ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

UDC 004.8

doi: 10.32620/reks.2025.2.02

Oleksandr PROKHOROV[1], Dmytro SHYMKO[2], Olena KUZMINSKA[3,4], Andrey CHUKHRAY[1], Oleksii SHATALOV[5], Oleksandr KHOLODNIAK[1]

[1] *National Aerospace University «Kharkiv Aviation Institute», Kharkiv, Ukraine*
[2] *Kharkiv National University of Radioelectronics, Kharkiv, Ukraine*
[3] *National University of Life and Environmental Sciences of Ukraine, Kyiv, Ukraine*
[4] *Kryvyi Rih State Pedagogical University, Kryvyi Rih, Ukraine*
[5] *Sorbonne University, Paris, France*

A SYSTEM FOR GENERATING CHATBOTS TO SUPPORT LEARNING IN THE FIELD OF EXACT SCIENCES USING GENERATIVE ARTIFICIAL INTELLIGENCE MODELS

The integration of generative artificial intelligence (AI) into education, especially for teaching exact sciences, represents an innovative opportunity to enhance student engagement and understanding. Chatbots such as ChatGPT can promote interactive learning, allowing students to explore complex scientific concepts through personalized support and real-time feedback. This approach not only transforms traditional pedagogical methods but also fosters deep curiosity and understanding among students. ***This study examines*** *the task of increasing the degree of automation in creating AI-powered chatbots and their integration into the learning process for exact sciences, particularly mathematics, for school students in online educational settings.* ***The relevance of the research*** *is driven by the need to improve the success rate of educational activities under wartime conditions, eliminate knowledge gaps, bridge disparities in knowledge and skills among school students in exact sciences when preparing for further studies in higher education institutions, enhance the efficiency of independent learning in online educational settings, and stabilize the socio-emotional state of children.* ***The purpose of this study was*** *to develop a web platform for generating various types of chatbots using artificial intelligence models to improve the quality of school students' preparation in exact sciences in online educational settings.* ***The objectives include****: analyzing the challenges and peculiarities of creating chatbots and preparing high-quality datasets; developing the structure and describing the functionality of the chatbot generation system; providing examples of creating various types of chatbots; conducting experiments to determine the effectiveness and cost of chatbot generation. The following results were obtained: A method to optimize the formation of datasets for chatbots was developed. A web platform for generating various types of chatbots was created, including assistant bots, bots for generating variants of mathematical problems and step-by-step explanations of their solutions, and testing bots using artificial intelligence models to improve the quality of school students' preparation in exact sciences in online educational settings.* ***Conclusions****. The scientific novelty of the research is associated with improving the method of creating structured chatbot datasets while maintaining the thematic integrity of text and context, ensuring a more accurate selection of relevant information by chatbots for responding to user queries. The effectiveness of the proposed approach is illustrated through examples of creating an assistant bot, a bot for generating variants of mathematical problems and step-by-step explanations of their solutions, and testing bots, which demonstrated cost optimization and more efficient resource utilization. A significant advantage is the convenience of tools for creating and configuring chatbots, as well as their use through a showcase of ready-made chatbots in messengers.*

Keywords: *chatbots; education; mathematics; datasets; web platform; chatbot generation system.*

1. Introduction

The development of artificial intelligence technologies has had a profound impact on various spheres of life, including reshaping models and methods of education [1, 2] in order to reorient modern students to work that is personally meaningful and that can only be performed by humans [3]. The careful integration of AI technologies and tools has the potential to improve the quality of educational environments in academic institutions [4] overall. The use of GAI (Generative Artificial Intelligence) systems has already demonstrated a positive effect on student performance [5], support for inclusive education [6], personalized learning, and foster student autonomy [7]. The latter is achieved primarily using personalized online assistants in the form of AI-

based chatbots [8, 9].

Moreover, the creation and implementation of such chatbots [10] mitigates educational losses in learning deficits caused by the absence of or limited access to education, which became a global concern during the COVID-19 pandemic and the widespread adoption of distance learning.

AI-powered chatbots can learn and adapt to the needs of educational programs, instructors, and students, thus enabling flexibility in educational processes and tailoring to specific learning contexts or areas of study.

These general trends in the use of AI technologies in education are particularly significant for Ukraine to ensure quality educational services during wartime conditions.

Challenges such as forced interruptions in education, destroyed educational institutions, the inability to conduct in-person learning, limitations in remote and hybrid formats, and disruptions in electricity and communication have been the main obstacles facing Ukrainian education in recent years. Such challenges in organizing learning during martial law create gaps in students' knowledge and skills, which are especially evident in the exact sciences.

Monitoring studies by the Ukrainian Center for Educational Quality Assessment [11] and international PISA evaluations [12, 13] have shown a long-term trend of declining performance in national multi-subject tests in exact sciences. Today, there is also a growing demand for interactive learning: students seek accessible and convenient ways to study complex subjects such as mathematics and physics, while teachers need effective tools to engage students and automate routine tasks.

1.1. Motivation

These challenges necessitate the search for new approaches and solutions based on artificial intelligence methods that generalize the experience of computerized learning of core competencies in exact sciences and solving algorithmic problems.

One such approach is the development of personalized online assistants in the form of AI-powered chatbots. Among the advantages of such chatbots is their ability to learn and adapt: leveraging artificial intelligence, chatbots can update their knowledge and skills based on new data, which makes them flexible and ready to adjust content and application conditions.

In our view, a partial solution to these significant problems and challenges could be the creation of a platform encompassing functionality for creating and fine-tuning chatbots (of various types [8] and themes) to provide individualized learning support in a convenient format, such as a website or messenger.

Considering the Ukrainian context, our research tested the developed platform for teaching mathematics; however, the proposed solutions are universal and can be applied to other subjects.

1.2. State of the art and problem statement

Considering the use and integration of AI tools in business processes, the following points should be noted: specific knowledge is required for interaction with API services, messengers, databases, interfaces, and more. Experience in formulating accurate and comprehensive bot prompts is crucial because each generation of a specific type may require an option that is unsuitable for another task. In addition, fine-tuning options are often unavailable to ordinary users but are accessible via APIs (Table 1).

Since the release of ChatGPT and the opening of the OpenAI API to the public, numerous services that partially address these issues have emerged. However, most of them are highly specialized. For instance, some services are focused specifically on creating AI assistants in chat form, while others are designed for generating characters (real or fictional) for entertainment or educational purposes [8].

Given that developing a chatbot as a modern AI assistant often requires significant resources, it is advisable to analyze the "Language Model for Development of Interactive Chat Bot for SaaS Solutions" [14]. Here, we discuss some commonly used chatbot creation services.

ChatGPT + Custom GPTs. The basic version of ChatGPT offers the minimal functionality required to send instructions to the bot, which is advantageous for ordinary users with limited configuration needs who can express their requirements in textual form. In the context of this research, attention should be paid to the analysis of OpenAI's GPTs marketplace for individual LLMs [15] and an example of using these tools to create conversational agents to enhance learning quality within the Advanced Placement Computer Science Principles (CSP) course in U.S. high schools [16]. This service operates on a subscription-based business model; thus, to use the more advanced ChatGPT-4 model, users must purchase a subscription, which provides additional features, such as image processing and generation based on textual prompts within the chat. Subscribers can also create custom bots ("GPTs") trained on diverse types of documents, and most of the configuration process is performed directly within the chat, thereby minimizing technical complexity. The primary drawback of using the official service is that a subscription is required not only by the creator of the bot but also by its users. Moreover, as noted in studies conducted by Harvard University representatives [17], the popularization of

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

this service raises concerns about user privacy and security.

Dropchat.co. This service allows users to upload a document or enter a website URL, after which a chat is generated that can respond to questions based on the provided materials. Overall, the service is useful for processing large volumes of content and retrieving specific relevant information through a chat format. However, the service does not address most of the issues associated with working with generative AI-based services and lacks the advantages of being used through messengers.

Quickchat.ai. This system allows the creation of a bot by providing a brief behavioral instruction in a text form or attaching additional data in the form of documents or web pages. However, the service does not allow users to preview the chatbot as a tool before its creation in their account, targeting a B2B model, where sharing a bot can only be performed manually. Considering the subscription cost, using this service for purposes other than automating business communication is impractical because there are cheaper alternatives. Nonetheless, its advantages include accurate translation and a user-friendly interface, which promote intercultural exchange [18].

chatbotkit.com. This service offers extensive capabilities for providing contextual data to the bot and configuring datasets. However, when creating datasets, the service not only fragments sentences but also truncates words, which significantly impacts the dataset quality. In addition, text fragments are limited to 260 tokens, which often prevents comprehensive topic representation; thus, only a few sentences are retained. The Xatkit experience is a model framework for chatbot development is detailed in [19]. In [20], the Xatkit structure was described as a means of providing a set of domain languages for defining chatbots (including voice bots and general bots) in a platform-independent manner.

character.ai. This service allows users to create any character for interaction and provides a public library of bots ranked based on the number of messages generated during interactions with other users. The proposed service provides the capability to immediately select the desired interaction instruction with a bot and start using it without the need for configuration, dataset provision, or additional setup. Research by Swiss scholars [21] has confirmed that the use of characters enhances individual client experiences through personalization.

Some of the services reviewed opt to split text into limited fragments based on a specific number of characters, leading to a lack of full context when processing user queries. This not only limits the bot's ability to provide adequate responses but also complicates user interaction for those seeking comprehensive and detailed answers.

Splitting text, even in the mid-sentence, results in parts of the context being in one data source and another part being in a different source. This leads to the loss of information from the sentence as a whole and may result in incorrect responses generated from fragmented information.

On the other hand, some services offer high-performance models by using expensive technological solutions. However, their costs are justified only in specific cases, and for many users who do not require specialized computational power, this becomes a financially inefficient solution.

Table 1

Comparison of Chatbot Creation Services

Functionality	dropchat.co	quickchat.ai	chatbotkit.com	character.ai	ChatGPT+ GPTs
Analysis of created datasets	Yes	No	Yes	No	No
Splitting text by topic, not just by a certain number of characters	No	?	?	No	?
Adding additional context for datasets	No	?	No	No	?
Availability of an open bot library	No	No	No	Yes	Yes
Message limit per month	6,000 messages	Unlimited	5k	?	40 messages in 3 hours
Minimum monthly cost	$24	$99	$65	$10	$20 + $20

The shortcomings of the analyzed systems can be summarized as follows:

- excessive costs for basic use and even higher costs for additional features;
- lack of clarity regarding the final output of the created bot;
- absence of search functionality among created bots;
- inability to analyze the generated dataset and its components, or splitting by a specific number of characters, which negatively impacts data quality;
- most systems are heavily focused on creating AI assistants for business purposes, particularly to enhance team productivity in software engineering [22], improve reporting in accounting and consulting activities [23], and enhance customer interactions in customer relationship management (CRM) [24].

From the perspective of using such services, there are drawbacks in the form of the need to configure the required service, which takes time and still requires certain skills and an understanding of the final goal. In addition, there are limitations in usage, such as difficulty in creating a specific tool, configuring it, and presenting it as a ready-made chatbot for another person to use intuitively. For this reason, some educational institutions develop their own systems based on chatbots. For instance, researchers at Queen Mary University of London developed a system called Q-Module-Bot, which is accessible to both technical and non-technical educators to leverage the benefits of e-learning [10]. Similarly, researchers in India created dynamic assistants using the Rasa open-source platform [9].

Regarding the thematic focus (in our study, we focus on mathematics education), the following analyses are noteworthy: recommendations for building AI bot scenarios for teaching mathematics in high schools [25]; experience in creating chatbots for adaptive mathematics learning [26]; testing chatbots on math and logic problems [27]; teachers' attitudes toward integrating AI-based chatbots into mathematics instruction [28]; the application of AI-based chatbots to enhance middle school students' mathematical thinking skills [29]; the use of ChatGPT versions in mathematics education to obtain basic information and supervised help [30]; and the advantages and limitations of applying GAI models in mathematics education for university students [31].

1.3. Objectives and methodology

The goal of this study is to develop a web platform for generating various types of chatbots using artificial intelligence models to improve the quality of school students' preparation in exact sciences within online educational environments. The reasons justifying our research stem from several problems (Fig. 1), such as: A decline in statistical indicators for completing tasks in the National Multidisciplinary Test (NMT); challenges in teaching students due to differing regional conditions caused by military aggression; a decrease in the percentage of participants prepared for higher education institutions; The Ministry of Education's requirement to increase the proportion of tasks related to real-life situations and tasks illustrated with graphs, tables, and diagrams. Therefore, the requirements we aim to meet with our solutions are as follows: improving the efficiency of independent learning in online educational conditions; enhancing the understanding of basic mathematical competencies; ensuring high adaptability so that chatbots can be used for various exact sciences and different educational levels; ensuring the personalization of the learning process; and providing accessibility on mobile devices within online educational conditions.

Our web platform will focus on creating various types of chatbots (Fig. 2):

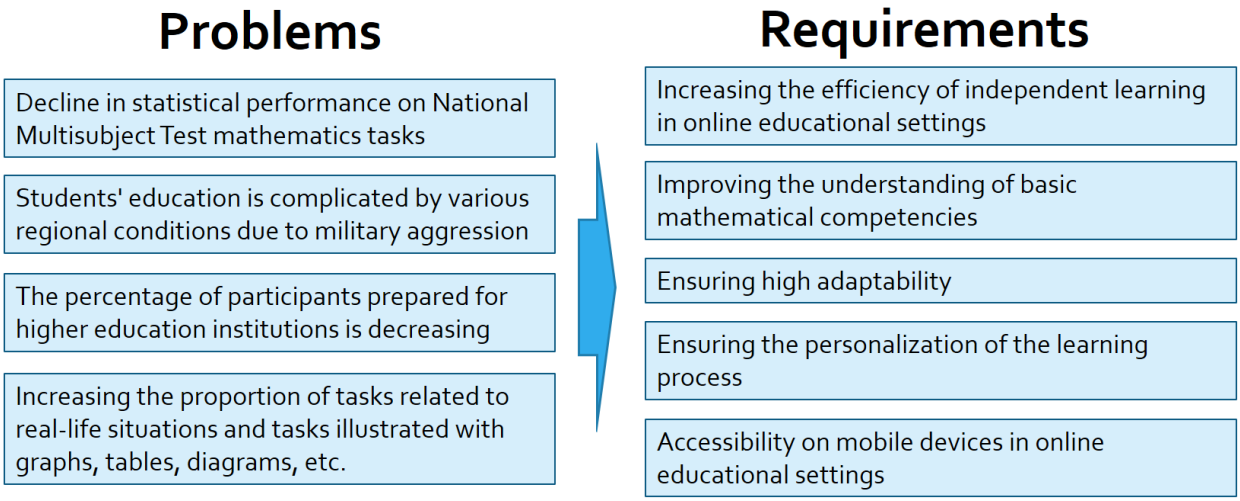

Fig. 1. Problems and Requirements for Preparing Students for Exact Sciences in Online Learning Environments

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

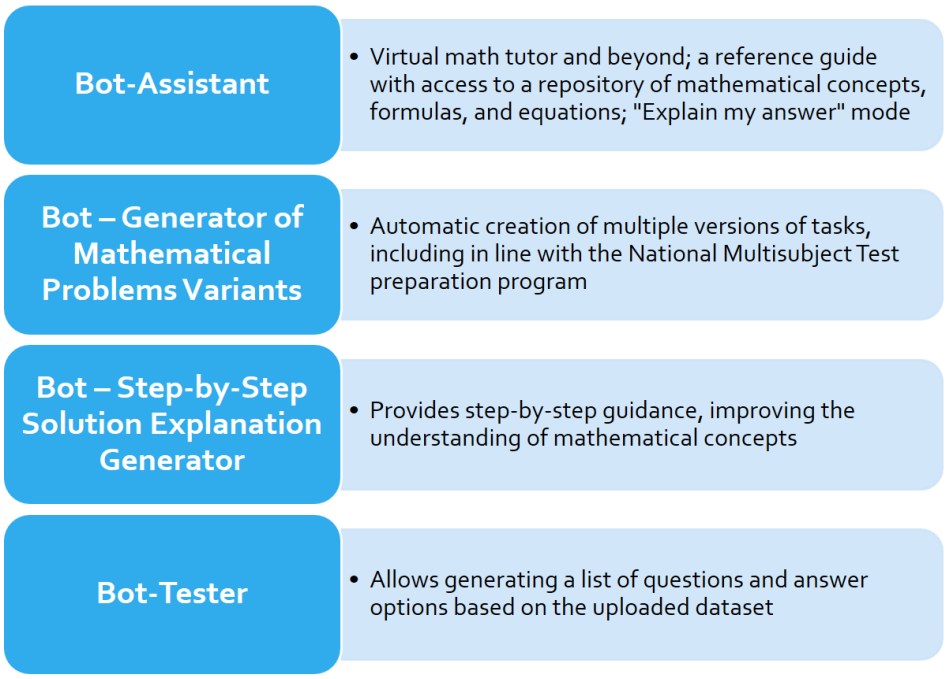

Fig. 2. List of bots created by the proposed web platform

- assistant bots, which act as virtual instructors in mathematics and other exact sciences, provide 24/7 access to an information repository of mathematical concepts with explanations of formulas and equations, and can work in modes like "Explain my answer," among others;
- task generator bots, capable of automatically creating multi-variant tasks for teachers' daily work, students' independent practice, and skill reinforcement, as well as generating new tasks in training mode according to the NMT preparation program;
- step-by-step solution generator bots provide step-by-step instructions to enhance the understanding of mathematical concepts;
- testing bots, which allow for quick and easy generation of question lists and answer options based on an uploaded dataset or simply a list of topics, and are also ready for students to use in messengers.

In accordance with the research goal, the following tasks were addressed:

1. Analyze the challenges and specifics of creating chatbots and preparing high-quality datasets.
2. Develop the structure and functionality of the chatbot generation system.
3. Provide examples of creating different types of chatbots.
4. We conducted experiments to determine the effectiveness and cost of chatbot generation.

This paper has the following structure. Section 2 describes the optimization method for forming high-quality chatbot datasets. Section 3 describes the structure and stages of the system's operation. In Section 4, we present examples and experiments to generate chatbots for various purposes. Section 5 discusses the findings and evaluates the effectiveness and cost of the developed solutions. Section 6 summarizes the key contributions and outlines potential directions for future research.

2. Methods for creating high-quality chatbot datasets

If the task involves creating chatbots trained on specialized data unavailable in standard ChatGPT models [32, 33], or if there is a need to provide clarification or data that should trigger unambiguous bot responses (with a volume reaching thousands of characters), this can significantly increase the cost of responses.

The simplest method to load a dataset is to enter data or provide a text file that is then processed automatically.

Requiring users to input critical data manually for each point of interest provides the advantage of producing precise and accurate bot responses. However, this would require significantly more time to process data into the correct format, which is unsuitable for most scenarios where a chatbot must quickly process and understand the material provided.

2.1. Existing methods to process large textual datasets

To provide a substantial amount of data for a bot, several data processing methods are available:

- additional data processing. This involves creating points that reflect the most relevant and concise information in the provided text. Additional text processing can reformat data into a specific structure while preserving the core context and supplementing the original dataset. The downside is the added time and resource demands, leading to increased costs;

- fine-tuning. This method allows us to modify the context of a bot's responses, tone, and quality of its messages. However, a universal approach is needed for possible user instructions to the bot, and fine-tuning would require a specific dataset to train the model;

- embeddings. This appears to be the most suitable option because it accounts for thousands of descriptions across various categories and effectively differentiates texts. However, experiments have shown that this method performs poorly on large text volumes (over 500 characters) and is unable to distinguish between texts with similar topics or individual sentences;

- manual search within text segments. By splitting the input query into words or their roots, a search can be conducted within the text to identify segments with the highest matches. These segments are then sent to the API to generate a bot response.

An analysis of these methods suggests that embeddings are an effective auxiliary tool, but they are best combined with additional text processing for structured context generation. Since the primary function of the project involves creating an assistant bot, the format for contextualization will be designed as a question-and-answer structure, allowing users to query the text.

2.2. Splitting text datasets into thematic sections

A critical step in chatbot development is splitting input text into thematic blocks. This process uses advanced natural language processing methods, such as clustering algorithms and topic modeling, to ensure effective text segmentation, key themes, and thematic groups.

The ChatGPT model, capable of recognizing and classifying text, is used to determine the boundaries of thematic blocks. Each block represents a specific aspect or topic of the original text. The process is based on sentence structure and specific keywords, ensuring precise thematic division.

The result of this step is structured text, where each block corresponds to a specific topic. This structure simplifies subsequent analysis and processing, thereby providing a foundation for improving interaction quality and response accuracy in future chatbot development stages.

Thematic text division is essential for maintaining the context and information structure, which is critical for chatbots to function effectively. Simply splitting text by character count can create information asymmetry, thereby degrading text comprehension and interaction quality.

Thematic division preserves logical information blocks, which improves the bot's understanding and response quality. It helps identify key concepts and ideas within each text segment and structure information for more convenient further analysis.

This approach is particularly significant when text is split mid-topic, leading to context loss and disrupted idea continuity, which are essential for accurate and clear responses. Thematic division avoids such issues and ensures information integrity, crucial for successful chatbot-user interactions.

To ensure effective topic separation, the text should be split at points at which new themes or logical contexts begin. This improves the dataset usability and makes it easier to search for relevant data to provide contextual responses.

The query formats of ChatGPT can vary. Users can request that the text be divided into specific formats or provide text sections indicating where new topics are beginning.

After receiving the results, the output must be reviewed for accuracy and adequacy. The task is marked as complete or sent for further processing. Based on experiments, 2–3 attempts with well-formed instructions are typically sufficient. Increasing the number of attempts to four is reasonable; afterward, the text should be split not arbitrarily but at sentence boundaries.

2.3. Creating additional individual questions and answers for text sections

Embeddings generated for text may not sufficiently reflect its essence when dealing with large volumes or unevenly structured text. Subsequent searches using the generated embeddings may yield results with minimal differences, thereby complicating the selection of the most relevant text segment.

To address this issue, additional descriptions in the form of questions and answers are proposed for each text section. The chatbot is being developed; thus, this format is most appropriate for user interaction, which often involves asking questions and receiving answers.

In addition, providing individual descriptions of text sections in the form of questions and answers carries additional value because this approach avoids the need to send the full text in the context of a chatbot query. Instead, only the question and its associated answer are used, which helps reduce API interaction costs and improve query processing speed.

In the modern context of natural language processing, creating embeddings for textual data can face

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

challenges, especially when the text is complex and contains a large volume of information. The generated embeddings may capture the general context but often fail to account for subtle nuances and the diversity of topics present in the text.

In cases in which the text is divided into different thematic blocks, the use of embeddings may lead to insufficient differentiation between them. Search results based on the created embeddings may differ only slightly, which makes it difficult to select the most appropriate text segment for a specific query.

To improve this process, it is proposed to provide an additional description of each text section in the form of questions and answers. This approach is particularly appropriate when developing chatbots because they interact with users primarily by asking questions and providing answers.

Creating individual questions and answers for each text section helps preserve the context and structure of the information, thereby ensuring more precise selection of relevant information during searches. In addition, this approach facilitates effective interaction with the chatbot, allowing users to send only questions and receive specific answers. This simplifies API usage and optimizes interaction cost.

2.4. Proposed approach to create a high-quality chatbot dataset

The justified conclusions of the analysis highlight the importance of cost optimization and ensuring more efficient resource utilization for chatbot development.

Thus, a combination of methods was chosen: splitting datasets into logical thematic sections, generating additional context by posing questions about the topics, and creating embeddings to identify similar text for chatbot response generation (Fig. 3).

After dividing the text into sections, each section is queried sequentially with ChatGPT to obtain a list of questions that users would most likely ask about the text. Based on the experiments, it was determined that using one question for every 150–200 characters in the text (approximately 1–2 sentences) was optimal.

If too many questions are requested, the answers may include repetitions, or the number of generated questions may significantly decrease when ChatGPT cannot create many questions for homogeneous text.

If the goal is not to create an assistant bot but rather a tester bot that quizzes the user and checks their answers, the query requests not only one correct answer but four options, with only one being correct.

Subsequently, when presenting the question, the system shuffles the answers to make the question-and-answer format appear even more randomized.

3. Structure and stages of system operation

3.1. Structure of the chatbot generation system

The following technologies were used to develop the web platform for chatbot generation: frontend – JavaScript + AJAX, backend – PHP, database – MySQL. External libraries used include: Smalot\PdfParser for decoding PDF files into text; danny50610/bpe-tokeniser for counting tokens in the text; Highcharts for displaying charts.

To create the described services, the system needs to be divided into specific modular components, each performing its own tasks (Fig. 4).

Since the project allows material uploads via URL or links to PDF files, the data may contain a significant amount of extraneous material, such as HTML tags, system symbols, page numbering, and other irrelevant elements.

To ensure a high-quality dataset, only useful text that is free of unnecessary symbols and informational noise should be obtained. At the same time, it is important to maintain a balance by retaining data that may be necessary when uploading datasets.

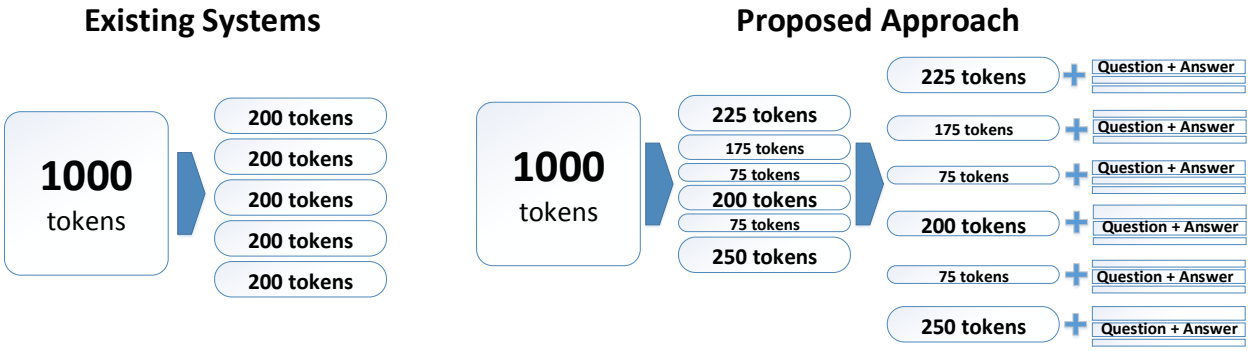

Fig. 3. Proposed Approach to Create a High-Quality Chatbot Dataset

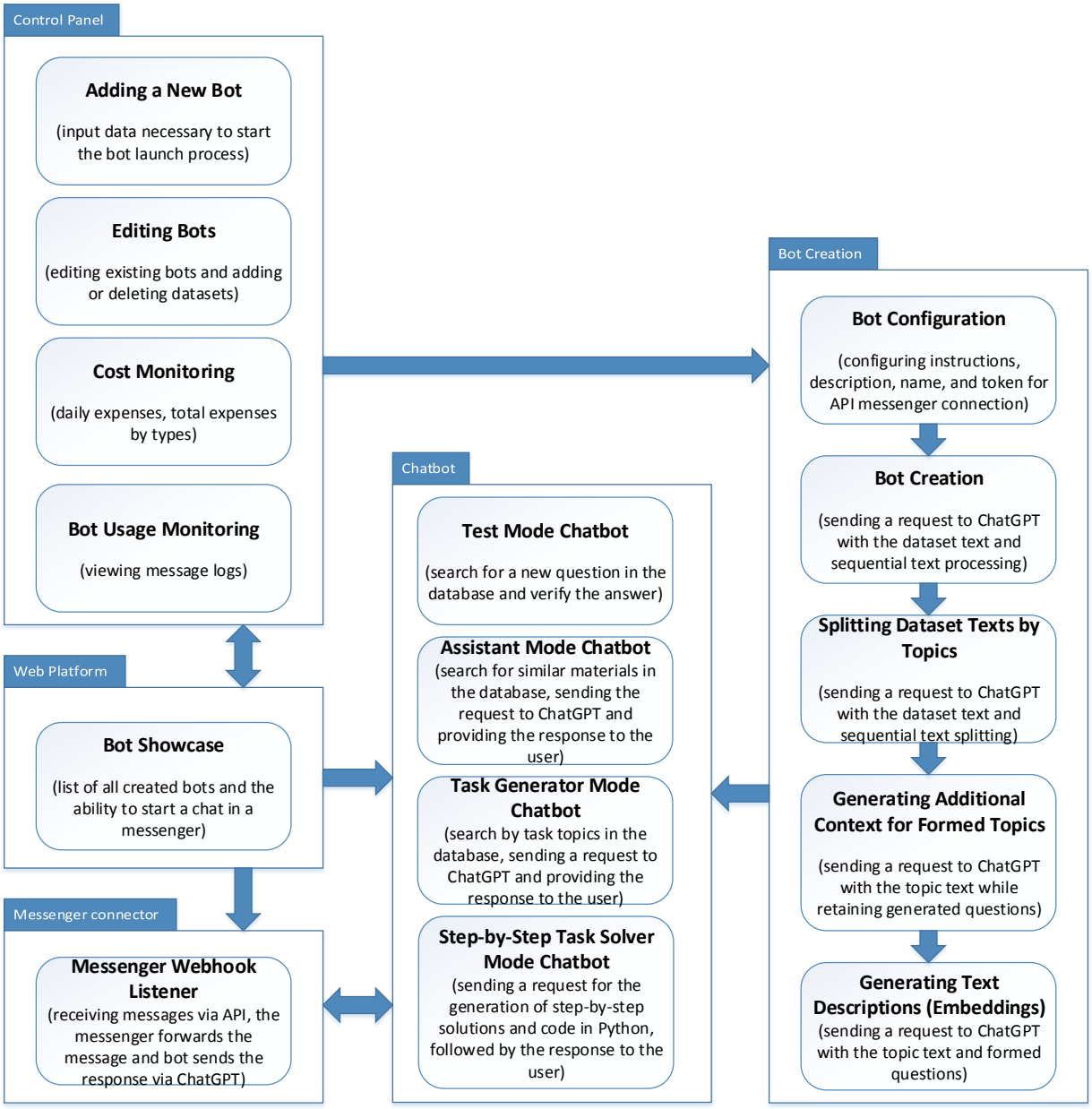

Fig. 4. Structure of the Chatbot Generation System

Therefore, users will be provided with the ability to edit the dataset after uploading it from an external resource.

This allows them to independently remove obvious extraneous parts of the text that are irrelevant to the materials intended for chatbot responses. Preliminary formatting is necessary to improve the text structure during subsequent dataset formation because there may be line breaks or text placed in separate blocks with tags such as “span,” among others.

3.2. Web interface of the chatbot generation system

The web interface is designed to allow users to browse and select created chatbots for further interaction. It serves as a showcase (Fig. 5), displaying a list of available bots along with their descriptions and functionality.

The control panel of the system for creating and managing chatbots is an interface that includes a set of features that are intended to provide convenient and efficient management of chatbot creation, configuration, and analysis processes.

This interface is designed to simplify routine tasks related to administration and optimize the functioning of chatbots (Fig. 6).

A key function of the control panel is the ability to create a new chatbot. In this window, the user can enter basic information about the bot, which allows for the creation of datasets and API configuration for its opera-

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

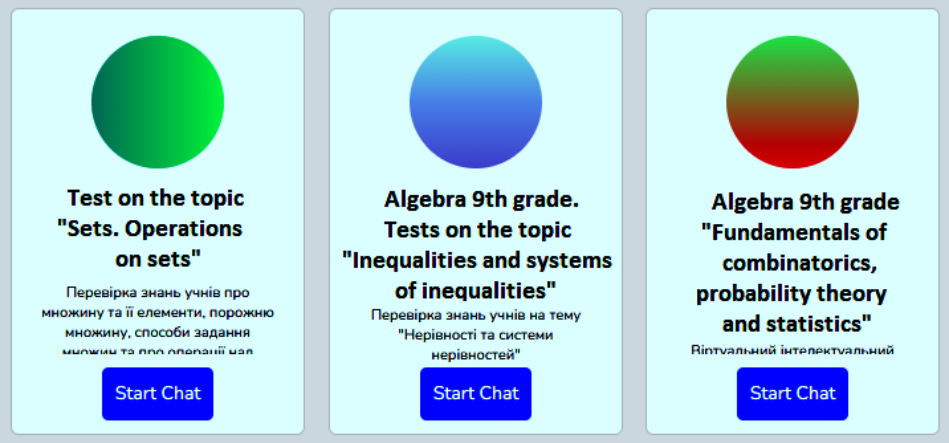

Fig. 5. Chatbot Showcase Fragment

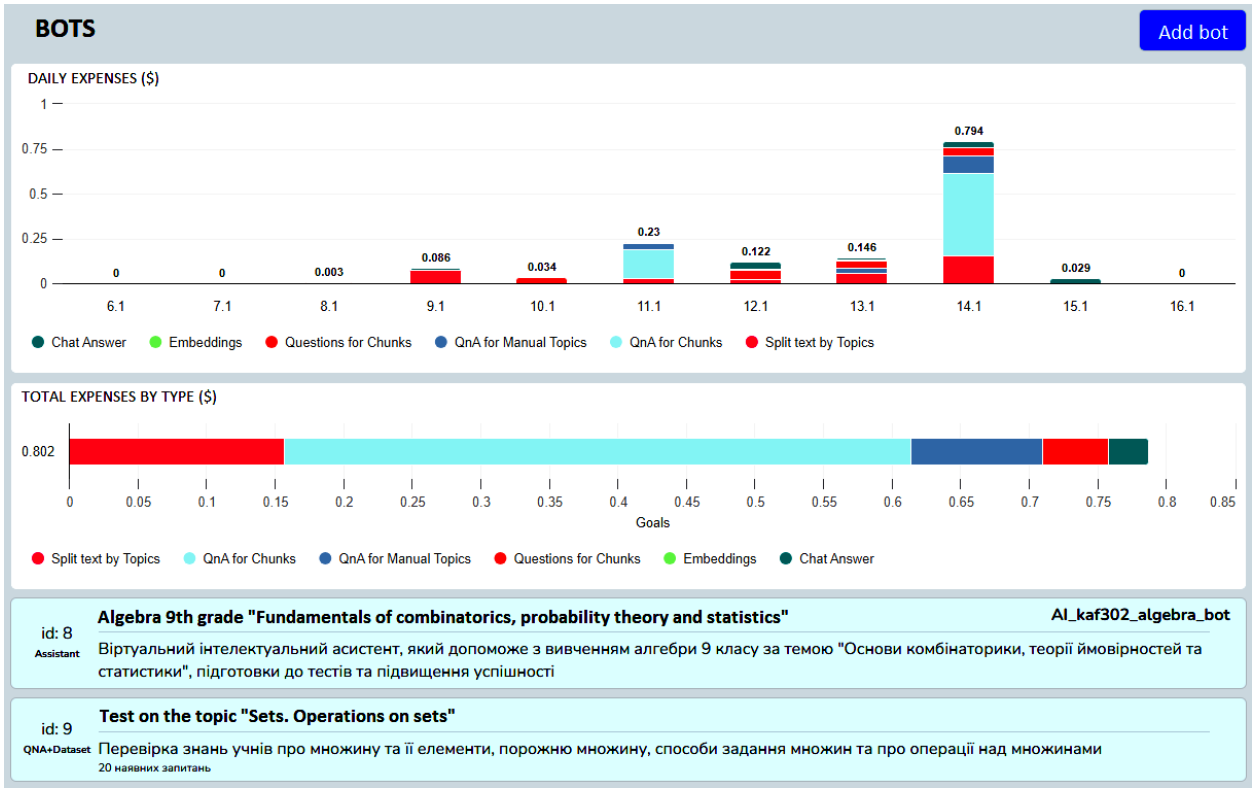

Fig. 6. Control Panel for Creating and Managing Chatbots

tion in the chat. This feature enables efficient setup of new bots and subsequent work with them.

The second important function is a module to review the created chatbots, which allows a detailed analysis of each one. This feature provides users with access to specific information about each bot and its settings.

The third function involves general settings for created bots, which allows the user to modify and maintain the chatbot operating parameters up to date.

A dedicated module is included in the system to provide the ability to upload datasets and view their details. This feature allows users to control and analyze datasets, which is a key aspect of chatbot creation.

Another component of the control panel is usage analysis functionality with the ability to generate graphs illustrating API usage costs. This feature keeps users informed about API-related expenses and helps optimize resource usage. The analytical graphs on the main page display the total API usage costs for the existing bots. Categorizing graphs by expense type and date allows users to identify the most popular usage types and days with the highest number of user interactions with bots.

A final but equally important feature is the ability to view user interactions logs. This feature enables the analysis of chat histories between users and bots, which

is a critical aspect of understanding user interactions with chatbots. The proposed control panel for creating and configuring chatbots offers several advantages that enhance the convenience and efficiency of bot management. One of the key benefits of this system is the centralized management of all aspects of chatbot creation, configuration, and operation. This reduces the time required to administer each bot individually and simplifies the interaction process.

4. Examples and experiments in chatbot generation

4.1. Example of configuring a bot in assistant mode

In the first operational mode, which was implemented as a classic teaching assistant, the user interacts with the bot by posing questions. The bot, in turn, uses models based on an extensive set of datasets to generate responses. This approach allows the bot to answer several questions based on the enriched knowledge obtained from the dataset.

Interaction occurs in real time, where the user inputs their questions the system generates appropriate responses based on the training data.

The tool created in the classic assistant mode stands out for its flexibility and versatility. It can answer a broad spectrum of questions and perform tasks in various fields by leveraging diverse datasets.

This makes it an effective tool for both obtaining information and completing various practical tasks. In addition to responding to specific questions, the bot is provided with a history of previous messages to ensure a more comprehensive interaction context. This enhances the bot's ability to understand and adapt to individual user needs, thereby helping to create more personalized and effective responses. In the same mode, the bot can also operate under specific instructions that define its behavior in different scenarios.

This functionality is essential for ensuring consistency and control in user interactions. Instructions provide the bot with structured directives, which allows it to respond appropriately to specific situations and avoids misunderstandings.

When creating a bot in its basic settings, general information about the bot can be configured, which is displayed to users in the web interface (Fig. 7). In addition to the bot's name and textual description, the basic settings of the assistant bot allow for the addition of a technical instruction to help it respond more effectively to questions and better understand the context of the data from the dataset.

The following instructions also enable the configuration of additional rules for specific interaction scenarios, such as cases where data from the dataset are missing or when there is a need to avoid answering particular questions. They can also specify a predefined response to specific inquiries.

A well-crafted context compensates for the shortcomings in the quality and quantity of data in the dataset and allows for the creation of behavioral rules using simple text-based queries.

This eliminates the need for code editing or development skills.

To provide relevant answers and understand the specifics of its functionality, the bot must have high-quality textual data to generate responses. Since our system can form datasets from external sources, users can upload links to ready-made resources (web pages or full-text PDF files) from which the necessary textual information can be extracted.

Alternatively, users can provide information as plain text copied from other pages or textbooks (Fig. 8).

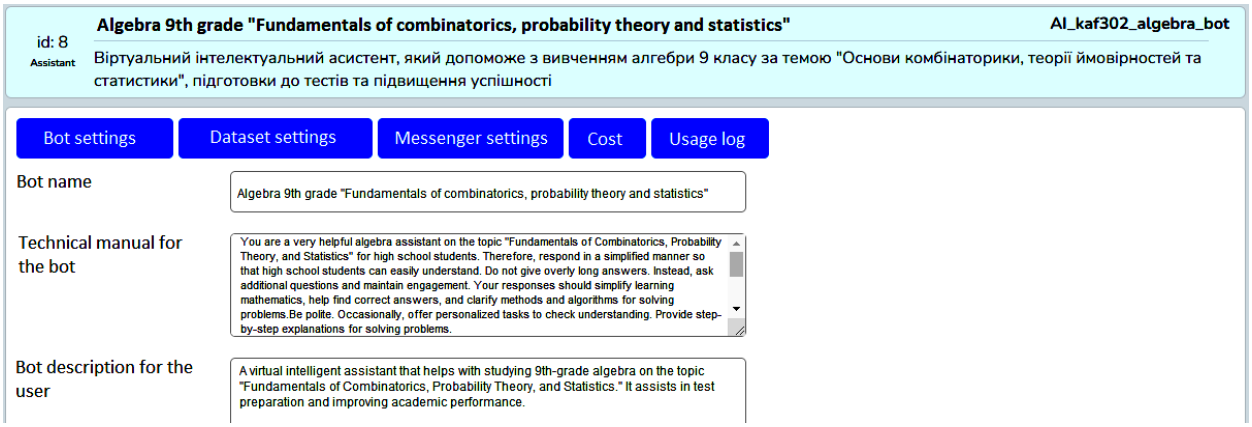

Fig. 7. Basic Settings of the Assistant Bot for the Topic "Fundamentals of Combinatorics, Probability Theory, and Statistics"

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

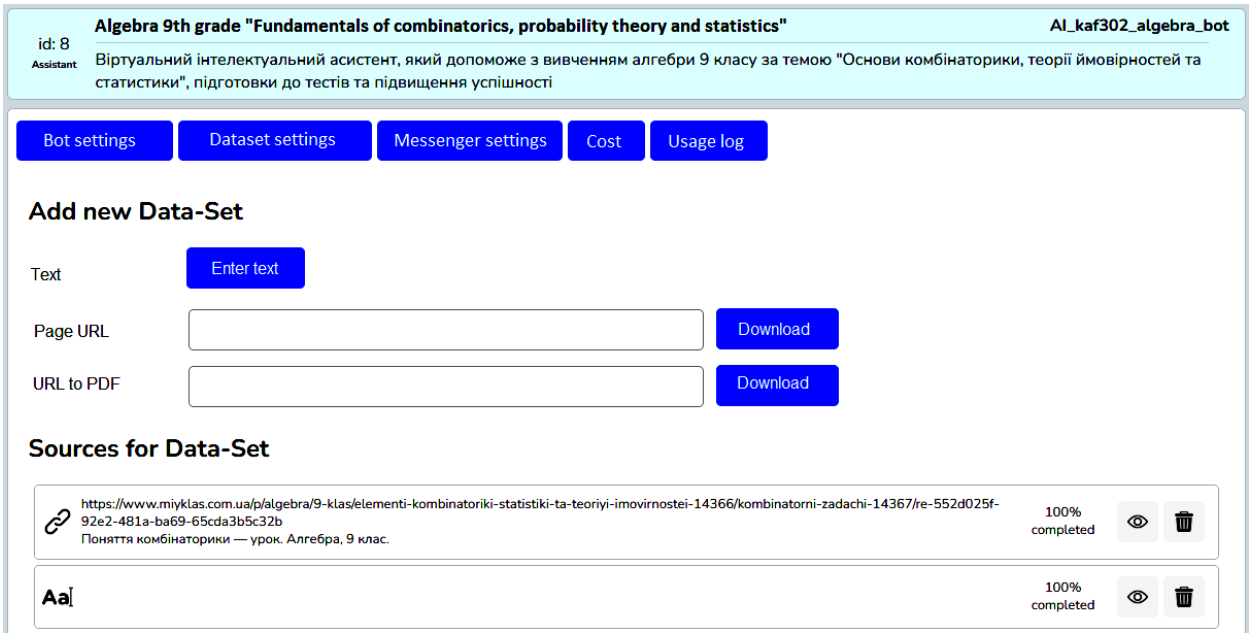

Fig. 8. Dataset Configuration for the Assistant Bot

To verify that the text has been properly divided, our system, unlike some similar services, provides the ability to view the created sections. This can be achieved simply by clicking the corresponding button next to the desired dataset. After pressing the button, the results of the selected dataset division into sections can be viewed (Fig. 9). The available examples clearly demonstrate the quality of text segmentation, forming the final dataset that is used to generate responses to queries. It can be observed that the questions accurately reflect the essence of the text and further divide it into smaller parts. This approach reduces the cost of bot responses when delivering materials from datasets while maintaining the quality and completeness of the answers.

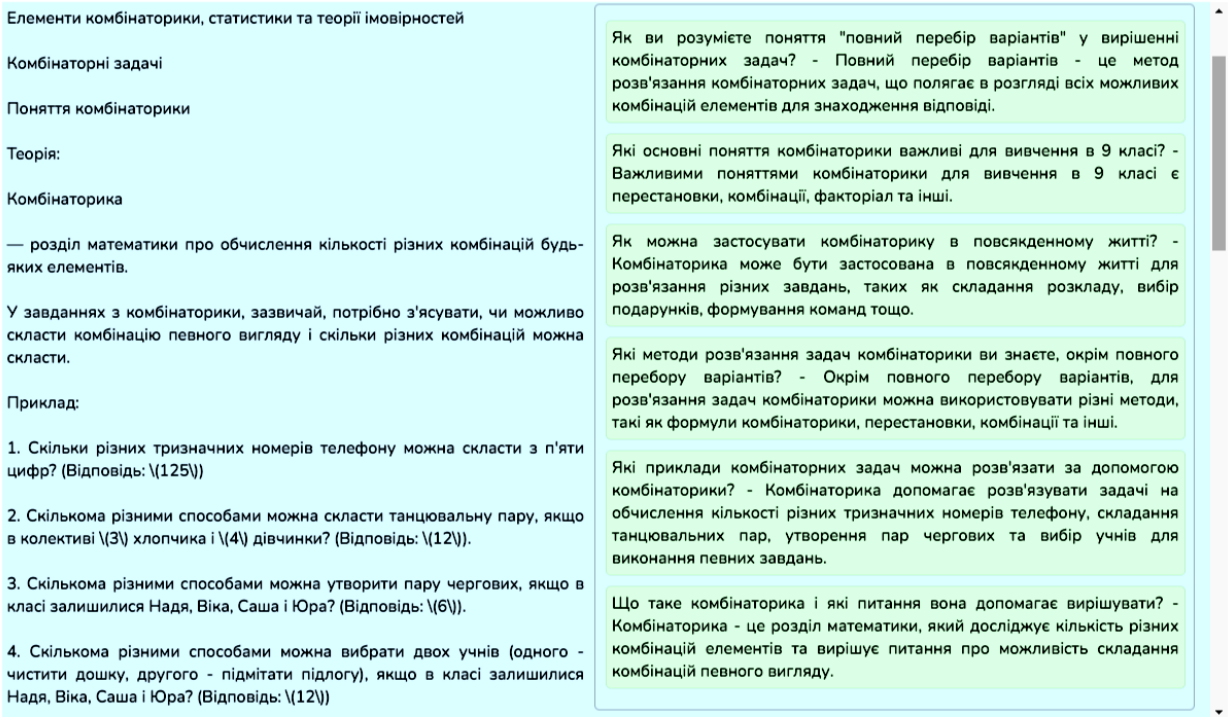

Fig. 9. Result of Web Page Processing and Automatic Dataset Formation

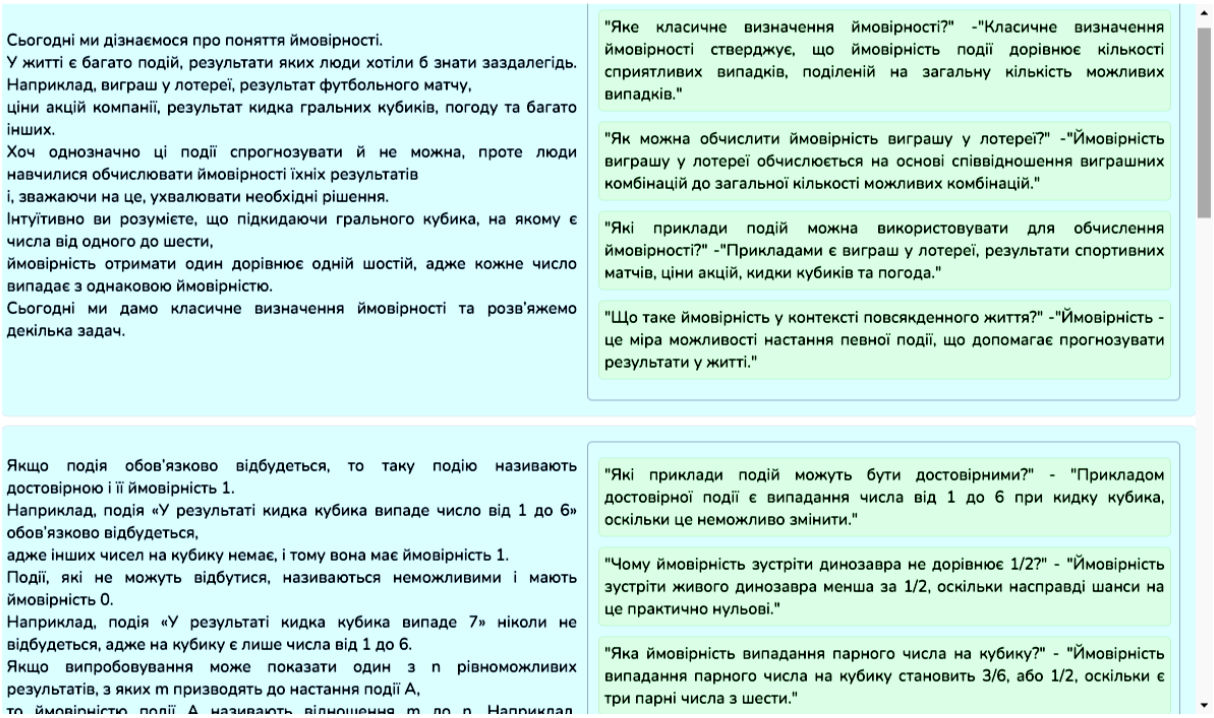

Fig. 10. Result of Dataset Segmentation into Sections and Generated Questions

Let us try asking the assistant chatbot a few questions (Fig. 11).

It can be observed that the bot performs excellently when generating responses to user queries. The bot was provided with no less than 1,000 tokens of material for generating answers, along with the complete history of previous messages. This helps the bot provide more accurate and user-specific responses.

The bot offers step-by-step solutions to problems, generates tasks for independent practice, identifies problematic areas, and provides recommendations (Fig. 12).

The total cost of using each bot is displayed in two graphs (Fig. 13):

- daily expenses – this allows analysis of the bot's usage daily. The division by days helps identify potential anomalies and monitor the bot's popularity;

- total expenses by type – categorizing expenses into groups allows examination of which specific operations with the bot incur the highest costs and how future optimization can be achieved.

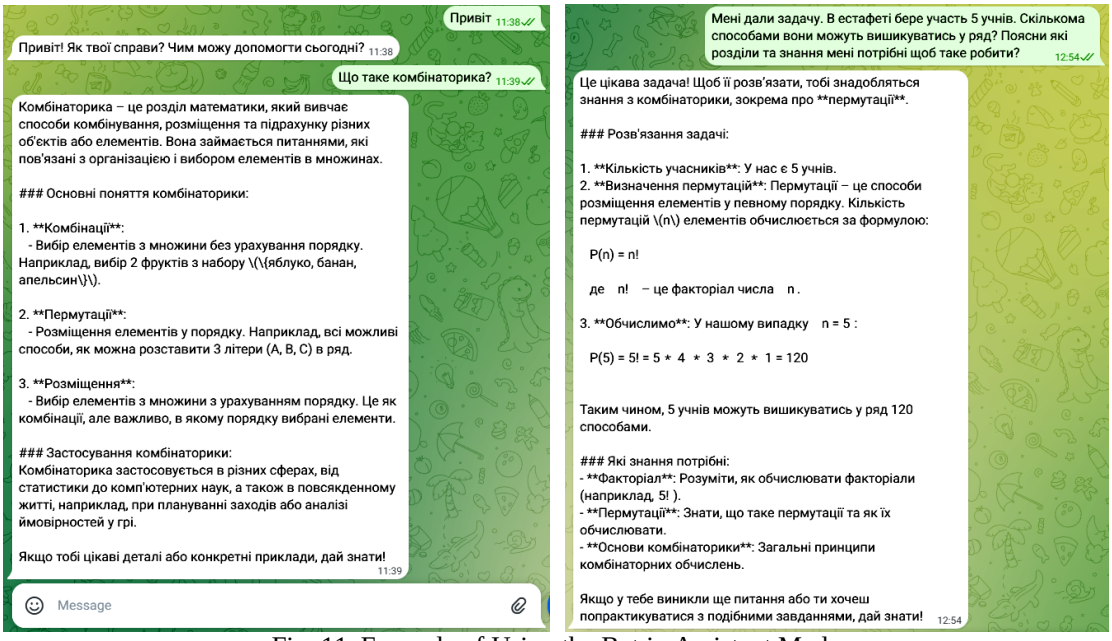

Fig. 11. Example of Using the Bot in Assistant Mode

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

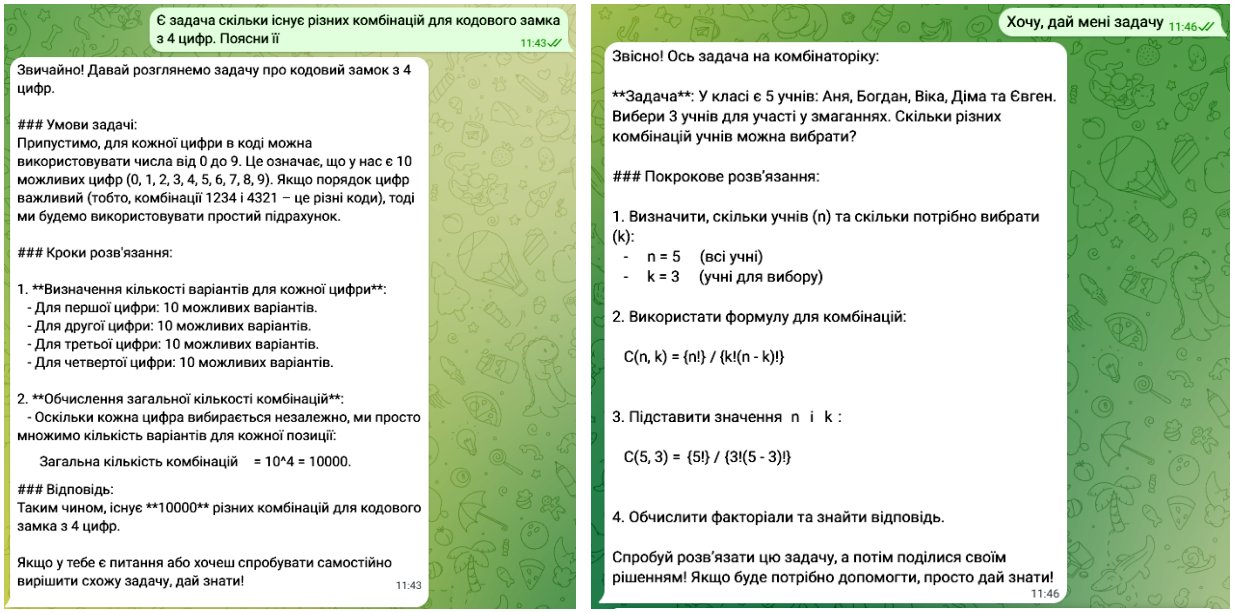

Fig. 12. Example of Using the Bot in Assistant Mode (Continued)

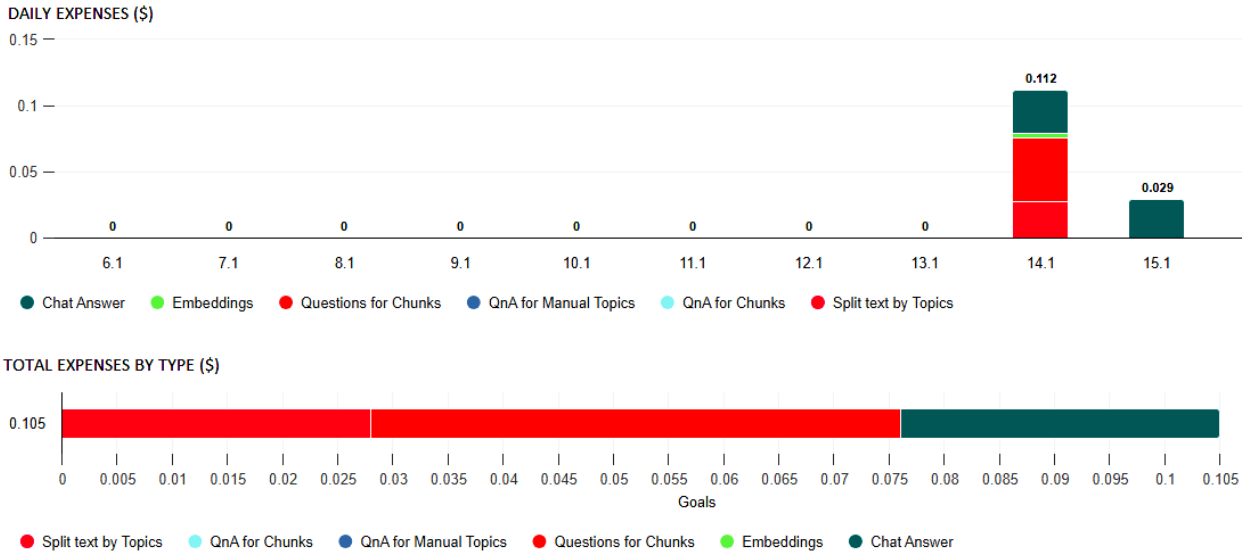

Fig. 13. Bot Usage Graphs

4.2. Example of configuring the bot in tester mode

The separate testing mode in the chatbot creation service allows users to generate a list of questions and answer options based on an uploaded dataset. This mode is designed to create a quiz bot that sends users questions with multiple-choice answers. The main functionality of this mode is that the bot checks the correctness of user answers and provides feedback on the test results.

The testing process began by uploading a dataset containing textual data with sequential information about the questions and their correct answers (Fig. 14).

The next step is the generation of questions and answer options based on the selected thematic lists. The system selects various question formulations considering their diversity and adjusts the probability of the correct answer appearing among the options.

To ensure question diversity, the bot can account for texts containing variables, which are later replaced with specific values.

Special attention is given to verifying the correctness of user answers. After sending a question, the system awaits a user response and compares the response to the proposed answer options.

In case of a match, the bot informs the user of the correct answer and may increase the user's ranking if such a feature is implemented.

In addition, test results can be stored in a database, where user response statistics are recorded. This enables tracking participant performance, comparing results, and creating leaderboards of top test performers. These data can also be used for statistical analysis and to improve the system based on the collected insights. The aforementioned functionalities and aspects of the separate testing mode are designed to enhance the effectiveness and capabilities of the chatbot creation service, as well as broaden user opportunities to utilize it in educational and entertainment contexts. The implementation of such features allows the service to be used more extensively to create custom trainers and remote tests, as well as to develop intellectual skills and knowledge in a continuous mode.

These quiz-based chatbots can be used to conduct thematic tests or evaluate the development of specific competencies.

After forming the dataset, the text sections were segmented by meaning or topics, and questions were generated for each section along with four answer options. Figure 15 shows an example of dataset segmentation and question generation for the test on the topic "Sets. Operations on Sets".

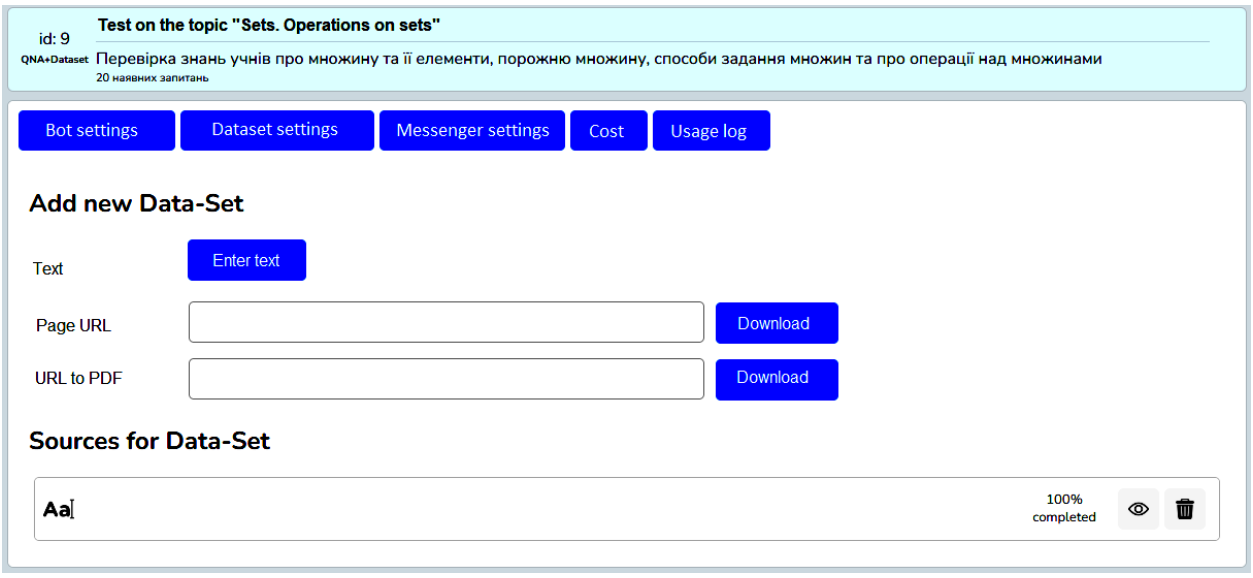

Fig. 14. Dataset Configuration for the Tester Bot

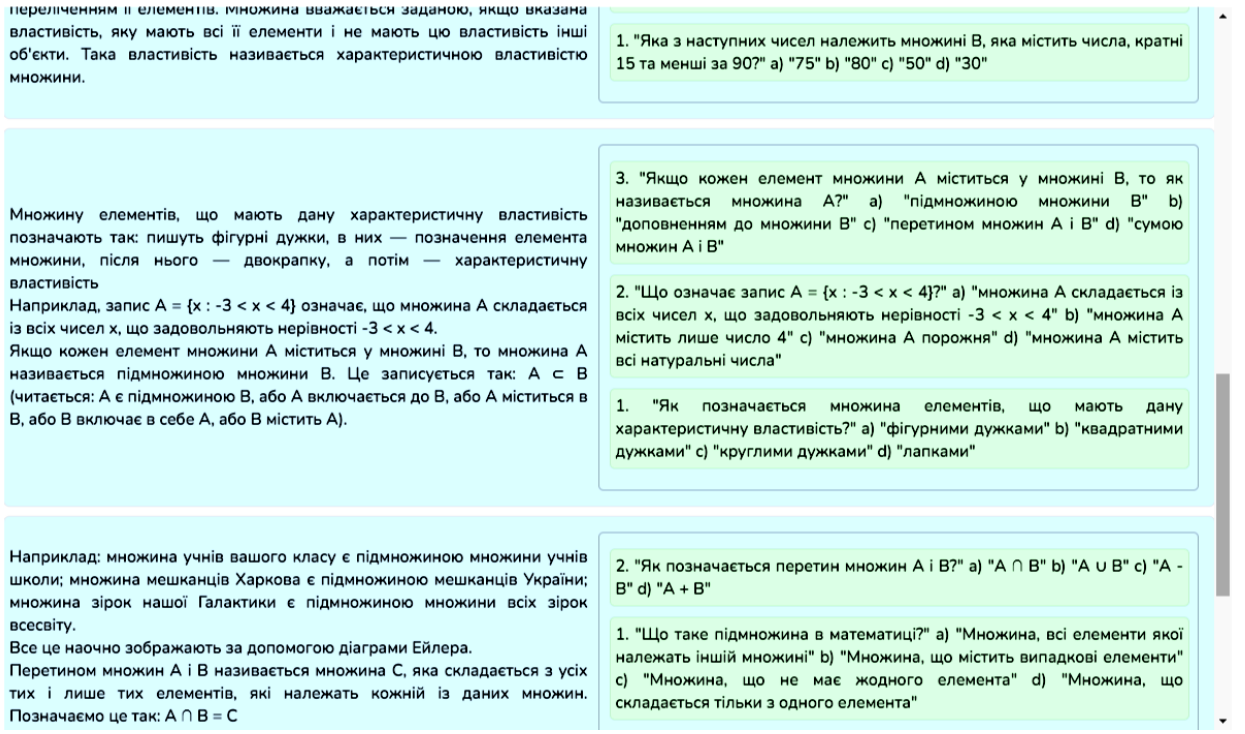

Fig. 15. Result of Dataset Segmentation into Sections and Generated Questions for the Tester Bot

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

The cost of dataset segmentation was $0.003, while generating questions and answer options cost $0.005 (Fig. 16).

As a result, we created a ready-to-use chatbot with 20 questions on the topic "Sets.” Operations on Sets".

When using the chatbot in a messenger, the user is presented with a question and four buttons to submit their answer. In case of an error, the bot will notify the user of the correct answer. An example of using the bot is shown in Fig. 17.

This number of questions can comprehensively cover almost any topic and provide users with the opportunity to test their skills on a large amount of unique material. If the bot is created in testing mode with questions and answers, it is possible to review which users interacted with the bot and how many questions they were able to answer correctly (Fig. 18).

This functionality helps create user rankings, which can motivate users to continue using the bot and compare their skills in a selected topic with other users.

There is also a chatbot generation mode for testing based on a list of topics. This mode, which allows tester bots to be created from a list of provided topics, realizes quick and efficient question and quiz generation.

Unlike the previous mode, which relies on dataset analysis, the proposed mode operates solely with the provided topics, which simplifies the question generation process and allows for the rapid creation of tester bots based on common topics.

The first step in creating a tester bot is to submit a list of topics that will be used to generate questions. The system then analyzes the provided topics and prepares for the question generation task. The generation process involves the bot creating a certain number of questions related to each given topic (Fig. 19).

In the example shown, a bot was created for the topic of 9th-grade algebra, specifically "Inequalities and Systems of Inequalities". The bot settings included the following topics: properties of numerical inequalities, coordinate line, numerical intervals, solving linear inequalities, rational inequalities, and systems of rational inequalities.

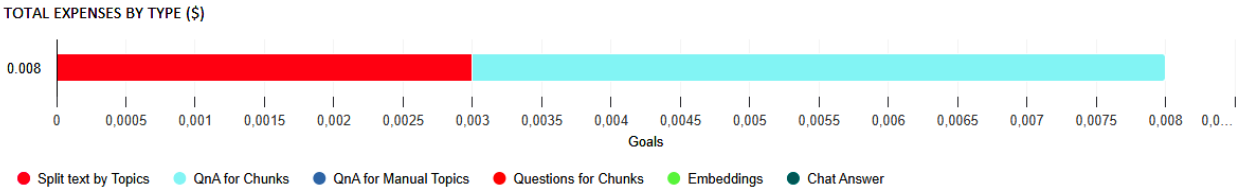

Fig. 16. Costs of Creating a Chatbot with 20 Questions on the Topic "Sets. Operations on Sets"

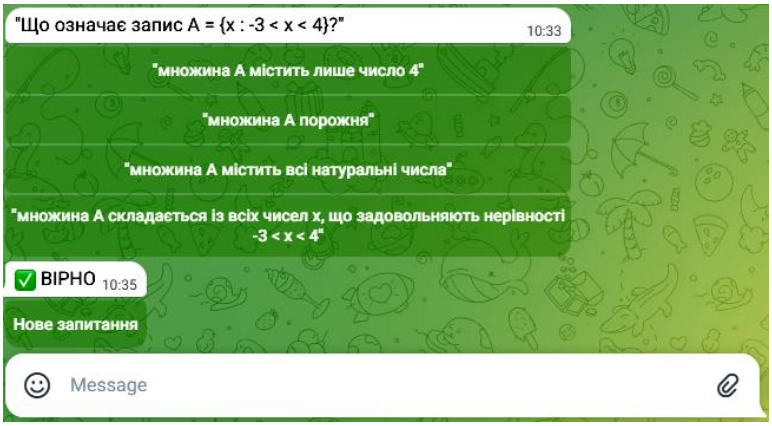

Fig. 17. Example of Using the Tester Chatbot on the Topic "Sets. Operations on Sets"

User id	Correctly	Questions	%	Last used
305807	11	14	79%	15 Jan 11:15
35238	2	9	22%	15 Jan 12:55
69545	2	4	50%	15 Jan 12:44

Fig. 18. Response Logs for the Bot in Testing Mode

After generating questions, the system creates corresponding answer options for each question (Fig. 20). Unlike the dataset mode, where the answer options can be derived from the text, in this mode, the answer options are generated according to the topic of the question. For each topic, 20 questions were generated.

Once the questions and answer options have been created, the bot is ready for use and testing. Users can interact with the bot, answer questions, and check their knowledge about the given topics.

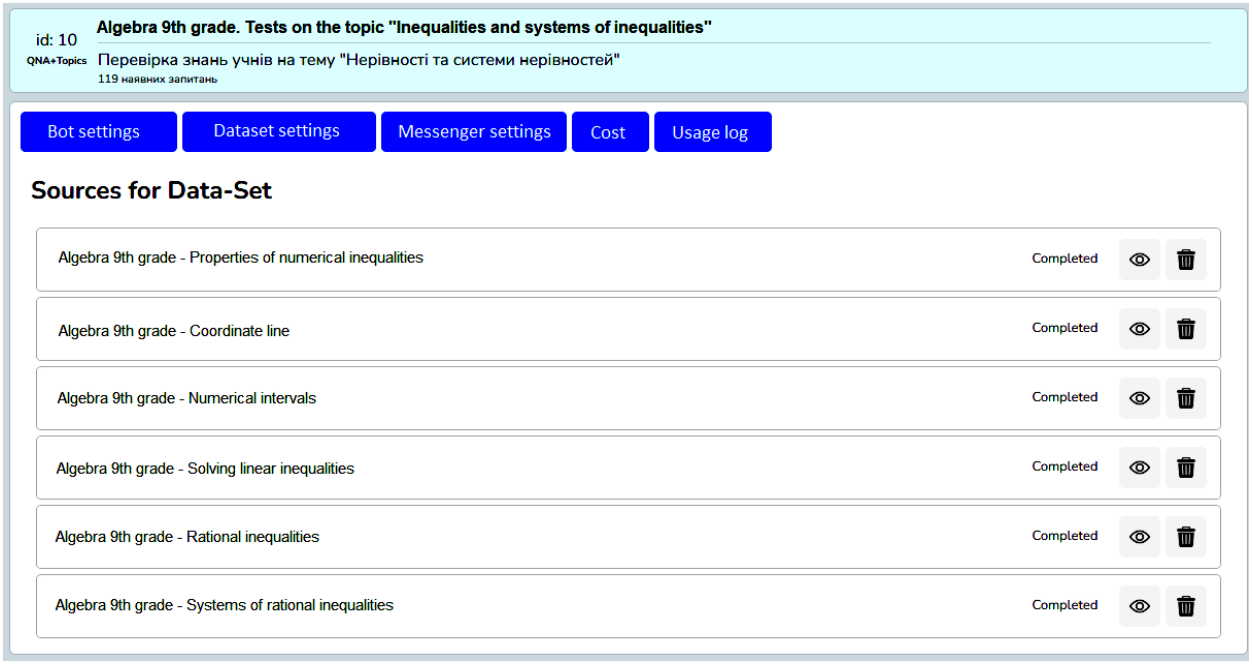

Fig. 19. Automatically Generated Datasets for the Tester Bot Based on a List of Topics

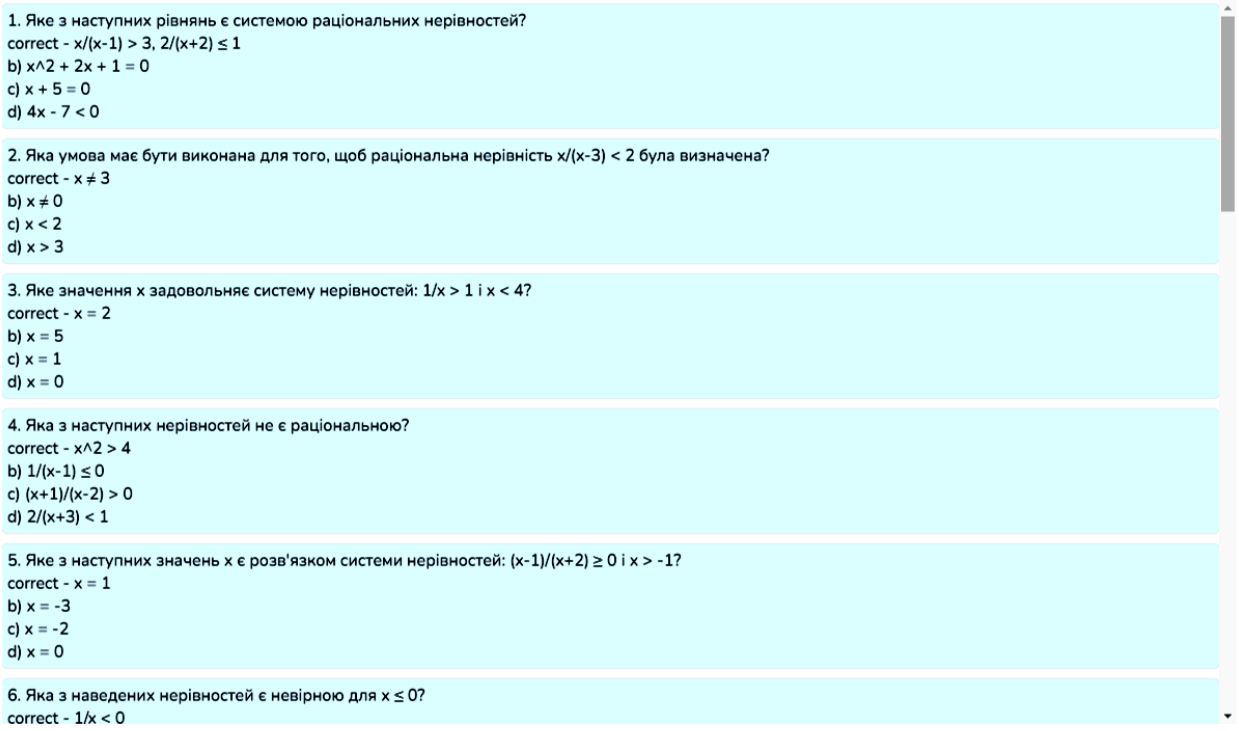

Fig. 20. Questions and Answers Generated by the System for the Tester Chatbot

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

This bot format does not require payment for operations other than generating questions for each text section; thus, it is the most cost effective bot mode.
Thus, this mode enables the quick and efficient creation of tester bots based on provided topics, making it a valuable and versatile tool for teachers creating quizzes and ready-to-use tester chatbots.

4.3. Chatbot for generating variants of mathematical tasks

To develop strong mathematical problem-solving skills, students must solve a series of tasks of a specific type, varying in conditions and numerical parameters. This approach enhances competency acquisition and prevents students from memorizing answers, thus ensuring better results when retaking tests. To achieve this, the following must be ensured:

- a set of meaningful formulations for individual classes of mathematical problems (where possible);
- The ability to generate random numerical parameters for all task classes.

Teachers often use the same tasks repeatedly, with solutions that can even be found online. As a result, instead of solving the tasks independently, some students search for answers online.

When preparing for the National Multidisciplinary Test (NMT) in exact sciences, there is also a need to generate tasks of a specific type and develop the necessary problem-solving skills. Our platform, powered by generative artificial intelligence, helps teachers address the following issues:

- generation of mathematical problem variants (automatic creation of multi-variant tasks);
- generation of possible answer options for tasks, considering the structure of the National Multidisciplinary Test;
- creation of step-by-step explanations for solving generated tasks and improving students' understanding of mathematical concepts;
- generation of templates for formulas, geometric figures, and function graphs in LATEX and PDF formats (LaTeX is a markup language and a macro package for TeX, commonly used for high-quality mathematical and technical document formatting, considered the de facto standard for scientific publications);
- generation of algorithms and Python code to solve the generated mathematical problems.

In the student mode, each test attempt generates a unique set of tasks that are specifically designed for the student.

The complexity of numerical parameterization lies in the fact that correct answers cannot be pre-calculated and stored in a database. In addition, random generation must ensure results produce whole numbers without requiring a calculator. When the input data changes, both the answer and all parameters of the step-by-step solution must be recalculated automatically. To achieve this, calculation functions need to be developed for each task class, which are triggered by data changes. At the same time, it is essential to preserve the system's scalability and ability to add new task classes.

Our approach allows for dynamic task generation. The teacher-user interacts with the system through a user-friendly web interface that provides access to an administrative panel to manage task template generation. This interface makes it easy to generate new mathematical task variants and, if needed, answer options, step-by-step solutions, Python code (for further experiments with numerical parameterization or computer science lessons), and templates for formulas, geometric figures, and function graphs in LATEX and PDF formats.

Fig. 21 shows the administrative panel interface for managing task template generation for a class of tasks, such as the "Problem-Oriented Task on Linear Equations," a task category included in the NMT.

Fig. 22 presents an example of task generation for the class "Simplify the Expression Using Factoring Formulas". In addition, a LATEX template of the formula used in the task was generated, along with a graphical representation in PDF.

Fig. 23 shows another example of task generation for geometry, specifically the "Parallelograms and Trapezoids" task class.

The tasks generated in this way can be added to the corresponding chatbot, which, for example, can serve as a practice tool for NMT preparation.

5. Discussion

The effectiveness and cost of the developed solutions were evaluated through experiments.

The developed tool, which is based on the classic assistant mode, stands out for its high flexibility and versatility. It can respond to several questions and perform tasks in various fields using diverse datasets. This makes it effective for both obtaining information and solving practical tasks. In addition to answering specific questions, the bot has access to the history of previous messages, providing a deeper interaction context. This feature enhances the bot's ability to understand and adapt to individual user needs, resulting in more personalized and useful responses. To test the assistant bot's functionality, multiple bots were created for high school algebra and geometry topics.

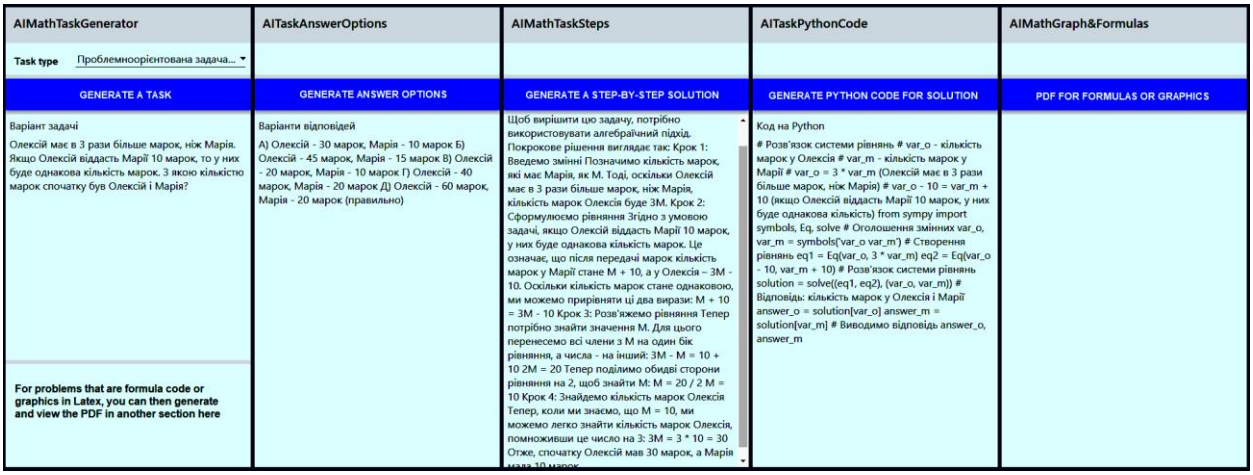

Fig. 21. Example of Using the Panel for Managing Task Template Generation for the Task Class "Problem-Oriented Task on Linear Equations"

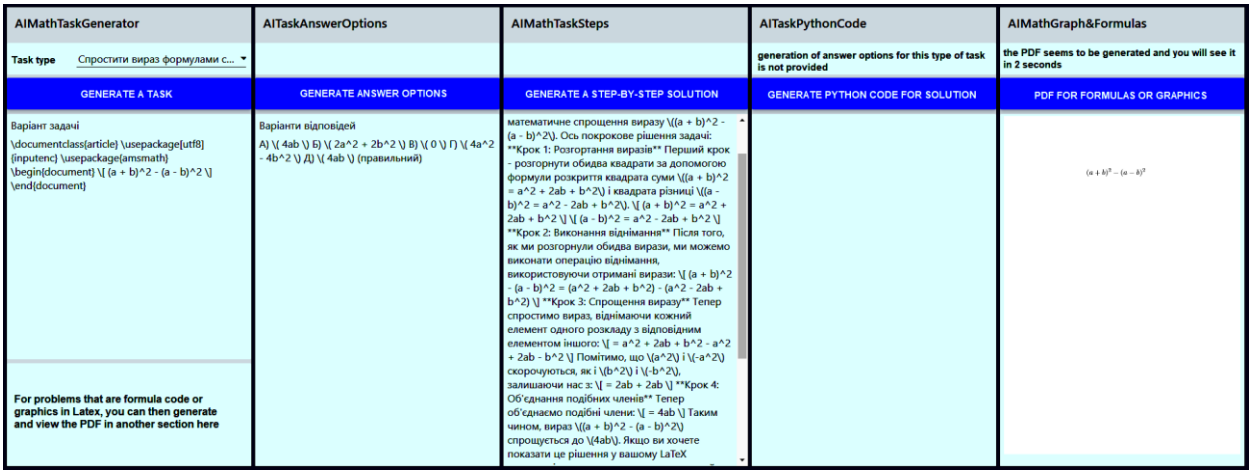

Fig. 22. Example of Using the Panel for Managing Task Template Generation for the Task Class "Simplify the Expression Using Factoring Formulas"

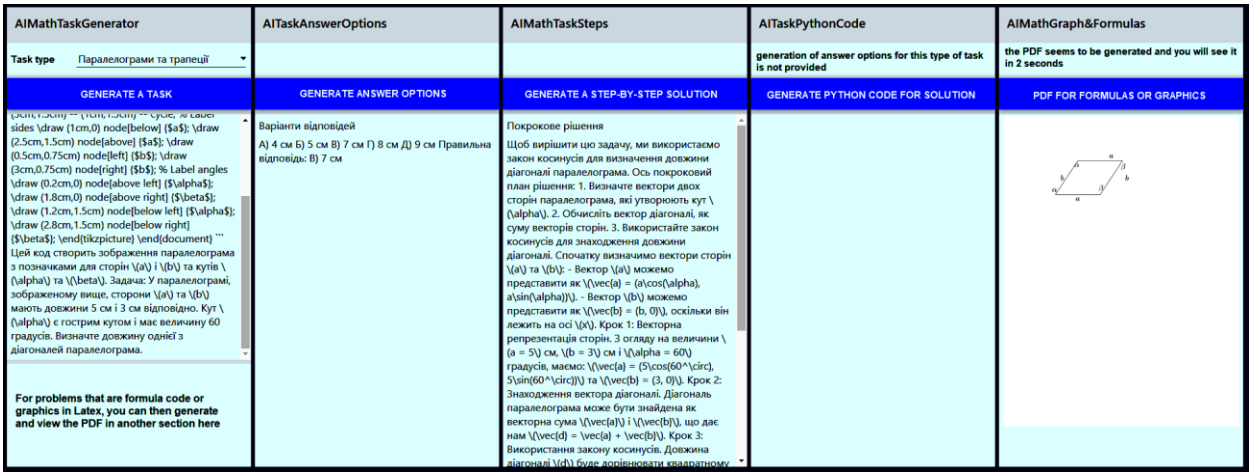

Fig. 23. Example of Using the Panel for Managing Task Template Generation for the Task Class "Parallelograms and Trapezoids"

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

As demonstrated in section 4.1 with the assistant bot for the topic "Fundamentals of Combinatorics, Probability Theory, and Statistics", the system effectively segments text into parts, forming a final dataset for generating responses.

Because the system can track expenses by query type, we can calculate the exact cost of dataset generation for the topic "Fundamentals of Combinatorics, Probability Theory, and Statistics", as shown in Table 2.

Table 2

Cost of Dataset Generation

Dataset	Input Tokens	Output Tokens	Cost Calculation	Cost ($)
Dataset 1	3343	1874	3343/1000 * 0.001 + 1874/1000 * 0.002	0.007
Dataset 1 (second attempt)	3343	580	3343/1000 * 0.001 + 580/1000 * 0.002	0.0045
Dataset 2	3084	1126	3084/1000 * 0.001 + 1126/1000 * 0.002	0.0053
Dataset 3	4841	674	4841/1000 * 0.001 + 674/1000 * 0.002	0.0061
Dataset 4	1725	63	1725/1000 * 0.001 + 63/1000 * 0.002	0.0018
Dataset 5	1506	1035	1506/1000 * 0.001 + 1035/1000 * 0.002	0.0035
Dataset 6	1857	452	1857/1000 * 0.001 + 452/1000 * 0.002	0.0027
Total	19699	5804	0.0309	0.0309

Based on these calculations, we provided sufficient comprehensive information for creating the bot, totaling 23,835 characters; however, the cost of segmenting the text into parts was only $0.03. Compared to similar services, this approach offers higher quality because it does not simply break the text in the middle of sentences but separates it thematically.

Furthermore, after the segmented parts are formed based on thematic meaning, the service initiates the process of generating primary questions for the text, which can be developed. This allows highly accurate answers to frequently asked questions.

The creation of additional context in the form of questions for every few sentences cost only $0.053, but increased the dataset size by 65% while structuring the data more effectively for chatbot interaction. This minor cost of bot configuration and question generation significantly improves the quality of responses by enabling more precise information retrieval from the datasets.

As shown in Figures 9 and 10, the generated questions accurately reflect the text's content and further break it down into smaller segments, which reduces the bot's response cost when delivering information while maintaining answer quality and completeness.

After the dataset text is divided into sections based on themes and supplementary question-answer descriptions are generated, each part is assigned an Embeddings description.

In the assistant bot example for 9th-grade algebra on the topic "Fundamentals of Combinatorics, Probability Theory, and Statistics", the length of the question texts was 15,173 tokens, while the length of the text segmented by themes was 23,303 tokens.

For the descriptions, the OpenAI model "text-embedding-ada-00" was used, and it was priced at $0.0001 per 1 K tokens, which equals $0.038 based on these calculations.

The total cost of configuring the assistant bot is summarized in Table 3.

Table 3

Total Cost of Dataset Generation for the Assistant Bot on the Topic "Fundamentals of Combinatorics, Probability Theory, and Statistics"

Task Type	Token Count	Cost ($)
Segmentation	19699 input + 5804 output	0.0309$
Question Generation	23303 input + 15173 output	0.0536$
Embeddings Generation	38476 output	0.038$
Total		0.1225$

Only $0.1225, considering that the configuration includes datasets totaling approximately 90,000 charac-

ters, which can be considered a comprehensive volume of information, makes this cost a clear advantage compared to other services reviewed in this article.

The implementation of such functionality as the creation of tester chatbots, expands the use of the service, which allows users to create their own training tools and remote tests while continuously developing their intellectual skills and knowledge.

One advantage of using the proposed mode without a dataset, relying instead on the topics specified in the bot's instructions, is the speed of creating tester bots based on commonly known topics.

This allows teacher-users to quickly prepare test questions and quizzes without the need to collect and analyze large datasets. Moreover, using such a bot does not incur additional costs because no API calls to ChatGPT are made when questions are displayed.

Another key point emphasizing the importance of generating new task variants during interactions with the assistant bot or through our separate task generation service is that when teachers use pre-existing tasks, such as those previously included in the NMT, solutions often eventually become available online (Fig. 24).

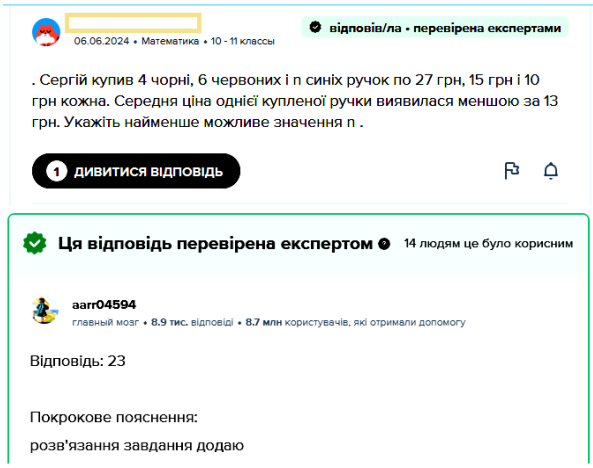

Fig. 24. Example of NMT Task Solutions Available Online

Students often use such resources without solving tasks independently, which reduces the effectiveness of the learning process. We understand that creating unique task variants for each topic adds extra workload for teachers. This is why our service has become an essential assistant to educators seeking to address this issue.

Using these tasks in our chatbot modes allows students to build strong problem-solving skills by completing a series of tasks of the same type. This approach promotes deeper material retention and competency development and prevents the mechanical memorization of answers, ultimately improving performance on repeated tests.

6. Conclusions

A study was conducted to describe the development of a web platform for generating various types of chatbots using artificial intelligence models to improve the quality of student preparation in exact sciences under online educational conditions.

1. An analysis of the problems and features of chatbot creation demonstrated the importance of cost optimization and more efficient resource utilization during the development process. Compared to similar services, there is a clear gap in the speed of accessing tools powered by generative AI models. This gap is evident from the moment the service page loads to the point where educators and students can start using the tool. Modern demands show a growing interest in interactive learning: students seek convenient and accessible ways to master complex subjects like mathematics and physics, while educators search for effective tools to engage learners and automate routine tasks. This highlights the importance of developing a web platform for generating various types of chatbots using AI models to enhance student preparation in exact sciences under online learning conditions.

As a result of the analysis of methods for processing large volumes of text while forming a quality dataset for chatbots, a combination of approaches was selected, including thematic segmentation, generating additional context through topic-specific questions, and generating embeddings to match similar texts when providing chatbot responses. Generating unique questions and answers for each text fragment helps preserve the context and structure of the data, thereby ensuring more accurate information retrieval during searches. This method also facilitates convenient interactions with the chatbot by allowing users to send only questions and receive clear answers, simplifying API usage and optimizing interaction costs.

2. The structure and functionality of the chatbot generation system were also developed and described. A breakdown of the system’s components and their functionalities was provided to ensure the implementation of the new data processing and dataset formation methods. A significant advantage over other services is the ability to provide ready-to-use chatbots can be provided for interaction. To achieve this, a web platform was developed where, on the homepage, a list of ready chatbots is available for immediate use. Users can start a conversation with the selected bot in a messenger by clicking a single button. In the chatbot creation mode, the user is redirected to the system's management panel, which

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

offers tools for creating, configuring, and manage chatbots effectively.

3. The experiments were conducted using various bots in different modes, each loaded with separate sets of thematic information. Examples of different types of chatbots created include:

- assistant bot;
- tester bot;
- math problem generator bot with step-by-step explanations.

4. Experiments were conducted to determine the efficiency and cost of chatbot generation. The results demonstrate that the developed bots successfully completed their tasks and were significantly cheaper compared to other similar systems. For instance, the cheapest subscription package of the services reviewed, which allowed the creation of a topic-unrestricted assistant bot, cost $25 per month and was limited to 5,000 total messages between the user and bot. In the presented example, a bot transmitting the full message history and providing comprehensive responses without artificial limitations costs approximately $0.00025 per user message + bot response. Compared with the cheapest alternative, the 5,000-message limit would cost only $6.25 (0.0025 * 2500) in operational costs. However, the dataset quality generated by the proposed system was significantly higher, demonstrating a clear competitive advantage.

These calculations confirm that the cost of generating responses with the developed system is four times lower than the cheapest comparable system. This validates not only the success of the new dataset formation methods, but also a significant reduction in operational costs. Additionally, the cheapest competing system had a limit of only 10 datasets, while the developed system created six datasets for $0.1225, allowing the creation of one dataset for an average of $0.02. This means that a user can create 50 datasets for only $1.

These calculations, the new dataset formation methods, conducted experiments, and cost analysis confirm that the proposed system offers clear qualitative and functional advantages, along with financial benefits.

Future studies will focus on further refining personalization algorithms to allow chatbots to adapt to individual learning styles, integrating models to assess student progress and automatically adjust the learning plan, using multimodal models (images, video) to explain complex topics, and analyzing the impact of chatbots on students' comprehension of complex subjects. In addition, metrics will be developed to evaluate the quality and effectiveness of chatbots in educational processes.

Contribution of authors: development of the method for optimizing dataset formation for chatbots, development of the structure of the chatbot generation system, development of the web platform for generating various types of chatbots – Prokhorov O., Shymko D., Kholodniak O., Shatalov O.; analysis of the problems and features of chatbot creation and preparation of quality datasets – Kuzminska O., Shatalov O., Chukhrai A.; development of math problem generation chatbots – Prokhorov O., Chukhrai A., Kholodniak O.; preparation and editing of the manuscript – Prokhorov O., Kuzminska O., Kholodniak O.

Conflict of interest

The authors declare that they have no conflict of interest in relation to this research, whether financial, personal, author ship or otherwise, that could affect the research and its results presented in this paper.

Financing

The research was conducted without financial support.

Data availability

The manuscript has no associated data.

Use of artificial intelligence

The authors confirm that they did not use artificial intelligence methods while creating the presented work.

All the authors have read and agreed to the published version of this manuscript.

References

1. Crompton, H., & Burke, D. Artificial intelligence in higher education: the state of the field. *International Journal of Educational Technology in Higher Education*, 2023, vol. 20, article no. 22. DOI: 10.1186/s41239-023-00392-8.

2. Almalawi, A., Soh, B., Li, A., & Samra, H. Predictive Models for Educational Purposes: A Systematic Review. *Big Data and Cognitive Computing*, 2024, vol. 8, no. 12, article no. 187. DOI: 10.3390/bdcc8120187.

3. Overono, A. L., & Ditta, A. S. The Rise of Artificial Intelligence: A Clarion Call for Higher Education to Redefine Learning and Reimagine Assessment. *College Teaching*, 2023, vol. 71, no. 4, pp. 1-4. DOI: 10.1080/87567555.2023.2233653.

4. Kuzminska, O., Mazorchuk, M., Morze, N., & Kobylin, O. Digital Learning Environment of Ukrainian Universities: The Main Components to Influence the Competence of Students and Teachers. *Communications in Computer and Information Science*, 2020, vol. 1175, pp. 210-230. DOI: 10.1007/978-3-030-39459-2.

5. Chen, J., Zhuo, Z., & Lin, J. Does ChatGPT Play a Double-Edged Sword Role in the Field of Higher

Education? An In-Depth Exploration of the Factors Affecting Student Performance. *Sustainability*, 2023, vol. 15, no. 24, article no. 16928. DOI: 10.3390/su152416928.

6. Pagliara, S.M., Bonavolontà, G., Pia, M., Falchi, S., Zurru, A.L., Fenu, G., & Mura, A. The Integration of Artificial Intelligence in Inclusive Education: A Scoping Review. *Information*, 2024, vol. 15, no. 12, article no. 774. DOI: 10.3390/info15120774.

7. Castañeda, R., Martínez-Gómez-Aldaraví, A., Mercadé, L., Gómez, V.J., Mengual, T., Díaz-Fernández, F.J., Sinusia Lozano, M., Navarro Arenas, J., Barreda, Á., Gómez, M., Pinilla-Cienfuegos, E. & Ortiz de Zárate D. Use of ChatGPT as a Virtual Mentor on K-12 Students Learning Science in the Fourth Industrial Revolution. *Knowledge*, 2024, vol. 4, no. 4, pp. 582-614. DOI: 10.3390/knowledge4040031.

8. Wölfel, M., Shirzad, M.B., Reich, A., & Anderer, K. Knowledge-Based and Generative-AI-Driven Pedagogical Conversational Agents: A Comparative Study of Grice's Cooperative Principles and Trust. *Big Data and Cognitive Computing*, 2024, vol. 8, no. 1, article no. 2. DOI: 10.3390/bdcc8010002.

9. Ramalakshmi, K., David, D.J., Selvarathi, M., & Jebaseeli, T.J. Using Artificial Intelligence Methods to Create a Chatbot for University Questions and Answers. *Engineering Proceedings*, 2023, vol. 59, no. 1, article no. 16. DOI: 10.3390/engproc2023059016.

10. Allen, M., Naeem, U., & Gill, S. Q-ModuleBot: A Generative AI-Based Question and Answer Bot for Module Teaching Support. *IEEE Transactions on Education*, 2024, vol. 67, no. 5, pp. 793-802. DOI: 10.1109/TE.2024.3435427.

11. *Ofitsiynyy zvit pro rezultaty NMT u 2024 rotsi. Tom1. Ukrayinskyy tsentr otsinyuvannya yakosti osvity* [Official report on the results of the National Education Quality Assessment in 2024. Volume 1. Ukrainian Center for Education Quality Assessment]. Kyiv, 2024, 148 p. Available at: https://testportal.gov.ua/wp-content/uploads/2024/10/Zvit_NMT_2024_Tom_I_gotovyj_onovlenyj.pdf (accessed December 20, 2024). (In Ukrainian).

12. *Natsionalnyi zvit za rezultatamy mizhnarodnoho doslidzhennia yakosti osvity PISA-2022. Ukrainskyi tsentr otsiniuvannia yakosti osvity* [National report on the results of the international study of the quality of education PISA-2022. Ukrainian Center for Educational Quality Assessment]. Kyiv, 2023, 395 p. Available at: https://pisa.testportal.gov.ua/wp-content/uploads/2023/12/PISA-2022_Naczionalnyj-zvit_povnyj.pdf (accessed December 20, 2024). (In Ukrainian).

13. Mazorchuk, M. S., Vakulenko, T. S., Bychko, A. O., Kuzminska, O. H., & Prokhorov, O. V. Cloud technologies and learning analytics: Web application for pisa results analysis and visualization. *Proceedings of the 8th Workshop on Cloud Technologies in Education*, Ukraine, Kryvyi Rih, 2020, vol. 2879, pp. 484-494. Available at: https://ceur-ws.org/Vol-2879/paper28.pdf (accessed December 20, 2024).

14. Cherednichenko, O., Sytnikov, D., Romankiv, N., Sharonova, N., & Sytnikova, P. Selection of Large Language Model for development of Interactive Chat Bot for SaaS Solutions. *Proceedings of 8th International Conference on Computational Linguistics and Intelligent Systems*, Lviv, Ukraine, 2024, vol. IV, pp. 66-87. DOI: 10.31110/COLINS/2024-4/006.

15. Zi, B., Zhao. H., Ikram. M., & Kaafar, M. GPTs Window Shopping: An analysis of the Landscape of Custom ChatGPT Models, *ArXiv*, 2024, DOI: 10.48550/arxiv.2405.10547.

16. Frazier, M., Damevski, K., & Pollock, L. Customizing ChatGPT to Help Computer Science Principles Students Learn Through Conversation. *Proceedings of Innovation and Technology in Computer Science Education*, 2024, vol. 1, pp. 633-639. DOI: 10.1145/3649217.3653570.

17. Antebi, S., Azulay, N., Habler, E., Ganon, B., Shabtai, A., & Elovici, Y. GPT in Sheep's Clothing: The Risk of Customized GPTs. *ArXiv*, 2024. DOI: 10.48550/arXiv.2401.09075.

18. Sharma V., Garg S., Goel U., Dass V. & Wadhera N. QuickAI-Chat: Introducing Multilingual Conversations Across English, Hindi, Bengali, Punjabi, and Tamil on iOS. *Indian Scientific Journal Of Research In Engineering And Management*, 2024, vol 8, no. 4. DOI: 10.55041/ijsrem31758.

19. Gwendal, D., & Cabot, J. Applying model-driven engineering to the domain of chatbots: The Xatkit experience. *Science of Computer Programming*, 2024, vol. 232, no. 5, article no. 103032. DOI: 10.1016/j.scico.2023.103032.

20. Daniel, G., Cabot, J., Deruelle, L., & Derras, M. Xatkit: A Multimodal Low-Code Chatbot Development Framework. *IEEE Access*, 2020, vol. 8, pp. 15332-15346. DOI: 10.1109/ACCESS.2020.2966919.

21. Hundertmar, S., Hafner, N., & Portmann, E. Dialog-Anpassung an den Charakter des Benutzers – Anwendung der Methoden des Charakter Computing. *Informatik Spektrum*, 2023, vol. 46, no. 4, pp. 210-219. DOI: 10.1007/s00287-023-01547-7.

22. Saklamaeva, V., & Pavlič, L. The Potential of AI-Driven Assistants in Scaled Agile Software Development. *Applied Sciences*, 2024, vol. 14, no. 1, article no. 319. DOI: 10.3390/app14010319.

23. Anton, C.E., Ciobanu, E., Brătucu, G., & Bucs, L. Using Chatbots to Enhance Integrated Reporting: Insights from Accounting and Consultancy Companies from Romania. *Electronics*, 2024, vol. 13, article no. 4801. DOI: 10.3390/electronics13234801.

24. Khneyzer, C., Boustany, Z., & Dagher, J. AI-Driven Chatbots in CRM: Economic and Managerial Implications across Industries. *Administrative Sciences*, 2024, vol. 14, no. 8, article no. 182. DOI: 10.3390/admsci14080182.

25. Giam, N. M., Doc, N. V., Nam, N. T. H., Thanh, N. T., & Giang, N. T. H. Process Of Building An AI Chatbot Scenario For Teaching Mathematics To High School Students. *International Journal of Advance Research in Mathematics Education*, 2023, vol. 1, no. 2,

ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

pp. 46–53. Available at: https://ejournal.papanda.org/index.php/ijarme/article/view/552 (accessed December 20, 2024).

26. Lee, D., & Yeo, S. Developing an AI-based chatbot for practicing responsive teaching in mathematics. *Computers & Education*, 2022, vol. 191, pp. 1–17. DOI: 10.1016/j.compedu.2022.104646.

27. Plevris, V., Papazafeiropoulos, G., & Jiménez Rios A. Chatbots Put to the Test in Math and Logic Problems: A Comparison and Assessment of ChatGPT-3.5, ChatGPT-4, and Google Bard. *AI*, 2023, vol. 4, no. 4, pp. 949-969. DOI: 10.3390/ai4040048.

28. Egara, F., & Mosimege, M. Exploring the Integration of Artificial Intelligence-Based ChatGPT into Mathematics Instruction: Perceptions, Challenges, and Implications for Educators. *Education Sciences*, 2024, vol. 14, no. 7, article no. 742. DOI: 10.3390/educsci14070742.

29. Van Doc N., Thi Hoai Nam N., Tu Thanh N., & Minh Giam, N. Teaching Mathematics with the Assistance of an AI Chatbot to Enhance Mathematical Thinking Skills for High School Students. *International Journal of Current Science Research and Review*, 2023, vol. 6, no. 12, pp. 8574-8580. DOI: https://doi.org/10.47191/ijcsrr/V6-i12-102.

30. Korkmaz Guler, N., Dertli, Z. G., Boran, E., & Yildiz, B. An artificial intelligence application in mathematics education: Evaluating ChatGPT's academic achievement in a mathematics exam. *Pedagogical Research*, 2024, vol. 9, no. 2, article no. em0188. DOI: 10.29333/pr/14145.

31. Remoto, J. P. ChatGPT and other AIs: Personal relief and limitations among mathematics-oriented learners. *Environment and Social Psychology*, 2024, vol. 9, no. 1, article no. 1911. DOI: 10.54517/esp.v9i1.1911.

32. Chang, D. H., Lin, M. P.-C., Hajian, S., & Wang, Q. Q. Educational Design Principles of Using AI Chatbot That Supports Self-Regulated Learning in Education: Goal Setting, Feedback, and Personalization. *Sustainability*, 2023, vol. 15, iss. 17, article no. 12921. DOI: 10.3390/su151712921.

33. Navas, G., Navas-Reascos, G., Navas-Reascos, G. E., & Proaño-Orellana, J. Exploring the Effectiveness of Advanced Chatbots in Educational Settings: A Mixed-Methods Study in Statistics. *Applied Sciences*, 2024, vol. 14, iss. 19, article no. 8984. DOI: 10.3390/app14198984.

Received 10.01.2025, Accepted 20.05.2025

СИСТЕМА ГЕНЕРАЦІЇ ЧАТ-БОТІВ ДЛЯ ПІДТРИМКИ НАВЧАННЯ В ГАЛУЗІ ТОЧНИХ НАУК З ВИКОРИСТАННЯМ МОДЕЛЕЙ ГЕНЕРАТИВНОГО ШТУЧНОГО ІНТЕЛЕКТУ

О. В. Прохоров, Д. І. Шимко, О. Г. Кузьмінська, А. Г. Чухрай, О. В. Шаталов, О. О. Холодняк

Інтеграція генеративного штучного інтелекту в освіту, особливо для викладання точних наук, представляє інноваційні можливості для підвищення залучення та розуміння студентів. Чат-боти, такі як ChatGPT, можуть сприяти інтерактивному навчанню, дозволяючи студентам досліджувати складні наукові концепції за допомогою персоналізованої підтримки та зворотного зв'язку в режимі реального часу. Цей підхід не тільки перетворює традиційні педагогічні методи, але й сприяє глибокій цікавості та розуміння серед учнів. **Досліджується** завдання підвищення ступеня автоматизації створення чат-ботів зі штучним інтелектом та їх інтеграції в процес навчання з точних наук, зокрема математики, школярів в онлайн-освітніх умовах. **Актуальність** дослідження викликана необхідністю підвищення рівня успішності навчальної діяльності в умовах воєнного стану, ліквідації прогалин в знаннях та вирівнювання розриву в знаннях і навичках серед школярів в точних науках при підготовці до продовження навчання в закладах вищої освіти, підвищення ефективності самостійного навчання в онлайн-освітніх умовах та стабілізації соціально-емоційного стану дітей. **Метою дослідження** є розробка веб-платформи генерації чат ботів різного спрямування, з використанням моделей штучного інтелекту, для підвищення якості підготовки школярів з точних наук в онлайн-освітніх умовах. **Завдання**: провести аналіз проблем та особливостей створення чат-ботів та підготовки якісних датасетів; розробити структуру та описати функціонування системи генерації чат-ботів; навести приклади створення різних типів чат-ботів; провести експерименти щодо визначення ефективності використання та вартості генерації чат-ботів. Отримані наступні результати. Розроблено метод оптимізації формування датасетів для чат-ботів. Розроблено веб-платформу генерації чат ботів різного спрямування, а саме ботів-асистентів, ботів-генераторів варіантів математичних завдань та покрокового пояснення їх вирішення, ботів-тестувальників, з використанням моделей штучного інтелекту, для підвищення якості підготовки школярів з точних наук в онлайн-освітніх умовах. **Висновки**. Наукова новизна дослідження пов'язана з удосконаленням методу створення структурованих датасетів чат-ботів зі збереженням тематичної цілісності тексту та контексту, що забезпечує більш точний вибір відповідної інформації чат-ботами для відповіді на запитання користувачів. Ефективність запропонованого підходу проілюстрована при створенні прикладів боту-асистенту, боту-генератору варіантів математичних завдань та покрокового пояснення їх вирішення, ботів-

тестувальників, що показало оптимізацію витрат і забезпечення більш ефективного використання ресурсів. Значною перевагою є зручність інструментів створення, налаштування чат-ботів, а також їх використання через вітрину готових чат-ботів у месенджерах.

Ключові слова: чат-боти; навчання; математика; датасети; веб-платформа; система генерації чат-ботів.

Прохоров Олександр Валерійович – д-р техн. наук, проф., проф. кафедри комп'ютерних наук та інформаційних технологій, Національний аерокосмічний університет ім. М. Є. Жуковського «Харківський авіаційний інститут», Харків, Україна.

Шимко Дмитро Ігорович – асп. каф. медіасистем та технологій, Харківський національний університет радіоелектроніки, Харків, Україна.

Кузьмінська Олена Геронтіївна – д-р пед. наук, проф., проф. кафедри інформаційних систем і технологій, Національний університет біоресурсів і природокористування України, Київ, Україна; проф. кафедри інформатики та прикладної математики, Криворізький державний педагогічний університет, Кривий Ріг, Україна.

Чухрай Андрій Григорович – д-р техн. наук, проф., проф. кафедри інженерії програмного забезпечення, Національний аерокосмічний університет ім. М. Є. Жуковського «Харківський авіаційний інститут», Харків, Україна.

Шаталов Олексій Вікторович – магістр комп'ютерних наук, асп. Лабораторії інженерії знань з медичної інформатики та електронної охорони здоров'я (LIMICS/UMRS-1142), Університет Сорбонна, Париж, Франція.

Холодняк Олександр Олександрович – асп. каф. комп'ютерних наук та інформаційних технологій, Національний аерокосмічний університет ім. М. Є. Жуковського «Харківський авіаційний інститут», Харків, Україна.

Oleksandr Prokhorov – Doctor of Technical Sciences, Professor, Professor at the Computer Science and Information Technologies Department, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine,
e-mail: o.prokhorov@khai.edu, ORCID: 0000-0003-4680-4082.

Dmytro Shymko – PhD Student of the Media Systems and Technologies Department, Kharkiv National University of Radioelectronics, Kharkiv, Ukraine,
e-mail: dmytro.shymko@nure.ua, ORCID: 0009-0005-3195-467X.

Olena Kuzminska – Doctor of Pedagogical Sciences, Professor, Professor at the Information Systems and Technologies Department, National University of Life and Environmental Sciences of Ukraine, Kyiv, Ukraine, Professor at the Computer Science and Applied Mathematics Department, Kryvyi Rih State Pedagogical University, Kryvyi Rih, Ukraine,
e-mail: o.kuzminska@nubip.edu.ua; ORCID: 0000-0002-8849-9648.

Andrey Chukhray – Doctor of Technical Sciences, Professor, Professor at the Software Engineering Department, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine,
e-mail: achukhray@gmail.com, ORCID: 0000-0002-8075-3664.

Oleksii Shatalov – M.Sc. of Computer Science, PhD Student at Laboratory of Medical Informatics and Knowledge Engineering in e-Health (LIMICS/UMRS-1142), Sorbonne University, Paris, France,
e-mail: oleksii.shatalov@sorbonne-universite.fr, ORCID: 0000-0002-7267-6718.

Oleksandr Kholodniak – PhD Student of the Computer Science and Information Technologies Department, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine,
e-mail: o.o.kholodniak@khai.edu, ORCID: 0009-0000-1012-7361.