# Prompt Strategies in Lesson Plan Assessment: Insights from Pre-Service Teachers' Prompt Dataset

**Wenting Sun[1], Jiangyue Liu[2]**

Humboldt-Universität zu Berlin, Germany[1]
Suzhou University, China[2]

**Abstract**

*Generative AI (GenAI) enhances personalized teaching materials and reduces teachers' workload by generating formative and summative feedback to improve learner's performance [1,2]. Prompt engineering, a skill that utilizes language and prior knowledge to construct prompts directing generative AI towards desired outcomes, encompasses basic knowledge of relevant language syntax and the strategic use of prompt modifiers [3].*
*Given that lesson planning is a time-intensive and labor-intensive task, GenAI tools can provide significant assistance. Some studies have explored the use of GenAI in designing lesson plans, demonstrating advantages in areas such as setting instructional objectives and identifying teaching priorities [1,4]. However, the role of GenAI in assisting users with lesson plan assessment tasks remains understudied.*
*In this study, we explored prompt strategies in lesson plan assessment tasks using prompts generated during human-GenAI interactions. With this goal, a GenAI tool was employed in lesson plan assessment activities by pre-service secondary school physics course teachers. We adopted a qualitative research approach. A prompt dataset from 45 pre-service teachers was collected and served as our data source.*
*Through interpretive analysis of the qualitative data, we found that different prompt approaches were employed when learners were stuck or dissatisfied with the outputs generated. Our findings both align with and differ from previous studies [5]. We summarized prompt strategies for lesson plan assessment activities, contributing to the effective use of GenAI in lesson planning tasks by providing practical suggestions for real-world educational contexts.*

**Keywords:** *Prompt Engineering, Lesson Plan Assessment, Qualitative Research, Pre-service Teachers*

## 1. Introduction

Generative artificial intelligence (GenAI)， esp. large language models (LLMs)-driven AI， enriches personalized teaching materials and reduces teachers' workload by generating formative and summative feedback to improve learners' performance [1, 2]. Some studies indicate that how learners use GenAI reflects their critical thinking skills [6]. Conversely, some research suggests that GenAI may foster poor learning behaviours and weaken critical evaluation [7,8]. Therefore, to investigate whether and how GenAI benefits learners, researchers need more detailed evidence to explore human-AI interaction process.
In the human-AI interaction process, prompting serves as the bridge linking human inquiry and AI feedback. As a new 21st century skill, prompting facilitates precise communication of problems to an AI assistant by articulating the problems, their context, and the constraints of the desired solution [9]. Prompt formulation requires users to self-aware of their task goals and converts their tasks into sub-tasks, verbalizing these as effective prompts followed by iterative output evaluation and adjustment [10]. Through practice and learning, effective prompting skills can be acquired [3].
As GenAI tools become increasingly accessible, their integration into education is reshaping teaching practices, including lesson planning and assessment. Prompt engineering, the skill of crafting effective inputs for generative AI, has emerged as critical for optimizing outputs in educational settings. GenAI has demonstrated benefits in course design and content generation. For example, LLMs excel in instructional objectives setting and teaching activities organizing [1]. GenAI demonstrated advantages in delivering adaptable information and saving time in course planning [11].

While prior studies have explored how teachers and students use GenAI for content generation, less attention has been paid to how pre-service teachers prompt GenAI to support complex teaching tasks like lesson plan assessment. In the GenAI era, teacher need to develop AI literacy to know when and how to critically use AI tools in class [12]. This study investigates the prompt strategies used by pre-service teachers to assess lesson plans using GenAI. We aim to uncover how these novice educators craft, adapt, and refine prompts to evaluate educational materials and receive actionable feedback from an GenAI assistant. Our guiding research question is: "What prompt strategies do pre-service teachers use when assessing lesson plans with GenAI tools?"

## 2. Related Works

### *2.1 Prompting in Education*

Prompt engineering is a skill that utilize language and prior knowledge to construct prompts that direct generative AI towards desired outcomes, encompassing basic knowledge of the relevant language syntax and the strategic use of prompt modifiers [3]. From a micro-level perspective, it unpacks human-AI interaction processes to understand how users construct and evaluate generated AI outputs. Effective prompting strategies require deep cognitive processing of generated outputs towards the desired goal [13, 14].

To assist non-AI experts in using GenAI for problem solving, some prompting strategies or patterns have been proposed. For example, at the prompt component level, [15] recommended incorporating several components into the written prompt, such as context, alignment, and constraints. The CLEAR framework introduces five core principles to facilitate more effective AI generated content evaluation and creation, such as being concise and logical [16]. As an attempt to extract regularities of prompts used in AI-assisted ill-defined complex tasks processing, TELeR has been proposed to design prompts with specific properties targeting a wide range of complex tasks [17]. [18] proposed that assigning AI different roles for classroom use would bring distinct pedagogical benefits and risks.

### *2.2 GenAI and Lesson Planning*

GenAI has demonstrated potential in course design and content generation. Utilizing GPT4, a high school mathematics teaching plan dataset was generated [1]. The evaluation of this dataset revealed that LLMs excel in instructional objectives setting, teaching priorities identifying, subjective content articulating, and teaching activities organizing. Algining with the classroom assessment framework required by the Council of Higher Education in Turkey, [11] developed course plans specifically for classroom assessment in science education. They pointed out that GenAI demonstrated advantages in developing implementable course plans, delivering adaptable information, and saving time. Meanwhile, communicating with GenAI is a challenge [19]. However, most studies did not explore the human-AI interaction process and how users iterate their prompting strategies during complex task processing.

### *2.3 This Study*

Using a dynamic process perspective, this study explored the cognitive process in GenAI-assisted lesson plan assessment task. The research question proposed in this study was: What prompt strategies do pre-service teachers use when assessing lesson plans with GenAI tools? On the one hand, this study contributes to understanding the human-AI interaction process from a dynamic perspective. On the other hand, the extracted prompting strategies from the human-AI interaction and the task design provide insights into pre-service teacher training in lesson plan design and assessment in the GenAI age.

## 3. Methods

Participants were 45 student teachers (third-year university students) in secondary school physics courses. They were required to assess their own lesson plans and their peers' lesson plans using a GenAI assistance tool. In this study, ERNIE driven GenAI, Wenxinyiyan (a popular local GenAI driven by local LLMs), was selected for the lesson plan revision task due to its unlimited access to file uploading in human-AI interaction without payment. To support students efficiently and critically evaluating and modifying lesson plans, a simple lesson plan evaluation guide was provided, including tips on word construction in prompting and suggested attitudes towards generated output for

assessment tasks. Students interacted with the GenAI in their mother language (Chinese). Human-AI interaction screen recordings from these students were collected, and prompt datasets were manually extracted from the screen recordings.
In this study, interpretive analysis was conducted as the qualitative method. Following the thematic analysis procedure proposed by [20], prompting strategies were extracted and synthesized. These prompts were analyzed using open coding to identify patterns in prompt strategies. Two coders independently reviewed the data, with one researcher coding 25% of the prompt data and another coding all of it. The code results achieved 90% consensus on thematic categories.
To visualize the prompt interation process, a Sankey diagram of the 1st prompts, students' responses to the generated outputs, and the 2nd prompts was drawn. Moreover, several prompt strategies or patterns were used as references to analyze the prompts used by the participants and provide more practical implementations.

**4. Results and Discussions**

For inquiry types in prompts, according to the scope of the lesson plan involved in the inquiry, they were categorized into specific inquiries and general inquiries. For example, "The teaching objectives of secondary school physics kinetic energy and kinetic energy theorem are based on core literacy" was regarded as specific inquiry, while "Please help me design a lesson plan about buoyancy in secondary school physics" (participant 15) was considered a general inquiry.
As shown in Figure 1, the most frequent prompting strategies used were Specific part (inquiry) (98), File+inquiry (67), Generated outputs (31), Prior prompting+new inquiry (24), Retry prior prompt (12). Some less frequently used but inspiring prompting strategies were GenAI recommended prompting, Ask GenAI self-reflection, Prior prompting+paragraphs, Combining prior-prompt, Search engine, and Another GenAI tool. More details of the prompting formation can be found in the section For Further Prompting.

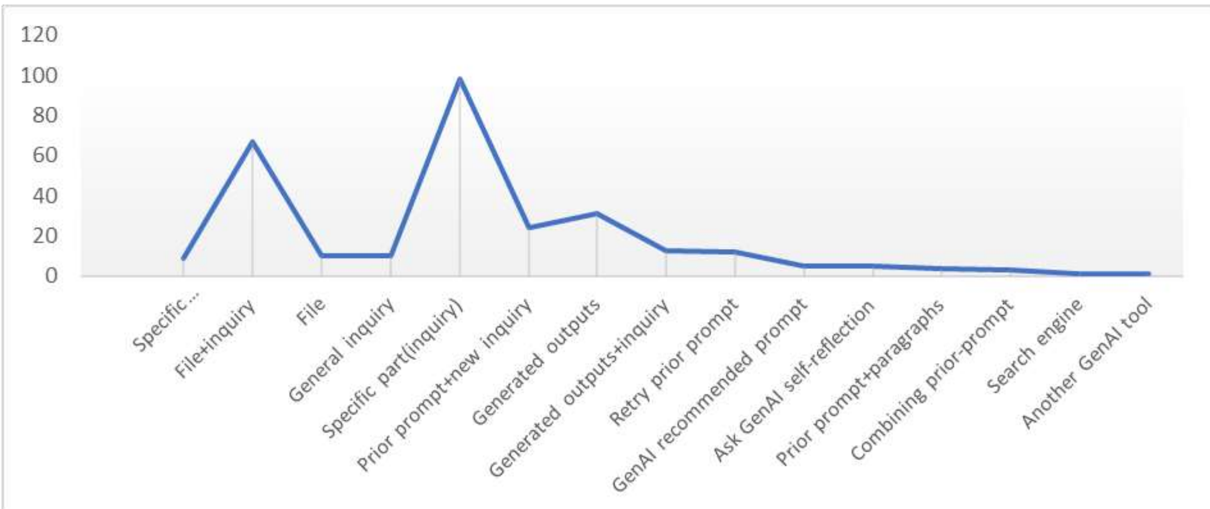


Figure 1. Frequently used prompting methods.
Note: the leftmost method is "Specific part(inquiry+paragraphs)"

Different prompt approaches were used when learners stuck or unsatisfied with the generated outputs. Specifically, the human-AI interaction patterns can be illustrated by integrating the prompt engineering strategies and prompt examples. The workflow can be visualized in Figure 2.

**For 1st Prompting**

At the starting point, most of participants initiated prompts in a combination form, including uploading lesson plan document through upload file button and providing a relatively general task requirement in the input, such as "Help me modify the lesson plan" or "Help me revise the lesson plan according to the teaching objectives".

**For Response to 1st Prompting**

Two main approaches were found to deal with the generated output from the first prompt. One approach involved a new document, pasting all the generated output in it, and then comparing the generated outputs with their own lesson plans. The other approach involved directly comparing the generated output with the lesson plan and then takes actions adoptively (either coping and enriching the pasted outputs or doing further prompting as shown in Figure 2).

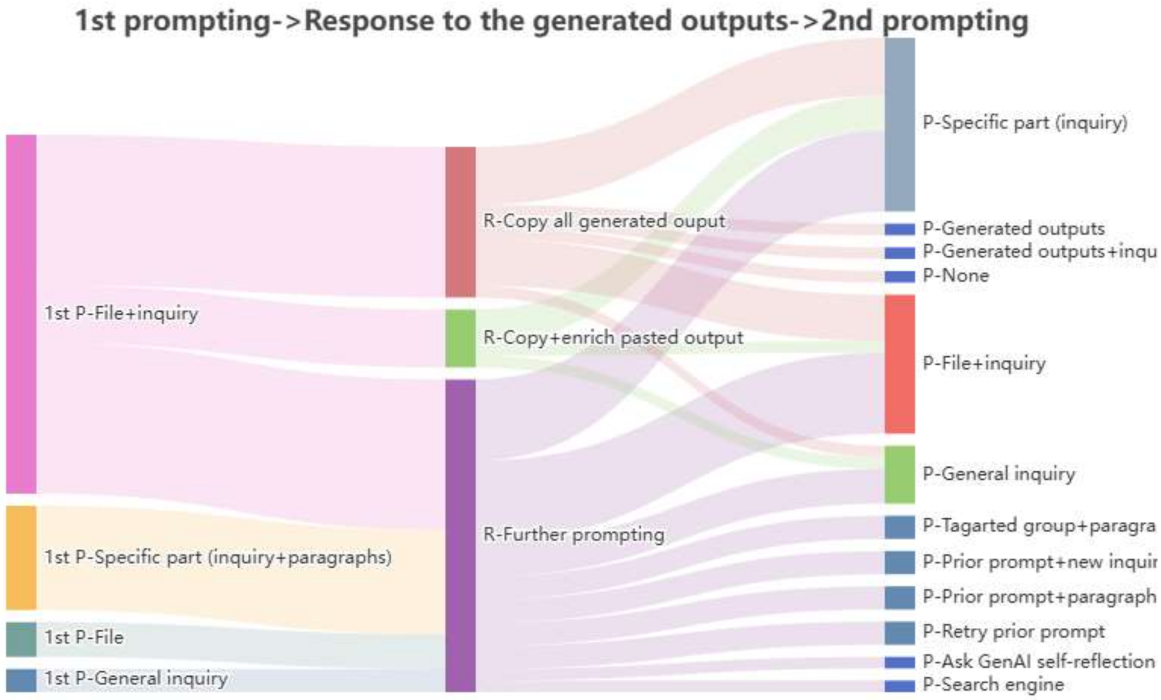


Figure 2. Sankey diagram of prompt strategy decision process workflow.
Note: “P-” = Prompt, “R-”= Response to the outputs generated from prior prompts.

**For Further Prompting**

When participants were stuck or unsatisfied the generated outputs, they took a few different approaches. 10 prompting forming methods were synthesized in Table 1. For example, “Special part”+inquiry/paragraphs shown in Figure 1 demonstrated the stepwise inquiry, dividing tasks into subtasks. More details can be seen in Table 1.
The findings show that pre-service teachers are capable of using a range of prompting strategies to improve GenAI’s feedback quality. These strategies mirror metacognitive skills such as monitoring output quality and adjusting inputs. This suggests that prompt engineering can serve as a bridge between human pedagogical judgment and AI affordances. Integrating prompt training into teacher education could empower novice educators to use GenAI tools more effectively and responsibly.

Table 1. Prompting forming methods and examples.

| Prompting forming | Description | Features | Prompt examples shown in this study |
|---|---|---|---|
| “Special part” +inquiry /paragraphs | Divide task A into task ($a_1$, $a_2$, … $a_n$): depart lesson plans into different sections and ask AI to modify them one by one | List details of knowledge points | Participant 43 used the prompt template “Three points of […] in the teaching of molecular thermal motion in high school physics should be specific” to provide extra content of teaching objectives. |
| | | List each section in different prompts | Participant 29 used the prompt template “Write a […] for the secondary school lesson "Temperature”” to generate different details about teaching objectives, student situation analysis, teaching process. |
| Ask GenAI self- | Pointing the gaps | Modify | “Why did I ask you to change the teaching |

| | | | |
|---|---|---|---|
| reflection | between the generated outputs and expected outputs | degree | process, but you changed the whole lesson plan?” (Participant 43) |
| | | Targeted student groups | “The double pendulum is not suitable for teaching high school students. Please change to a specific case of complex system capacity conservation analysis.” (Participant 45) |
| | | Use of personal pronouns | “Don’t use “I” to describe something, just keep the details.” (Participant 45) |
| Search engine | Search engine as additional data source and then copy the searched result in prompt to specify detailed desired generated outputs. | Pedagogical knowledge | "The four dimensions of physics teaching objectives." (Participant 5) |
| | | Domain knowledge | "Differentiation" (Participant 4). |
| Another GenAI tool | Using the same prompt to generate outputs in another AI tools and compare the results | | "Help me refine the teaching plan to make its structure complete. The teaching objectives should be written based on the core literacy of high school physics." (Participant 28) |
| Prior prompt+new inquiry | Copy prior prompts and add new inquiry to construct new prompt | | "Help me revise the teaching plan and change the teaching objectives to the four dimensions: 'Physical Concepts,' 'Scientific Thinking,' 'Scientific Inquiry,' and 'Scientific Attitude and Responsibility.'" (Participant 5) |
| Combining prior-prompt | Combing prior prompts to form new prompts | | "Please expand the teaching process section, emphasizing a student-centered approach and focusing on cultivating students' scientific thinking." (Participant 3) |
| Generated outputs+new inquiry | Copy the generated outputs based on prior prompts and add new inquiry forming new prompt | | "Please generate an image based on the above text." (Participant 2) |
| Generated outputs | Copy the generated outputs based on prior prompts to form new prompt | | "Tell some interesting timekeeping stories or riddles." (participant 6) |
| Retry prior prompting | Participant primarily used the “retry” button or copy prior prompt to generate iterative outputs | | "Please provide a classroom introduction to work and power for first-year high school students, accompanied by an image." (Participant 2) |
| Using GenAI recommended prompting | Try recommended prompts to get ideas | Teaching activities | “When introducing a new lesson, what historical or cultural background can the teacher introduce?” (Participant 27) |
| | | Application scenarios of domain knowledge | “Why do astronauts lie flat when the rocket takes off?” (Participant 37) |

Note: Items in “Prompting forming” column can be found in Figure 1.

The prompting patterns identified in this study were compared with existing prompting strategies or patterns to link the concrete prompt examples with abstract prompting strategies or patterns. As shown in Table 2, to some degree, the selected prompting strategies or patterns could be used to categorize prompt examples within the prompt structure.

Table 2. Prompt examples demonstrated referring to existing prompting patterns.

| Prompting features | Items | Prompt examples shown in this study |
|---|---|---|

| | | |
|---|---|---|
| Prompt components [15] | Verb, focus, context, focus and condition, alignment, constraints and limitations | Verb: “polish”, “perfect”, “add”, “generate”, “how to guide…to understand…” (Participant 2) |
| CLEAR Framework [16] | Concise, logical, explicit, adaptive, reflective | Reflective:<br>“Can the modified teaching process reflect the characteristics of physical modeling?” (Participant 39) |
| TELeR [17] | Turn, expression, level of details, role | Level of detials (level 3):<br>” I am a high school physics teacher. Please help me revise my lesson plan. It is required to meet the high school physics curriculum standards. The teaching objectives should be based on core literacy. The teaching methods should use experimental methods, lectures, intuitive demonstrations, group discussions, and group exploration methods. The course content should be practical, interesting, and exploratory.” (Participant 17) |
| Assign AI roles in learning [18] | Mentor, tutor, coach, teammate, student, simulator, tool | Tutor:<br>“Based on the uploaded high school physics lesson plan file, provide a brief overview of the lesson plan, and make some suggestions for revisions to this lesson plan file from the perspective of an excellent teacher” (Participant 24) |

For prompt engineering strategies in human-AI collaboration, our findings demonstrate both alignments and differences with previous studies:
Human-human interaction phenomena in human-AI interaction: As found in the article by [5], our results also detected behaviour expectations drawn from human-human interaction phenomena in human-AI interaction. Participants 17 seemed to believe that LLM-driven AI can understand the difference of display channels without detailed descriptive explanation, resulting in prompts like “Display blackboard design in the form of pictures”.
Unlike the challenges of struggling getting started (a design-stage barrier) mentioned in the study by [5], students in our study formed their 1st prompts using diverse strategies. This might be related to the their prior knowledge of the tasks and the interface design and interaction modes provided by GenAI tools. In our study, students completed their own lesson plans before this experiment, which led to high familarity with the artifact during problem solving. Compared to the text-oriented interaction modes in the study by [5], the GenAI tool we used supports document input, text-to-image generation, text-to-text generation, and recommended prompts that users might be interested in. These diverse interface deisgn features enrich human-AI interaction channels, inspiring more exploratory actions.
Compared to the “over generalization from limited experience” phenomenon shown in the study by [5], most of the students in this study performed multiple turns and used diverse strategies to utilize GenAI feedback. This might be related to the task completion duration (authentic classroom vs. lab setting) and support for task completion. In our study, students were in their normal authentic classroom and had a clear time limitation to finish the GenAI-assisted lesson plan assessment task. Additionally, to support students effectively communicating with GenAI, a simple lesson plan evaluation guide was designed, including tips on word construction in prompting and suggested attitudes towards generated outputs for assessment tasks. The task in their study required participants to recreate a professional chef that walks an amateur through various steps of cooking a recipe. The requirements are abstract and lack clear domain knowledge points to assist participants in evaluating the performance of GenAI. It was also not mentioned whether participants needed to finish the task within a limited time.

**5. Conclusion and Implementations**

Our findings provide further evidence that learners tend to ask GenAI at the beginning to gather information for subsequent questions and actions adaptively. In general, inquiries to GenAI about specific parts of the lesson plan or directly uploading the whole lesson plan document with modification inquiries were dominant. Different combinations can be found in prompting formation, such as paragraphs from the lesson plan document, prior prompts, prior generated outputs, and new inquiries. Besides the GenAI tool provided in this experiment, students also used other tools like search engines and other GenAI tools for information verification. In addition to text-to-text generation mode, text-to-image generation mode was found in some examples.
Our study offers two key implications.
First, AI tool designers should support scaffolded prompting by providing templates or guided interactions tailored to educational contexts. To effectively utilize GenAI in complex problems solving process, prompt formulation, prompt iteration, output evaluation, workflow understanding, workflow adapting need to be considered [10]. Therefore, task decomposition and self-awareness are recommended to be supported in GenAI-assisted complex problem solving. Moreover, task design in GenAI-assisted contexts is recommended to provide domain knowledge points or objective evaluation criteria. This would help users quicky know how to evaluate the generated outputs and stimulate GenAI generate more domain-specific vocabulary. The prompt forming methods (details in Table 1) identified can be referred to in GenAI-assisted problem solving to scaffold cognitive and metacognitive process.
Second, teacher training programs should consider including prompt engineering as part of AI literacy education to know when and how to critically use AI tools in class. For this context, GenAI tools with diverse interface design and interaction modes are suggested to be employed. This would increase exploration opportunities and motivation for non-AI expert users, esp. beginners. Additionally, multiple roles can be assigned to GenAI to explore complex tasks from different perspectives [18]. This would enrich the diverse generated outputs, providing idea pools for users to synthesize information targeting goals. Furthermore, according to the prompt examples in this study, utilizing private datasets would customize the generated output from LLMs, such as lesson plan examples, stored historical lesson plans, and national curriculum standards.
There are several limitations in our study. Firstly, the samples size in our study is relatively small, potentially affecting the generalization of the results. Secondly, the prompting features were synthesized according to the relationships among the prompting series and prompting components at a coarse granularity. Future studies can analyze the prompting strategies from sentence and word level from a finer granularity perspective. Moreover, more complex problems can be set to stimulate human-AI interaction. Other prompting strategies like Zero-shot, Few-shot, Chain of Thought (CoT) were not considered in this study. This is related to the purpose of this study, which aims to uncover how these novice educators craft, adapt, and refine prompts to evaluate educational materials and receive actionable feedback from an AI assistant. If we use these prompting strategies as an analysis framework, many meaningful details of the prompting components would not be demonstrated.

## REFERENCES

[1] Hu B., Zheng L., Zhu J., Ding L., Wang Y., Gu X., "Teaching Plan Generation and Evaluation With GPT-4: Unleashing the Potential of LLM in Instructional Design", IEEE Transactions on Learning Technologies, 2024.
[2] Meyer J., Jansen T., Schiller R., Liebenow L. W., Steinbach M., Horbach A., Fleckenstein J., "Using LLMs to bring evidence-based feedback into the classroom: AI-generated feedback increases secondary students' text revision, motivation, and positive emotions", Computers and Education: Artificial Intelligence, 2024, 6, 100199.
[3] Oppenlaender J., Linder R., Silvennoinen J., "Prompting AI art: An investigation into the creative skill of prompt engineering", International Journal of Human–Computer Interaction, 2024, 1-23.
[4] Hasan Z. O., Muslu N., "Designing a course for pre-service science teachers using ChatGPT: what ChatGPT brings to the table", Interactive Learning Environments, 2024.
[5] Zamfirescu-Pereira J. D., Wong R. Y., Hartmann B., Yang Q., "Why Johnny can't prompt: how non-AI experts try (and fail) to design LLM prompts", Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, 2023, 1-21.
[6] Liu M., Zhang L. J., Biebricher C., "Investigating students' cognitive processes in generative ai-assisted digital multimodal composing and traditional writing", Computers & Education, 2024, 211, 104977.

[7] Lo C. K., Hew K. F., Jong M. S. Y., "The influence of ChatGPT on student engagement: A systematic review and future research agenda", Computers & Education, 2024, 105100.
[8] Zirar A., "Exploring the impact of language models, such as ChatGPT, on student learning and assessment", Review of Education, 2023, 11(3), e3433.
[9] Federiakin D., Molerov D., Zlatkin-Troitschanskaia O., Maur A., "Prompt engineering as a new 21st century skill", Frontiers in Education, 2024, 9, 1366434.
[10] Tankelevitch L., Kewenig V., Simkute A., Scott A. E., Sarkar A., Sellen A., Rintel S., "The metacognitive demands and opportunities of generative AI", Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems, 2024, 1-24.
[11] Okulu H. Z., Muslu N., "Designing a course for pre-service science teachers using ChatGPT: what ChatGPT brings to the table", Interactive Learning Environments, 2024, 1-18.
[12] Ding A. C. E., Shi L., Yang H., Choi I., "Enhancing teacher AI literacy and integration through different types of cases in teacher professional development", Computers and Education open, 2024, 6, 100178.
[13] Dang H., Mecke L., Lehmann F., Goller S., Buschek D., "How to prompt? Opportunities and challenges of zero-and few-shot learning for human-AI interaction in creative applications of generative models", arXiv preprint arXiv:2209.01390, 2022.
[14] Knoth N., Tolzin A., Janson A., Leimeister J. M., "AI literacy and its implications for prompt engineering strategies", Computers and Education: Artificial Intelligence, 2024, 6, 100225.
[15] Eager B., Brunton R., "Prompting higher education towards AI-augmented teaching and learning practice", Journal of University Teaching and Learning Practice, 2023, 20(5), 1-19.
[16] Lo L. S., "The CLEAR path: A framework for enhancing information literacy through prompt engineering", The Journal of Academic Librarianship, 2023, 49(4), 102720.
[17] Santu S. K. K., Feng D., "Teler: A general taxonomy of llm prompts for benchmarking complex tasks", arXiv preprint arXiv:2305.11430, 2023.
[18] Mollick E., Mollick L., "Assigning AI: Seven approaches for students, with prompts", arXiv preprint arXiv:2306.10052, 2023.
[19] Moundridou M., Matzakos N., Doukakis S., "Generative AI tools as educators' assistants: Designing and implementing inquiry-based lesson plans", Computers and Education: Artificial Intelligence, 2024, 7, 100277.
[20] Braun V., Clarke V., "Using thematic analysis in psychology", Qualitative research in psychology, 2006, 3(2), 77-101.