

# Leveraging Generative AI for Educational Feedback in Data Modeling within Information Systems

MATTHAIOS PETROU, University of Twente, The Netherlands

The demand for IT talent has increased compoundly in recent years across many industries, leading to the rise of citizen developers. Citizen or novice developers are individuals without formal expertise in IT, aiming to fill this gap. This research explores the capabilities of generative AI technologies, specifically ChatGPT-4, to evaluate the work of novices and provide feedback in the context of data and behavior modeling. The study assessed ChatGPT-4's feedback based on four predefined criteria which are the following: accuracy, relevance, comprehensiveness, and adherence to standards. After conducting a survey and receiving the opinion of many experienced participants in data modeling, results showed that ChatGPT-4 can provide valuable and relevant feedback, helping inexperienced developers such as the novices enhance their models. However, in some cases it struggled to identify the correct relationships of classes within data models and sometimes offered feedback that was too generic. Future studies relevant to this topic should involve more experts from across different contexts and use models that are more complex to challenge further the limitations of this tool. The enormous potential of this tool requires continuous improvement and research around it, to maximize effectiveness in helping novices learn.

Additional Key Words and Phrases: Generative AI, ChatGPT-4, novice developers, data modeling, behavior modeling, citizen developers, feedback assessment, information systems design, AI in education.

## 1 INTRODUCTION

In today's fast paced technological evolution, the need for IT talent grows bigger continuously and daily reaching unprecedented levels. Companies across all sectors struggle to keep up with this technological evolution and fill technical roles, as many professionals fail to keep the pace of the growing needs. This problem led to the shift towards citizen developers, who may lack formal IT knowledge but are eager to learn and contribute to the development process. By leveraging low-code and no-code platforms, these developers can create applications and models, even if they have a shortage of technical skills.

Generative AI is an emerging technology which can help support citizen developers through their endeavors in technology development. This tool includes advanced AI models, such as ChatGPT, which have the capabilities to automate the assessment of work done by citizen developers and provide constructive feedback. In this manner, the learning curve of an individual who has limited knowledge on information systems design is enhanced and the possibility of them reaching a deadend becomes very limited.

Despite the potential benefits of generative AI, it is a very recent topic and limited research has took place on how effectively it can be used to evaluate feedback on data and behavior models created

by novices. This study seeks to address this, by investigating the capabilities of AI models on the work done by citizen developers and the feedback offered.

## 2 PROBLEM STATEMENT

There are already many studies about the effectiveness of generative AI in providing feedback and in general education. However, there is a broad gap in research targeting specifically data modeling done by citizen developers. These studies mostly focus on the general educational impacts of generative AI and do not give specific results on guiding non-expert developers.

This research aims to address this gap by exploring how generative AI can be tailored and utilized to support novice developers in the endeavors of quality and consistency. The integration of AI-driven technologies can contribute significantly in the model design process, as the main aspect of these technologies include analyzing, understanding and evaluating information. AI technologies can also be the root of the feedback mechanisms resulting in a more efficient and accurate development process.

### 2.1 Research Question

The above problem statement led to the following research question:

*In what ways can generative AI technologies be exploited to automate the assessment and provision of feedback on data and behavior models created by novices/learners in the context of information systems design?*

The sub-questions that guide this research are:

- (1) How accurately can ChatGPT-4 identify the classes, attributes, methods, and relationships within data and behavior models?
- (2) To what extent does ChatGPT-4 adhere to standard modeling practices?

## 3 METHODOLOGY

In this study a mixed-methods approach takes place, combining both qualitative and quantitative analysis to evaluate the effectiveness of generative AI tools in automating the assessment and delivery of feedback on data and behavior models created by novice developers. The research starts by providing textual descriptions of two data models and two behaviors to ChatGPT-4, one simple and one more complex example for each to gain a deeper understanding on the feedback provision. These models include classes, attributes, methods and relationships for the data models and process flows for the behavior models. ChatGPT-4 is then used to generate feedback on these models.

After reviewing the ISO/IEC 25010:2011 [17] standard for software quality models criteria have been set to evaluate the performance

---

TScIT 41, July 5, 2024, Enschede, The Netherlands

© 2024 University of Twente, Faculty of Electrical Engineering, Mathematics and Computer Science.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

of ChatGPT-4: Accuracy, Relevance, Comprehensiveness, and Standards.

- (1) **Accuracy:** *The extent to which ChatGPT-4 correctly identifies the classes, attributes, methods, and relationships or process flows within the models. This criterion aligns with the correctness aspect of the Functional Suitability characteristic in the ISO/IEC 25010:2011 standard.*
- (2) **Relevance:** *The relevance of the feedback to the functionality of the system being modeled. This criterion is associated with the functional appropriateness sub-characteristic of the Functional Suitability characteristic in software quality models (ISO/IEC 25010:2011).*
- (3) **Comprehensiveness:** *The degree to which ChatGPT-4 covers all aspects of the diagrams and provides detailed suggestions for improvement. This aligns with the functional completeness aspect of the Functional Suitability characteristic (ISO/IEC 25010:2011).*
- (4) **Standards:** *The adherence of ChatGPT-4's feedback to standard modeling practices, including validation of relationships and multiplicity. This criterion is related to the compliance characteristic, ensuring adherence to industry standards and best practices (ISO/IEC 25010:2011).*

To assess further the capability of ChatGPT-4 on feedback provision, responses are collected from 11 students from various studies in University of Twente who are familiar with knowledge in data modeling. The above students review the feedback given by the AI model and complete a survey evaluating the helpfulness of this tool based on the four predefined criteria. This procedure involves generating feedback through ChatGPT-4 and passing a survey to the selected students to evaluate it.

After collecting the data from every student, the data is analyzed using both quantitative and qualitative methods. Quantitative analysis involves the use of mathematics, more specifically descriptive statistics such as mean scores to measure the central tendency of the values and standard deviations to check whether the dispersion of the answers is near the mean score. This procedure will happen for each criterion to measure the effectiveness of ChatGPT-4's feedback. For qualitative analysis open-ended questions will be examined to understand the students' perception on scoring the criteria.

Following this methodology, this study aims to answer the research question and provide a comprehensive evaluation of the potential of AI tools in providing feedback to novice developers in learning and improving their skills in data modeling.

## 4 LITERATURE REVIEW

In this section, literature review will be discussed to present the findings on the capabilities of generative AI technologies to provide feedback in data modeling and its purpose on empowering citizen developers. By examining current research, this section will depict the benefits, risks and potential of generative AI in an educational setting.

### 4.1 AI Feedback

The use of generative AI technologies and more specifically of large language models (LLM) like ChatGPT-4 which is used for these research study has been explored very thoroughly in educational settings the last 3 years. This tool provides automated and instant feedback to learners making it instantly accessible which can enhance the learning experience.

In a study by Thomas et al. (2023) [6], the use of generative AI in delivering real-time feedback and explanatory feedback to tutors during scenario-based lessons is highlighted. This study also demonstrates how AI can help teachers show their approval to students and receive immediate feedback on their performance by large language models (LLMs). The research reached a conclusion that using AI for feedback provision is an immediate, cost-effective, and efficient choice, but it needs to be implemented with great caution as there are some practical and ethical considerations.

Another study that was conducted by Baidoo-Anu and Owusu Ansah (2023) [2] to examine the potential pros and cons of using ChatGPT in education. In this study, researchers found that ChatGPT adapts continuously to the user's prompts to provide personalized and interactive prompts, by such means providing continuous feedback to enhance teaching and learning. Again, these researchers also noted that some limitations occur such as false information generation, and some that were more rare such as biases in training data and privacy issues.

Dai et al. (2023) [11] conducted research to explore the capabilities of ChatGPT to provide feedback to students. After extensive testing of models like ChatGPT, the study revealed that those models could offer valuable explanatory feedback which can aid the process of learning. They also realized that the capability of these models to provide consistent and detailed feedback to learners enhances drastically existing educational practices.

A different perspective of generative AI has been presented by the study of Ji et al. (2023) [12] which discusses the so-called phenomenon of "hallucination" in prompt generation, where AI models produce confidently incorrect information misleading users. This phenomenon depicts the need for continuous validation of AI-generated feedback to ensure the reliability and accuracy of these models. Additionally, the ethical concern of bias in AI algorithms is a critical concern that requires careful use and needs to be addressed.

Many resources regarding the automated feedback by generative AI exist, but there is still a gap in the modeling process and the quality for novices. This gap indicates that there is a need to investigate the capabilities of AI in this specific context.

In their research Sedrakyan and Snoeck (2014) [18] introduced a lightweight semantic prototyper which facilitated conceptual modeling and provided automated feedback on model quality. Their work shows the potential of generative AI as an automated tool to improve the modeling process for novices. In a similar research again Sedrakyan and Snoeck (2017) [19] discussed cognitive feedback which reasons the how and why the desired outcome is achieved. Also, feedforward automation perspectives were depicted for modeling

and validation in a learning context which discusses how things can be done better in future. These two studies show the importance and justify the need for more research in the topic of providing feedback on data modeling practices by novices.

## 4.2 Citizen Developers

Citizen developers are people without expertise on IT, who leverage various tools to contribute to the development process. This concept became popular in recent years to address the shortage of IT professionals and to empower many people to take part in software development.

Based on a study by Chhabra et al. (2022) [15], it has been proven that the use of low-code and no-code platforms is a way to make the development process available to the broader audience without the need of any IT expertise. These platforms provide interface templates and pre-built components which makes development doable for many people.

The combination of citizen developers and generative AI is revolutionary in the sector of technology as the capabilities of novices are enhanced. AI tools are able to provide real-time consistent guidance and adapt to any instance to assure the accurate provision of suggestions to improve. The integration of these tools can also assist new novices to identify and correct errors, ensuring robust and reliable work. The combination of all these benefits significantly reduces the learning curve for novices.

While the potential of AI is enormous and it is highlighted in many articles, it is essential to not overlook the possible errors and limitations of these tools. Some issues such as the accuracy of the AI algorithms and the possible bias in some topics. These potential issues outline only problems on the technical level. A serious long term risk is the over-reliance of people on AI, which requires ongoing checks for human oversight.

## 5 RECEIVING FEEDBACK FROM CHATGPT-4

In this section, the process of obtaining feedback from ChatGPT-4 for our data and behavior models is discussed. The purpose of this procedure is to test and assess the effectiveness of AI models such as ChatGPT-4 to provide real-time feedback based on the previously discussed criteria.

Initially, to receive feedback from ChatGPT-4 it is essential to give an all around explanatory description of the models we want to assess. Therefore, the first step of this procedure is to describe textually the models of the experiment in order to provide the necessary context to ChatGPT-4. This includes the classes, attributes, methods and relationships for data models, as well as the process flows of the behavior models. These models can be found in Appendix A1.

Then, before submitting the textual descriptions it is essential to inform and explain to the AI model what to assess and provide in order to adapt its behavior and provide the information relevant to the experiment. Prompts with requests are carefully crafted to ensure that ChatGPT-4 understands the models and provides comprehensive feedback. In this case, two different prompts are used

for data and behavior models accordingly, and can be found in Appendix B.

Elaborating on the questions asked to assess the feedback responses of ChatGPT-4, first the standard prompts for data and behavior models accordingly were used. Using these prompts as a starting point general feedback was received for the whole model which then helped to proceed with more specific questions. Refinement loops were employed to improve the queries and extract more information through the processing power of ChatGPT-4, ensuring optimal feedback, the chat history and responses can be found in Appendix A2.

For data models, after the initial query a question such as: "Can you verify if the relationships between these specific classes are correctly identified?" were employed which lead to the verification of relationships in comparison with the initial given relationships. Also, after deploying the question "Do all the classes, attributes and methods recognised relate to each other? Is there anything missing?", the AI model checked the existing specifications of the model and gave further suggestions for changes in classes, attributes or methods.

For behavior models, after receiving a response to the initial query a question such as: "Please review this behavior model and identify any errors or missing steps in the process flows.", which resulted in ChatGPT-4 explaining again the steps, giving suggestions and alternatives which was a very insightful response. Another question such as "Are there any potential improvements or additional steps needed.", here it responded with suggestions for possible improvements on every process flow.

## 6 SURVEY DESIGN

The purpose for conducting a survey is to assess the effectiveness of the feedback generated by ChatGPT-4 on data and behavior models. Specifically, the survey aims to evaluate how well AI models recognize the classes, attributes, methods and relationships for data models, as well as the process flows of the behavior models. Also, utilizing this information how well it identifies the correctness, relevance, comprehensiveness and adherence to standard modeling practices which are the four predefined criteria that were set at the start of the research. This is done by gathering responses from participants that have sufficient knowledge in data modeling to evaluate the performance of such models.

Initially, participants were given the data and behavior models, but also the provided feedback from ChatGPT-4. Participants were then asked to carefully review the materials and answer questions designed to assess the AI feedback based on the four predefined criteria. The survey included questions rated on a Likert scale (1-5), followed by an open-ended question asking why they graded the question in this way. The survey can be found in Appendix C. All the responses collected from the participants were organized for each criterion and with the use of descriptive statistics means and standard deviations were calculated to identify strengths and areas for improvement in ChatGPT-4's feedback.

The participants that were chosen for this research project were University of Twente students who are part of a technical study and have background knowledge in data modeling. To ensure the quality of the feedback, participants were chosen based on their level in data modeling. This selection criterion was crucial as the accuracy of the feedback assessment was directly associated with their ability to understand models and use their experience to assess the helpfulness and effectiveness of the feedback. For this reason the students that were targeted were mostly senior students who had already completed modules involving data modeling.

The participants that were recruited were students enrolled in a technical study. Most of the students have completed Module 2 of their study in University of Twente, which involved working with complex data and behavior diagrams to build a group project. Because of this experience and many other assignments in data modeling, students became very familiar with building and reading diagrams which lead them to understand classes, attributes, methods, relationships and process flows. Also some senior students had far more experience with data modeling as they were required to construct models for their projects in most modules to build a website or solve real world problems. This hands-on experience ensured that the participants were well equipped with knowledge to assess the feedback provided by ChatGPT-4.

Before starting to collect responses from the participants, a request was made to the ethics committee to obtain an approval for the study. The purpose, study process and intended information to gather were all explained with detail to the committee to clear any ethical considerations. Each participant was given an information sheet explaining the purpose of the study, procedures involved and information gathering and usage. Participants indicated their consent to be part of the study by ticking a box in the Google Forms survey, acknowledging that they have read and accepted to be part of the study as volunteers. This process ensured that all participants were informed properly and no ethical issues occurred.

## 7 RESULTS

After collecting the results of the survey conducted, it is time to assess the effectiveness of ChatGPT-4's feedback on data and behavior models. The survey was filled by 14 participants with proficient knowledge in data modeling, and the results depict the helpfulness of this tool based on the four predefined criteria: Accuracy, Relevance, Comprehensiveness, and Standards.

To understand how each criterion was rated by the participants, the mean and standard deviation was calculated for each question. The results are shown in the table below:

### 7.1 Accuracy

Starting off with the accuracy the participants were asked how well ChatGPT-4 identified the classes, attributes, methods and relationships for data models and process flows for behavior models. The mean rating for accuracy in classes, attributes and methods was 4.36, which score indicates the high ability of the AI model to identify

Criterion	Mean	Std Dev
Accuracy (Classes, Attributes, Methods)	4.36	0.48
Accuracy (Relationships)	3.43	0.49
Accuracy (Process Flows)	4.07	0.63
Relevance (Functionality)	4.21	0.64
Comprehensiveness (Data Models)	4.19	0.53
Comprehensiveness (Behavior Models)	4.14	0.49
Comprehensiveness (Suggestions)	4.07	0.76
Standards (Adherence)	4.14	0.36

Table 1. Mean and Standard Deviation of Survey Responses

these elements. The relationships of the data models received a lower mean score of 3.43, depicting the occasional struggle of ChatGPT-4 to identify the relationships between classes in data models. Lastly, the ability to identify process flows in behavior models had a mean score of 4.07 which also indicates the fine ability of AI models.

### 7.2 Relevance

Next, the relevance of the feedback to the functionality of the system received a mean rating of 4.21. This score suggests that participants found the feedback highly relevant and useful for the purpose of improving the functionality of the models. The calculated low standard deviation indicates that there was a strong common consent of the relevance of the feedback.

### 7.3 Comprehensiveness

For the criterion of comprehensiveness, the participants that took the survey rated the ability of ChatGPT-4 to take into account all the aspects of data and behavior models and provide detailed suggestions for improvement. The mean rating for this criterion was 4.19 in the context of covering all aspects and being comprehensive and 4.14 for giving detailed suggestions. These ratings suggest that ChatGPT-4 has the ability to understand and include all the aspects of both data and behavior models in feedback and suggestions in most cases but it still requires some human involvement to resolve any oversights.

### 7.4 Standards

The adherence of ChatGPT-4's with standard modeling practices has been rated with a mean rating 4.14. This indicates that the participants felt that the feedback and suggestions provided were aligned with the established standard modeling practices.

For the open-ended questions, most of the participants wrote comments about the effectiveness of ChatGPT-4, but also some problems they have experienced in the real-time provided feedback. Overall, participants found the feedback helpful, highlighting the clear and understandable explanation of aspects, but also the constructive feedback and suggestions that lead to a deeper understanding of the models and some elements that needed optimisation. Some of the negative feedback received by several participants in the Google Forms was the weakness of ChatGPT-4 to identify relationships

of data models with one participant commenting, "Some of the relationships were not correctly identified," highlighting an area where improvement is needed. Also, some participants found the feedback and suggestions too generic in some cases with one participant mentioning, "I found the feedback relevant but sometimes too generic," suggesting a need for more tailored and specific feedback.

The results from the survey indicate that ChatGPT-4, is a very valuable tool which can reduce the learning curve for citizen developers in the data and behavior modeling process. Despite the overall positive perception, there are some aspects that are not yet perfect such as identifying relationships between elements in data models and narrowing down on feedback to give tailored and specific feedback. It is now visible that AI shows a very promising ability in several areas, but ongoing development is required to address its limitations.

## 8 CONCLUSION/FUTURE STUDIES

This study had a purpose to assess the effectiveness of AI-generated feedback, specifically provided by ChatGPT-4 on data and behavior models created by novices. The results of this research indicate that ChatGPT-4 can provide valuable feedback always relevant to the aspects of each model, with mean ratings above 4.0 for all criteria except the accuracy of relationships identification. The feedback generated by this AI model has been proven to be consistent, real-time and adaptive to any request given by novices, resulting in an upgraded learning experience. However, there are some areas where ChatGPT-4 is less robust and needs improvement, particularly in relationship identification between the elements of a data model and giving more comprehensive feedback in some cases where it is more generic.

To enhance this research and make it more robust and applicable, future studies should aim to have a larger and more diverse participant pool. Also, it is important to include more experts with knowledge in data modeling to make sure that the material that is reviewed includes more insightful evaluations of ChatGPT-4's feedback. Diversity in participants can upgrade this research as having experts from multiple backgrounds can result in a deeper understanding of AI's effectiveness in many contexts.

In addition, testing ChatGPT-4 with more data and behavior models with different difficulties can be crucial in understanding its limitations and capabilities to provide feedback better. The current study focused on simpler data models which may differ from some models found in the industry which may not have challenged AI fully.

In conclusion, while ChatGPT-4 shows a very promising tool to enhance and reduce the learning process of novices, there are clear pathways to enhance this research. As mentioned before a broader range of experts challenging AI with more complex models, future studies can better assess the full potential of AI and contribute to the continuous improvement of generative AI technologies in educational and developmental contexts.

## REFERENCES

- [1] Brynjolfsson, Erik Li, Danielle Raymond, Lindsey R. (2023). Generative AI at Work. 10.3386/w31161 <http://www.nber.org/papers/w31161>

- [2] Baidoo-anu, D., Owusu Ansah, L. (2023). Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning. *Journal of AI*, 7(1), 52-62. <https://doi.org/10.61969/jai.1337500>
- [3] Euan D Lindsay Aditya Johri Johannes Bjerva. (2023) A Framework for Responsible Development of Automated Student Feedback with Generative AI. arXiv preprint arXiv:2308.15334.
- [4] Su, Jiahong Yang, Weipeng. (2023). Unlocking the Power of ChatGPT: A Framework for Applying Generative AI in Education. *ECNU Review of Education*. 6. 1-12. 10.1177/20965311231168423.
- [5] Wongvorachan, Tarid Bulut, Okan. (2022). Feedback Generation Through Artificial Intelligence. *The Open/Technology in Education, Society, and Scholarship Association Conference*. 2. 10.18357/otessac.2022.2.1.125.
- [6] Thomas, Danielle Gatz, Erin Gupta, Shivang Lin, Jionghao Tipper, Cindy Koedinger, Kenneth. (2024). Using Generative AI to Provide Feedback to Adult Tutors in Training and Assess Real-life Performance. 10.35542/osf.io/n9w4b.
- [7] Carroll, Noel Mórán, Liam Garrett, Dave Jammadass, Arjun. (2021). The Importance of Citizen Development for Digital Transformation. *Cutter IT Journal*. 34. 5-9.
- [8] Lebens, Mary Finnegan, Roger Sorsen, Steven Shah, Jinal. (2021). Rise of the Citizen Developer. *Muma Business Review*. 5. 101-111. 10.28945/4885.
- [9] Rokis, Karlis Kirikova, Marite. (2023). Exploring Low-Code Development: A Comprehensive Literature Review. *Complex Systems Informatics and Modeling Quarterly*. 68-86. 10.7250/csimq.2023-36.04.
- [10] Hevner, A. R., March, S. T., Park, J., Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, 28(1), 75-105.
- [11] Dai, W., Lin, J., Jin, H., Li, T., Tsai, Y. S., Gašević, D., Chen, G. (2023). Can large language models provide feedback to students? A case study on ChatGPT. In 2023 IEEE International Conference on Advanced Learning Technologies (pp. 323-325).
- [12] Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., ... Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12), 1-38.
- [13] Kizilcec, R. F., Lee, H. (2022). Algorithmic fairness in education. In *The ethics of artificial intelligence in education* (pp. 174-202). Routledge.
- [14] U.S. Department of Education, Office of Educational Technology, Artificial Intelligence and Future of Teaching and Learning: Insights and Recommendations, Washington, DC, (2023).
- [15] Chhabra, P., Chine, D. R., Adeniran, A., Gupta, S., Koedinger, K. R. (2022). An evaluation of perceptions regarding mentor competencies for technology-based personalized learning. In E. Langran (Ed.), *Proceedings of Society for Information Technology Teacher Education International Conference. Association for the Advancement of Computing in Education (AACE)*, San Diego, USA, pp. 1812-1817.
- [16] Baidoo-Anu, D., Owusu Ansah, L. (2023). Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning. *Journal of AI*, 7(1), 52-62.
- [17] ISO/IEC. (2011). ISO/IEC 25010:2011 Systems and software engineering—Systems and software Quality Requirements and Evaluation (SQuaRE)—System and software quality models. International Organization for Standardization.
- [18] Sedrakyán, G., Snoeck, M. (2014). Lightweight semantic prototyper for conceptual modeling. In *Advances in Conceptual Modeling: ER 2014 Workshops, ENMO, MoBiD, MReBA, QMMQ, SeCoGIS, WISM, and ER Demos*, Atlanta, GA, USA, October 27-29, 2014. *Proceedings 33* (pp. 298-302). Springer International Publishing.
- [19] Sedrakyán, G., Snoeck, M. (2017). Cognitive feedback and behavioral feedforward automation perspectives for modeling and validation in a learning context. In *Model-Driven Engineering and Software Development: 4th International Conference, MODELSWARD 2016, Rome, Italy, February 19-21, 2016, Revised Selected Papers 4* (pp. 70-92). Springer International Publishing.

## A APPENDIX A: MODELS

### A.1 Model Descriptions

The detailed textual descriptions of the data and behavior models used in this study can be accessed [here](#).

### A.2 Model Feedback

The chat history for all the models individually can be accessed through these links:

- **Data Model 1:** Link to chat history
- **Data Model 2:** Link to chat history
- **Behavior Model 1:** Link to chat history
- **Behavior Model 2:** Link to chat history

B APPENDIX B: PROMPTS FOR CHATGPT-4 FEEDBACK

B.1 Data Model Prompt

// Here is a textual description of a data model //

Please provide feedback on the following criteria:

- (1) Accuracy: Have I correctly identified all the classes, attributes, and relationships?
- (2) Relevance: Is the feedback relevant to the functionality of the e-commerce system?
- (3) Comprehensiveness: Have I covered all aspects of the model? Are there any missing elements or improvements you can suggest?
- (4) Standards: Does this model adhere to standard modeling practices?

B.2 Behavior Model Prompt

// Here is a textual description of a behavior model //

Please provide feedback on the following criteria:

- (1) Accuracy: Have I correctly identified all the steps in the user login process?
- (2) Relevance: Is the feedback relevant to the functionality of the user login process?
- (3) Comprehensiveness: Have I covered all aspects of the process flow? Are there any missing steps or improvements you can suggest?
- (4) Standards: Does this behavior model adhere to standard modeling practices?

C APPENDIX C: SURVEY

The survey conducted for this study can be accessed [here](#).

D APPENDIX D: SURVEY DEMOGRAPHICS

14 responses

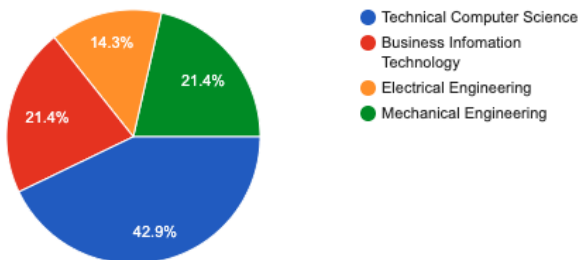


Fig. 1. Study Fields

14 responses

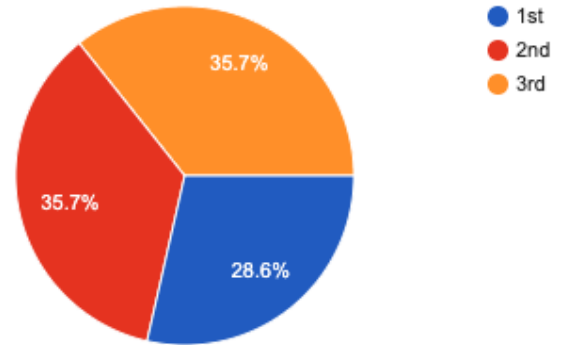


Fig. 2. Year of Study

14 responses

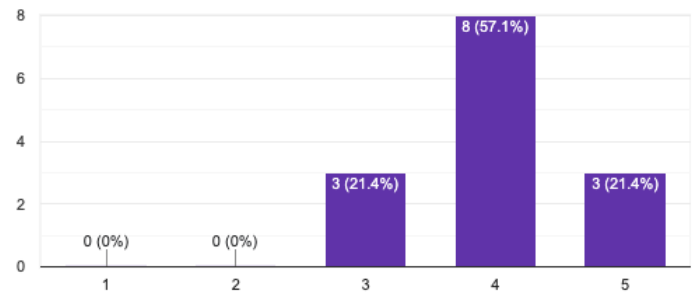


Fig. 3. Familiarity with Data Modeling