Analytical article
UDC 37.01:004
DOI: 10.25688/2076-9121.2023.17.4.02

EXPERIMENTAL GENERATION OF EDUCATIONAL TASKS IN NATURAL SCIENCE DISCIPLINES USING ARTIFICIAL INTELLIGENCE[1]

Evgeny D. Patarakin[1], Vasiliy V. Burov[2], Dmitry V. Soshnikov[3]

[1, 2] *Moscow City University, Moscow, Russia*
[1, 2, 3] *Higher School of Economics, Moscow, Russia*
[3] *Moscow Aviation Institute, 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] *dmitri@soshnikov.com, https://orcid.org/0000-0003-1021-091X*

Abstract. This study investigates the suitability of modern generative models for the automatic generation of educational task texts. In the first part of the study, we conducted a bibliometric mapping of the research field related to automatic question generation, utilizing three databases: Lens, Dimensions, and the ACM Digital Library. In the second part, we compared the capabilities of three generative systems (ChatGPT-3.5, YaGPT, GigaChat) to formulate various types of assignments based on a textbook content: multiple-choice questions, open-ended questions, and essay topics based on a given text fragment. The source material was a fragment of a fifth-grade biology textbook describing the difference between living and non-living things. The evaluation encompassed an assessment of the models' ability to generate diverse question variants, their proficiency in recording these questions in JSON format for integration into digital platforms, and the correctness of the questions in terms of grammar, relevance, and pedagogical appropriateness.

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

[1] Статья публикуется в авторской редакции.

Аналитическая статья
УДК 37.01:004
DOI: 10.25688/2076-9121.2023.17.4.02

ЭКСПЕРИМЕНТАЛЬНАЯ ГЕНЕРАЦИЯ ЗАДАНИЙ ПО ЕСТЕСТВЕННО-НАУЧНЫМ ДИСЦИПЛИНАМ ПРИ ПОМОЩИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Евгений Дмитриевич Патаракин[1], Василий Владимирович Буров[2], Дмитрий Валерьевич Сошников[3]

[1, 2] *Московский городской педагогический университет, Москва, Россия*
[1, 2, 3] *Высшая школа экономики, Москва, Россия*
[3] *Московский авиационный институт, 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] *dmitri@soshnikov.com, https://orcid.org/0000-0003-1021-091X*

Аннотация. В работе исследовалась пригодность современных генеративных моделей для автоматического создания текстов учебных задач. В первой части работы мы провели библиометрическое картирование поля исследовательских работ, связанных с автоматической генерацией вопросов. В качестве источников были использованы три базы данных: Lens, Dimensions и Digital Library ACM. Во второй части работы мы сравнивали возможности трех генеративных систем (ChatGPT-3.5, YaGPT, GigaChat) формулировать на основе текста учебника задания различных видов: вопросы с вариантами ответа, вопросы с открытым ответом, темы эссе по заданному фрагменту текста. В качестве исходного материала был взят фрагмент текста учебника по биологии для пятого класса, в котором описывалось различие живого и неживого. Для каждой из поставленных задач оценивалась способность генеративной модели формулировать разнообразные варианты вопросов, записывать вопросы в формате JSON, корректность создаваемых моделями вопросов.

Ключевые слова: контент, генерируемый искусственным интеллектом, AIGC, ChatGPT-3.5, YaGPT, GigaChat

For citation: Patarakin, E. D., Burov, V. V., & Soshnikov, D. V. (2023). Experimental generation of educational tasks in natural science disciplines using artificial intelligence. *MCU Journal of Pedagogy and Psychology, 17*(4), 28–41. https://doi.org/10.25688/2076-9121.2023.17.4.02

Для цитирования: Патаракин, Е. Д., Буров, В. В., и Сошников, Д. В. (2023). Экспериментальная генерация заданий по естественно-научным дисциплинам при помощи искусственного интеллекта. *Вестник МГПУ. Серия «Педагогика и психология», 17*(4), 28–41. https://doi.org/10.25688/2076-9121.2023.17.4.02

Introduction

The topic of automatic generation of learning tasks has gained significant relevance in recent years, particularly due to the advent of advanced generative artificial intelligence models (GenAI). This research area merges the techniques of automatic task generation or educational material generation with computer programming methods and technologies and has the potential to revolutionize the way we approach education. The automatic generation of learning tasks can be applied across a wide range of educational fields, including but not limited to mathematics, physics, languages, and computer science. The importance of this technology becomes evident when we consider its potential applications. One of the most significant applications is in the realm of adaptive learning. This involves the creation of personalized learning tasks that are tailored to suit the individual needs and knowledge level of each student. This not only makes learning more efficient, but also ensures that the student is engaged and challenged at an appropriate level. Another key application is in mass learning. The technology can provide a large number of students with learning materials that can be dynamically generated. This could potentially revolutionize the way education is delivered, particularly in large-scale educational institutions or online learning platforms. The technology also has significant implications for testing and assessment. It can be used for the creation of test tasks and the evaluation of student results. This could streamline the assessment process and make it more objective and efficient. Furthermore, the technology can be used for content creation automation. This involves the automatic generation of training materials, tutorials, and other educational content. This could significantly reduce the workload of educators and allow them to focus more on teaching and less on content creation. Finally, generative models and large language models in particular present new opportunities for the creation of personal educational assistants, which will support students in their learning process through natural language interactions.

The automatic task generation technology employs machine learning, artificial intelligence, and optimization algorithms to create a variety of tasks. These tasks take into account various parameters, such as the complexity of the task, the student's knowledge level, and the learning context. However, it is important to note that the creation of automatic learning tasks is not a straightforward process. It is a complex task that may require significant effort to ensure the quality and effectiveness of the educational process. It is also important to understand that automatic task generation is not a panacea. It should be viewed as an additional tool in the educational process, rather than a replacement for professional teachers and experts.

Automatic question generation: A bibliometric mapping of the research literature

To identify the most promising direction for development, we conducted a mapping of the bibliometric field related to the topic at hand. Specifically, we mapped bibliometric information on the topic of generating educational tasks and questions. To do this, we utilized three of the largest and currently accessible in Russia bibliographic databases: Lens, Dimensions, and the ACM Digital Library.

Lens (www.lens.org) is a free bibliographic database that provides access to scientific information, with a particular emphasis on patents and patent citations. This makes it a valuable resource for identifying trends and developments in the field of educational task generation.

Dimensions (app.dimensions.ai) is a scientific database that covers various types of scientific publications, grants, and patents. This comprehensive coverage allows us to gain a broad understanding of the research landscape in the area of educational task generation.

The ACM Digital Library (dl.acm.org) is a platform that hosts a variety of articles and reports related to computer science and technology. This makes it particularly relevant for our research, as the automatic generation of educational tasks often involves the use of such technologies.

For each of these databases, we formulated identical queries requesting publications on the topic of automatic generation of educational tasks. The general schema of the query is as follows:

[All: "automatic question generation"] AND [E-Publication Date: (01/01/2010 TO 12/31/2023)]

As a result of our research, we collected a substantial body of research on the topic of automatic generation of educational tasks. We obtained a sample of 60 articles from the ACM Digital Library, a sample of 152 articles from Lens, and a sample of 342 articles from Dimensions. These numbers provide a rich dataset for our bibliometric analysis. All records were saved and verified in Zotero, a bibliographic manager (Winslow et al., 2016). This tool was instrumental in organizing and managing the large volume of data we collected. It allowed us to efficiently sort and categorize the articles, making the subsequent analysis more manageable. Following the verification process, the data was presented in the form of maps using VOSviewer, a software tool for constructing and visualizing bibliometric networks (Ginting, 2023; Al Husaeni, & Nandiyanto, 2022). These networks made up researchers or individual publications, and they provide a visual representation of the relationships between these entities.

The mapping of materials from the ACM Digital Library was particularly illustrative, as the library's materials are detailed and structured using keywords. This allowed us to identify clusters of related articles and visualize the landscape of research on the automatic generation of educational tasks.

The division of words into clusters is presented in the figure 1.

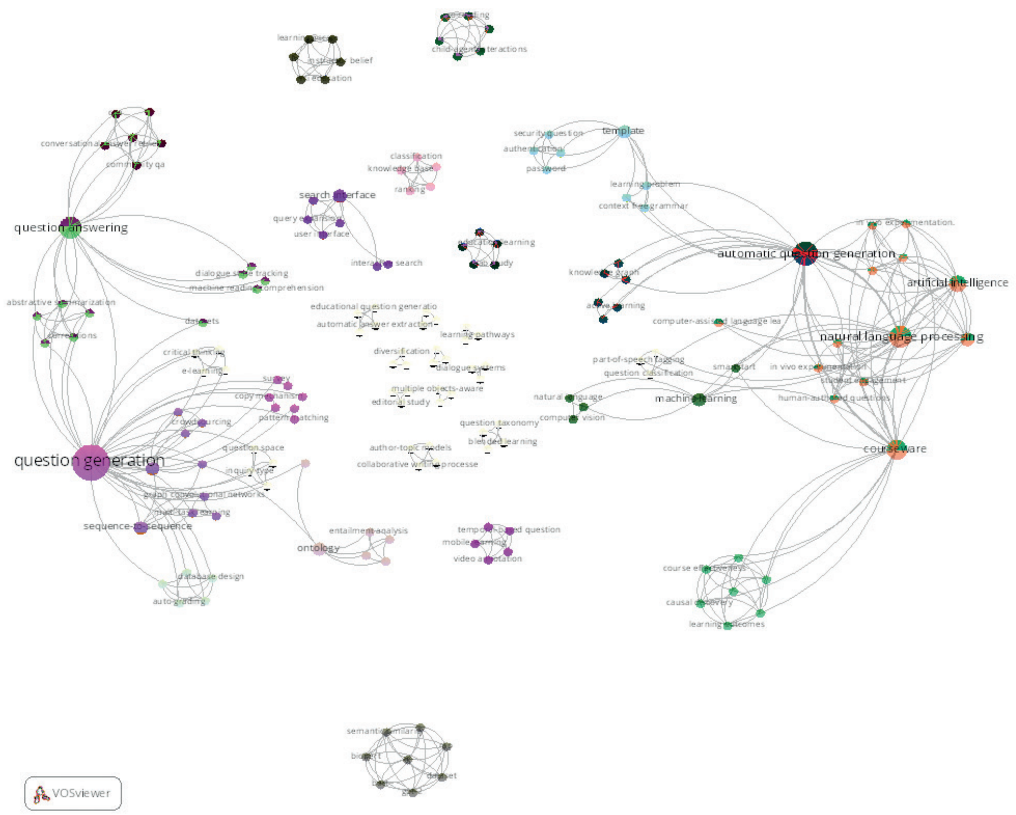

Fig. 1. ACM DL keyword clusters on learning tasks generation

Рис. 1. Кластеры ключевых слов ACM DL по формированию учебных задач

To provide a more detailed view of the data, with the ability to navigate and examine individual groups, we have provided a link to an interactive feature (https://app.vosviewer.com/?json=https://drive.google.com/uc?id=1T1jR3taER-r1Ax-JMuqSg-8enzJ6rFFQL).

This feature allows for a deeper exploration of the research landscape, enabling users to delve into specific areas of interest and uncover patterns and trends that may not be immediately apparent from a high-level overview.

Figure 1 clearly distinguishes two clusters of research. On the left side, we can see a cluster related to crowdsourcing for question creation and verification, which includes topics such as question answering, data-driven text generation, and crowdsourcing. On the right side, there is a cluster related to automatic generation, which includes artificial intelligence, automatic generation, and student engagement. In our subsequent analysis of the publications, we paid particular attention to the relationship between automatic question generation and the subsequent crowdsourcing verification of their quality. This is a significant area of interest as it combines the power of artificial intelligence with the collective intelligenceourcing, potentially leading to more effective and efficient educational task generation. Content generation systems, as reviewed in references (Wu et al., 2023; Cao et al., 2023), offer options for content generation either with the support of artificial intelligence (AI-assisted writing, AIAW) or entirely by artificial intelligence (AI-generated writing, AIGW). The authors emphasize that content generation programs can be configured to adhere to formal rules for creating educational and assessment materials. The strength of these systems lies in their ability to from text format into other formats, expanding the range of potential applications.

Of practical interest is the review of automatic question generation systems (Mulla, & Gharpure, 2023), in which the authors formulate the problem of question generation, group question creation systems, and highlight individual question generation systems related to specific knowledge domains. This review provides valuable insights into the current state of the field, offering a comprehensive overview of the various systems and methodologies being employed. It also identifies potential areas for future, particularly in the context of domain-specific question generation. This could lead to more targeted and effective educational tasks, enhancing the learning experience for students in those specific domains.

Among the works dedicated to question generation in the ACM digital library, it is important to highlight a review (Zhang et al., 2021) that discusses the challenges of creating questions for learning systems and the various levels of text that can be utilized for question generation. This review provides a comprehensive examination of the complexities involved in question creation, including the consideration of different text levels, from simple sentences to complex paragraphs, as potential sources for question generation. This multi-level approach to question generation offers a more nuanced understanding of the process, potentially leading to more effective and engaging learning materials. Another notable work is a study on the crowdsourcing evaluation of multiple-choice questions in mathematics and chemistry (Moore et al., 2023). This study explores the potential of crowdsourcing as a tool for assessing the quality of automatically generated questions. By tapping into the collective intelligence of a crowd, this approach could offer a more robust and reliable evaluation method, ensuring that the generated questions are both accurate and effective in assessing student understanding.

A third work of interest is a study on question generation based on image analysis (Patil, & Patwardhan, 2020). This approach could be particularly significant in generating questions in fields such as biology, chemistry, and medicine, where visual information plays a crucial role. Incorporating image analysis into the question generation process, this approach could lead contextually relevant and engaging questions, enhancing the learning experience in these visually oriented fields.

Among the publications found in the Dimensions database, noteworthy are a review of automatic multiple-choice question generation systems (Madri, & Meruva, 2023) reviews of methods for automatic generation of texts, questions, and answers

(Goyal et al., 2023a; Goyal et al., 2023b), and a study on the evaluation of automatically created tests for medical education (Falcão et al., 2023). These publications offer valuable insights into the current state of automatic question generation, highlighting various methodologies and their applications in different fields.

In the Lens database, a description of a framework for generating multiple-choice questions (Kumar et al., 2023) and a study on the generation of multiple-choice questions, one of the results of which was the creation of a training chatbot (Panchal et al., 2021), were highlighted. These works demonstrate the versatility of automatic question generation, showcasing its potential applications in diverse contexts, from traditional assessment methods to more innovative approaches like chatbots.

The application of GenAI in education is observed to be uneven, with the most significant breakthroughs seen in the realm of computer science education. This is largely attributed to the fact that the neural network was trained on materials from the GitHub repository, and the potential for using artificial intelligence in teaching programming has been explored for quite some time. In the last two years, several studies have been conducted that highlight the capabilities of GenAI in this field. One study (Finnie-Ansley et al., 2022) investigated the ability of robot-agents to solve problems in the field of computer science. This research valuable insights into how AI can be used to automate problem-solving, potentially freeing up more time for educators to focus on other aspects of teaching. Another study (Kim et al., 2021) focused on how a program can explain to a student the process of obtaining solutions. This research underscores the potential of GenAI as a teaching tool, capable of providing detailed explanations and step-by-step guidance to students, thereby enhancing their understanding of complex concepts. A third study (Suh, & An, 2022) explored the use of GenAI in creating conditions for learning computational thinking through comics. This innovative approach demonstrates how GenAI can be used to make learning more engaging and interactive, potentially increasing student motivation and interest in the subject matter. In another study GenAI was used to simulate an educational microworld similar to the well-known Boxer learning system (Lewis, 2022). This research showcases the potential of GenAI in creating immersive and interactive learning environments, which can provide students with a more hands-on and engaging learning experience. Finally, a study (Jonsson, & Tholander, 2022) focused on generating situations where students are required to solve problems together with GenAI. This research highlights the potential of GenAI in promoting collaborative problem-solving, a key skill in today's increasingly interconnected and complex world. By working together with AI, students can develop their problem-solving skills while also gaining a deeper understanding of how AI works.

The conducted study resulted in the selection of the following main directions according to the types of content and the logic of its use in the educational process. When considering possible approaches to implementation, we imposed a restriction on the possibility of using Russian implementations of large language models. This is related both to certain limitations on the use of such popular international products as ChatGPT, and to the fact that it is important for us to have a high-quality interaction in the Russian language, which is not yet well-possessed by the newly available open-source language models such as LLaMA or similar ones. However, possibly due to their low hardware resource requirements, this latter class of models may potentially be the most promising for use in mass practical implementations in case satisfactory Russian language implementations become available. This effectively narrows the spectrum of possible solutions today to solutions from two Russian providers: Yandex (YaGPT) or Sberbank (GigaChat and ruGPT-3).

Possible Approaches to Generating Educational Tasks

Two potential methods for creating educational tasks using generative artificial intelligence include using an existing task dataset or textbook content. When using an existing task dataset, a large language model, like ruGPT-3, is fine-tuned on the dataset to generate tasks on specific topics (see Fig. 2).

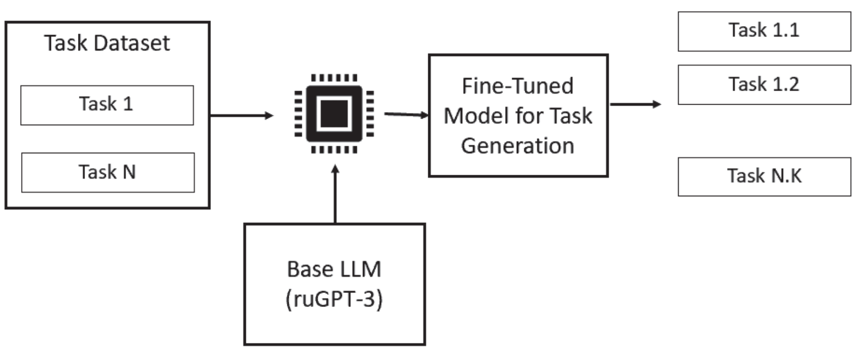

Fig. 2. Generation of tasks using fine-tuned LLM

Рис. 2. Генерация задач с использованием тонкой настройки LLM

However, this approach has several limitations. Firstly, the quality and diversity of the generated tasks are dependent on the original dataset. If the dataset is flawed or lacks diversity, the tasks generated may be incorrect or inadequate. Secondly, the automatic generation of tasks may limit creative thinking as it's based on set algorithms and rules. While AI models can mimic patterns, they may not replicate the creativity a human teacher can bring. Lastly, automatically generated tasks may not always consider the learning context or curriculum needs. Some tasks may not be relevant to the current learning material, potentially hindering students' educational progress.

The proposed research will conduct an experiment with dialogue language models YaGPT and GigaChat to evaluate their ability to generate tasks from textbook excerpts automatically (so-called content-augmented generation). This method doesn't require re-training the model but does require careful evaluation of the tasks' appropriateness (see fig. 3).

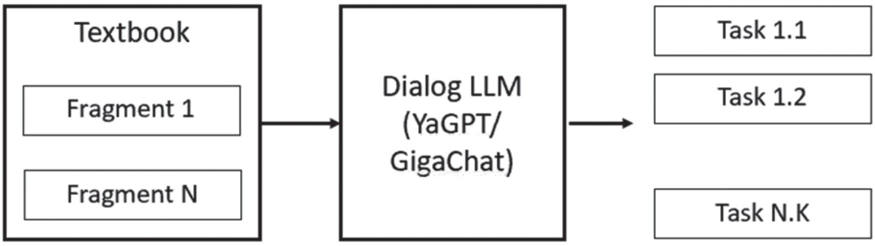

Fig. 3. Using standard LLM for content-augmented task generation

Рис. 3. Использование стандартного LLM для генерации задач с расширенным содержанием

The approach's benefits include better consistency with educational material, improved understanding of the course context, and the ability to generate more organized tasks due to the structured nature of textbooks. However, the method has limitations, including limited task diversity, incomplete textbook coverage, interpretation difficulties for machine learning algorithms, and a potential lack of creativity in task creation.

A Comparative Study of ChatGPT-3.5, YaGPT, and GigaChat

In the second part, we compared the capabilities of three generative systems (ChatGPT-3.5, YaGPT, GigaChat) to formulate various types of assignments based on textbook fragments: multiple-choice questions, open-ended questions, and essay topics based on a given text fragment. The source material was a fragment of a fifth-grade biology textbook. The selected passage discusses the differences and similarities between living and non-living things, that both are composed elements, but in different proportions. It emphasizes that living organisms have unique characteristics like cellular structure and metabolism, which are absent in non-living entities. The text also explains that life processes in living organisms, including nutrition, respiration, and excretion, ensure a continuous flow of matter and energy. The fragment size was 3404 characters including spaces.

In our study, we conducted a comprehensive evaluation of the capabilities of various generative models in several key pedagogical areas. This assessment was designed to provide a robust understanding of the potential and limitations of these generative models in the context of automated educational task generation, a burgeoning field with significant implications for the future of education. The first parameter we assessed was the models' ability to formulate a diverse range of question variants. This aspect is of paramount importance in maintaining student engagement and ensuring a thorough understanding of the educational material. The ability to generate a variety of questions from a single piece of text allows for a more comprehensive exploration of the topic at hand, thereby promoting a deeper level of understanding among students. This parameter was rated on a scale from 1 to 5, with 5 indicating excellent performance in question diversity.

The second parameter we evaluated was the models' capability to record these questions in JSON format. This feature is essential for the seamless integration of the generated questions into digital learning platforms. As education increasingly moves towards digital platforms, the ability to easily incorporate generated questions into these systems becomes increasingly important. This parameter was also rated on a scale from 1 to 5, with 5 indicating a high degree of compatibility with digital learning platforms. Lastly, we examined the correctness of the questions generated by the models. This involved an in-depth analysis of the grammatical accuracy, relevance to the source material, and the pedagogical appropriateness of the questions. Ensuring that the generated questions are grammatically correct, contextually relevant, and pedagogically sound is crucial for their effective use in an educational setting. This parameter was rated on a scale from 1 to 5, with 5 indicating a high degree correctness.

Multiple-choice questions are a popular form of assessment in many academic fields. They consist of a question or statement, followed by several possible answers, typically four or five, from which the student must choose the correct one. The main advantage of this type of assignment is its objectivity, as the answers are either right or wrong, leaving no room for interpretation. It allows for easy grading and is efficient for testing a wide range of knowledge in a short period. However, it may not fully assess a student's depth of understanding or critical thinking skills.

The following prompt was used for generating multiple-choice questions in the study:

«Imagine that you are a biology teacher for younger students. Formulate 5 questions, each with five answer options, pertaining to the following text: <<TEXT>>»

Table 1 / Таблица 1

Comparative Evaluation of Models for Multiple Choice Question Generation

Сравнительная оценка моделей генерации вопросов с множественными вариантами ответов

	ChatGPT-3.5	YaGPT	GigaChat
Question Diversity	5	3	3
Digital Platform Compatibility	5	2	1
Question Correctness	5	3	4

Open-ended questions, on the other hand, require students to formulate their own responses. These questions are designed to evaluate a student's ability to apply, analyze, and synthesize information. They encourage critical thinking and allow students to express their understanding in their own words. these questions can provide a deeper insight into a student's comprehension, they can be time-consuming to grade due to the need for individual evaluation and feedback.

The following query was used for generating open-ended questions in the study:

«*Envision yourself as a biology teacher for younger students. Formulate 5 open-ended questions pertaining to the text enclosed in square brackets. Provide the correct answer for each question. [<<TEXT>>]*»

Table 2 / Таблица 2

Comparison of Generative Models for Open-Ended Question Generation

Сравнение генеративных моделей для генерации открытых вопросов

	ChatGPT-3.5	**YaGPT**	**GigaChat**
Question Diversity	5	1	3
Digital Platform Compatibility	5	2	1
Question Correctness	5	3	2

Essay topics based on a text fragment are a common assignment in literature and language courses. In this type of assignment, students are given a piece of text and are asked to write an essay related to it. The essay could be an analysis of the text, a discussion of themes, or a response to the ideas presented. This type of assignment encourages close reading, critical analysis, and thoughtful interpretation. It allows students to deep into the text and demonstrate their understanding and analytical skills. However, it can be challenging as it requires a high level of comprehension, writing skills, and the ability to construct a coherent and persuasive argument.

We asked each generative model to devise 5 essay topics based on a given text fragment. The following request was used:

«*Envision yourself as a biology teacher for younger students. Formulate 5 short essay topics related to the text enclosed in square brackets. For each topic, list the key points (a couple of words each) that should be reflected in the student's response. [TEXT>>]*»

Table 3 / Таблица 3

Comparison of Generative Models for Short Essay Topic Generation

Сравнение генеративных моделей для генерации темы короткого эссе

	ChatGPT-3.5	**YaGPT**	**GigaChat**
Topic Diversity	5	2	3
Digital Platform Compatibility	5	2	1
Topic Correctness	5	3	2

Conclusion

Our study has highlighted the potential of GenAI for automating the generation of educational tasks and their subsequent validation, marking a significant departure from traditional algorithmic methods. The emergence of large language models has ushered in a new era in educational task generation, enabling innovative approaches.

We explored various generative models, including ChatGPT-3.5, YaGPT, and GigaChat, assessing their capabilities in terms of question diversity, compatibility with digital platforms, and question correctness. Our analysis showed that these models have the potential to enhance student engagement and understanding by formulating a wide range of question variants. They also demonstrated the ability to record these questions in JSON format, ensuring easy integration into digital learning platforms However, the quality of the questions varied in terms of grammatical accuracy, relevance to source material, and pedagogical appropriateness. Among the models evaluated, ChatGPT-3.5 showed superior performance in all characteristics. Importantly, these generative models offer the potential to create educational tasks across a broad spectrum.

These outcomes underscore the feasibility of leveraging GenAI to autonomously generate test tasks using a content-augmented approach. After comprehensive evaluation, we have determined that this approach stands out as the most effective means to mitigate potential drawbacks associated with automatically generated content, while offering extensive practical utility.

Furthermore, it provides a clear trajectory for further research and development aimed at practical implementation, promising enhanced educational experiences for students and educators alike.

References

1. Winslow, R. R., Skripsky, S. L., & Kelly, S. L. (2016). Chapter 14. Not Just for Citations: Assessing Zotero While Reassessing Research. In: D'Angelo, B. J., Jamieson, S., Maid, B., & Walker, J. R. (Eds.), *Information Literacy: Research and Collaboration across Disciplines* (pp. 287–304). The WAC Clearinghouse; University Press of Colorado. https://wac.colostate.edu/books/perspectives/infolit/

2. Ginting, S. L. B. (2023). A Computational Bibliometric Analysis of Esport Management using VOSviewer. *International Journal of Informatics, Information System and Computer Engineering (INJIISCOM), 4*(1), 31–48. https://doi.org/10.34010/injiiscom.v4i1.9570

3. Al Husaeni, D. F., & Nandiyanto, A. B. D. (2022). Bibliometric using Vosviewer with Publish or Perish (using google scholar data): From step-by-step processing for users to the practical examples in the analysis of digital learning articles in pre and post Covid-19 pandemic. *ASEAN Journal of Science and Engineering, 2*(1), 19–46.

4. Wu, J., Gan, W., Chen, Z., Wan, S., & Lin, H. (2023). *AI-Generated Content (AIGC): A Survey* (arXiv:2304.06632). arXiv. https://doi.org/10.48550/arXiv.2304.06632

5. Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P. S., & Sun L. (2023). *A Comprehensive Survey of AI-Generated Content (AIGC): A History of Generative AI from GAN to ChatGPT* (arXiv:2303.04226). arXiv. https://doi.org/10.48550/arXiv.2303.04226

6. Mulla, N., & Gharpure, P. (2023). Automatic question generation: a review of methodologies, datasets, evaluation metrics, and applications. *Progress in Artificial Intelligence*, *12*(1), 1–32. https://doi.org/10.1007/s13748-023-00295-9

7. Zhang, R., Guo, J., Chen, L., Fann, Y., & Chend, X. (2021). A review on question generation from natural language text. *ACM Trans. Inf. Syst., 40*(1). https://doi.org/10.1145/3468889

8. Moore, S., Nguyen, H. A., Fang, T., & Stamper, J. (2023). Crowdsourcing the evaluation of multiple-choice questions using item-writing flaws and bloom's taxonomy. *Proceedings of the Tenth ACM Conference on Learning @ Scale.* New York, NY, USA: Association for Computing Machinery, 25–34. https://doi.org/10.1145/3573051.3593396

9. Patil, C., & Patwardhan, M. (2020). Visual question generation: The state of the art. *ACM Comput. Surv., 53*(3). https://doi.org/10.1145/3383465

10. Madri, V. R., & Meruva, S. (2023). Acomprehensive review on MCQ generation from text. *Multimedia Tools and Applications,* 1–20. https://doi.org/10.1007/s11042-023-14768-5

11. Goyal, R., Kumar, P., & Singh, V. P. (2023a). Automated question and answer generation from texts using text-to-text transformers. *Arabian Journal for Science and Engineering,* 1–15. https://doi.org/10.1007/s13369-023-07840-7

12. Goyal, R., Kumar, P., & Singh, V. P. (2023b). A Systematic survey on automated text generation tools and techniques: application, evaluation, and challenges. *Multimedia Tools and Applications,* 1–56. https://doi.org/10.1007/s11042-023-15224-0

13. Falcão, F., Pereira, D. M., Gongalves, N., De Champlainn, A., Costa, P., & Pego J. M. (2023). A suggestive approach for assessing item quality, usability and validity of Automatic Item Generation. *Advances in Health Sciences Education,* 1–25. https://doi.org/10.1007/s10459-023-10225-y

14. Kumar, A. P., Nayak, A., Shenoy, K. M., Chaitanya, & Ghosh, K. (2023). A Novel Framework for the Generation of Multiple Choice Question Stems Using Semantic and Machine-Learning Techniques. *International Journal of Artificial Intelligence in Education.* Netherlands: Springer Science and Business Media LLC. https://doi.org/10.1007/s40593-023-00333-6

15. Panchal, P., Thakkar, J., Pillai, V., & Patil, S. (2021). Automatic Question Generation and Evaluation. *Journal of University of Shanghai for Science and Technology, 23*(5), 751–761. https://doi.org/10.51201/jusst/21/05203

16. Finnie-Ansley, J., Denny, P., Becker, B. A., Luxton-Reilly, A., & Prather, J. (2022). The Robots Are Coming: Exploring the Implications of OpenAI Codex on Introductory Programming. *Proceedings of the 24th Australasian Computing Education Conference.* New York, NY, USA: Association for Computing Machinery, 10–19. https://doi.org/10.1145/3511861.3511863

17. Kim, C., Lin, X., Collins, C., Taylor, G. W., & Amer, M. R. (2021). Learn, Generate, Rank, Explain: A Case Study of Visual Explanation by Generative Machine Learning. *ACM Transactions on Interactive Intelligent Systems, 11*(3–4), 23:1–23:34. https://doi.org/10.1145/3465407

18. Suh, S., & An, P. (2022). Leveraging Generative Conversational AI to Develop a Creative Learning Environment for Computational Thinking. In: *27th International Conference on Intelligent User Interfaces.* New York, NY, USA: Association for Computing Machinery, 73–76. https://doi.org/10.1145/3490100.3516473

19. Lewis, C. (2022). Automatic Programming and Education. In: *Companion Proceedings of the 6th International Conference on the Art, Science, and Engineering of Programming* (pp. 70–80). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3532512.3539664

20. Jonsson, M., & Tholander, J. (2022). Cracking the code: Co-coding with AI in creative programming education. In: *Creativity and Cognition* (pp. 5–14). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3527927.3532801

Статья поступила в редакцию: 11.07.2023;
одобрена после рецензирования: 27.08.2023;
принята к публикации: 15.09.2023.

The article was submitted: 11.07.2023;
approved after reviewing: 27.08.2023;
accepted for publication: 15.09.2023.

Информация об авторах:

Евгений Дмитриевич Патаракин — доктор педагогических наук, доцент, профессор департамента информатики, управления и технологий, Институт цифрового образования, Московский городской педагогический университет; профессор департамента образовательных программ, Институт образования, НИУ ВШЭ, Москва, Россия,
patarakined@mgpu.ru, https://orcid.org/0000-0002-1216-5043

Василий Владимирович Буров — научный сотрудник Института цифрового образования, Московский городской педагогический университет; советник Московского института электроники и математики им. А. Н. Тихонова, НИУ ВШЭ, Москва, Россия,
burovvv@mgpu.ru, https://orcid.org/0000-0001-5571-5693

Дмитрий Валерьевич Сошников — кандидат физико-математических наук, доцент, доцент Московского авиационного института; доцент департамента программной инженерии, факультет компьютерных наук, НИУ ВШЭ, Москва, Россия,
dmitri@soshnikov.com, https://orcid.org/0000-0003-1021-091X

Information about the authors:

Yevgeny D. Patarakin — Doctor of Education Sciences, Associate Professor, Professor of the Department of Informatics, Management and Technology at the Institute of Digital Education, Moscow City University; Professor of the Department of Educational Programs at the HSE Institute of Education, Moscow, Russia,
patarakined@mgpu.ru, https://orcid.org/0000-0002-1216-5043

Vasiliy V. Burov — Research Fellow at the Institute of Digital Education, Moscow City University; advisor to the HSE Tikhonov Moscow Institute of Electronics and Mathematics, Moscow, Russia,
burovvv@mgpu.ru, https://orcid.org/0000-0001-5571-5693

Dmitry V. Soshnikov — PhD in Physics and Mathematics, Associate Professor at the Moscow Aviation Institute; Associate Professor of the Department of Software Engineering, Faculty of Computer Science, Higher School of Economics, Moscow, Russia,
dmitri@soshnikov.com, https://orcid.org/0000-0003-1021-091X

Все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

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