Journal of Information Systems Education

Volume 37
Issue 1
Winter 2026

Generative AI-Augmented Human Judgment: A Task-Technology Fit Perspective

John Xavier Volker and Al Tilooby

Recommended Citation: Volker, J. X., & Tilooby, A. (2026). Generative AI-Augmented Human Judgment: A Task-Technology Fit Perspective. *Journal of Information Systems Education*, 37(1), 151-166. https://doi.org/10.62273/STSP3767

Article Link: https://jise.org/Volume37/n1/JISE2026v37n1pp151-166.html

Received: May 17, 2025
First Decision: August 18, 2025
Accepted: September 15, 2025
Published: March 15, 2026

ISSN: 2574-3872 (Online) 1055-3096 (Print)

Journal of Information Systems Education, 37(1), 151-166, Winter 2026
https://doi.org/10.62273/STSP3767

Generative AI-Augmented Human Judgment: A Task-Technology Fit Perspective

John Xavier Volker
Al Tilooby
College of Business
Austin Peay State University
Clarksville, TN 37044, USA
volkerj@apsu.edu, tiloobya@apsu.edu

ABSTRACT

This study examines how generative artificial intelligence can augment human judgment in assurance of learning assessments within business education, using the task-technology fit framework as a guiding lens. A case study in a college of business – where the Management Information Systems program served as a central unit in the assurance of learning cycle – compared generative artificial intelligence-driven evaluations of student writing with traditional faculty assessments. The results demonstrate that when mediated by human-based prompt engineering and moderated by human oversight, generative artificial intelligence markedly improves assessment efficiency and scoring consistency while providing more in-depth feedback without compromising evaluation accuracy. These findings indicate that generative artificial intelligence is most effective as a complement to rather than a replacement for human evaluators. The study extends task-technology fit theory to generative artificial intelligence-driven educational assessment and introduces a human-integrated, generative artificial intelligence-augmented theoretical model for assurance of learning assessments. In this model, human expertise acts as an iterative mediator (via prompt engineering) to strengthen task-technology alignment, while human oversight serves as a moderator ensuring contextual fidelity and output quality. Beyond its theoretical contribution, the study highlights practical implications for information systems educators and curriculum designers.

Keywords: Generative AI, Task technology fit (TTF), Assurance of learning, Learning goals & outcomes, Assessment

1. INTRODUCTION

Generative artificial intelligence (GenAI) is a disruptive information systems (IS) artifact rapidly permeating multiple sectors (Hill-Yardin et al., 2023; Liebrenz et al., 2023; Lund & Wang, 2023; Sun & Deng, 2025), generating growing interest in its potential to transform educational processes (Shahid & Mishra, 2024). In business schools, a promising application is in the assurance of learning (AOL) process, which encompasses evaluating student assignments – often centered on writing – and addressing gaps (Attaway et al., 2011; Beard et al., 2008). It entails systematic, rubric-based assessments to verify that students meet program-level competencies (Borschbach & Mescon, 2021), but these assessments are labor-intensive and difficult to scale as student numbers grow. GenAI large language models, with their capacity to analyze text and produce content (Andriole, 2024; Bahi et al., 2024; Davazdahemami et al., 2024; Henkin & Zangrilli, 2024; Marimon et al., 2025), offer a potential solution by providing rapid, consistent evaluations of student writing at scale. However, it remains unclear how such GenAI systems integrate with the nuanced judgment of human evaluators in assessing complex, qualitative criteria. To address this uncertainty, we ask the following research question: *How can GenAI effectively augment humans in assurance of learning assessments to achieve a strong fit with the task of evaluating student writing?*

In this paper, we argue that GenAI can meaningfully augment human judgment (defined here as enhancing, supporting, and extending human evaluators' expertise without displacing their critical role) in assurance of learning assessments when its integration is guided by strong task-technology fit, with human expertise iteratively mediating prompt design, and human oversight moderating output quality.

Recent literature reflects both optimism and caution about GenAI in academic assessment. On one hand, GenAI is hailed as a means to transform assessment practices by enabling timely, unbiased feedback at scale (Dwivedi et al., 2023; Susarla et al., 2023). On the other hand, GenAI systems pose reliability and transparency issues: they function as opaque "black boxes" and can occasionally produce incorrect or nonsensical outputs (Nishant et al., 2024; Schlagwein & Willcocks, 2023). Such errors or "hallucinations" (Susarla et al., 2023) underscore whether GenAI can honestly evaluate complex student work with human-like nuance. This tension between GenAI's promise and pitfalls highlights the need to investigate how well these tools fit AOL's task requirements.

We use the Task-Technology Fit (TTF) framework (Goodhue & Thompson, 1995) to determine whether GenAI meets AOL's needs. TTF posits that technology is most effective when its capabilities align closely with the task demands: a good fit leads to better performance and user acceptance (Goodhue & Thompson, 1995; Ulfa et al., 2024), whereas a poor fit results in suboptimal outcomes or resistance (Howard & Rose, 2019). This theory provides a valuable lens for our context, where the task is to evaluate student writing with attention to assessment duration, accuracy, consistency and objectivity, and depth of feedback. We hypothesize that GenAI will substantially aid the AOL process only if its capabilities – such as language understanding, speed, and consistency – match these task requirements. Moreover, we propose that human expertise and oversight are essential conditions for sustaining high task-technology fit, especially in addressing areas where GenAI's performance alone may fall short. By applying TTF, we can pinpoint where the GenAI aligns well with AOL's needs and where it falls short, providing a theory-driven explanation for our findings.

Our case study was conducted within the College of Business's AOL cycle, during which the Management Information Systems (MIS) program – a key unit within the College – was evaluated alongside other core programs as part of the college-wide assessment process. We collected a sample of student assignments, which were evaluated both by business educators using the standard AOL rubric, and by GenAI. We then compared the outcomes of the AI-assisted evaluations to the traditional human assessments. Our results indicate that while GenAI enhances multiple facets of the AOL assessment process, its effectiveness is contingent on human intervention at two critical levels: expertise that iteratively mediates task-technology fit through prompt engineering and oversight that moderates GenAI outputs to ensure contextual accuracy and rigor. Thus, GenAI appears most effective as an assistive tool that offloads routine assessment tasks while experts help with prompt engineering and evaluators handle complex judgments and ensure academic rigor.

This paper offers contributions to both IS theory and educational practice. Theoretically, we extend TTF into AI-driven assessment, introducing the human-integrated GenAI-augmented theoretical model for AOL, which explains when and why GenAI succeeds or fails in augmenting human judgment. Practically, our findings offer guidance for integrating GenAI into competency assessment processes. Provided GenAI complements rather than replaces evaluators, we demonstrate that human-GenAI collaboration significantly reduces assessment duration by streamlining the evaluation process. It maintains accuracy and ensures consistent, objective application of rubric criteria across evaluators. At the same time, it enhances the depth of feedback by providing more detailed, structured, and actionable comments for students. We underscore the need for structured human expertise in prompt engineering and sustained human oversight to achieve and maintain a strong task-technology fit when deploying GenAI in AOL. Finally, by demonstrating a successful GenAI deployment in an authentic AOL setting, our work helps build trust in the responsible use of GenAI for enhancing educational assessments.

2. GENERATIVE ARTIFICIAL INTELLIGENCE

GenAI is a class of deep learning models designed to analyze or produce original content across modalities, including text, images, and audio (Marimon et al., 2025; Sun & Deng, 2025; Susarla et al., 2023). Leveraging large-scale transformer architectures, GenAI systems generate human-like outputs that enable more intuitive and accessible human-machine interactions (Heyder et al., 2023; Schlagwein & Willcocks, 2023), signaling GenAI's potential to reshape the future of work (Basole et al., 2024; Dwivedi et al., 2023). GenAI's versatility, including its ability to perform new tasks via natural language instructions, underscores its transformative potential.

Despite its promise, GenAI raises several challenges. Its "black box" nature limits interpretability and complicates efforts to audit or explain outcomes (Schlagwein & Willcocks, 2023). Concerns also arise around intellectual property, content accuracy, and the coherence of outputs, especially in high-stakes contexts (Nishant et al., 2024; Susarla et al., 2023). Moreover, biases embedded in training data may be unintentionally reinforced, and phenomena such as "hallucinations" – plausible but inaccurate outputs – remain insufficiently understood, partly due to proprietary training corpora (Susarla et al., 2023).

GenAI ushers in a transformative shift within educational systems (Shahid & Mishra, 2024), signaling profound changes in academic assessment (Susarla et al., 2023). Yang et al. (2023) note that GenAI alignment with human judgment in ranking preferences underscores its value in tasks related to content assessment. Its role in assessments extends beyond essential pattern recognition, enabling the analysis and interpretation of text data with remarkable depth.

3. THEORETICAL FRAMEWORK: TASK-TECHNOLOGY FIT (TTF)

TTF (Goodhue & Thompson, 1995) provides a foundational framework for evaluating how well an information technology supports a particular task. According to TTF, the effectiveness of technology, in terms of improving performance or outcomes, is determined by the degree of match between the technology's capabilities and the demands of the task at hand. A high degree of fit implies that the technology can adequately meet user needs for that task, leading to greater utilization and better performance outcomes (Ulfa et al., 2024). Conversely, a misalignment (poor fit) means the technology may not support the task effectively, resulting in suboptimal outcomes or user resistance (Howard & Rose, 2019). Figure 1 provides an overview of the TTF model.

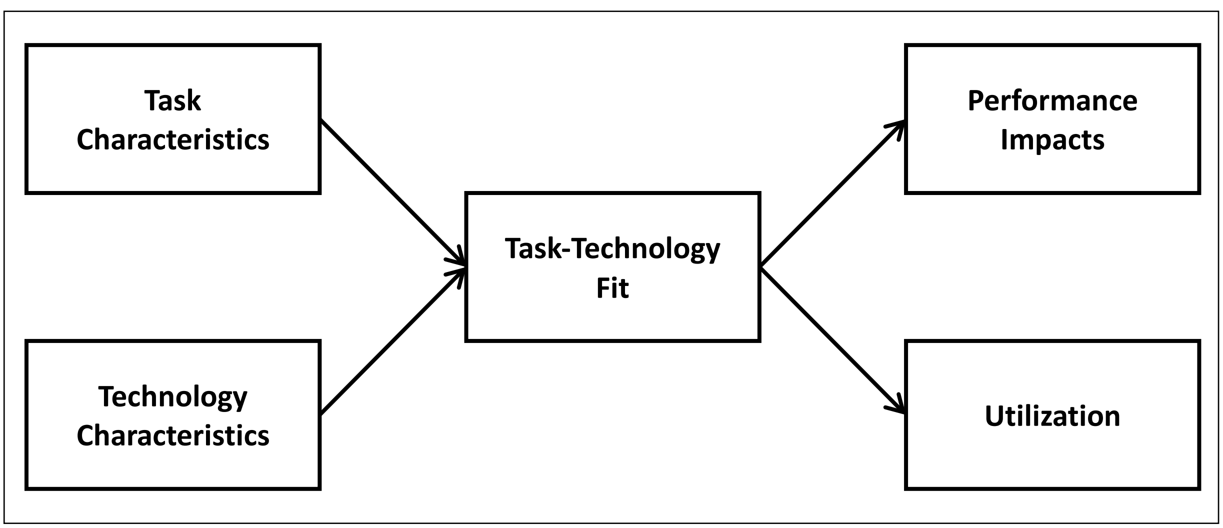

Figure 1. Task-Technology Fit Model (Goodhue & Thompson, 1995)

In our study, the task involves evaluating student essays in business education. The key requirements of this task align with four evaluation metrics: assessment duration (efficiency in processing potentially large volumes of work); accuracy (correct identification of content and organization); consistency and objectivity (stability of rubric-based judgments across essays and evaluators); and depth of feedback (providing meaningful strengths, weaknesses, and improvement suggestions). The technology in focus is ChatGPT, a GenAI tool with advanced natural language processing capabilities. Its relevant capabilities include rapid text analysis, generation of detailed responses, recognition of patterns in language, and adaptability to user prompts (Dwivedi et al., 2023; Hessari et al., 2024; Kim et al., 2024; Sundberg & Holmström, 2024).

4. METHODOLOGY

We conducted this case study (Yin, 2013) at a regional public university's business college during the 2024 AOL cycle, which included undergraduate and graduate MIS program learning outcomes. In alignment with AACSB accreditation requirements, educator-led AOL Committees had previously established written communication as a core learning goal and developed rubrics to assess student writing in selected courses. Traditionally, student submissions were sampled from designated courses and assessed by evaluators using a binary "go/no-go" scale across multiple criteria, including organization, content, and writing quality, culminating in an overall rating for the learning outcome. We leveraged this established AOL infrastructure for this paper by integrating GenAI into the evaluation process, positioning it alongside human raters.

We assessed the performance of GenAI-assisted evaluation relative to traditional evaluator-only assessment. A total of 153 student writing samples were collected – 28 from a graduate-level course and 125 from an undergraduate course – each chosen by the AOL Committees to assess written communication. The submissions, which included project reports and essays, varied in length from a few paragraphs to several pages and focused on topics aligned with course content in management and strategy.

4.1 GenAI Integration Process

We selected ChatGPT-4 as the GenAI tool for this assessment task. The choice of ChatGPT was motivated by its strong natural language understanding and analysis capabilities, as well as its demonstrated potential in content evaluation tasks. Prior studies indicate that although ChatGPT can align with human judgment in tasks such as content quality ranking, its performance remains constrained by inaccurate reasoning and prediction instability, often triggered by sensitivity to minor prompt variations (Yang et al., 2023). However, the rapid pace of GenAI development suggests that some of these limitations may become less salient over time.

4.1.1 Prompt Development. We developed purpose-built prompts for ChatGPT as a critical component in achieving task-technology fit. Prompts serve as the primary communication channel with GenAI, guiding it to produce outputs that adhere to specified rules, automate evaluation steps, and maintain the required quality and depth (White et al., 2023). Figure 2 exemplifies one of the prompts we used during the process.

To enhance response quality, we drew on the prompt engineering literature and incorporated "chain-of-thought" techniques, defined by Wei et al. (2022) as a sequence of intermediate natural-language reasoning steps that progressively lead to the final output. In practice, this meant encouraging step-by-step reasoning through each rubric category. As such, the model was directed to proceed sequentially: (1) identify specific strengths and weaknesses for each rubric dimension, (2) briefly justify its evaluation with reference to rubric definitions, and (3) assign a "go" or "no-go" judgment for each category and the overall outcome. This structured reasoning process fostered transparency and consistency, reducing the risk of unsubstantiated single ratings.

Prompts explicitly instructed the GenAI to adhere strictly to rubric criteria and to provide constructive, criterion-based feedback. Before deployment, the prompts were piloted on a subset of student essays not included in the primary dataset and iteratively refined to ensure completeness and fidelity to rubric definitions. This intentional design process was central to optimizing TTF, ensuring that the GenAI's outputs were both relevant and aligned with the evaluative requirements of the AOL task.

The development process involved translating rubric criteria and scoring guidelines into structured, machine-readable instructions. ChatGPT was instructed to assess writing samples across three dimensions – organization, content, and quality – using rubric-specific definitions embedded directly in the prompt. For example, under organization, the model evaluated whether the writing included a clear opening, logical flow, and coherent transitions. Under content, it assessed whether students correctly identified and applied relevant theories, frameworks, or tools introduced in class. Under quality, it examined grammatical accuracy, tone, and professionalism of expression.

> You will be evaluating samples of undergraduate student writing. You will be considering the writing samples on their organization, content, and quality. For quality, you are looking for a clear opening statement, there is structure, and there is a clear summary or closing statement. These will be scored as go or no-go. For content, you seek relevance, a depth of analysis expected of an undergraduate university student, and the use of sources and other evidence. These will be scored as a go or no-go. Your final evaluation area is quality; here, you are looking for few to zero grammatical errors, few to zero spelling errors, and a professional air to the writing. You will score these as a go or no-go. At the end of each evaluation, provide an overall go or no-go score based on your scoring. You will always use this rubric and scoring methodology. If you understand, respond with a "yes."
>
> ChatGPT
>
> **ChatGPT**
>
> Memory updated
>
> Yes.

Figure 2. One of the Prompts Used for ChatGPT

4.1.2 Assessment Procedure. Using TTF as our lens, we examined the alignment between task requirements and ChatGPT's capabilities. We assessed ChatGPT's fit in AOL writing assessments across three key metrics: (1) its accuracy in interpreting the content and organization of student essays, (2) its consistency and objectivity in applying rubric criteria, and (3) its ability to generate feedback with sufficient depth. Table 1 outlines the operational definitions of the three metrics, clarifying how ChatGPT's fit was evaluated. If these metrics are performed well, then from a TTF perspective, the technology's features align closely with the task requirements. In such cases, integrating ChatGPT into assessment processes can improve accuracy and scoring consistency, enrich the quality of feedback, and ultimately allow evaluations to be completed more quickly. Conversely, if these metrics are not met, the resulting misfits underscore the continued importance of human judgment in maintaining assessment quality and contextual relevance.

Following the finalization of the prompts, the GenAI-assisted assessment was conducted using the complete set of 153 student writing samples. Each submission was converted into plain text and sequentially input into ChatGPT, accompanied by the corresponding prompt – graduate or undergraduate – based on the course level of origin. Upon receiving each input, ChatGPT generated an evaluation in near real time, typically within seconds. Each output included: (a) a summary or acknowledgment of the submission's content, (b) categorical judgments ("go" or "no-go") for each rubric dimension – such as organization, content, and quality, (c) an overall determination of the written communication outcome, and (d) qualitative feedback detailing rationale and suggestions for improvement.

Metric	Operational Definition
Assessment duration	Time required by the evaluator (human or GenAI) to complete the evaluation of an essay, recorded in minutes or seconds, and used to compare efficiency across evaluators.
Accuracy	The degree to which the evaluator's identification of key ideas and structural elements in student essays aligns with those identified by expert human graders. Accuracy of interpreting the content and organization of student essays.
Consistency and objectivity	Extent to which rubric-based scores remain stable across different essays, evaluators, and repeated assessments, as indicated by inter-rater reliability and reproducibility of results.
Depth of feedback	Level of detail and coverage in evaluator feedback, measured by the inclusion of strengths, weaknesses, relevant concepts, and improvement suggestions across rubric dimensions.

Table 1. Operational Definitions of the Three Key Metrics of Evaluation

We completed the GenAI-based evaluation in approximately 1.5 hours. This substantial reduction in assessment time highlighted a key benefit of effective TTF: a process that traditionally demands several hours or even days of evaluators' labor was executed within an afternoon using GenAI.

We incorporated a validation step into the process to ensure consistency and reliability of ChatGPT's evaluations. While the whole test was run by one of the two researchers, the second researcher randomly selected a subset of 10 student samples (five graduate, five undergraduate) and reprocessed them using ChatGPT on a separate computer, applying the same prompts used by the first researcher. This served as a test of output reproducibility across independent instances of the GenAI model. The resulting evaluations were then compared to the originals. In all ten cases, ChatGPT generated identical category judgments and substantively equivalent feedback, with only minor lexical variations. This consistency reinforced confidence in the precision of the prompt design and confirmed that GenAI's stochastic elements did not interfere with the evaluation outcomes when prompts and inputs remained constant.

4.2 Human Educators' Assessments

In parallel with the GenAI-assisted assessments, the AOL Committees conducted their standard evaluation process, which served as the control condition for this endeavor. Educators independently reviewed and scored the student writing samples using the established rubrics, after which the committees met to discuss results and identify potential learning gaps. We shielded evaluators from GenAI evaluations until they completed their assessments to avoid bias. The educators' evaluation process occurred over several weeks, interspersed with regular academic responsibilities, and culminated in a formal AOL report. From this report, we collected summary data on student performance (e.g., number of students meeting expectations, prevalent areas of weakness) and recommendations made by the evaluators to improve the curriculum.

5. RESULTS

Our findings (Table 2) show that employing GenAI to assist in AOL assessment of student writing leads to notable improvements in efficiency, consistency, and the depth of analysis, while maintaining a level of accuracy comparable to traditional methods. We organize the results below around these key metrics, followed by the evaluators' adoption decision outcomes. Throughout this section, we interpret the results in light of task-technology fit, noting how well the technology's performance aligned with the task's needs.

Beyond these core metrics, two additional impacts were noted based on the experimental work conducted by the researchers (rather than the committee's full human evaluation effort). First, human workload was significantly reduced, as GenAI shifted involvement from labor-intensive scoring to prompt design and oversight. Second, task-technology fit proved strong: while humans excel in nuanced judgment, GenAI delivered clear advantages in large-scale, repetitive evaluation tasks, with human oversight ensuring necessary refinement.

Aspect	Traditional Educators' Assessment	GenAI-Augmented Assessment
Assessment duration	Slow: Evaluating ~153 (distributed among 5 evaluators) took between 2 to 4 weeks for each of the committee members while balancing their ongoing academic responsibilities.	Fast: All 153 samples evaluated in ~1.5 hours total.
Accuracy	Generally accurate (experts applying rubric), but occasional human error or oversight is possible.	High adherence to rubric instructions (prompt-guided); however, GenAI might misinterpret if the prompt is ambiguous (prompt design mitigated).
Consistency and objectivity	Potential variation between different evaluators; subject to human fatigue and bias.	High consistency – the same rubric applied uniformly by GenAI across all samples; reproducible results on re-run.
Depth of feedback	Variable depth; often minimal due to time constraints (e.g., brief comments or just scores).	Extensive feedback for each sample, addressing each rubric dimension with suggestions for improvement.

Table 2. AOL Assessment: Traditional Educators-Only vs. GenAI-Augmented

5.1 Dramatic Gains in Efficiency

Integrating GenAI drastically accelerated the assessment process. The GenAI evaluated all 153 writing samples in roughly 1.5 hours, compared to an estimated several dozen hours of cumulative work if done by evaluators (even when spread among multiple evaluators). The time savings observed align with the initial premise that GenAI can facilitate the scaling of the AOL process.

5.2 Improved Consistency and Reliability

Consistency was another area of improvement. GenAI applied the rubric criteria uniformly across all samples, judging each student's work against the same standards without drifting expectations over time. Our reproducibility check (re-running samples to see if ChatGPT produced the same results) demonstrates high reliability in GenAI's outputs. Moreover, we find a high agreement rate when comparing GenAI vs. evaluators' overall decisions (go/no-go) for each sample. In over 90% of the cases, ChatGPT's overall judgment matched that of the human evaluators. In the remaining instances, discrepancies generally involved borderline cases. For example, ChatGPT might flag a "no-go" for a paper that barely missed some criteria, whereas an educator might have given it a pass (or vice versa). This indicates that the AI's strict rubric adherence can sometimes catch issues that human graders might overlook or hesitate to penalize. Notably, we observed no instances of completely incorrect or nonsensical feedback, suggesting that our prompt design successfully constrained ChatGPT to the task context (mitigating the risk of GenAI "hallucinations"). These consistency gains align well with TTF expectations: GenAI maintained consistent performance on repetitive tasks without fatigue or bias, exactly as we had hoped from fitting GenAI capabilities to AOL needs.

5.3 Depth and Richness of Feedback

One of the most striking results is the depth of feedback provided by ChatGPT (Figure 3). We analyzed the GenAI-generated evaluations for each rubric category and concluded with an overall assessment and suggestions (e.g., to improve, the student should better integrate key business terms and ensure the conclusion ties back to the introduction). In contrast, many evaluators provided accurate scores but did not include elaborate feedback for each student, which is understandable given that offering detailed written feedback for over a hundred samples exceeds the typical scope of AOL assessments. The consistency in

GenAI depth was notable, where every sample – whether strong or weak – received a thorough treatment from the GenAI. We note that GenAI sometimes misses contextual subtleties – areas where human expertise can complement GenAI's feedback.

Overall Evaluation: NO-GO

Organization: Score: GO

The document opens with a clear statement introducing the subject of analysis and frames the purpose of the report effectively. Structurally, the writing moves from background information to a high-level overview of activities and concludes with a personal reflection. While the progression of ideas is generally logical, transitions between sections are abrupt, and the shift from company analysis to personal learning is not sufficiently signposted. A professional executive summary would benefit from cleaner sectioning and more deliberate framing of transitions. Nonetheless, the document has an identifiable structure with a beginning, middle, and end.

Content: Score: NO-GO

The writing remains thematically aligned with the analysis of the corporation's operations, growth, and strategic positioning. However, the depth of analysis falls short of undergraduate expectations. Assertions about the organization's achievements, market share, and environmental efforts are made without substantiating evidence, cited sources, or specific data. Critical concepts such as the S.W.O.T. analysis and financial breakdowns are referenced but not developed — they are merely gestured toward. At this level, a more rigorous engagement with factual details, comparative data, or external validation would be expected. The narrative leans too heavily on broad summaries and personal impressions rather than an analytical dissection of the firm's competitive dynamics and challenges.

Quality: Score: NO-GO

The document suffers from multiple avoidable issues. Grammatical inconsistencies — such as misplaced spaces before punctuation and spelling errors (e.g., "Eath" instead of "Earth") — diminish the professionalism of the writing. Stylistic choices occasionally slip into an informal register ("Tool 11 was very challenging to me but I hope to have provided some answers"), which is inappropriate for an executive summary intended for a formal audience. Capitalization errors and inconsistent terminology further detract from the polish expected at the undergraduate level. Overall, the lack of careful proofreading compromises the authority and credibility of the analysis.

Commentary and Forward Guidance

The student demonstrates a promising understanding of what an executive summary should aim to accomplish: presenting a concise overview of a company's activities and challenges. However, moving forward, it will be crucial to:

- **Deepen the analysis**: Avoid high-level summaries and instead drill down into concrete details — financial data, competitive comparisons, strategic initiatives — substantiated with evidence.
- **Elevate the professionalism**: Maintain a formal, third-person tone throughout; remove casual language and expressions of personal struggle unless explicitly required by the assignment.
- **Proofread meticulously**: Even minor mechanical issues erode the professionalism of the work; treat error-free writing as a non-negotiable standard.
- **Strengthen structural transitions**: Signal shifts between sections more clearly and maintain thematic coherence between company analysis and any personal reflection components.

The student is close to producing work at the expected standard but must attend to both the depth and the mechanics of execution to achieve a "go" in future evaluations.

Figure 3. Example of the ChatGPT Evaluation

5.4 Educators' Acceptance and Decision to Adopt

A pivotal outcome of this paper was the AOL Committee's decision to formally adopt ChatGPT for future written communication assessments. This decision followed a thorough review of GenAI outputs, demonstrations, and extensive discussions addressing academic integrity and reliability concerns.

6. DISCUSSION

Our successful integration of ChatGPT into the AOL assessment process offers several insights into GenAI's role in educational institutions and the conditions needed to effectively augment human judgment. In this discussion, we examine the findings in terms of both practical and theoretical implications, using TTF theory as the primary lens for theoretical interpretation.

6.1 Interpreting the Findings Through TTF

A central takeaway from our paper is that GenAI's positive impact on AOL was largely due to a strong task-technology fit. The tasks involved in AOL writing assessment (applying a rubric systematically across many essays, providing formative feedback, and doing so efficiently) are well-matched to the GenAI capabilities (natural language processing, consistency in following instructions, and speed). Our use of TTF theory helps explain why the integration worked: the technology is able to meet the task's demands to a high degree, thereby improving performance outcomes. This resonates with the core premise of TTF that when users perceive a high fit, they are more likely to use the technology and see improvements in effectiveness (D'Ambra et al., 2013; Dishaw & Strong, 1999; Zigurs & Buckland, 1998). Indeed, in our case, the AOL Committee's willingness to adopt GenAI resulted directly from perceiving a good fit.

Examining which aspects of the task align well with the technology and which do not yield valuable insights. Routine, high-volume, and criteria-based evaluations fit GenAI assistance nearly perfectly. GenAI excelled at the repetitive elements of grading without losing concentration or changing standards over time. This finding aligns with other research noting GenAI's advantage in pattern recognition and consistency for structured tasks (Susarla et al., 2023). On the other hand, where did we observe limitations or the need for human judgment? There were a few subtle areas, such as understanding contextual nuance. These elements of the AOL task are less explicit in rubrics and more in the realm of expert judgment. Not surprisingly, GenAI, by following the rubric literally, sometimes diverges slightly from a human's holistic judgment. TTF theory frames this as part of the task where current GenAI technology achieves incomplete fit. However, rather than treating this as a failure, we demonstrate a complementary relationship where humans focus on nuances and exceptions. In our model, we conceptualize human intervention as two distinct mechanisms: (1) Human Expertise, which iteratively mediates the task-technology fit by improving prompt engineering and task fit, and (2) Human Oversight, which moderates the final outputs to ensure contextual quality. At the same time, GenAI covers the broad base of routine analysis. This dual-role structure shows that introducing GenAI can redefine the task by combining automation with targeted human augmentation – a key feature of the Human-Integrated GenAI-Augmented Theoretical Model for AOL we propose.

6.2 Alignment With and Contribution to Existing Literature

Our findings align with and extend the emerging literature on GenAI in education. Many scholars have discussed the potential of GenAI to transform educational practices. For example, Dwivedi et al. (2023) present multidisciplinary perspectives on generative conversational AI, highlighting opportunities for increased efficiency and personalized feedback, alongside challenges in ensuring academic integrity. Our paper provides concrete evidence of the efficiency gains and illustrates a way to harness GenAI for evaluation and feedback.

Another thread in the literature addresses ethical and reliability concerns when using GenAI in assessment. Schlagwein and Willcocks (2023) and Susarla et al. (2023) touch on what could be called the "Janus Effect" of GenAI – the idea that it has two faces: one of great promise and one of potential peril. Our results support this notion: ChatGPT significantly improved the assessment process; however, we had to remain mindful of issues such as GenAI's opaque reasoning process and possible biases. By implementing structured prompt engineering (enabled by human expertise) and validation checks through human oversight, we took steps toward responsible use, as recommended by those authors. Notably, Susarla et al. (2023) warn that GenAI models may inadvertently present biases in their training data or produce plausible sounding but incorrect information. In our context, because we constrained GenAI to evaluate given content rather than retrieve external facts, the risk of misinformation remained low. However, GenAI could manifest bias in how it judges writing style or tone. We did not observe any biased behavior during our trials (e.g., no evidence that GenAI favored a writing style based on cultural or linguistic bias), but future implementations could proactively include bias checks in feedback language.

Our integration of TTF into the analysis fills a theoretical gap. Much of the early commentary on GenAI in academia has been either exuberantly optimistic or cautiously skeptical. Still, few researchers have provided a structured, theory-driven analysis of when and why the tool will be effective. Applying TTF, we introduce a way to generalize from our case: educators and researchers can evaluate other educational uses of GenAI by systematically asking, "What are the task's requirements? What are the technology's capabilities? How well do they fit?" Further, our model highlights that human expertise actively shapes the initial task-technology fit, while human oversight sustains trust and output quality, both of which are critical for reliable integration. For example, when grading programming assignments or evaluating presentations, educators can use the TTF framework to predict if a tool is suitable or if misfits will cause problems. We contribute by marrying a classic information systems theory with a cutting-edge application in education, extending the applicability of TTF theory, and offering a more principled way to assess technologies in a specific context.

6.3 Implications for Practice in IS Education

For IS educators and curriculum designers, GenAI presents both challenges and opportunities in shaping assessment and teaching practices. Faculty development is critical, as training on AI tools can help instructors integrate GenAI productively, rather than view it solely as a threat. At the same time, frameworks such as TTF provide a lens for identifying where GenAI adds genuine value, such as enhancing consistency, efficiency, and depth of feedback in outcome measurement. At the program level, these improvements can support accreditation processes and continuous improvement cycles, positioning IS programs to not only safeguard standards but also harness GenAI to strengthen learning outcomes.

Our results support strategically integrating GenAI into AOL and similar assessment processes, particularly for large volumes of student work. Institutions, especially those facing resource constraints, can benefit from improved efficiency and consistency. Offloading routine evaluative tasks to GenAI allows evaluators to focus on higher-order responsibilities such as interpreting results and implementing curricular improvements. Importantly, we do not imply removing human judgment. Our findings affirm that a hybrid model – where human expertise guides prompt design and task framing, and human oversight ensures contextual appropriateness – strikes a productive balance between automation and judgment.

Institutions must invest in training and process development to achieve successful implementation. We found that prompt engineering as an IS skill is central to aligning GenAI outputs with rubric standards. Training initiatives should focus both on developing human expertise in prompt engineering and reinforcing practices for effective oversight of GenAI outputs. Educators likely need targeted training to translate rubric language into machine-readable prompts and critically interpret GenAI responses. Educators can mitigate risk by establishing contingency protocols, such as requiring secondary human reviews for unexpected outputs.

Finally, the scalability and transferability of this approach merit consideration. While our pilot focused on written communication in two courses, GenAI is well-suited for text-based outputs commonly found in business education. Institutions could start by piloting GenAI in one task and expand incrementally, evaluating TTF at each stage. Not all outcomes (e.g., teamwork or leadership) may lend themselves to GenAI-based assessment, so task fit should remain a guiding criterion.

6.4 Implications for Research

From a research perspective, we open several avenues in this paper. One immediate need is to replicate and quantify our findings under more controlled conditions. We conducted an exploratory study with a relatively small sample size. Future research could involve controlled experiments where researchers assess multiple course sections either with GenAI-assisted methods or purely human evaluations, then compare assessment outcomes and subsequent student performance to see if feedback mechanisms affect learning gains. Additionally, while we observe improved consistency, conducting rigorous statistical analysis (e.g., inter-rater reliability between GenAI and humans or among multiple GenAI runs) would bolster the case. Research could also assess evaluators' workload and stress – does adopting GenAI meaningfully reduce evaluators' burnout in assessment periods? These measurable outcomes can make a strong case for or against GenAI adoption.

Another research implication concerns the evolution of GenAI capabilities. These models are rapidly improving. It would be valuable to track how newer models (or fine-tuned education-specific models) perform on AOL tasks. Does the fit improve (e.g., the GenAI starts to handle nuances it could not before)? Or do diminishing returns set in? The answers will help inform how much we can lean on GenAI versus needing human input. Future extensions of the Human-Integrated GenAI-Augmented Theoretical Model for AOL can examine dynamic shifts in the iterative mediating role of expertise and the moderating role of oversight as GenAI capabilities evolve.

Lastly, on the theoretical front, using TTF in this context suggests synergy with other models such as the technology acceptance model (TAM) (Davis, 1985) or unified theory of acceptance and use of technology (UTAUT) (Venkatesh et al., 2003) regarding evaluators' adoption of GenAI. While TTF explains performance impact, TAM or UTAUT could complement it by explaining intention to use. In this case, evaluators chose to adopt the tool primarily for two reasons: (1) its high perceived usefulness, attributed to a strong task–technology fit, and (2) its moderate perceived ease of use, as interaction was conducted through natural language and did not require specialized technical skills. Studying more implementations could validate this and see if factors such as trust in GenAI and perceived risk play roles in acceptance. There is an emerging literature on GenAI acceptance in academia (e.g., concerns about fairness or about GenAI replacing jobs). Our paper's collaborative acceptance outcome could be a positive data point in that conversation, illustrating that augmentation rather than replacement is a palatable approach.

6.5 Proposed Theoretical Model: AI-Augmented Assurance of Learning

Drawing on the TTF framework and empirical findings, we propose the GenAI-augmented AOL fit model (Figure 4), a theoretical model illustrating how GenAI can enhance, but not replace, human judgment in AOL processes. The model integrates five core constructs: Task Characteristics, Technology Characteristics, Task-Technology Fit, GenAI Outcomes, and Final Outcomes, with Human Expertise and Oversight serving as mediating and moderating influences.

Task Characteristics refer to the specific demands of AOL assessment, such as scalability, consistency, rubric complexity, and the need for timely, actionable feedback. Technology Characteristics capture GenAI's capabilities – including natural language processing, rapid response generation, rubric adherence, and known limitations in areas such as factual verification and contextual sensitivity. The alignment between these two domains determines the degree of task-technology fit, which, when high, facilitates superior assessment outcomes.

Human Expertise iteratively mediates the task-technology fit by enabling effective prompt engineering. Expertise ensures that task requirements are translated into prompt designs that optimize GenAI performance, thereby enhancing the quality of the fit. The GenAI Outcomes include gains in assessment efficiency, scoring consistency, and feedback depth and relevance. Additional benefits, such as increased educator satisfaction and a greater willingness to integrate GenAI into AOL practices, may also emerge as indirect outcomes. The Human Oversight acts as a moderating factor that ensures quality assurance, addresses borderline or ambiguous cases, and contextualizes GenAI-generated outputs within the broader educational goals. Even when TTF is high, human judgment remains critical for validating, refining, and situating GenAI contributions appropriately. In our study, evaluators provided vital oversight, particularly in interpreting nuanced assessment cases and addressing GenAI's limitations.

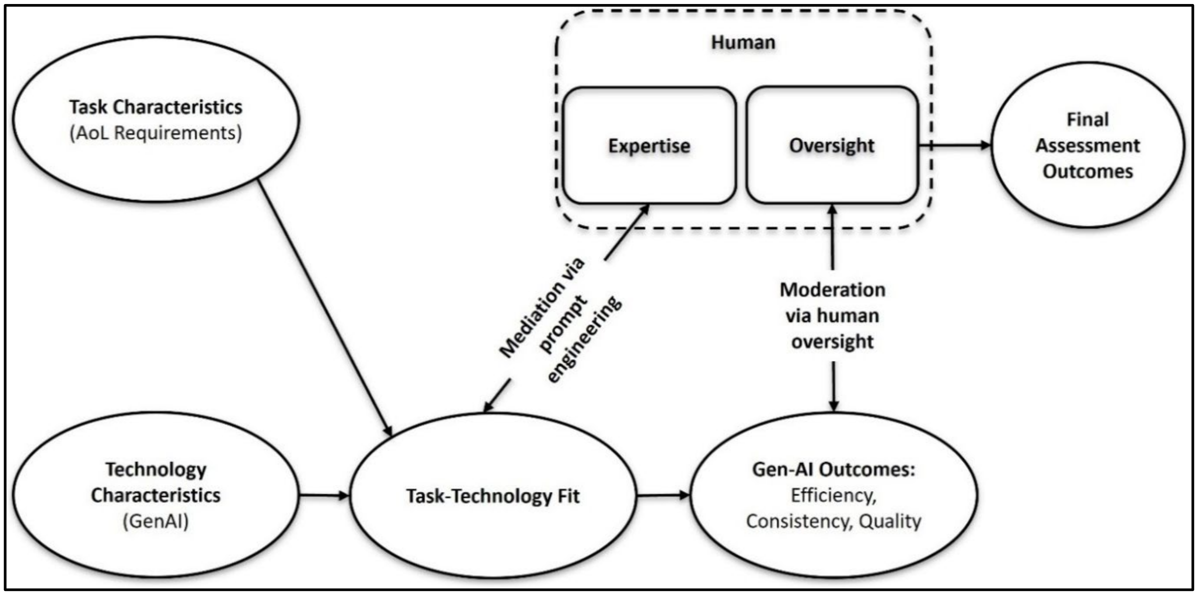

Figure 4. The Human-Integrated GenAI-Augmented Theoretical Model for AOL

The model thus emphasizes augmentation rather than automation, demonstrating that technological capabilities must be aligned – and continuously refined – through human judgment to achieve meaningful educational outcomes. This model serves both as a guide for practice and as a foundation for future research. It posits that effective AOL integration of GenAI occurs when: (1) task and technology characteristics align to produce a high task-technology fit, (2) human expertise iteratively mediates prompt engineering to optimize the fit, (3) human oversight moderates GenAI outcomes to ensure quality and contextual appropriateness, and (4) outcomes are actively monitored, validated, and used to drive continuous improvement. While developed in the context of written communication assessment, the model is readily extensible to other domains where GenAI is applied in educational evaluation.

7. LIMITATIONS

While our paper is based on sound methodology, several limitations must be acknowledged. First, the scope is limited to a single institution and a narrow set of course contexts, which may affect generalizability. The institutional culture, support for innovation, and nature of the assignments and structured writing tasks are conducive to GenAI integration. Outcomes may differ in other settings, particularly where tasks are less structured or institutional attitudes toward GenAI are more conservative.

Second, our comparative evaluation between GenAI and human assessments is observational rather than statistically rigorous. We did not employ blinded rating or formal inter-rater reliability measures; thus, while we observed strong fit, we cannot definitively claim parity in scoring accuracy. Future studies should incorporate experimental controls to assess agreement more robustly.

Third, the effectiveness of GenAI was highly dependent on prompt design. While our prompts were carefully engineered and tested, we treated them as fixed once deployed. This underscores the iterative mediating role of human expertise in the proposed model, where prompt engineering directly affects the task-technology fit. Variations in prompt wording or structure may significantly influence evaluation quality. Systematic investigation of prompt sensitivity remains an open area for research.

Fourth, our findings represent a snapshot in time, based on the then-current version of ChatGPT. As GenAI models evolve rapidly, their capabilities – and potential limitations – will shift. Continuous re-evaluation is essential to ensure the sustained relevance of the empirical findings and the proposed model. Future studies should explore how advances in GenAI impact both the necessity and the nature of human expertise and oversight over time.

Fifth, some student submissions may have been partially or fully generated with AI tools. This raises the possibility of GenAI evaluating content produced by another instance of the same technology. While we cannot confirm the extent of such cases in our dataset, we recognize this as an important limitation. We note that this possibility underscores the need for ongoing oversight by human evaluators and future research into how GenAI assessment tools interact with AI-generated student work.

Sixth, the notion of task-technology fit is a moving target given the rapid and ongoing advances in GenAI capabilities. What constitutes a good fit today may evolve quickly as these tools improve, requiring continual reassessment of GenAI's role in educational contexts. This point highlights the dynamic nature of GenAI adoption and underscores the need for longitudinal research.

Finally, while our model presents a static framework, we recognize that the broader socio-technical system is dynamic. Feedback loops are likely: positive outcomes may lead to expanded use of GenAI, improved technological design, or enhanced human oversight through training. The moderating role of human oversight may itself evolve, becoming either more critical or more supervisory as GenAI capabilities advance. These dynamics warrant a longitudinal study.

Despite these constraints, we contend that the observed trends are robust: when carefully aligned with task demands and supported by active human expertise and oversight, GenAI can play a meaningful role in advancing assurance of learning. TTF remains a valuable lens for understanding and guiding such integration.

8. CONCLUSION

We explore the role of GenAI in augmenting human judgment within AOL processes in business education, including IS programs. Guided by the TTF framework, we examined the integration of GenAI into AOL's student writing assessment. We find that GenAI can serve as an effective ally in assessing student learning outcomes, delivering significant gains in efficiency, consistency, and feedback depth. These benefits were realized without compromising the quality of assessment, provided that GenAI outputs are enhanced through human expertise and subject to appropriate human oversight. The observed fit between GenAI capabilities and AOL task requirements validates the practical utility of TTF in this context.

The institutional decision to formally adopt GenAI-assisted assessment confirms both the feasibility and perceived value of the approach. Yet, this marks the beginning rather than the end of the journey. We plan to develop a more robust, iterative AOL protocol incorporating continuous validation, prompt refinement, and oversight mechanisms to sustain task-technology fit over time. Future phases will continue to operationalize the Human-Integrated GenAI-augmented theoretical model for AOL, adjusting the balance between human expertise and oversight as GenAI technologies evolve.

We contribute to the discourse on GenAI in IS and business education by framing its role as augmentative rather than substitutive. The model highlights how task-technology fit emerges at the intersection of assessment requirements and GenAI capabilities, but the value materializes only when mediated by human expertise and oversight. Prompt engineering channels human knowledge into the system, ensuring relevance and alignment, while oversight moderates GenAI outcomes – efficiency, consistency, and quality – before they inform final assessment outcomes. This symbiotic human-GenAI partnership strengthens Assurance of Learning processes, advancing both quality assurance and curriculum improvement.

9. DECLARATION OF GenAI USAGE

During the preparation of this work, the authors used ChatGPT-4 to conduct the comparative analysis between GenAI and human evaluators, as described in the Methodology section. In addition, ChatGPT-4 was used to assist with editing, specifically to check the consistency of terminology usage across the Abstract, Introduction, and Conclusion. The prompt used was "check for terminology consistency between the attached three sections and provide suggestions for improving the consistency." All ChatGPT-generated outputs were carefully reviewed by the authors, who take full responsibility for the final content of the manuscript.

10. REFERENCES

Andriole, S. J. (2024). The Big Miss: AI Will Replace Just About Everything. *Communications of the Association for Information Systems*, 55(1), 29. https://doi.org/10.17705/1CAIS.05530

Attaway, A. N., Chandra, S., Dos Santos, B. L., Thatcher, M. E., & Wright, A. L. (2011). An Approach to Meeting AACSB Assurance of Learning Standards in an IS Core Course. *Journal of Information Systems Education*, 22(4), 355-366. https://jise.org/volume22/n4/JISEv22n4p355.html

Bahi, A., Gharib, J., & Gahi, Y. (2024). Integrating Generative AI for Advancing Agile Software Development and Mitigating Project Management Challenges. *International Journal of Advanced Computer Science & Applications*, 15(3), 54-61. https://doi.org/10.14569/IJACSA.2024.0150306

Basole, R. C., Park, H., & Seuss, C. D. (2024). Complex Business Ecosystem Intelligence Using AI-Powered Visual Analytics. *Decision Support Systems*, 178, 114133. https://doi.org/10.1016/j.dss.2023.114133

Beard, D., Schweiger, D., & Surendran, K. (2008). Integrating Soft Skills Assessment Through University, College, and Programmatic Efforts at an AACSB Accredited Institution. *Journal of Information Systems Education*, 19(2), 229-240. https://aisel.aisnet.org/jise/vol19/iss2/11

Borschbach, A., & Mescon, T. (2021). Transforming Assurance of Learning for Lasting Impact. https://www.aacsb.edu/insights/articles/2021/05/transforming-assurance-of-learning-for-lasting-impact

D'Ambra, J., Wilson, C. S., & Akter, S. (2013). Application of the Task-Technology Fit Model to Structure and Evaluate the Adoption of E-Books by Academics. *Journal of the American Society for Information Science and Technology*, 64(1), 48-64. https://doi.org/10.1002/asi.22757

Davazdahemami, B., Zolbanin, H., & Zadeh, A. (2024). Revitalizing Scholarly Compass: Harnessing GPT-Powered Automation for Dynamic Aims & Scope Evolution in Academic Journals. *Communications of the Association for Information Systems*, 55, 1029-1068. https://doi.org/10.17705/1CAIS.05539

Davis, F. D. (1985). *A Technology Acceptance Model for Empirically Testing New End-User Information Systems: Theory and Results* (Doctoral Dissertation, Massachusetts Institute of Technology).

Dishaw, M. T., & Strong, D. M. (1999). Extending the Technology Acceptance Model With Task-Technology Fit Constructs. *Information & Management*, 36(1), 9-21. https://doi.org/10.1016/S0378-7206(98)00101-3

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., & Ahuja, M. (2023). Opinion Paper: "So What If ChatGPT Wrote It?" Multidisciplinary Perspectives on Opportunities, Challenges and Implications of Generative Conversational AI for Research, Practice and Policy. *International Journal of Information Management*, 71, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Goodhue, D. L., & Thompson, R. L. (1995). Task-Technology Fit and Individual Performance. *MIS Quarterly*, 19(2), 213-236. https://doi.org/10.2307/249689

Henkin, D., & Zangrilli, C. (2024). Enhancement, Not Elimination: AI's Impact on Jobs in the Indirect Tax Domain. *Communications of the Association for Information Systems*, 55(1), 31. https://doi.org/10.17705/1CAIS.05532

Hessari, H., Bai, A., & Daneshmandi, F. (2024, October 25, Published Online). Generative AI: Boosting Adaptability and Reducing Workplace Overload. *Journal of Computer Information Systems*, 1-14. https://doi.org/10.1080/08874417.2024.2417672

Heyder, T., Passlack, N., & Posegga, O. (2023). Ethical Management of Human-AI Interaction: Theory Development Review. *Journal of Strategic Information Systems*, 32(3), 101772. https://doi.org/10.1016/j.jsis.2023.101772

Hill-Yardin, E. L., Hutchinson, M. R., Laycock, R., & Spencer, S. J. (2023). A Chat (GPT) About the Future of Scientific Publishing. *Brain, Behavior, and Immunity*, 110, 152-154. https://doi.org/10.1016/j.bbi.2023.02.022

Howard, M. C., & Rose, J. C. (2019). Refining and Extending Task-Technology Fit Theory: Creation of Two Task-Technology Fit Scales and Empirical Clarification of the Construct. *Information & Management*, 56(6), 103134. https://doi.org/10.1016/j.im.2018.12.002

Kim, D., Kim, S., Kim, S., & Lee, B. H. (2024, December 14, Published Online). Generative AI Characteristics, User Motivations, and Usage Intention. *Journal of Computer Information Systems*, 1-16. https://doi.org/10.1080/08874417.2024.2442438

Liebrenz, M., Schleifer, R., Buadze, A., Bhugra, D., & Smith, A. (2023). Generating Scholarly Content With ChatGPT: Ethical Challenges for Medical Publishing. *The Lancet Digital Health*, 5(3), e105-e106. https://doi.org/10.1016/S2589-7500(23)00019-5

Lund, B. D., & Wang, T. (2023). Chatting About ChatGPT: How May AI and GPT Impact Academia and Libraries? *Library Hi Tech News*, 40(3), 26-29. https://doi.org/10.1108/LHTN-01-2023-0009

Marimon, F., Mas-Machuca, M., & Akhmedova, A. (2025). Trusting in Generative AI: Catalyst for Employee Performance and Engagement in the Workplace. *International Journal of Human-Computer Interaction*, 41(11), 7076-7091. https://doi.org/10.1080/10447318.2024.2388482

Nishant, R., Schneckenberg, D., & Ravishankar, M. (2024). The Formal Rationality of Artificial Intelligence-Based Algorithms and the Problem of Bias. *Journal of Information Technology*, 39(1), 19-40. https://doi.org/10.1177/02683962231176842

Schlagwein, D., & Willcocks, L. (2023). 'ChatGPT et al.': The Ethics of Using (Generative) Artificial Intelligence in Research and Science. *Journal of Information Technology*, 38(3), 232-238. https://doi.org/10.1177/02683962231200411

Shahid, A. R., & Mishra, S. (2024). A Framework for a Master's in Applied Artificial Intelligence Program in Computer and Information Systems Discipline. *Journal of Information Systems Education*, 35(4), 495-511. https://doi.org/10.62273/EQZE3625

Sundberg, L., & Holmström, J. (2024). Innovating by Prompting: How to Facilitate Innovation in the Age of Generative AI. *Business Horizons*, 67(5), 561-570. https://doi.org/10.1016/j.bushor.2024.04.014

Sun, R., & Deng, X. (2025). Using Generative AI to Enhance Experiential Learning: An Exploratory Study of ChatGPT Use by University Students. *Journal of Information Systems Education*, 36(1), 53-64. https://doi.org/10.62273/ZLUM4022

Susarla, A., Gopal, R., Thatcher, J. B., & Sarker, S. (2023). The Janus Effect of Generative AI: Charting the Path for Responsible Conduct of Scholarly Activities in Information Systems. *Information Systems Research*, 34(2), 399-408. https://doi.org/10.1287/isre.2023.ed.v34.n2

Ulfa, S., Surahman, E., Fatawi, I., & Tsukasa, H. (2024). Task-Technology Fit Analysis: Measuring the Factors That Influence Behavioural Intention to Use the Online Summary-With Automated Feedback in a MOOCs Platform. *Electronic Journal of e-Learning*, 22(1), 63-77. https://doi.org/10.34190/ejel.22.1.3094

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27(3), 425-478. https://doi.org/10.2307/30036540

Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., Yogatama, D., Bosma, M., Zhou, D., & Metzler, D. (2022). Emergent Abilities of Large Language Models. arXiv preprint arXiv:2206.07682.

White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., & Schmidt, D. C. (2023). A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT. arXiv preprint arXiv:2302.11382.
Yang, K., Ji, S., Zhang, T., Xie, Q., Kuang, Z., & Ananiadou, S. (2023). Towards Interpretable Mental Health Analysis With Large Language Models. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing* (pp. 6056-6077). Singapore Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.emnlp-main.370
Yin, R. K. (2013). *Case Study Research: Design and Methods*. Sage Publications.
Zigurs, I., & Buckland, B. K. (1998). A Theory of Task/Technology Fit and Group Support Systems Effectiveness. MIS Quarterly, 22(3), 313-334, https://doi.org/10.2307/249668

AUTHOR BIOGRAPHIES

John X. Volker is a Professor of Management in the College of Business at Austin Peay State University, where he teaches entrepreneurship and creativity. A Teaching & Learning Fellow, he is recognized for his innovative, student-centered pedagogy that emphasizes critical thinking, real-world application, and ethical use of emerging technologies. His research explores GenAI in business education, entrepreneurship as a method, and the intersection of creativity and management. Dr. Volker is committed to preparing students for leadership in today's dynamic business environment, blending academic rigor with practical insights that inspire innovation and lifelong learning.

Al Tilooby is an Assistant Professor of Management Information Systems in the College of Business at Austin Peay State University. He earned his doctorate from the J. Mack Robinson College of Business at Georgia State University. His research explores technology-driven shifts in business, including generative artificial intelligence, blockchain, and cloud computing, with a focus on how these innovations reshape decision-making, risk management, and organizational practices. He teaches courses such as Management Information Systems and Strategic Deployment of AI for Business, preparing students to navigate the evolving intersection of technology and business strategy.

INFORMATION SYSTEMS & COMPUTING ACADEMIC PROFESSIONALS

STATEMENT OF PEER REVIEW INTEGRITY

All papers published in the ***Journal of Information Systems Education*** have undergone rigorous peer review. This includes an initial editor screening and double-blind refereeing by three or more expert referees.

ISSN: 2574-3872 (Online) 1055-3096 (Print)