED narrative, while Project 3 enumerated a greater number of elements but with much less elaboration on each.

**Conclusions**

This pilot study endeavors to address a critical gap in AEC education, specifically in the realm of technical report writing. The lack of timely and comprehensive feedback on technical writing skills poses a significant challenge for both students and instructors, hindering holistic skill development and creating disparities among learners. Despite the pivotal role of writing in the professional development
of AEC professionals, the focus on teaching these skills remains inadequate, particularly as students progress through their academic journey. While the integration of AI-based feedback can be proposed to address these challenges, providing timely, specific, and actionable feedback to enhance student's writing proficiency, the capability of these tools to generate accurate feedback is still under current scrutiny.

Results from the presented experiment 1 demonstrate the model's ability to evaluate sustainability aspects in different projects, providing specific and relevant feedback. From the results, it was possible to observe that (1) GPT-4 was able to give good categories for evaluation of the project in sustainability, matching current sustainability criteria and changing the criteria as it adjusted to the project that it was evaluating. (2) The ranking provided did not match the one obtained by the human graders. In addition, in all the trials, there was disagreement between either the above-average and average or below-average and average (see Table 4). This may indicate that the ability to rank projects is a more complex task that is not well supported by the current state of the art in LLMs. Moreover, ideas and content are more important than the form of writing for the human grader, while LLMs seem to prioritize the form of writing. Nonetheless, this issue requires further exploration, specifically when looking at large projects (i.e., more than 25 pages each) such as the ones used in this experiment. (3) The summary and feedback provided by the model shed some light on the points of improvement for each project. While a more detailed feedback prompt was not given to the model, it is possible that if asked, the model can give more specific feedback that can be significantly valuable for students. (4) The prompt given to the model was critical, and multiple runs of the experiment were done at the beginning to have a prompt that indicated ranking but that did not give any detail. In this same line, the format of the project was relevant, as some of the documents were not analyzed due to formatting (i.e., Project 8). The exploration of this issue revealed that it is important to take into consideration the way documents are submitted to these platforms and communicate to students clearly so that they can take advantage of the feedback. Overall, this experiment reveals the potential that LLMs have for classroom help in supporting general feedback but not grading or ranking of projects, at least not in the context of construction management. Nonetheless, ultimately, the integration of AI-based feedback agents has the potential to elevate writing proficiency, foster critical thinking, and enhance communication skills across diverse AEC disciplines, aligning with broader educational frameworks and pedagogical principles.

Acknowledgements:
We want to thank the University of Florida's strategic funding process for providing the funds to develop this research. This research was conducted under the IRB Protocol # ET00021975. Finally, we would like to acknowledge our student, Elizabeth Lourcey, for her help in validating the results of the experiment.

References


