Research article
UDC 378.091.64:004.8
DOI: 10.25688/2076-9121.2024.18.1-1.04

EXPERIMENTAL USE OF EDUCATIONAL MATERIALS DEVELOPED USING ARTIFICIAL INTELLIGENCE IN NATURAL SCIENCE EDUCATION[1]

Yevgeny D. Patarakin [1] ✉, ***Vasily V. Burov*** [2], ***Karim D. Salimullin*** [3], ***Dmitry V. Soshnikov*** [4]

[1, 2, 3] *Moscow City University, Moscow, Russia*
[4] *Moscow Aviation Institute, Moscow, Russia*
[4] *Higher School of Economics, Moscow, Russia*

[1] *patarakined@mgpu.ru* ✉, *https://orcid.org/0000-0002-1216-5043*
[2] *burovvv@mgpu.ru, https://orcid.org/0000-0001-5571-5693*
[3] *salimullinkd@mgpu.ru*
[4] *dmitri@soshnikov.com, https://orcid.org/0000-0003-1021-091X*

Abstract. This study explores the potential of modern generative models for automatically generating educational task texts. Building on prior research in educational task generation, we focused on leveraging generative artificial intelligence to create tasks based on textbook content, leading to the development of a multiple-choice educational task generator. This tool, powered by a large language model, empowers educators to independently craft tasks for their courses. In the experiments, teachers from various disciplines were involved in selecting topics for the generation of educational materials. The results demonstrate the capability of modern large language models to generate simple text-based multiple-choice questions suitable for use. While the current need for manual verification and refinement of distractors by educators presents a challenge, it is anticipated that generative AI will address this soon. The study sheds light on the potential of generative AI in education.

Keywords: artificial intelligence-generated content, AIGC, ChatGPT-3.5, YaGPT, GigaChat

[1] The article is published in the author's edition.

Научная статья
УДК 378.091.64:004.8
DOI: 10.25688/2076-9121.2024.18.1-1.04

ЭКСПЕРИМЕНТАЛЬНОЕ ИСПОЛЬЗОВАНИЕ УЧЕБНЫХ МАТЕРИАЛОВ, РАЗРАБОТАННЫХ С ПРИМЕНЕНИЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА, В ЕСТЕСТВЕННОНАУЧНОМ ОБРАЗОВАНИИ[2]

***Евгений Дмитриевич Патаракин*[1] ✉, *Василий Владимирович Буров*[2], *Карим Дамирович Салимуллин*[3], *Дмитрий Валерьевич Сошников*[4]**

[1, 2, 3] *Московский городской педагогический университет, Москва, Россия*
[4] *Московский авиационный институт, Москва, Россия*
[4] *Высшая школа экономики, Москва, Россия*

[1] *patarakined@mgpu.ru ✉, https://orcid.org/0000-0002-1216-5043*
[2] *burovvv@mgpu.ru, https://orcid.org/0000-0001-5571-5693*
[3] *salimullinkd@mgpu.ru*
[4] *dmitri@soshnikov.com, https://orcid.org/0000-0003-1021-091X*

Аннотация. Целью исследования было изучение пригодности современных генеративных моделей для автоматического создания текстов учебных задач. Мы разработали генератор учебных задач множественного выбора, который действует на основании большой языковой модели и позволяет учителю самостоятельно создавать задания к своему учебному курсу. В экспериментах к выбору тем для генерации учебных материалов привлекались преподаватели различных дисциплин. Результаты демонстрируют способность современных больших языковых моделей генерировать простые текстовые вопросы с несколькими вариантами ответов, пригодные для использования. Хотя текущая потребность преподавателей в ручной проверке и доработке отвлекающих факторов и представляет собой проблему, но ожидается, что генеративный ИИ решит эту проблему в ближайшем будущем.

Ключевые слова: artificial intelligence-generated content, AIGC, ChatGPT-3.5, YaGPT, GigaChat

For citation: Patarakin, E. D., Burov, V. V., Salimullin, K. D., & Soshnikov, D. V. (2024). Experimental use of educational materials developed using artificial intelligence in natural science education. *MCU Journal of Pedagogy and Psychology, 18*(1-1), 78–90. https://doi.org/10.25688/2076-9121.2024.18.1-1.04

Для цитирования: Патаракин, Е. Д., Буров, В. В., Салимуллин, К. Д., и Сошников, Д. В. (2024). Экспериментальное использование учебных материалов, разработанных с применением искусственного интеллекта, в естественно-научном образовании. *Вестник МГПУ. Серия «Педагогика и психология», 18*(1-1), 78–90. https://doi.org/10.25688/2076-9121.2024.18.1-1.04

[2] Статья публикуется в авторской редакции.

Introduction

A multiple-choice question, or more objectively, a selection question, is a form of objective assessment in which respondents are presented with a list of correct answers and asked to choose only the accurate ones. This format is used frequently in educational testing, as it allows teachers to quickly assess students' understanding of the material. In recent years, there has been a growing interest in the automatic generation of multiple-choice questions from text, which has led to the development of various algorithms and techniques. The use of these methods has the potential to revolutionize the way questions are generated and assessments are conducted in educational settings.

The process of generating multiple-choice questions (MCQs) from educational textbooks has been the focus of several studies (Kurdi et al., 2020; Ch, & Saha, 2020; Matos, 2022). These studies have mainly focused on the automatic generation of MCQs and the subsequent selection of high-quality questions based on specific criteria. Some notable works in this area include the use of Bloom's taxonomy for crowd-sourced selection of MCQs (Moore et al., 2023), and the development of a high-quality MCQ dataset called EduQG (Hadifar et al., 2023). However, there has been less attention given to how teachers can use these generated questions and what scenarios can be created for their use. One study by Laban et al. (2022) addressed this gap by proposing an automatic MCQ generation and selection pipeline that allows teachers to filter questions based on multiple criteria. The proposed pipeline comprises four core modules: preprocessing, sentence selection, key selection, and distractor selection. The study demonstrated that the proposed pipeline can generate high-quality MCQs that meet the criteria set by teachers.

In our previous work on educational task generation (Patarakin, Burov, & Sochnikov, 2023), we explored various approaches using generative artificial intelligence and ultimately focused on generating tasks based on textbook content. The advantages of this approach include a higher level of alignment with the educational material, as the generated tasks will be closely linked to the textbook content, enabling students to better understand and deepen their knowledge. Additionally, using the textbook allows for better contextual understanding and alignment with the educational course's objectives, enabling the creation of tasks that more fully correspond to the curriculum. Furthermore, textbooks typically present material in a structured manner, which can assist the generation system in creating tasks of varying levels of complexity and on different topics in a more organized fashion. This approach holds promise for improving the quality and relevance of educational tasks through the strategic use of generative artificial intelligence, but it also comes with certain drawbacks. Some of the main challenges include:

– Limited information: Generation of tasks exclusively based on the textbook may limit the variety of tasks and reduce the creativity in creating educational materials.

– Incomplete coverage of the textbook: The textbook may not cover all topics or necessary nuances, leading to limited task generation.

– Difficulty in interpretation: Understanding the textbook may be a challenging task for machine learning algorithms, especially when the text contains ambiguities or requires contextual interpretation.

– Lack of creativity: When using the textbook as the basis for generating tasks, innovative and creative approaches to task creation may be lacking.

Despite these drawbacks, the use of textbooks as a source of context for generating tasks can help the model better understand and adapt to the specific educational context, improving the quality of task assessment. However, it is essential to consider the potential limitations and challenges associated with this approach and strive to address them to ensure the effectiveness of the generated tasks in enhancing student learning.

Experiments with a Multiple-Choice Educational Task Generator

Based on the presented rationale, we have devised a multiple-choice educational task generator. This generator operates on the basis of a large language model, empowering educators to independently craft tasks for their educational courses. At the outset of the process, the system prompts the educator to upload a text excerpt from the textbook. Upon receiving this original text, the system generates multiple-choice questions from it and presents the user with the option to select the questions they wish to receive. The selection process is illustrated in the Figure 1.

At the next stage, the system offers the teacher to download the created list of ready-made questions and correct and false answers.

The general scheme of organizing the educational process using generative artificial intelligence is illustrated in Figure 2.

The generative artificial intelligence performs a merely auxiliary function. We can see only the final part of the process on the figure. The teacher has already presented the educational material to the students and organized their work with this material. Then, the teacher uses the support of artificial intelligence to form a variety of test questions. The interaction between the teacher and the question generation system is as follows:

The teacher uploads a text fragment from the textbook.

The system proposes a list of questions that it is capable of generating based on the uploaded text.

If the suggested questions do not suit the teacher's preferences, the teacher can request new question variants.

If the teacher selects a question variant, the system generates a set of answers to these questions, one of which is correct and several of which are incorrect.

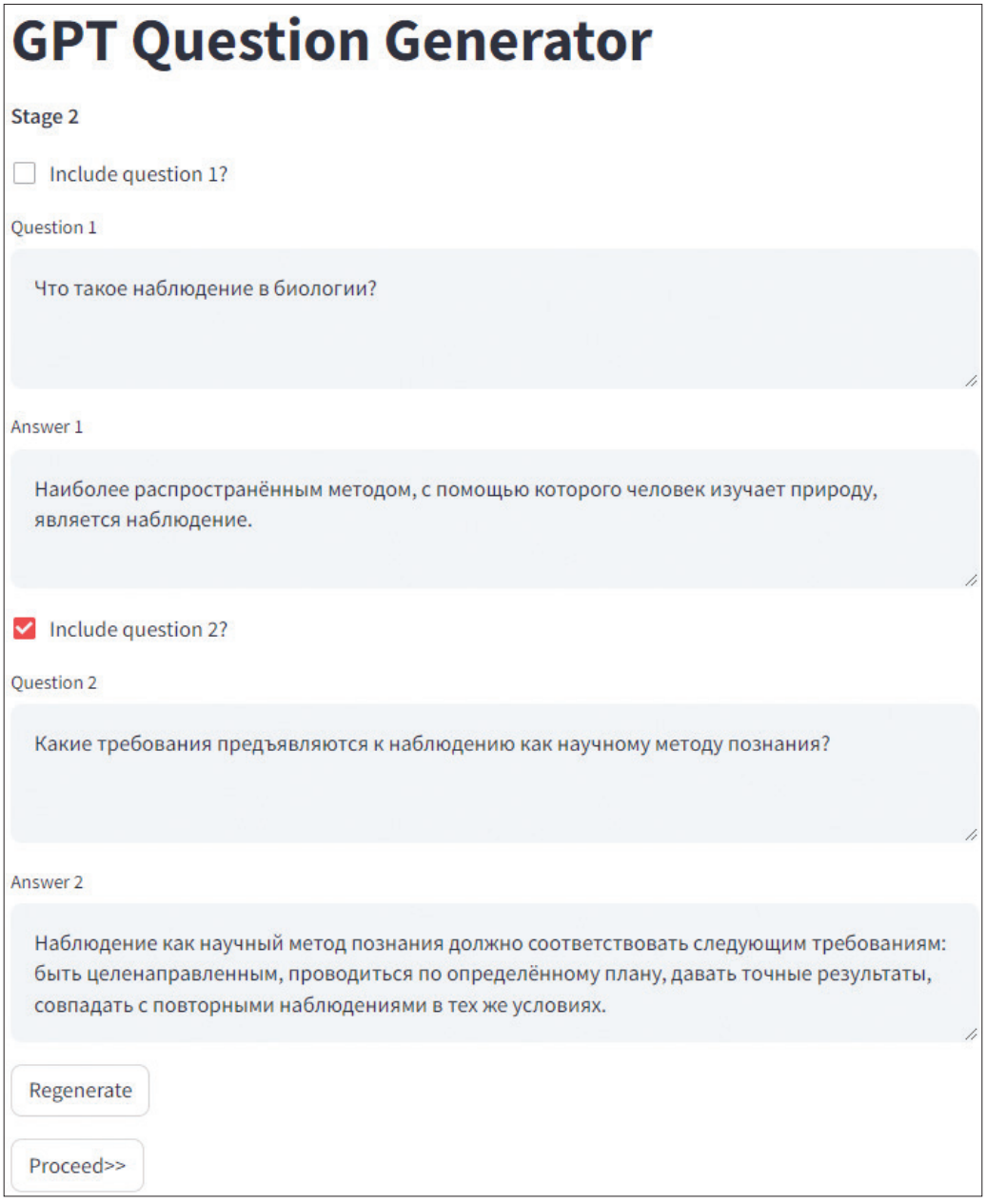

Fig. 1. The process of selection

Рис. 1. Процедура отбора

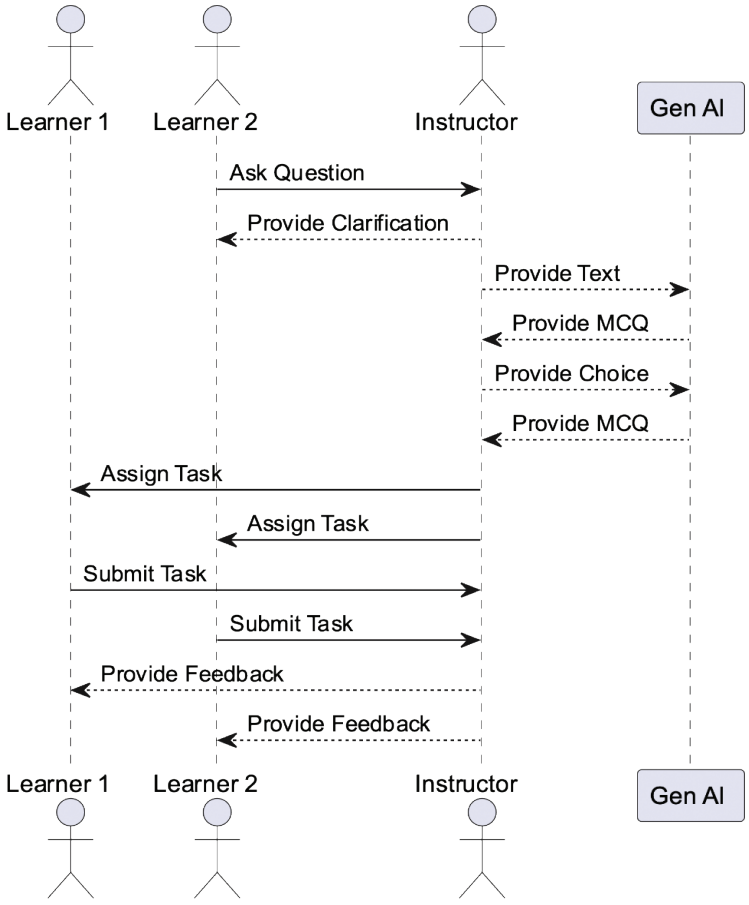

Fig. 2. The general scheme of organizing the educational process using generative artificial intelligence

Рис. 2. Общая схема организации образовательного процесса с помощью генеративного искусственного интеллекта

The teacher receives a set of questions and answers, which they can edit and use in their practice.

The further scenarios of using the questions depend on the teacher's preferences. They can be used during face-to-face work with students in their class. However, we assume that the most likely scenario is their inclusion in electronic lessons, as part of the Moscow Electronic School course scenarios to improve their quality (Vachkova, Patarakin, & Petryaeva, 2020). In this regard, we separately considered the ability of generative artificial intelligence to generate questions in JSON format, which can easily be included in the scenarios of electronic lessons of the Moscow Electronic School. However, before embarking on the development of including questions in the scenarios of educational activities, it is necessary to consider the possible problems that a teacher may encounter when working with automatically generated questions.

The teacher's interaction with automatically generated questions presents several challenges. After the teacher uploads a text fragment from the textbook, the language model generates a question and a correct answer, along with several incorrect answers. The text of the question should be accurately written, assuming the possibility of answering, corresponding to the studied topic, and not containing hints to the correct answer. It is important that the correct answer is genuinely correct and does not contain contradictions. It is very important that the incorrect answers appear plausible, and their rejection requires the student's effort and understanding. The general scheme of checking the questions and answers is presented on Figure 3. We used this scheme as a hint for teachers on what to pay attention to and as an instrument for assessing the quality of questions.

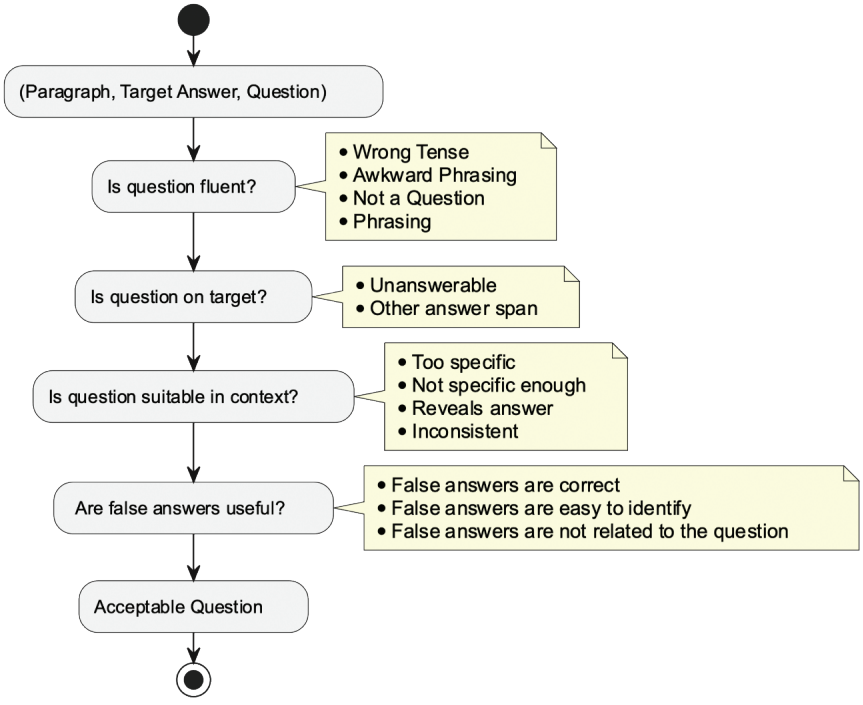

Fig. 3. The general scheme of checking the questions and answers

Рис. 3. Общая схема проверки вопросов и ответов

In the experiments, teachers of computer science, biology, and history were involved in the independent selection of topics for the generation of educational materials. For instance, the computer science teacher chose the theoretical topic "Modeling" for question generation, justifying this choice by the limited availability of theoretical questions with non-ambiguous applications in the field of computer science. While ready-made problem sets and tests are often available for most educational topics, teachers of computer science frequently find themselves having to devise theoretical questions on their own.

Which of the following statements best describes a model?

A. A new entity that reflects essential features of the studied subject, process, or phenomenon for the purpose of modeling.

B. A new entity that reflects non-essential features of the studied subject, process, or phenomenon for the purpose of modeling.

C. A new entity that reflects all features of the studied subject, process, or phenomenon.

D. A new entity that reflects only some features of the studied subject, process, or phenomenon.

The experiment described in this paper focuses on the subject of history, with the theme "Great Geographical Discoveries." This theme was chosen deliberately, as the text is abundant in dates and facts, which can be effectively verified using tests. The purpose of this experiment is to explore the potential of using multiple-choice questions to assess students' understanding of historical events and their ability to verify the accuracy of the information provided in the text. The chosen theme also highlights the importance of evaluating the effectiveness of various assessment methods in history education, thereby contributing to the ongoing discourse on the role of testing in historical studies.

Who was the first explorer to discover North America?

Christopher Columbus

Ferdinand Magellan

Vasco da Gama

Amerigo Vespucci

The field of biology presented a different scenario. Here, the test generator was not used for regular lessons but was necessary to prepare students for exams. Students taking biology are faced with the need to know a vast amount of theoretical material across various biology topics, starting from the 5th grade. One of the most significant features of biology is the extensive conceptual framework within the subject, the knowledge of which is essential for passing the biology exam. Technical terms such as chromatid, bivalent, metaphase plate, crossing over, and many others are abundant in the school subject of biology. These terms are essential for passing the biology exam, similar to the necessity of knowing facts and event dates in history. In this context, the test generator becomes a valuable tool that can be used to assimilate this conceptual framework within the field of biology. Multiple-choice questions (MCQs) can be utilized to enhance student learning in biology through various mechanisms. Well-constructed MCQs can prompt students to engage in higher-order thinking, such as analysis and evaluation, thereby fostering a deeper understanding of the subject matter. Additionally, MCQs can be used formatively to provide ongoing feedback to students, enabling them to identify and address areas of weakness. Furthermore, the generation of MCQs can be automated using artificial intelligence, allowing educators to focus on instructional activities while ensuring a steady supply of high-quality assessment items. Moreover, the use of MCQs can facilitate efficient

grading, particularly in large classes, and can be instrumental in preparing students for standardized examinations. Therefore, the strategic deployment of MCQs in biology education can contribute to the cultivation of critical thinking skills, the provision of timely feedback, and the optimization of instructional resources.

The differences between objects in the living and non-living nature are as follows:

A. Living and non-living objects differ in the complexity of their structure and the high level of organization of the life processes occurring within them.

B. Living and non-living objects differ in the complexity of their structure and the low level of organization of the life processes occurring within them.

C. Living and non-living objects differ in the complexity of their structure and the high level of organization of the processes occurring within them.

The correct answer is A. Living and non-living objects differ in the complexity of their structure and the high level of organization of the life processes occurring within them. This distinction is fundamental in the study of biology and is essential for understanding the characteristics of living organisms and their interactions with the environment.

The conducted experiments demonstrate the capability of modern large language models to generate simple text-based multiple-choice questions suitable for use. While the current need for distractors requires some manual verification and refinement by educators, this challenge is expected to be addressed by generative AI in the near future. The process of creating questions and their correct and incorrect solutions is no longer purely a creative task. Experiments conducted by school teachers using artificially generated materials show that hierarchical structures can be formed around such materials, similar to those that emerged around other technologies in the past, such as personal computers and internet access (see Figure 4).

During those times, the role of the computer science teacher was instrumental in guiding the adoption of these groundbreaking innovations.

Conclusion

The experiments conducted on the generation of educational questions using generative artificial intelligence currently demonstrate that teachers can create multiple-choice questions with the technical tools available. The quality of the generated questions largely depends on the quality of the original material. The application of the teaching methodology, which includes AI-generated tasks, brings us back to the key theme of fostering computational thinking among all participants in the educational process (Parandekar, Patarakin, & Yayla, 2023). This includes not only students but also teachers and authors of modern textbooks. From this perspective, the experimental process of obtaining multiple-choice questions from generative AI serves as a simultaneous test of the capabilities of Gen AI and a validation of the textbook's unambiguity and consistency of the presented knowledge. It is possible that

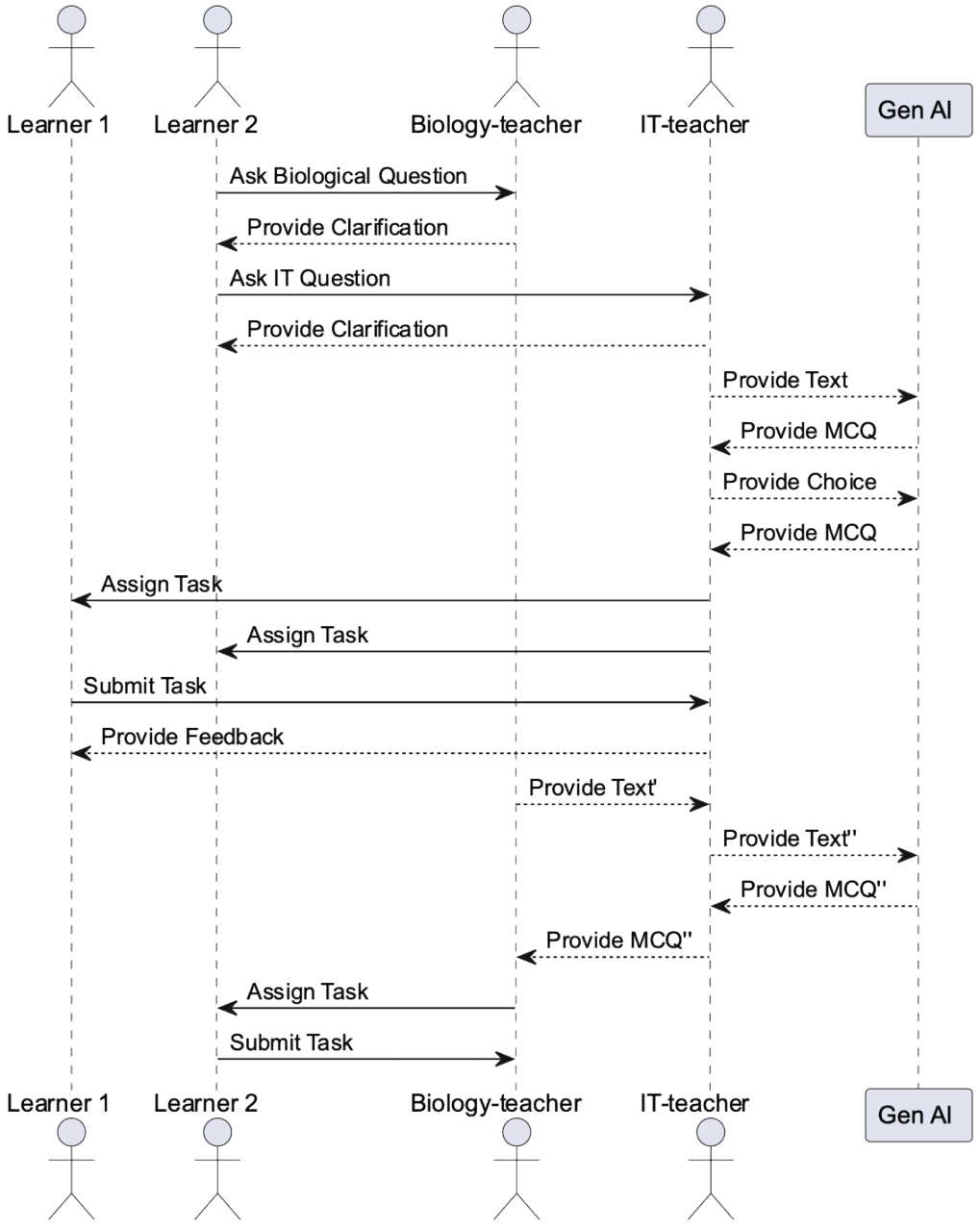

Fig. 4. Hierarchical structures

Рис. 4. Иерархические структуры

the process also serves as a test of the textbook's ontological integrity. From the perspective of computational thinking, a good modern textbook is one from which generative AI can formulate coherent and unambiguous questions. The creation of questions based on the textbook may serve as a test of the quality and ontological integrity of the material. If incorrect questions are generated by the AI during the translation process, there is a high probability that the textbook itself engenders confused perceptions in the student's mind. The links between generative artificial intelligence and computational thinking are evident in the potential for generative artificial intelligence to validate the coherence and integrity of educational materials through the generation of unambiguous questions.

The potential scenario in which generative artificial intelligence (AI) serves as distinct actors for both students and teachers raises concerns about the future of education. This framework can be likened to the long-standing "1 student: 1 computer" model, where the emphasis is not only on each student having their own computer, but also on the same computer being used by the student for both educational and everyday tasks, without a division between home and school computer use. This potential future scenario, as depicted in Figure 5, presents an uncomfortable yet highly probable situation where generative artificial intelligence for students and teachers are separate actors.

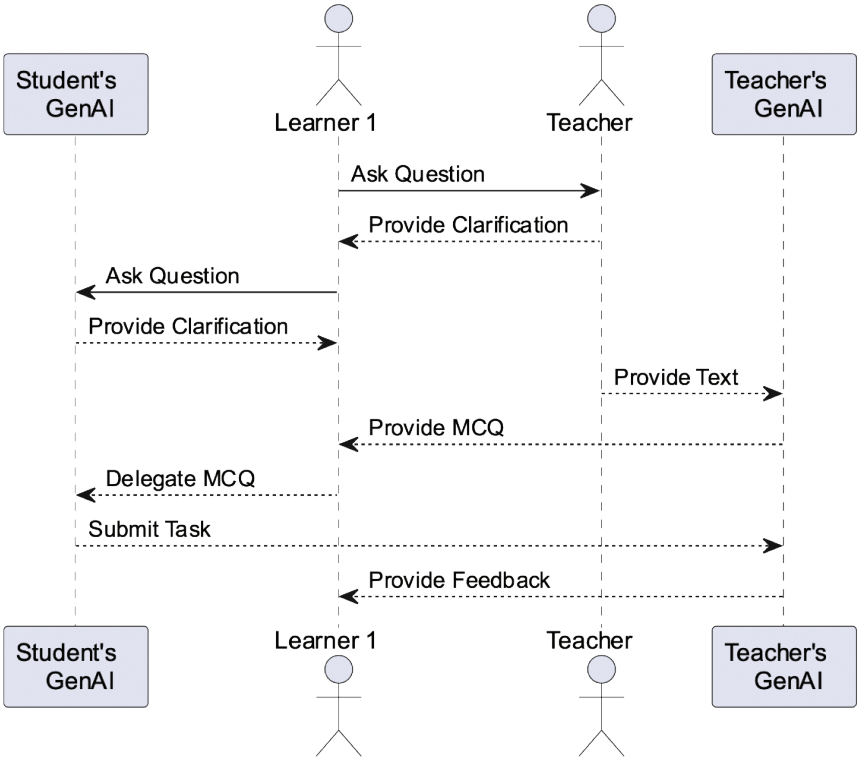

Fig. 5. Separate generative artificial intelligence actors

Рис. 5. Отдельные акторы генеративного искусственного интеллекта

Educators and administrators must thoroughly understand the risks and benefits associated with the use of generative AI in education to effectively navigate this evolving landscape.

References

1. Kurdi, G., Parsia, B., Leo, J., & Sattler, U. (2020). A Systematic Review of Automatic Question Generation for Educational Purposes. *Int J Artif Intell Educ, 30*(1), 121–204.

2. Ch, D. R., & Saha, S. K. (2020). Automatic Multiple Choice Question Generation From Text: A Survey. *IEEE Transactions on Learning Technologies, 13*(1), 14–25.

3. Matos, R. M. F. de. (2022). *Automatic Generation of Multiple Choice Questions.* Dissertation in the context of the Master in Informatics Engineering, specialization in Intelligent Systems, advised by Prof. Hugo Oliveira and Hugo Amaro and presented to the Department of Informatics Engineering of the Faculty of Sciences and Technology of the University of Coimbra.

4. Moore, S., Nguyen, H. A., Fang, E., & Stamper, J. (2023). *Crowdsourcing the evaluation of multiple-choice questions using item-writing flaws and bloom's taxonomy.* In: arXiv:2205.01730v1 [cs.CL].

5. Hadifar, A., Bitew, S. K., Deleu, J., Develder, C., & Demeester, T. (2023). EduQG: A multi-format multiple-choice dataset for the educational domain. *IEEE Access, 11,* 20885–20896.

6. Laban, P., Wu, Ch.-Sh., Murakhovska, L., Liu, W., & Xiong, C. (2022). Quiz Design Task: Helping Teachers Create Quizzes with Automated Question Generation. In: *Findings of the Association for Computational Linguistics: NAACL 2022,* July 10–15, 2022 (pp. 102–111). Association for Computational Linguistics.

7. Patarakin, E. D., Burov, V., & Sochnikov, D. (2023). Experimental generation of educational tasks in natural science disciplines using artificial intelligence. *MCU Journal of Pedagogy and Psychology, 17*(4), 38–58.

8. Vachkova, S. N., Patarakin, E. D., & Petryaeva, E. Y. (2020). Content Quality of Lesson Scenarios in Moscow E-School. *SHS Web Conf, 79,* 01017.

9. Parandekar, S., Patarakin, E., & Yayla, G. (2023). A Modern Aspect of Instrumental Literacy: Coding. In: Dobryakova, M., Froumin, I., Barannikov, K., Moss, G., Remorenko, I., & Hautamäki, J. (Eds.). *Key Competences and New Literacies: From Slogans to School Reality* (pp. 367–390). UNIPA Springer Series. Springer, Cham.

Статья поступила в редакцию: 25.11.2023;
одобрена после рецензирования: 19.12.2023;
принята к публикации: 15.01.2024.

The article was submitted: 25.11.2023;
approved after reviewing: 19.12.2023;
accepted for publication: 15.01.2024.

Information about the authors:

Yevgeny D. Patarakin — Doctor of Education Sciences, Professor at Institute of Digital Education of Moscow City University and Institute of Education of Higher School of Economics, patarakined@mgpu.ru ✉, https://orcid.org/0000-0002-1216-5043

Vasiliy V. Burov — scientist at Institute of Digital Education, Moscow City University, Moscow, Russia,
burovvv@mgpu.ru, https://orcid.org/0000-0001-5571-5693

Karim D. Salimullin — graduate student at Institute of Digital Education, Moscow City University, Moscow, Russia,
salimullinkd@mgpu.ru

Dmitry V. Soshnikov — PhD, Associate Professor at Moscow Aviation Institute and Faculty of Computer Science of Higher School of Economics, Technical Director of Generative AI Lab at HSE Art and Design School, Moscow, Russia,
dmitri@soshnikov.com, https://orcid.org/0000-0003-1021-091X

Информация об авторах:

Евгений Дмитриевич Патаракин — доктор педагогических наук, доцент, профессор департамента информатики, управления и технологий Института цифрового образования Московского городского педагогического университета; профессор департамента образовательных программ Института цифрового образования НИУ ВШЭ школы дизайна, Москва, Россия,
patarakined@mgpu.ru ✉, https://orcid.org/0000-0002-1216-5043

Василий Владимирович Буров — научный сотрудник Института цифрового образования Московского городского педагогического университета, Москва, Россия,
burovvv@mgpu.ru, https://orcid.org/0000-0001-5571-5693

Салимуллин Карим Дамирович — аспирант департамента информатики, управления и технологий Института цифрового образования Московского городского педагогического университета, Москва, Россия,
salimullinkd@mgpu.ru

Дмитрий Валерьевич Сошников — кандидат физ.-мат. наук, доцент Московского авиационного института, доцент НИУ ВШЭ, технический руководитель лаборатории генеративного ИИ Школы дизайна НИУ ВШЭ, Москва, Россия,
dmitri@soshnikov.com, https://orcid.org/0000-0003-1021-091X

Contribution of the authors: the authors contributed equally to this article. The authors declare no conflicts of interests.

Вклад авторов: все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.