ISSN: 2473-4901
v9 n5948

A Comparative Analysis of Large Language Models for Writing Teaching Cases

Guido Lang
guido.lang@quinnipiac.edu

Tamilla Triantoro
tamilla.triantoro@quinnipiac.edu

Business Analytics & Information Systems
Quinnipiac University
Hamden, CT 06518, USA

Jason H. Sharp
jason.sharp@uvu.edu
Information Systems & Technology
Utah Valley University
Orem, UT 84058, USA

Abstract

This study explores the potential of large language models (LLMs), specifically GPT-4 and Bard, in generating teaching cases for information systems (IS) courses. A unique prompt for writing three different types of teaching cases (i.e. a descriptive case, a normative case, and a project-based case) on the same IS topic (i.e. the introduction of blockchain technology in an insurance company) was developed and submitted to each LLM. The generated teaching cases from each LLM were subsequently assessed using subjective content evaluation measures (i.e. relevance and accuracy, complexity and depth, structure and coherence, and creativity) as well as objective readability measures (i.e. Automated Readability Index, Coleman-Liau Index, Flesch-Kincaid Grade Level, Gunning Fog Index, Linsear Write Index, and SMOG Index). The findings suggest that while both LLMs perform well on objective measures, GPT-4 outperforms Bard on subjective measures, indicating a superior ability to create content that is more relevant, complex, structured, coherent, and creative. This research provides initial empirical evidence and highlights the promise of LLMs in enhancing IS education while also acknowledging the need for careful proofreading and further research to optimize their use.

Keywords: Bard, GPT-4, comparative analysis, large language model, llm, teaching case

1. INTRODUCTION

Recent advancements in artificial intelligence have led to the development of large language models (LLMs) capable of generating coherent and contextually relevant text. OpenAI's GPT-4 and Google's Bard are among the leading LLMs demonstrating these capabilities. One of the key challenges affecting broader adoption of LLMs is that they sometimes generate text that is nonsensical or factually incorrect – an outcome commonly referred to as a hallucination (Ziwei et al., 2023). However, hallucinations are less problematic in the context of creative writing, which focuses more on imagination and invention over factual correctness (Teubner et al., 2023).

©2023 ISCAP (Information Systems and Computing Academic Professionals)

In this context, teaching cases may present a potentially fruitful area for LLM-generated content.

This study aims to explore and compare the efficacy of LLMs in generating teaching cases for information systems (IS) courses without direct human involvement. Specifically, the main objectives of this study are to (1) determine the ability of different LLMs (i.e. GPT-4 and Bard) in generating coherent and educationally relevant teaching cases, (2) compare the quality and educational suitability of teaching cases generated by different LLMs, and (3) understand the limitations and potential strengths of the different LLMs in creating teaching cases.

In the context of IS education, teaching cases can be broadly classified into three types (Cappel & Schwager, 2000): (1) descriptive non-project-based cases, (2) normative non-project-based cases, and (3) project-based cases. Non-project-based cases facilitate class discussion over one or a limited number of class periods and typically involve a question-and-answer format. They can be descriptive (i.e. presenting past events or decisions and asking students to analyze what went wrong, why it happened, and how it could've been prevented or improved) or normative (i.e. presenting a current situation where the outcome is not provided, thus asking students to develop and evaluate alternatives, and choose and defend the best course of action). Project-based cases require a more extended and sustained effort than non-project-based cases as the solutions to these cases usually involve systems requirements (e.g. creating diagrams, charts, models, documentation, screen prototypes, or an entire system) rather than simply answering questions.

This paper focuses on the ability of different LLMs (i.e. GPT-4 and Bard) to write all three types of IS teaching cases, i.e. descriptive cases, normative cases, and project-based cases. The remainder of this paper is structured as follows. The next section presents an overview of related literature. Section three describes the methodology. Section four presents the findings. Finally, the last section discusses practical implications for IS educators and researchers.

2. RELATED LITERATURE

With the launch of OpenAI's ChatGPT on November 30, 2022 (Ortiz, 2023b), institutions of higher education found themselves quickly scrambling to address its potential impact, both positive and negative, on teaching and learning. A common concern arose over a possible increase in cheating and plagiarism as educators feared that students would simply use ChatGPT to complete assignments. In the span of a few short weeks, numerous websites, webinars, panels, presentations, and discussion forums were up and running.

While ChatGPT initially garnered the most early attention, in the subsequent time since its launch, OpenAI released GPT-4, March 14, 2023, which is considered a more advanced version of ChatGPT (Ortiz 2023a), Google unveiled its own AI ChatBot named Bard on February 6, 2023 (Ortiz, 2023c), and not to be outdone, also in February 2023, Microsoft released an updated version of Bing which included its own AI ChatBot, referred to as Bing Chat, using the same technology used to power ChatGPT (Ortiz, 2023d), and these are just the "big" names. With so much increased interest in LLM a growing stream of publications related to its use in education has followed. Considering the short timeframe since the launch of various LLM, a considerable portion of current writing is found in commentaries, editorials, conference proceedings, trade publications, and personal, professional, and educational websites. However, more and more peer-reviewed, empirically based articles are sure to come as the popularity of LLM grows.

In a sweeping editorial opinion paper led by Yogesh Dwivedi (Dwivedi et al., 2023), leading experts and academicians from education, business, and society share their perspective on the topic of generative conversational AI. This group includes a good number of recognized scholars in the field of IS including Varun Grover, Kai Larsen, Jan Pries-Heje, Frantz Row, Suprateek Sarker, and Michael Wade. A notable contribution of this paper is its treatment of LLMs, specifically ChatGPT, within the realm of academics and education. While the paper is not empirically based, it along with other review type papers (e.g., Jeon, Lee, & Choi, 2023; Zamfiroiu, Vasile, & Savu, 2023), is a good starting point for discussion and provides nine contributions specifically related to educational impact and teaching and learning. The paper is also beneficial in providing a sound set of references for beginning one's journey into the world of LLMs and education.

Similarly, Kasneci et al. (2023), provide a commentary discussing the opportunities and challenges of LLMs for education. With so much attention given to the negative aspects of LLMs in terms of cheating and other academic integrity issues (Cotton, Cotton, & Shipway, 2023), this

paper presents a balancing perspective by examining ways LLMs can be used for creating educational content, improving student engagement and interactions, and personalizing learning experiences. As the authors note, "LLMs, such as ChatGPT, have the potential to revolutionize teaching and assist in teaching processes" (Kasneci et al., 2023, p. 2).

Since the initial release of ChatGPT, a primary concern regarding its use revolves around the issue of academic integrity. Considering that LLMs have the capabilities to assist students in generating papers, responses to short answer questions, and programming code, just to name a few, a reasonable question among educators is whether cheating and plagiarism will increase in this new era of LLM. As such Cotton, Cotton, and Shipway (2023) address opportunities and challenges of using LLMs in higher education as well as the potential risk and rewards of LLMs. The opportunities afforded by LLMs are varied, including communication, student engagement, collaboration, assessment, and feedback. In particular, "one potential opportunity for GPT-3 in higher education is the creation of personalized assessments" (Cotton, Cotton, & Shipway, 2023, p. 2). Providing customized, real-time feedback and grading assistance to educators are also potential opportunities. However, LLMs do not come without their challenges. As noted earlier, plagiarism is a prime concern. However, according to Cotton, Cotton, and Shipway (2023), with appropriate strategies, administrators and instructors can mitigate the impact as follows: (1) ensure that students receive clear and detailed instructions upfront, (2) implement rubrics, and (3) use a variety of assessment methods.

The papers examined thus far provide a solid theoretical and policy-related perspective on the use of LLMs in education, but how do students feel about its use? Strzelecki (2023) examined students' acceptance and use of LLMs – ChatGPT specifically – using the Unified Theory of Acceptance and Use of Technology (UTAUT2). UTAUT2 comprises seven predictors including: (1) performance expectancy, "the degree to which an individual believes that using the system will help him or her attain gains in job performance" (Venkatesh et al., 2003 p. 447), (2) effort expectancy, "the degree of ease associated with the use of the system" (Venkatesh et al., 2003, p. 450), (3) social influence, "the degree to which an individual perceives that important others believe he or she should use the new system" (Venkatesh et al., p. 451), (4) facilitating conditions, "the degree to which an individual believes that an organizational and technical infrastructure exists to support use of the system" (Venkatesh et al., 2003, p. 453), (5) hedonic motivation, "the fun or pleasure derived from using a technology" (Venkatesh et al. 2012, p. 161), (6) price value, "consumers' cognitive tradeoff between the perceived benefits of the applications and the monetary cost for using them" (Venkatesh et al., 2012, p. 161), and (7) habit, "the extent to which people tend to perform behaviors automatically because of learning" (Venkatesh et al., 2012, p. 161). Because at the time of writing, ChatGPT was free, Strzelecki decided to drop "price value" and add an alternative predicator referred to as "personal innovativeness" which is defined as the "willingness of an individual to try out any new information technology" (Agarwal & Prasad, 1998). Based upon self-reported data from 534 students and using Partial Least Squares Structural Equation Modeling, Strzelecki (2023) found that the best predictor of behavioral intention was habit followed by performance expectancy and then hedonic motivation. Behavioral intention was the dominant determinant of use behavior followed by personal innovativeness the alternative predictor added to the model. In sum, the study correlated with prior research in that a strong association exists between performance expectancy and habit. This indicates that students are comfortable with the adoption of technologies such as ChatGPT and that the fact that technologies like ChatGPT are fun to use makes students even more likely to adopt them. Given the characteristics of current students, these findings are not surprising.

Beyond the theory, policy, and acceptance of LLMs, Jacques (2023) provides a more practical application of how LLMs can impact the actual teaching of an introductory programming course. According to Jacques, there are some who contend that it is no longer necessary for students to learn to program, as LLMs can now handle that task. However, Jacques suggests that a "rethinking" of how to teach students to program is the new challenge. As such, Jacques argues that computing disciplines (computer science, in particular) look to math education for guidance with working with multiple representations, experimenting with different approaches, and explaining someone else's solution. She then provides sample assignments and assessment strategies. In conclusion she asserts "specifically, we need to consider how to develop good programmers while still acknowledging and engaging with these new tools for programming" (p. 46).

While the existing literature provides a growing base that examines LLMs in education, from a theoretical, policy, acceptance, and practical standpoint, more empirically-based studies are needed. To this point, this paper helps to fill this gap by presenting a mixed-method approach for comparing two LLMs (GPT-4 and Bard) in the generation of IS teaching cases and addressing the opportunity suggested for the creation of course content (Kasneci et al., 2023).

3. METHODOLOGY

This study uses a mixed-method approach combining automated and human elements in a four-step process. First, a set of teaching cases was created using each of the LLMs (i.e. OpenAI's GPT-4 and Google's Bard). To do so, the LLMs were accessed on May 12, 2023, at OpenAI.com using the paid ChatGPT Plus version of GPT-4) and Google Bard respectively. Each LLM was provided with the same prompt to write three different types of teaching cases (i.e. a descriptive case, a normative case, and a project-based case) on the same topic. This was done to ensure comparability of the generated content. The topic was chosen to be the introduction of blockchain technology in an insurance company. This topic was chosen because the authors noted an absence of IS teaching cases on the topic of blockchain technology in the *Journal of Information Systems Education* (JISE) and other IS education outlets and thus presented an unmet need for IS teaching cases. However, the generated teaching cases have not yet been used in a university course. Thus, the specific topic of blockchain technology simply aims to serve as an illustrative example of a relevant IS topic. The prompt was developed based on Cappel and Schwager's (2000) guidelines for writing IS teaching cases along with best practices for prompt writing (also known as prompt engineering) found in DAIR.AI (2023), Fulford and Ng (2023), Learn Prompting (2023), and Weng (2023). The final prompt can be found in Appendix A. The zero-shot prompt approach (DAIR.AI, 2023), consisting of a single prompt without practice examples or additional follow-up prompts, was used in order to further ensure comparability of the generated output across the LLMs. The resulting three teaching cases (i.e. a descriptive case, a normative case, and a project-based case) for each LLM can be found in Appendix B.

Second, the generated teaching cases were subjected to a detailed content evaluation, which is a type of expert-focused content quality evaluation (De Jong & Schellens, 1997) This involved evaluating the teaching cases from each LLM based on multiple subjective criteria, including:

- **Relevance and Accuracy**: The extent to which the content of the generated teaching case adheres to the prompt (see Appendix A).
- **Complexity and Depth**: The level of detail and complexity of the concepts discussed in the generated teaching case.
- **Structure and Coherence**: The overall organization, logical flow, and readability of the generated teaching case.
- **Creativity**: The originality of the scenarios and solutions proposed in the generated teaching case.

The authors of this study independently evaluated the generated teaching cases using the above criteria. Specifically, the first author, who generated the teaching cases using the different LLMs (i.e. GPT-4 and Bard), shared the cases without revealing which LLM generated the cases with the other two authors. The authors then independently evaluated the teaching cases on each criterium using a scale where 1 = poor, 2 = fair, 3 = good, 4 = very good, 5 = excellent. The authors used their extensive experience with traditional teaching cases as a frame of reference when evaluating the generated teaching cases.

Third, the generated teaching cases were assessed using natural language processing (specifically the Python package "py-readability-metrics" version 1.4.5) to obtain objective measures of readability. The chosen measures of readability are among the most widely used readability measures (Pitler & Nenkova, 2008) and all produce an approximate representation of the U.S. grade level needed to comprehend the text (1 = First Grade, 2 = Second Grade, ..., 12 = Twelfth Grade, 13 = First-year undergraduate, 14 = Second-year undergraduate, etc.). While there are differences between the various readability measures, most of them calculate ratios representing word difficulty (e.g. number of letters per word) and sentence difficulty (e.g. number of words per sentence). The objective measures of readability used include:

- **Automated Readability Index** (Kincaid, 1975)
- **Coleman-Liau Index** (Coleman & Liau, 1975)
- **Flesch-Kincaid Grade Level** (Kincaid, 1975)
- **Gunning Fog Index** (Gunning, 1952)
- **Linsear Write Index** (Klare, 1974)
- **SMOG Index** (McLaughlin, 1969)

Finally, the results from the subjective content evaluation and the objective readability metrics were then compared across the different LLMs (i.e. GPT-4 and Bard) to understand their strengths and weaknesses in generating IS teaching cases.

4. RESULTS

Subjective Content Evaluation

Both LLMs generated teaching cases that followed roughly the same outline, which is typical and appropriate for teaching cases (cf. Cappel & Schwager, 2000): Introduction, background, situation, followed by questions. GPT-4 generated three unique and specific teaching cases that adequately reflected the prompt for being descriptive (i.e. focusing on a past event, see Appendix B-1), normative (i.e. focusing on a current or future event, see Appendix B-2), or project-based (i.e. requiring the creation of more complex deliverables, see Appendix B-3). However, Bard generated three teaching cases that were more or less similar and general and did not clearly differ in their focus (see Appendix C-1 to C-3).

Before proceeding with the subjective content evaluation, Fleiss' Kappa was computed to assess interrater reliability (Fleiss, Levin, & Paik, 2003). Its value of $\kappa = 0.36$ ($p < 0.001$) suggests there was significant agreement between the raters. Table 1, below, provides the subjective content evaluation aggregated across the three teaching cases for each LLM. The unaggregated results for each teaching case (i.e. descriptive, normative, and project-based) can be found in Appendix C.

A two-way repeated measures ANOVA was performed to analyze the effect of LLM (i.e. GPT-4 vs. Bard) and teaching case type (i.e. descriptive case vs. normative case vs. project-based case) on the subjective content evaluation. There was a statistically significant difference in the subjective content evaluation between the LLMs ($F(1, 3) = 48.86$, $p < 0.001$), but not between the teaching case types ($F(2, 6) = 1.25$, $p = 0.35$). As shown in Table 1 above, GPT-4 had higher subjective content evaluation scores than Bard. Thus, based on the subjective content evaluation, GPT-4 appears to be consistently better than Bard at writing teaching cases. However, it shall be noted that GPT-4's output was not error-free, as it made a mistake when calculating the sum of several fictitious values in one of the teaching cases (see Appendix B-1: Exhibit A).

	GPT-4	**Bard**
Relevance and Accuracy	4.89 (0.33)	2.89 (1.05)
Complexity and Depth	4.89 (0.33)	2.22 (0.83)
Structure and Coherence	5.00 (0.00)	3.67 (0.71)
Creativity	4.56 (0.53)	2.00 (0.50)
M (SD)	4.83 (0.38)	2.69 (1.01)

Table 1: Subjective Content Evaluation of Each LLM Aggregated Across Teaching Cases

Objective Measures of Readability

While GPT-4 generated slightly longer teaching cases ($M_{words} = 691$, $SD = 154$) than Bard ($M_{words} = 412$, $SD = 73.5$), that difference was not significant ($t(2) = 2.67$, $p = 0.12$). Table 2, below, provides the objective readability scores aggregated across the three teaching cases for each LLM. The unaggregated results for each teaching case (i.e. descriptive, normative, and project-based) can be found in Appendix D.

The overall average readability measures of 13.6 and 13.7 for GPT-4 and Bard, respectively, suggest that the teaching cases generated by both LLMs were appropriate for second-year undergraduate students. A two-way repeated measures ANOVA was performed to analyze the effect of LLM (i.e. GPT-4 vs. Bard) and teaching case type (i.e. descriptive case vs. normative case vs. project-based case) on the objective readability measures. The interaction between LLM and teaching case type was significant ($F(2, 10) = 54.81$, $p < 0.001$). Pairwise post-hoc comparisons with Bonferroni correction were conducted, which found significant differences in the readability measures between the LLMs for normative cases ($\Delta M = 2.15$, $t(2) = 6.67$, $p_{corr} < 0.01$) and for project-based cases ($\Delta M = -3.15$, $t(2) = -8.14$, $p_{corr} < 0.001$), but not for descriptive cases ($\Delta M = 0.58$, $t(2) = 1.38$, $p_{corr} > 0.05$). In other words, while GPT-4 had slightly higher required reading levels for the normative case, Bard had slightly higher required reading levels for the project-based case. However, there was no significant difference in the required reading levels between the LLMs for the descriptive case. Thus, based on the objective readability measures, both LLMs were able to generate content that is generally appropriate for the intended audience of undergraduate and graduate students.

	GPT-4	Bard
Automated Readability Index	13.1 (1.8)	13.3 (1.9)
Coleman-Liau Index	15.7 (1.0)	15.5 (1.8)
Flesch-Kincaid Grade Level	13.2 (1.6)	12.6 (1.6)
Gunning Fog Index	15.4 (1.2)	16.4 (1.9)
Linsear Write Index	11.9 (3.1)	12.5 (1.5)
SMOG Index	12.2 (1.2)	12.0 (1.2)
M (SD)	13.6 (1.5)	13.7 (1.7)

Table 2: Objective Readability Measures for Each LLM Aggregated Across Teaching Cases

5. DISCUSSION

LLMs are generally well-suited for creative writing, as issues such as hallucinations and outdated information are less concerning in creative writing tasks. This study aimed to better understand how different LLMs (i.e. GPT-4 and Bard) perform in a specific type of creative writing: writing teaching cases. After a zero-shot prompt was developed combining best practices for prompt writing with best practices for teaching case writing, the same prompt was administered to GPT-4 and Bard. Subsequently, each LLM generated three different types of teaching cases on the same topic: a descriptive case, a normative case, and a project-based case. The teaching cases were then rated blindly across four dimensions by the study's authors to determine a subjective content evaluation. Moreover, the teaching cases were compared using six different objective measures of readability. The results found that, in terms of content, GPT-4 created teaching cases that were more specific and overall better across all dimensions: relevance and accuracy, complexity and depth, structure and coherence, and creativity. However, in terms of readability, the results were less clear cut: Both LLMs were able to generate cases that are appropriate for undergraduate and graduate students.

Based on the results, it appears that both GPT-4 and Bard are able to generate coherent and educationally relevant teaching cases. However, GPT-4 appears to write teaching cases that are of better quality than Bard. However, also GPT-4 is not perfect, as its output included a factual error when summing up several fictitious values in an exhibit to one of the teaching cases. Taken together, it is clear that LLMs, such as GPT-4, are very well-suited to writing teaching cases for IS courses. As such, IS educators would be well-advised to leverage LLMs to quickly develop custom teaching cases for their courses – as long as they carefully proofread and double-check any calculations included in the teaching cases generated by an LLM.

Contributions

The rapid evolution of technology and the dynamic nature of the IS field necessitate the continuous creation of new learning materials. Traditional teaching cases, while effective, can become outdated quickly due to the fast-paced changes in technology and business environments. Moreover, the development of these cases is often time-consuming and labor-intensive, which can limit the ability of educators to keep their teaching materials current and relevant. In this context, the ability to generate new, high-quality teaching cases quickly and efficiently becomes crucial. It allows educators to keep pace with the changing landscape of the IS field and ensures that students are learning from cases that reflect the current state of the industry. This not only enhances the educational experience but also better prepares students for their future careers in the IS field.

This research contributes to addressing the need for fresh, relevant learning materials by evaluating the potential of LLMs, specifically GPT-4 and Bard, in generating three types of teaching cases for IS courses: descriptive cases, normative cases, and project-based cases. By doing so, the study provides empirical evidence on the ability of LLMs to generate coherent and educationally relevant teaching cases, offering a promising solution to the challenge of creating new learning materials.

Furthermore, this research contributes to the field by developing a comprehensive evaluation framework to assess the quality of teaching cases generated by LLMs. This framework includes both objective measures, such as readability scores, and subjective measures based on expert opinion. Objective measures, such as readability scores, assess the clarity, complexity, and comprehensibility of the text. These measures are typically based on quantifiable factors such as sentence length, word length, and vocabulary difficulty. Subjective measures, on the other hand, rely on expert opinion and involve a more

nuanced evaluation of the content. These measures assess aspects such as relevance and accuracy, complexity and depth, structure, coherence, and creativity. Unlike objective measures, these are not easily quantifiable and require a deep understanding of the subject matter and the educational goals of the teaching cases. The study's findings indicate that while both GPT-4 and Bard perform well on objective measures and can produce readable and comprehensible text, GPT-4 outperforms Bard on subjective measures, suggesting a superior ability to create content that is more relevant, complex, structured, coherent, and creative.

Another significant contribution of this study is the development of a prompt for writing teaching cases. This prompt, designed by combining best practices for prompt writing with best practices for teaching case writing, serves as a unique tool to guide LLMs in generating educationally relevant teaching cases. This development is an important step in leveraging LLMs for educational purposes, as it provides a structured way to direct the LLMs' creative writing capabilities towards generating content that is not only coherent and readable but also educationally valuable.

The research also contributes to the broader discussion on the role of artificial intelligence in education. By demonstrating the potential of LLMs in generating teaching cases, the study adds to the growing body of evidence supporting the use of AI tools in enhancing teaching and learning experiences. This could pave the way for further exploration of AI's potential in various educational contexts.

Limitations and Future Research

While the findings of this study offer valuable insights into the potential of LLMs in generating teaching cases, it is important to note that the results are context-specific and may not be universally applicable. The study's scope was limited to specific LLMs, prompts, and case selections, which could influence the generalizability of the findings. As such, different LLMs or prompts might yield different results, underscoring the need for future research to explore a variety of designs and prompts.

Moreover, although the evaluation of the teaching cases relied on both objective and subjective measures, the latter was based on the authors' perspectives. Future studies could enhance the robustness of the evaluation by involving a more diverse panel of evaluators, thus broadening the range of considered perspectives.

This study did not incorporate student feedback or assess learning outcomes, which are important elements in evaluating the educational value of teaching cases. Future research could provide a more comprehensive evaluation by integrating these elements, offering a more holistic view of the impact of LLM-generated teaching cases on student learning.

Finally, it is worth mentioning that some content generated by LLMs may fall outside the purview of copyright protection based on the human authorship requirement (USCO 2023). Thus, LLM generated cases may be vulnerable to duplication and dissemination without legal recourse if published.

6. CONCLUSION

By suggesting an alternative approach to writing teaching cases using LLMs, we tested the potential of LLMs, specifically GPT-4 and Bard, in generating teaching cases for IS courses. The findings suggest that LLMs can generate coherent, educationally relevant, and creatively engaging teaching cases, offering a promising solution to the challenge of creating fresh and relevant learning materials. While both GPT-4 and Bard perform well on objective measures, GPT-4 outperforms Bard on subjective measures, indicating a superior ability to create content that is more relevant, complex, structured, coherent, and creative.

The findings suggest that using LLMs not only reduces the time and effort required to develop new cases but also allows for the generation of a diverse range of cases that can cater to different learning objectives and student needs. However, the findings also acknowledge the need for careful proofreading of the generated cases, as LLMs are not without limitations. In light of these findings, LLMs hold great promise for enhancing IS education and can further enrich learning experience and contribute to the advancement of IS education.

7. REFERENCES

Agarwal, R., & Prasad, J. (1998). A Conceptual and Operational Definition of Personal Innovativeness in the Domain of Information Technology. *Information Systems Research*, 9(2), 204-215.

Cappel, J. J., & Schwager, P. H. (2020). Writing IS Teaching Cases: Guidelines for JISE *Journal of Information Systems Education*, 13(4), 287-294.

Coleman, M., & Liau, T. L. (1975). A Computer Readability Formula Designed for Machine Scoring. *Journal of Applied Psychology*, 60, 283–284.

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and Cheating: Ensuring Academic Integrity in the Era of ChatGPT. *Innovations in Education and Teaching International*, 1-12.

Crossley, S. A., Greenfield, J., & McNamara, D. S. (2008). Assessing Text Readability Using Cognitively Based Indices. *TESOL Quarterly*, *42*(3), 475–493.

DAIR.AI (2023). Prompt Engineering Guide. Available online at https://www.promptingguide.ai/

De Jong, M., & Schellens, P. J. (1997). Reader-Focused Text Evaluation: An Overview of Goals and Methods. *Journal of Business and Technical Communication*, *11*(4), 397-513.

Dwivedi, Y. K., et al. (2023). "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, 1-63.

Fleiss, J. L., Levin, B., & Paik, M. C. (2003). *Statistical Methods for Rates and Proportions*, 3rd Edition. New York: John Wiley & Sons.

Flesch, R. (1948). A New Readability Yardstick. *Journal of Applied Psychology*, *32*(3), 221–233.

Fulford, I. & Ng, A. (2023). ChatGPT Prompt Engineering for Developers. Available online at https://www.deeplearning.ai/short-courses/chatgpt-prompt-engineering-for-developers/

Gunning, R. (1952). *The Technique of Clear Writing*. McGraw-Hill: New York, NY.

Jacques, L. (2023). Teaching CS-101 at the Dawn of ChatGPT. *ACM Inroads*, 14(2), 40-46.

Jeon, J., Lee, S., & Choi, S. (2023). A Systematic Review of Research on Speech-Recognition Chatbots for Language Learning: Implications for Future Directions in the Era of Large Language Models. *Interactive Learning Environments*, 1-19.

Kasneci, E., et al. (2023). ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education. *Learning and Individual Differences*, 103, 1-9.

Kincaid, J. P., et al. (1975). *Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel*. Institute for Simulation and Training. 56.

Klare, G. R. (1974). Assessing Readability. *Reading Research Quarterly*, *10*(1), 62-102.

Learn Prompting: Your Guide to Communicating with Artificial Intelligence (2023). Available online at https://learnprompting.org/

McLaughlin, G. H. (1969). SMOG Grading: A New Readability Formula. *Journal of Reading*, *12*(8), 639-646.

Ortiz, S. (2023a, March 15). *What is GPT-4? Here's everything you need to know*. ZDNET. https://www.zdnet.com/article/what-is-gpt-4-heres-everything-you-need-to-know/

Ortiz, S. (2023b, May 30). *What is ChatGPT and why does it matter? Here's what you need to know*. ZDNET. https://www.zdnet.com/article/what-is-the-new-bing-heres-everything-you-need-to-know/

Ortiz, S. (2023c, June 1). *What is Google Bard? Here's everything you need to know*. ZDNET. https://www.zdnet.com/article/what-is-google-bard-heres-everything-you-need-to-know/

Ortiz, S. (2023d, June 9). *What is Bing Chat? Here's everything you need to know.* ZDNET. https://www.zdnet.com/article/what-is-the-new-bing-heres-everything-you-need-to-know/

Pitler, E., & Nenkova, A. (2008). Revisiting Readability: A Unified Framework for Predicting Text Quality. *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*, 186-195.

Štajner, S., et al. (2012). What can Readability Measures Really Tell Us About Text Complexity? *Proceedings of the Workshop on Natural Language Processing for Improving Textual Accessibility (NLP4ITA)*, 14-21.

Strzelecki, A. (2023). To Use or Not to Use ChatGPT in Higher Education? A Study of Students' Acceptance and Use of Technology. *Interactive Learning Environments*, 1-14.

Teubner, T., et al. (2023). Welcome to the Era of ChatGPT et al.: The Prospects of Large Language Models. *Business & Information Systems Engineering*, 65, 95-101.

ISSN: 2473-4901
v9 n5948

United States Copyright Office (USCO). (2023). Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence. Available online at https://copyright.gov/ai/ai_policy_guidance.pdf

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.

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology. *MIS Quarterly*, 36(1), 157-178.

Weng, L. (2023). Prompt Engineering. Available online at https://lilianweng.github.io/posts/2023-03-15-prompt-engineering/

Zamfiroiu, A., Vasile, D., & Savu, D. (2023). ChatGPT - A Systematic Review of Published Papers. *Informatica Economica*, 27(1), 5-16.

Ziwei, J., et al. (2023). Survey of Hallucination in Natural Language Generation. *ACM Computing Surveys*, *55*(12), 1–38.

©2023 ISCAP (Information Systems and Computing Academic Professionals)

Appendix A
Final Prompt

Act as an information systems (IS) teaching case writer.

While all types of teaching cases are designed to promote analysis and problem solving in a university classroom, IS teaching cases differ somewhat in terms of their focus and presentation. IS teaching cases can be classified into three types: (1) descriptive non-project-based cases, (2) normative non-project-based cases, and (3) project-based cases. Non-project-based cases facilitate class discussion over one or a limited number of class periods and typically involve a question-and-answer format. They can be descriptive (i.e. presenting past events or decisions and asking students to analyze what went wrong, why it happened, and how it could've been prevented or improved) or normative (i.e. presenting a current situation where the outcome is not provided, thus asking students to develop and evaluate alternatives, and choose and defend the best course of action). Project-based cases require a more extended and sustained effort than non-project-based cases as the solutions to these cases usually involve systems requirements (e.g. creating diagrams, charts, models, documentation, screen prototypes, or an entire system) rather than simply answering questions.

A good IS teaching case has four characteristics: (1) It provides realism: The case should feel authentic, whether it's based on real or fictional events. This can be achieved through the use of quotes, actual events, and visual aids like tables or figures. (2) It is of appropriate length: The case should be long enough for readers to perform the required analysis and address the issues raised. The case's complexity and objectives will determine its length. The case should strike a balance between providing too much or too little information and include both relevant and irrelevant details. (3) It is objective in presentation and tone: The writer should present the case facts without expressing opinions to maintain objectivity and encourage students' analysis and interpretation. (4) It has a hook: The case should start with a statement or short paragraph that grabs readers' interest and defines the case's focus, often presenting a decision-making dilemma.

Write a {*descriptive non-project-based*|*normative non-project-based*|*project-based*} case for the introduction of blockchain technology in an insurance company. Make sure it adheres to the characteristics of a good IS teaching case.

Note: The above prompt was submitted three times to each LLM. Each time, only one of the three options presented in curly brackets ({...|...|...}) was used. Thus, the first prompt ended with "Write a descriptive non-project-based case for...", the second prompt ended with "Write a normative non-project-based case for...", and the third prompt ended with "Write a project-based case for..."

Appendix B-1
Descriptive Case Written by GPT-4

Case Title: Bridging Trust with Technology: The Blockchain Initiative at InsureCo

Introduction:

In late 2023, the executive board of InsureCo, a medium-sized but growing insurance company, was contemplating a significant decision. The company's CIO, Sara Jackson, proposed an ambitious plan to introduce blockchain technology into their existing systems to enhance transparency, reduce fraud, and improve customer trust. However, this proposal faced significant resistance from some senior executives who doubted the utility and necessity of this technology in the insurance sector.

Background:

Founded in 2010, InsureCo had grown to become a significant player in the insurance industry. They were well-known for their innovative approach to business and their commitment to integrating technology to streamline operations and improve customer service. Despite their successful track record, the company was constantly looking for ways to improve their business and stay ahead of competitors.

The Blockchain Proposal:

Sara Jackson saw an opportunity in blockchain technology. A decentralized, transparent ledger system, blockchain would offer numerous benefits to InsureCo, including reduced fraud, streamlined claim settlement processes, and enhanced trust with customers. Jackson proposed a pilot project, which would involve the integration of blockchain technology into InsureCo's existing systems, with a budget of $2 million and an estimated timeline of 18 months.

The Opposition:

However, Jackson's proposal was met with skepticism from a significant portion of the executive board. They argued that the risks and costs associated with implementing a relatively untested technology in the insurance sector outweighed the potential benefits. Additionally, they voiced concerns about cybersecurity, the potential for job loss due to automation, and the challenge of integrating the technology with existing systems.

The Dilemma:

While Jackson and her supporters believed that blockchain could be a game-changer for InsureCo, the opposition was equally adamant about the potential drawbacks. The executive board had to decide whether to approve the project, despite the risks and challenges. This decision was made more complicated by the fact that InsureCo's major competitors had not yet adopted blockchain technology, leaving little industry precedent or benchmark for such an undertaking.

Case Questions:

1. Analyze the potential benefits and drawbacks of implementing blockchain technology at InsureCo. What factors should the executive board consider in their decision-making process?
2. Based on the details presented in the case, make an argument for or against the proposed blockchain project. Justify your position.
3. How might the introduction of blockchain affect InsureCo's relationship with its clients and competitors? Consider the potential impacts on trust, reputation, and market share.
4. If the blockchain project is approved, what steps should InsureCo take to mitigate the potential risks and challenges associated with this technology?
5. What role should employees' opinions (as reflected in Exhibit C) play in the executive board's decision? Should their concerns about job loss and cybersecurity be considered?

Exhibit A: Proposed Budget Breakdown for Blockchain Project

Item	Cost (USD)
Blockchain Technology Infrastructure	$500,000
Software and Licensing	$200,000
System Integration	$400,000
Staff Training	$100,000
Project Management	$150,000
Contingency Funds	$150,000
Legal & Compliance Review	$100,000
Cybersecurity Measures	$300,000
Total	$2,000,000[1]

Exhibit B: Potential Benefits of Blockchain in Reducing Fraudulent Claims

The following charts show a hypothetical representation of how blockchain could potentially reduce fraudulent insurance claims at InsureCo. It's important to note that these figures are estimates and are based on research done on blockchain technology's capability to reduce fraud.

Chart 1: Fraudulent Claims at InsureCo (Current State)

- Number of Claims Filed: 10,000
- Number of Fraudulent Claims Detected: 1,000 (10%)

Chart 2: Estimated Fraudulent Claims at InsureCo (After Blockchain Implementation)

- Number of Claims Filed: 10,000
- Estimated Number of Fraudulent Claims Detected: 1,500 (15%)

This potential increase in detected fraudulent claims could lead to significant savings for InsureCo.

Exhibit C: Internal Survey of Employee Attitudes Towards Adoption of Blockchain Technology

The company conducted an anonymous survey of 500 employees to understand their attitudes towards the proposed blockchain project. The following are the summarized results:

Question: Do you believe the adoption of blockchain technology will benefit InsureCo?

- Yes: 68%
- No: 15%

[1] Note: The LLM generated an incorrect total. It should be $1,900,000 (not $2,000,000).

- Unsure: 17%

Question: Are you concerned about job loss due to automation and the introduction of blockchain technology?

- Yes: 52%
- No: 32%
- Unsure: 16%

Question: Are you worried about potential cybersecurity threats related to the adoption of blockchain technology?

- Yes: 57%
- No: 28%
- Unsure: 15%

Appendix B-2
Normative Case Written by GPT-4

Case Study: Adapting to Innovation - Blockchain Technology in SecureLife Insurance

The Hook:

The senior management of SecureLife Insurance is contemplating implementing blockchain technology across its operations. They are convinced about the potential of this disruptive technology but are also aware of the considerable challenges involved. It is not about whether to adopt blockchain, but when and how.

Background:

SecureLife Insurance, established in 1978, is a reputable insurance company in the United States that provides various types of insurance services such as life, health, auto, and property insurance. The company has always been a pioneer in adopting technology solutions to enhance customer experience and operational efficiency.

The digital transformation journey of SecureLife started in the early 2000s when the company adopted an insurance core system to automate its critical processes. In recent years, the company has started leveraging big data analytics for risk assessment and fraud detection. Despite these advancements, the company faces several challenges, including lack of transparency, slow claim processing, data security issues, and high operational costs.

The Dilemma:

As the company's Chief Information Officer, you've come across blockchain technology and believe it could revolutionize the way SecureLife operates. The technology can provide enhanced security, faster transactions, and greater transparency, transforming the customer experience. But, implementing it could pose significant challenges - the technology is still evolving, the regulatory landscape is uncertain, and the required investment is substantial.

To move forward, you've organized a meeting with the company's top executives to propose the introduction of blockchain technology in SecureLife's operations. They are intrigued but unsure about the risks and benefits associated with this new technology. They want a detailed plan - a feasible strategy that minimizes risks while maximizing the benefits of blockchain technology.

The Questions:

1. Analyze the potential benefits and risks associated with the implementation of blockchain technology in SecureLife Insurance. What could be the impacts on the various stakeholders (e.g., policyholders, employees, regulatory bodies, and partners)?
2. Considering the potential benefits and risks, should SecureLife Insurance move forward with the implementation of blockchain technology? If yes, how should the company phase its implementation to manage risks effectively?
3. How can SecureLife communicate its new strategic plan to its various stakeholders to ensure a smooth transition?
4. What measures should SecureLife put in place to deal with the possible challenges during the implementation of blockchain technology?
5. What type of personnel and expertise might be required for such a project? Consider both in-house talent and external expertise.

Your job as the CIO is to address the concerns raised, craft a viable strategy, and help the company make an informed decision about adopting blockchain technology. This case challenges you to think critically about the application of emerging technologies in established industries, and how to navigate the complexities of such a change.

Exhibit 1: SecureLife's Organizational Structure

SecureLife's organizational structure is hierarchical, with the CEO at the top followed by several vice presidents overseeing different departments such as Marketing, Operations, Underwriting, Claims, Customer Service, and IT. Each department is headed by a manager and consists of several teams working in specialized roles.

Exhibit 2: SecureLife's Operational Processes

1. Underwriting Process: Policies are issued after risk evaluation. Complex algorithms and data analysis help in making decisions. This process is largely automated.
2. Claim Management Process: This involves filing, processing, and settling claims. Although there are automated systems in place, several steps still require manual intervention, leading to slower processing times.
3. Customer Service Process: It includes managing customer interactions across various channels. The department uses a CRM system for customer service management.

Exhibit 3: Overview of Blockchain Technology

Blockchain technology is a type of distributed ledger technology. In essence, it's a chain of blocks, where each block records a collection of transactions. Once a block is added to the blockchain, the information it contains is very difficult to change, lending to its security and transparency. These features make it suitable for a variety of applications, including secure transactions, supply chain management, and secure sharing of records.

Exhibit 4: Summary of SecureLife's Current Challenges

1. Lack of Transparency: Current systems do not provide a clear, real-time view of claim status leading to customer dissatisfaction.
2. Slow Claim Processing: While automation has increased processing speed, several manual interventions still slow down the process.
3. Data Security: Concerns over the security of policyholder's sensitive data.
4. High Operational Costs: Despite automation, operational costs remain high due to inefficiencies in claim management and customer service.

Exhibit 5: Trends in Insurance Industry Technology Adoption

The insurance industry has seen a significant uptake in technology adoption over the last decade. From core system automation to the use of AI and big data analytics, insurance companies are exploring innovative technologies to improve efficiency and customer experience. More recently, blockchain has emerged as a technology with significant potential for transforming the insurance industry, with its promise of enhanced security, transparency, and efficiency. However, it also brings challenges, with regulatory uncertainties and implementation complexities being the major hurdles.

©2023 ISCAP (Information Systems and Computing Academic Professionals)

Appendix B-3
Project-Based Case Written by GPT-4

Teaching Case: Blockchain Transformation in United Insurance

Introduction

United Insurance is a traditional insurance company that has been operating for over 50 years. It boasts a large customer base spread across multiple states. Recently, United Insurance has seen the need to innovate their business processes and is considering integrating blockchain technology into their operations. The executive team believes that blockchain may streamline processes, decrease fraudulent claims, and ultimately increase customer satisfaction. However, the implementation of such a system is fraught with challenges, especially given the age and scale of the firm.

Company Overview

United Insurance is a diversified insurance company offering auto, home, life, and commercial insurance products. It is recognized for its commitment to quality service and customer trustworthiness. However, in an era of digital transformation, the company acknowledges that maintaining this reputation requires a paradigm shift in their operating models.

The Situation

The Chief Information Officer (CIO), Mary Jones, is tasked with evaluating the feasibility of blockchain technology within United Insurance. She recognizes that this could be a game-changer for the company but is also aware of the challenges it poses.

Fraud detection and prevention is a key area of concern for United Insurance. Each year, fraudulent claims cost the company millions of dollars. Blockchain, with its decentralized, transparent, and immutable nature, could potentially tackle this issue head-on.

Moreover, Mary believes that blockchain could expedite claim processing times, reduce administrative costs, and improve the customer experience. Yet, she must consider the potential roadblocks such as legal and regulatory hurdles, technology compatibility, security concerns, costs, and the required cultural shift within the organization.

Your Role

As a group, you are part of the United Insurance's Information Systems team reporting to the CIO. You have been asked to develop a comprehensive plan to implement blockchain technology within the firm. The following questions may guide your analysis:

1. What are the potential benefits of implementing blockchain technology at United Insurance?
2. What could be the potential hurdles to its successful implementation?
3. How could you address these potential hurdles?
4. What type of blockchain technology (public, private, or consortium) would best fit the needs of United Insurance? Why?
5. How will you ensure the security of data and transactions on the blockchain?
6. How would you manage the necessary change within the organization to ensure the successful implementation of blockchain technology?

Deliverables

By the end of the project, your team is expected to provide:

1. A detailed project plan outlining how you intend to implement blockchain technology at United Insurance.
2. A risk analysis covering potential issues and mitigation plans.
3. A change management strategy to ensure acceptance of the new technology by the staff.
4. A prototype or model demonstrating how the blockchain system might work within the context of United Insurance.

Conclusion

United Insurance stands at a crossroads, with the potential to change the face of their business forever. The decisions you make will not only determine the success of this project but the future of the company. Your recommendations should be innovative, comprehensive, and firmly grounded in the realities of the business. The fate of United Insurance's blockchain transformation is now in your hands.

Appendix C-1
Descriptive Case Written by Bard

Introduction

Blockchain technology is a distributed ledger technology that has the potential to revolutionize the insurance industry. By providing a secure, transparent, and tamper-proof way to store and share data, blockchain could help insurers to improve efficiency, reduce costs, and increase customer satisfaction.

Background

The insurance industry is a $5 trillion dollar industry that is ripe for disruption. The industry is facing a number of challenges, including rising costs, increasing competition, and changing customer demands. Blockchain technology could help insurers to address these challenges and stay ahead of the competition.

Benefits of Blockchain

Blockchain technology offers a number of benefits that could be of value to insurers, including:

- Security: Blockchain is a secure technology that uses cryptography to protect data from unauthorized access. This could help insurers to reduce the risk of fraud and data breaches.
- Transparency: Blockchain is a transparent technology that allows all participants in a network to view the same data. This could help to improve trust and collaboration between insurers and their customers.
- Efficiency: Blockchain can automate many of the manual processes involved in insurance, such as underwriting and claims processing. This could help insurers to reduce costs and improve efficiency.

Challenges of Blockchain

While blockchain technology offers a number of benefits, there are also a number of challenges that insurers will need to address before they can fully adopt the technology. These challenges include:

- Cost: Blockchain is a new technology and there is still a lack of expertise in the insurance industry. This could make it expensive for insurers to implement blockchain solutions.
- Complexity: Blockchain is a complex technology and it can be difficult to implement and manage. This could make it difficult for insurers to adopt blockchain solutions without significant investment in training and resources.
- Regulation: Blockchain is a new technology and there is still a lack of regulatory clarity around its use in the insurance industry. This could make it difficult for insurers to adopt blockchain solutions without risking regulatory compliance issues.

Decision

The decision of whether or not to adopt blockchain technology is a complex one. Insurers will need to weigh the benefits and challenges of the technology before making a decision. If an insurer decides to adopt blockchain, it will need to carefully consider its implementation strategy and address the challenges that have been identified.

Questions

- What are the specific benefits that blockchain technology could offer to the insurance industry?
- What are the specific challenges that insurers will need to address before they can fully adopt blockchain technology?
- What is the best way for insurers to implement blockchain technology?
- What are the risks associated with using blockchain technology in the insurance industry?

ISSN: 2473-4901
v9 n5948

Conclusion

Blockchain technology is a promising new technology that has the potential to revolutionize the insurance industry. However, there are a number of challenges that insurers will need to address before they can fully adopt the technology. Insurers that are able to successfully implement blockchain technology could gain a significant competitive advantage.

©2023 ISCAP (Information Systems and Computing Academic Professionals)

Appendix C-2
Normative Case Written by Bard

Introduction

Blockchain technology is a distributed ledger technology that allows for secure, transparent, and tamper-proof transactions. It has the potential to revolutionize a wide range of industries, including insurance.

Insurance companies are constantly looking for ways to reduce costs and improve efficiency. Blockchain technology can help them do both. For example, blockchain can be used to automate claims processing, which can save time and money. It can also be used to create a more secure and transparent claims process, which can reduce fraud.

The Case

ABC Insurance Company is considering the introduction of blockchain technology. The company is concerned about the potential benefits of blockchain, but it is also aware of the risks.

The company's management team is divided on the issue. Some members of the team believe that blockchain is the future of insurance and that ABC Insurance Company should be at the forefront of the adoption of this new technology. Other members of the team are more cautious. They are concerned about the security of blockchain technology and the potential for fraud.

The Decision

The company's management team has decided to conduct a pilot project to test the use of blockchain technology in the claims process. The pilot project will involve a small number of claims and will be monitored closely. The results of the pilot project will help the company decide whether to implement blockchain technology on a wider scale.

Questions for Discussion

- What are the potential benefits of using blockchain technology in the insurance industry?
- What are the risks of using blockchain technology in the insurance industry?
- How can the risks of using blockchain technology be mitigated?
- What are the challenges of implementing blockchain technology in the insurance industry?
- How can these challenges be overcome?

Conclusion

Blockchain technology has the potential to revolutionize the insurance industry. However, there are risks associated with using this new technology. ABC Insurance Company is conducting a pilot project to test the use of blockchain technology in the claims process. The results of the pilot project will help the company decide whether to implement blockchain technology on a wider scale.

Appendix C-3
Project-Based Case Written by Bard

Introduction

Blockchain technology is a distributed ledger technology that allows for secure, transparent, and tamper-proof transactions. It has the potential to revolutionize many industries, including insurance.

Background

The insurance industry is ripe for disruption by blockchain technology. The industry is characterized by high costs, low transparency, and long settlement times. Blockchain technology can help to address all of these issues.

Benefits of Blockchain Technology for Insurance

Blockchain technology can offer a number of benefits for the insurance industry, including:

- Reduced costs: Blockchain technology can help to reduce costs by eliminating the need for intermediaries and by streamlining the claims process.
- Increased transparency: Blockchain technology can increase transparency by providing a secure and tamper-proof record of all transactions.
- Faster settlement times: Blockchain technology can speed up settlement times by eliminating the need for manual processing.

Challenges of Blockchain Technology for Insurance

While blockchain technology offers a number of benefits, there are also a number of challenges that need to be addressed, including:

- Scalability: Blockchain technology is still in its early stages of development and it is not yet clear how scalable it will be for large-scale use.
- Regulation: The regulatory environment for blockchain technology is still evolving and it is not yet clear how insurance companies will be regulated.
- Security: Blockchain technology is a relatively new technology and there are still some security concerns that need to be addressed.

Project

The project is to develop a plan for the introduction of blockchain technology in an insurance company. The plan should address the following:

- The benefits of blockchain technology for the insurance industry.
- The challenges of blockchain technology for the insurance industry.
- How to address the challenges of blockchain technology.
- A timeline for the implementation of blockchain technology.
- A budget for the implementation of blockchain technology.

Conclusion

Blockchain technology has the potential to revolutionize the insurance industry. By addressing the challenges of blockchain technology and developing a plan for its implementation, insurance companies can reap the benefits of this disruptive technology.

Questions for Discussion

- What are the specific benefits of blockchain technology for the insurance industry?

- What are the specific challenges of blockchain technology for the insurance industry?
- How can the challenges of blockchain technology be addressed?
- What is a realistic timeline for the implementation of blockchain technology in the insurance industry?
- What is a realistic budget for the implementation of blockchain technology in the insurance industry?

ISSN: 2473-4901
v9 n5948

Appendix D
Subjective Content Evaluation

	Descriptive Case		**Normative Case**		**Project-Based Case**	
	LLM A	**LLM B**	**LLM A**	**LLM B**	**LLM A**	**LLM B**
Relevance and Accuracy	4.67 (0.58)	3.33 (1.15)	5.00 (0.00)	2.67 (1.15)	5.00 (0.00)	2.67 (1.15)
Complexity and Depth	4.67 (0.58)	2.33 (1.15)	5.00 (0.00)	2.00 (1.00)	5.00 (0.00)	2.33 (0.58)
Structure and Coherence	5.00 (0.00)	4.00 (1.00)	5.00 (0.00)	3.67 (0.58)	5.00 (0.00)	3.33 (0.58)
Creativity	4.67 (0.58)	2.00 (1.00)	4.67 (0.58)	2.00 (0.00)	4.33 (0.58)	2.00 (0.00)
M	4.75	2.92	4.92	2.58	4.83	2.58
SD	0.17	0.92	0.17	0.79	0.33	0.57

©2023 ISCAP (Information Systems and Computing Academic Professionals)

ISSN: 2473-4901
v9 n5948

Appendix E
Objective Measures of Readability

	Descriptive Case		**Normative Case**		**Project-Based Case**	
	LLM A	**LLM B**	**LLM A**	**LLM B**	**LLM A**	**LLM B**
Gunning Fog Index	15.9	17.0	16.3	14.2	14.0	17.9
Flesch-Kincaid Grade Level	14.5	13.1	13.6	10.8	11.4	13.9
Coleman-Liau Index	15.6	15.7	16.7	13.6	14.7	17.2
SMOG Index	13.3	12.4	12.4	10.7	11.0	13.0
Automated Readability Index	14.7	14.0	13.4	11.1	11.2	14.8
Linsear Write Index	15.1	13.4	11.6	10.7	8.9	13.3
M	14.9	14.3	14.0	11.9	11.9	15.0
SD	0.9	1.7	2.1	1.6	2.1	2.1

©2023 ISCAP (Information Systems and Computing Academic Professionals)