Модель формирования индивидуального маршрута освоения лексического материала с применением методов и инструментов обработки естественного языка

А. П. АВРАМЕНКО, М. В. ХВАН, В. В. ТЕРНОВСКИЙ

АННОТАЦИЯ

Введение. Технологии искусственного интеллекта (ИИ) предоставляют из себя средство персонализации обучения иностранным языкам. Актуальность исследования обусловлена необходимостью разработки методик интеграции постоянно обновляющихся технологий в учебный процесс. Одновременно, технологии ИИ представляют из себя актуальное техническое средство обучения для реализации персонализации обучения в парадигме непрерывного образования. *Цель настоящего исследования* заключается в разработке алгоритма вступительного тестирования, основанного на применении инструментов обработки естественного языка, которые позволяют автоматически анализировать словарный запас индивида на его родном языке.

Материалы и методы. Исследование проводилось в рамках курса «Иностранный язык в гуманитарной сфере» для студентов бакалавриата, находящихся на уровнях C1-C2. Количественные данные были подвергнуты статистическому анализу с использованием t-теста Стьюдента и U-теста Манна-Уитни.

Результаты исследования. Результаты запуска алгоритма и его тестирования подтверждают гипотезу о возможности автоматического подбора лексического диапазона. В экспериментальной группе средний прирост составил 10%, в контрольной – 2% ($U = 0$; $p<0,05$).

Заключение. Предложен инновационный подход к созданию персонализированных списков лексики, который может быть использован для выявления грамматических и прагматических особенностей общения студентов.

КЛЮЧЕВЫЕ СЛОВА

обучение иностранным языкам, цифровизация образования, персонализированное обучение, словарный запас, обработка естественного языка, искусственный интеллект

Для цитирования: Авраменко А. П., Хван М. В., Терновский В. В. Модель формирования индивидуального маршрута освоения лексического материала с применением методов и инструментов обработки естественного языка // Перспективы науки и образования. 2025. № 5. С. 618–633. https://doi.org/10.32744/pse.2025.5.40

Поступила: 11.04.2024 | Одобрена: 14.06.2025 | Опубликована: 31.10.2025

Designing personalized vocabulary learning path based on natural language processing tools

A. P. AVRAMENKO, M. V. KHVAN, V. V. TERNOVSKI

ABSTRACT

Introduction. The educational system is still lacking the necessary methodologies to effectively harness the rapidly developing artificial intelligence (AI) technologies. The relevance of the research is determined by the necessity to develop methods of integrating technologies that are subject to constant updating into the educational process. Concurrently, AI technologies signify a tangible tool for the facilitation of personalised learning within lifelong learning. *The objective of this study* is to develop an algorithm for placement testing based on NLP tools that enable the automatic analysis of an individual's vocabulary range in their native language, with the aim of facilitating the creation of a personalised vocabulary range in a foreign language.

The research methods. The study was conducted as part of a “Foreign language in the humanitarian field” course for Bachelor of Arts degree students (C1-C2) at the Faculty of Foreign Languages and Area Studies at Lomonosov Moscow State University. The quantitative data was then subjected to statistical analysis using Student's t-test and Mann-Whitney U-test.

The results of the launched algorithm and its piloting prove the hypothesis of the possibility of automatic vocabulary range selection. The statistical analysis indicates that the comparison of the final test results in the experimental and control groups reveals the grade growth in the experimental group by on average 10% (statistically significant, according to t-test) compared to 2% (statistically insignificant, according to t-test) in the control group ($U = 0$, $p<0.05$).

Conclusion. The study suggests an innovative automatic approach for the creation of personalised vocabulary lists. Further research might include the use of the proposed method to identify the grammar and pragmatic peculiarities of the students' communication.

KEYWORDS

foreign language teaching, digitalisation of education, personalised language learning, vocabulary range, natural language processing, artificial intelligence

For Citation: Avramenko, A. P., Khvan, M. V., & Ternovski, V. V. (2025). Designing personalized vocabulary learning path based on natural language processing tools. *Perspektivy nauki i obrazovania = Perspectives of Science and Education*, (5), 618–633. https://doi.org/10.32744/pse.2025.5.40

Received: 11 April 2025 | Approved: 14 June 2025 | Published: 31 October 2025

INTRODUCTION

Artificial intelligence (AI) has become increasingly embedded in educational platforms, promising to transform personalized learning experiences. In language education, AI-driven systems can adapt content and feedback to individual learner needs in real-time, potentially improving engagement and outcomes. The UNESCO document "Artificial Intelligence Technologies in Education: Prospects and Consequences" aims to establish a collective understanding of the opportunities and challenges that AI presents for education, as well as its implications for the essential competencies required in the age of AI.

This paper expands on current research by examining the pedagogical foundations of AI-driven personalized learning, recent advances in natural language processing (NLP) models, real-world implementations in various contexts, research methodologies and findings, as well as the challenges and ethical considerations associated with AI-powered adaptive learning. By synthesizing theoretical insights and empirical studies from 2019 onwards, we aim to provide a comprehensive overview of how AI is reshaping language learning and what this means for educators and learners [1]. The analysis of 71 empirical studies conducted between 2006 and 2019 identifies the impact of personalized learning on student educational outcomes as a key area of future research and development in this field [2]. However, the educational system still lacks the requisite methodologies to effectively harness these tools [3].

THEORETICAL BACKGROUND

Personalized language learning path. In a broader context, personalized learning is referred to as an educational approach that aims to tailor materials and tasks for each student according to their specific needs. These incorporate not only the goals and skills of a student but also *'learner profiles and attitudes, previous knowledge and beliefs'* [4, p. 485]. Personalized learning pathways are designed to cater to the specific needs, preferences and learning styles of individual learners. The approaches entail the creation of bespoke learning experiences, informed by data-driven insights and learner profiles.

Considered the founder of modern education, the Czech philosopher and educator Jan Amos Komensky introduced a number of innovations and ideas in the 17th century, such as an approach to teaching that gradually progressed from basic to more complex ideas; the requirement that the subjects taught have real-world applications, and that they encourage lifelong learning, as well as others. As we can see, each of these concepts takes into account the goals, requirements, and individual skills of the learners. One of the greatest Russian educators, Konstantin Ushinsky, initially brought up the issue of the necessity of individualized education. He thought that since each child is different, education should be tailored to meet his or her requirements and skills. Each student must be thoroughly assessed by the teacher in order to develop a personalized learning plan. The concepts of differentiated learning, which Ushinsky established, served as the foundation for numerous contemporary teaching strategies. When developing concepts for individualized learning, Vygotsky's method is crucial. According to Lev Vygotsky's concept of the zone of proximal development, learning should concentrate on subjects that students are currently unable to complete on their own but will be able to do in the near future with minimal assistance from a teacher. We can provide a credible forecast and useful psychological and pedagogical recommendations on the best methods of study based on the concept of the zone of proximal growth, which gives an idea of a student's prospective development opportunities. It is interesting to note that each student's zone of immediate development varies, depending on their needs, abilities, and work ethic. The zone of immediate development notion has a wide range of applications in pedagogy, and it serves as the foundation for the concepts of student autonomy and personalized learning paths [5].

V.A. Sukhomlinsky, another exceptional Russian educator, valued individualization highly and saw it as a teacher's capacity to adjust to the various demands of each pupil and take charge of their growth. According to Sukhomlinsky, a teacher should thoroughly analyse each student's objectives and aptitudes. He also recommended tailoring training materials and study time accordingly. Sukhomlinsky says that a crucial aspect of educational wisdom is the capacity to accurately assess each student's current level of ability and how to best foster his mental capacities going forward [6].

One of the top priorities in higher education is the individualization of the educational process, which includes establishing the conditions for students to have access to educational opportunities, their academic mobility, and the design of their own educational trajectory. However, it is not just the individual's needs and ambitions that need to be taken into account. National curriculums dictate what subjects and materials are taught according to what our governments see as societal requirements. So, in other words, individualisation can serve not only the needs of the individual, but also those of the society they are in. Individual educational trajectories (IET), are among the most cutting-edge methods for implementing individualization in education. IET is currently a top priority for schools when it comes to creating the learning environment [7]. The following tenets serve as the foundation for the individual educational trajectory: independence, flexibility, reflexivity, goal-setting, individual orientation, and scientific validity. The steps taken in IET are diagnostics, programming, implementation, and control [8].

The continued development of digital technologies has enabled the wide-spread use of IET. This approach is being adopted by a growing number of schools worldwide, who are using artificial intelligence and data analysis to improve their outcomes. Digital technologies make it possible to automate an individual learning trajectory based on individuals' assessment results and in turn caters to that individual's learning needs as they go through the learning process, which encompasses the idea of adaptive learning. In the context of the digital world, adaptive learning refers to the capacity of contemporary technology to apply a personalized approach by using machine learning to customize material for the user. A. K. Clouder, Gordon Pask, Berres Frederick Skinner, and other scientists have made significant contributions to the development of adaptive learning.

Thus, the concept of a personalized language learning path has transformed language acquisition by acknowledging the individuality of every student and customizing education to suit their need. It opens the door to more successful language learning outcomes by offering customizable learning pathways, focused skill development, and consequently increased motivation. Russia and many other nations are actively working to create AI-powered adaptive learning platforms as it becomes more and more popular.

Theoretical Frameworks for Personalized Learning

AI-driven personalized learning is underpinned by established pedagogical theories that emphasize learner-centered education. **Constructivism** posits that learning is an active process where learners build new knowledge upon their prior experiences. AI systems align with constructivist principles by allowing learners to explore and discover content at their own pace, constructing understanding through adaptive feedback loops. For example, an AI tutor can present progressively challenging tasks as the learner demonstrates mastery, mirroring Vygotskian *scaffolding* in the learner's Zone of Proximal Development. **Social learning theory** (Bandura's framework) highlights learning through interaction and observation; modern AI platforms incorporate social features like peer discussion forums or collaborative problem-solving tasks, enabling learners to learn from one another even in online environments. **Self-directed learning (SDL)** is another critical framework: learners take initiative in diagnosing their learning needs, setting goals, and evaluating progress. AI-driven tools can foster SDL by providing dashboards and analytics that help learners monitor their performance, reflect on errors, and self-regulate their study plans. The ability of AI tutors to give immediate feedback and suggest resources supports learners in taking control of their learning trajectory [9].

Differentiated instruction, traditionally achieved by teachers tailoring lessons to different student groups, is greatly enhanced by AI personalization. Instead of one-size-fits-all teaching, AI systems dynamically adjust difficulty, modality, and pace for each learner. This is consistent with **differentiated instruction principles**, which aim to address individual learning styles and levels. AI can handle such differentiation at scale by analyzing each learner's interactions and performance in real time. For instance, an AI language app might offer more grammar exercises to a learner weak in grammar, while another learner strong in grammar but weak in listening comprehension receives more audio exercises. These practices echo mastery learning theory – the idea that all learners can achieve mastery if given enough time and the right support. AI tutors can implement mastery learning by allowing students to repeat practice until they reach proficiency, offering targeted remediation when errors persist. In sum, the pedagogical foundations of AI-driven personalized learning are rooted in constructivist, social, and self-regulated learning theories, combined with practical approaches like differentiated instruction and mastery learning. Aligning AI tools with these theories is essential; a recent systematic review emphasizes that successful AI integration in education requires alignment with sound pedagogical principles (e.g., constructivism, connectivism, self-directed learning) to truly enhance learning experiences [10].

Advances in NLP Models for Education

At the heart of AI-driven language learning are sophisticated **NLP models**. Over the past few years, NLP has been revolutionized by the advent of transformer-based architectures and large language models. These models, exemplified by OpenAI's GPT-4, Meta's LLaMA, and Anthropic's Claude, are pre-trained on massive text corpora and can understand and generate human-like language. Such models underpin conversational agents and intelligent tutors that interact with learners in adaptive ways. Modern large language models (LLMs) like GPT-4 represent a transformative approach to personalized education through real-time adaptability and dynamic content generation. Unlike earlier rule-based or narrowly trained systems, GPT-4 can engage in open-ended dialogue, provide instant explanations, and adjust its responses based on the learner's input. This enables Socratic-style tutoring interactions, where the AI asks guiding questions or prompts critical thinking, much as a human tutor would. The generative capabilities of these models mean that learners can receive nuanced, context-aware feedback and even full explanations or examples on any topic, improving the personalization of support.

Transformer models (originating from Vaswani et al.'s 2017 architecture) have advanced language learning by handling long-range dependencies in text and supporting attention mechanisms that weigh the importance of different input tokens. The evolution from models like BERT to GPT-3 and GPT-4 has brought significant improvements in language understanding. For instance, GPT-4 exhibits greater coherence over multi-turn interactions and can maintain context over a dialogue, which is crucial for a tutoring system that tracks a student's progress through a lesson.

Additionally, GPT-4's multimodal abilities allow it to process text and images together, hinting at future language learning applications where a student could, say, describe an image in the target language and get feedback. Other models like **reinforcement learning-based dialogue systems** have also emerged: these use reinforcement learning (RL) to fine-tune responses based on educational reward signals (e. g., student correctness or engagement). In education, RL can optimize when to give hints or what difficulty of question to present next, treating the tutoring process as a sequential decision problem. Studies have demonstrated the promise of RL in instructional policy adaptation – for example, an RL-enhanced tutor that learned to sequence math problems led to students spending less time on each problem while achieving the same or better learning gains. By 2023, RL techniques were being extensively applied to reduce the time needed to acquire knowledge and to improve the efficiency of intelligent tutoring systems [11].

Another cutting-edge approach is **transformer-based feedback generation**, where models not only present content but also analyze student responses (essays, spoken language, etc.) and generate individualized feedback. For instance, large language models can evaluate a learner's free-text answer to a comprehension question and offer hints or corrections, functioning as automated writing coaches or conversation partners in language learning. Research into models like GPT-4, as well as open-source variants like LLaMA, is ongoing to refine their use in educational contexts, ensuring they align with curriculum standards and pedagogical goals. Overall, the rapid advancement of NLP models has created powerful tools for AI-driven personalized learning: they enable adaptive dialogue, content generation tailored to each learner, and intelligent feedback mechanisms that were not feasible just a few years ago. Education-focused adaptations of these models, often through fine-tuning or prompt engineering, are at the forefront of AI in language education today

Case Studies of AI-Powered Adaptive Learning

Real-world implementations of AI-driven personalized learning illustrate how these technologies are applied in diverse educational contexts. This section highlights case studies from higher education, online language learning, and corporate training to demonstrate the breadth of applications.

University Mathematics Tutoring (Cognitive Tutor): One of the earliest and most cited successes of AI in education is the Cognitive Tutor for math, developed at Carnegie Mellon University. **Cognitive Tutor** is an intelligent tutoring system that personalizes instruction and feedback in subjects like algebra. It uses a cognitive model to track each student's mastery of specific skills and adapt problem sequences accordingly. In practice, as a student works through problems, the system selects subsequent problems that target the student's weaknesses and prerequisite gaps, effectively differentiating the curriculum for each learner. Studies have found that the Cognitive Tutor significantly improves learning outcomes in mathematics, especially for struggling students. For example, a controlled study reported that classes using Cognitive Tutor showed higher gains on algebra post-tests compared to traditional classes. This evidence underscores the potential of AI tutors to enhance STEM education through personalization. Many universities and high schools have integrated such AI-based math tutors into their curricula to supplement classroom instruction with tailored practice, demonstrating positive impacts on student performance and retention of material.

Adaptive Learning in Higher Education (Knewton): Another case comes from adaptive learning platforms like **Knewton**, which has been used in university courses and massive open online courses (MOOCs). Knewton's platform analyzes learner data (quiz results, time on tasks, etc.) and adjusts the content delivered – difficulty, topics, and pacing – to optimize learning for each student. In subjects ranging from college biology to economics, Knewton provides recommendations and resources based on individual learner profiles. Research indicates that using Knewton can lead to improved course grades and increased student engagement at both K-12 and university levels. For instance, institutions that piloted Knewton in large introductory classes observed that students who spent more time with the adaptive platform performed better on final exams than those who only used static textbooks. These findings illustrate AI's role in scaling differentiated instruction beyond what a single instructor could manage, by continuously tailoring learning pathways for hundreds of students simultaneously.

Online Language Learning (Duolingo): In the realm of language education, **Duolingo** stands out as a widely used AI-driven platform. Duolingo offers courses in over 100 languages and employs AI algorithms to personalize each learner's experience. It uses techniques such as spaced repetition (to review vocabulary at optimal intervals) and adaptive lesson sequencing (adjusting the next lesson based on past mistakes). A recent quasi-experimental study examined Duolingo's impact on English as a Foreign Language (EFL) learners in an online class setting. The results showed significant gains in learners' engagement and willingness to communicate in the target language for those using Duolingo compared to a control group [9].

Specifically, students using the AI-driven app demonstrated improved vocabulary and grammar skills and increased confidence in speaking and writing, with large effect sizes reported for gains in communication willingness.

The study concluded that Duolingo's adaptive and gamified approach (e.g., immediate feedback, tailored practice, and motivational elements like points and streaks) had a beneficial impact on learners' attitudes and performance, and it encouraged teachers to integrate such AI tools to enhance language learning experiences.

These findings reinforce earlier observations that Duolingo can effectively support vocabulary and grammar development, making language learning more accessible and engaging for diverse learners. Beyond formal studies, Duolingo's own data science team has reported that its AI-driven personalization keeps users more engaged and helps them progress faster through content, highlighting the value of data-driven refinement in educational apps.

Corporate Training Programs: AI-powered personalized learning is not confined to academic settings; it has also been embraced in corporate training and professional development. Large organizations have begun using AI-driven learning platforms to upskill employees in a targeted way. A notable example is IBM's use of its Watson AI for employee training. IBM built an AI-powered training platform that analyzes each employee's role, current skills, and learning history to recommend personalized learning modules. By tailoring content to individual learning styles and paces, IBM reported improved training engagement and better retention of new skills among employees.

This personalized approach in corporate e-learning has been linked to boosts in workforce productivity, as training is more closely aligned with each employee's needs. Another case comes from Schneider Electric (an energy management company), which implemented an AI-based adaptive learning system for its global workforce. After adoption, the company observed a 60% increase in training efficiency, as the AI system could identify employees' knowledge gaps and serve appropriate content to fill those gaps swiftly.

These corporate case studies demonstrate that AI-driven learning is effective beyond traditional classrooms – adaptive learning systems can accelerate onboarding, continuous education, and skill development in business environments. They also show how ROI (return on investment) in training can improve when learning is personalized: time spent in training is reduced while mastery is increased, and employees appreciate training that is relevant to their personal growth goals.

Across these case studies – from university courses and online platforms to workplace training – a common thread is the ability of AI to handle large-scale personalization. Each learner (or employee) receives a custom experience, which in turn leads to measurable improvements: higher engagement, better skill acquisition, and often greater satisfaction with the learning process. These real-world implementations underscore the promise of AI-powered adaptive learning, while also providing lessons on best practices (such as aligning AI tools with curriculum objectives and providing instructors with insight into the AI's decisions).

Methodology in AI-Personalized Learning Research

Researchers studying AI-driven personalized learning have employed a range of methodologies, combining quantitative and qualitative approaches to evaluate effectiveness. A common approach is the **experimental or quasi-experimental design**, where one group of learners uses an AI-powered tool and their outcomes are compared to a control group using traditional methods or a non-adaptive system. For instance, in the EFL case study mentioned above, two classes were compared with and without the Duolingo intervention, using pre- and post-tests to measure gains in engagement and communication skills [12].

Such designs often involve **pretest-posttest control group** setups, enabling researchers to attribute observed differences in learning outcomes (e. g., test scores, skill proficiency, engagement metrics) to the use of AI personalization. Randomized controlled trials (RCTs) are considered a gold standard; however, in educational settings RCTs can be challenging to implement, so researchers sometimes resort to quasi-experimental designs (e. g., using intact classes or volunteer participants) and then apply statistical controls to account for initial group differences.

Another methodological approach is **learning analytics and data mining**. Given that AI learning platforms collect fine-grained data on learner interactions (clicks, time spent, question attempts, etc.), researchers often analyze this log data to understand how personalized learning unfolds. Techniques like sequence analysis or clustering can reveal different learner pathways and behaviors in an adaptive system. For example, researchers might analyze whether students who follow the AI-recommended personalized path achieve mastery faster than those who deviate or how often the AI needs to intervene with hints for learners of varying proficiency. Such data-driven studies help validate the adaptivity of the system and can illuminate which features of personalization (e. g., frequency of feedback or difficulty adjustment) contribute most to learning gains.

Qualitative methods are also employed to capture learner and teacher perspectives on AI-driven personalized learning. Interviews, focus groups, and observations allow researchers to explore experiences beyond test scores. For instance, a qualitative study might interview language learners about their autonomy and motivation when using an AI tutor at home, probing how the tool supports self-directed learning or where it falls short. Self-reported measures, such as surveys on learner satisfaction or perceived effectiveness, complement usage data. In some studies, mixed-methods designs are used, where quantitative results (like improved test performance) are triangulated with qualitative feedback (like student reflections) to provide a holistic assessment of the AI tool's impact [13].

In evaluating AI tutors or platforms, researchers also use **iterative design-based research**. Here, an AI learning system is deployed in a real educational setting, studied, then refined in cycles. This approach acknowledges the importance of context – the same AI system might need adjustments when used in a university course versus a corporate training program. Design-based research yields rich insights into how AI-driven personalization can be integrated into curricula, how teachers mediate its use, and how learners engage with it over time. For example, one design-based study might iteratively improve an AI writing assistant in a college class by collecting student feedback each semester and tweaking the feedback algorithms accordingly.

Finally, systematic literature reviews and meta-analyses have become valuable for synthesizing findings across studies. Reviews of AI in education research (e.g., a 2019 systematic review by

Zawacki-Richter [14] et al.) have highlighted prevalent themes and gaps, such as an emphasis on certain subjects or the need for more studies on long-term efficacy and teacher roles. By surveying a broad range of empirical studies (sometimes numbering in the hundreds), such reviews can generalize which aspects of personalized AI learning consistently show benefits (like increased engagement) and which areas show mixed results or require caution. In summary, the methodology of studying AI-driven personalized learning is multidisciplinary, blending experimental rigor with data analytics and user-centered inquiry to fully understand both the outcomes and the processes of personalized learning in action.

The focus of this research is the acquisition of vocabulary. **The objective** of this research is to develop an algorithm for placement testing based on natural language processing tools with the aim of facilitating the development of personalized language learning path. It is hypothesized that the automatic analysis of a person's vocabulary range in their native language will facilitate the creation of a personalized vocabulary range in a foreign language. This paper presents a theoretical analysis of the fundamental principles of NLP, an overview of the semantic fields of lexemes from an NLP perspective, and an algorithm for the automatic selection of personalized vocabulary ranges. In particular, open libraries and large language models are employed for the automatic processing of the user's texts in their native language with the objective of creating their personal vocabulary lists in the target language, utilizing the vectorization approach.

MATERIALS AND METHODS

The paper is interdisciplinary in nature, situated at the intersection of pedagogy, with a particular focus on teaching English as a Foreign Language, and computer studies, particularly computational linguistics. In order to design a personalised vocabulary learning path, a number of methods were employed, including literature analysis and synthesis, as well as comparison, clustering, and classification of technologies. These included open libraries from GitHub and large language models such as BERT, GPT, and so forth. Secondly, a pedagogical experiment was conducted in order to pilot the aforementioned personalised vocabulary learning path. The study was conducted as part of a "Foreign language in the humanitarian field" course for Bachelor of Arts degree students (C1-C2) at the Faculty of Foreign Languages and Area Studies at Lomonosov Moscow State University. Finally, the quantitative data was subjected to statistical analysis using Student's t-test and Mann-Whitney U-test.

In order to facilitate the automated selection of materials, students are invited to respond to the queries posed in the diagnostic testing, which incorporates elements of questionnaires within a chatbot. The Telegram data collection bot @Lex_lang_bot differs from its nearest analogues that are designed to ascertain only language proficiency levels (e.g., Pearson Versant Test, MyVocab, Preply, VST, Games & Quizzes by Merriam-Webster, English Vocabulary Level Test by Oxford). However, these alternatives fail to consider the user's interests and the idiosyncrasies of their native language. In other words, the questionnaire and diagnostic testing take into account individual speech features in the native language, which lead to both destructive and constructive outcomes in the target language. These outcomes manifest at all levels of language and in all types of speech activity. The most typical interferent language errors of native speakers of Russian when learning English include so-called "wrong word" mistakes (e. g., the translation of concepts by means of calcification without consideration of the semantic differences between synonyms) and grammatical errors, including morphological (e.g., incorrect use of the article), syntactic (e.g., translation of participle and de-participle turns through adjectival sentences, word order in a sentence), and punctuation errors (e.g., separation of an adjectival sentence from the main one by a comma in a postposition; punctuation marks in direct speech).

A glossary (utilising English as a case study) serves as a dictionary and preliminary database for the pre-training of the language model. The glossary comprises phrases, phrasal verbs, idioms and individual lexemes at levels C1 to C2. The glossary was compiled by us in accordance with the principles set forth by the Oxford and Cambridge lexical minimum, the 18 thematic glossaries that were subjected to analysis, and with the addition of illustrative examples of word usage drawn from the Merriam-Webster corpus. As part of the research process, a comprehensive examination of the current state of computer lexicography was conducted, specifically focusing on the analysis of electronic resources with modular dictionary databases. A sub-corpus was constructed based on the analysis of the modular vocabulary databases for the purpose of training a model for the selection of personalised language material. The vocabulary database is updated on an ongoing basis within the programme as a result of machine learning on new user data and feedback.

The algorithm for processing the user's native language speech comprises several stages. First, the text preparation stage (text cleaning and tokenisation) was carried out using NLTK and string libraries, while lemmatisation was performed using Pymorphy2. Second, a personal list of language units (200-300 words and expressions) was compiled using modern large language models through vectorisation based on 30-50 most frequent content words usage from the user's texts. Third, the latent semantic analysis (LSA) is used to identify the relationships between a given set of documents and the terms that occur within them. Consequently, the output provides the answers to two key questions: firstly, how often a term occurs in a document (frequency embedding score), and secondly, how important the word is (frequency inversion, or inverse frequency score). A number of common technologies are employed in conjunction with pre-trained transformers:

- The Python programming language, which was invented in the early 1990s and has been developed since 2019 in conjunction with the advancement of neural networks, is one of the simplest programming languages and is therefore particularly well-suited to machine learning.
- The integrated development environment (IDE) in which code can be written. Such a Jupyter notebook may take the form of software for a computer, such as PyCharm, or an online equivalent, such as Google Colab or Yandex Datasphere, which are designed for researchers.
- A database (DB) for natural language processing tasks is a linguistic corpus or subcorpus consisting of documents. Here, a document is defined as a set of tokens that belong to one semantic unit. This may take the form of a sentence, a comment, or a user post. The most famous datasets in Russian are the Deepavlov datasets. For example, SQLite is a database management system (DBMS) that is used for relational databases (i.e., tables) based on SQL (Structure Query Language).
- The open-source libraries Pandas (a basic library for data preparation) and NumPy (a basic library for model training, on which the ScikitLearn package is developed) are typically used in conjunction for machine learning in general. The most popular international library for text processing tasks is NLTK (Natural Language Toolkit), which offers modules for the Russian language and can be used with Yandex software for MyStem natural language processing.

Hence, the automation of the selection of materials in accordance with the specified criteria involves the collection of user data through the utilisation of questionnaire elements, followed by the processing of this data through the access and utilisation of open large language models based on the author's modular corpus. In order to process the collected big data, an algorithm was developed for use in our study. As illustrated in Fig. 1, the algorithm initially accesses text data preparation libraries via application programming interfaces (APIs). These include the Python string library, the NLTK library modules for tokenisation, stop word and frequency cleaning, and Pymorphy2 for text normalisation. Subsequently, it utilises large open-source language models from GitHub for text vectorisation and natural language analysis (Figure 1):

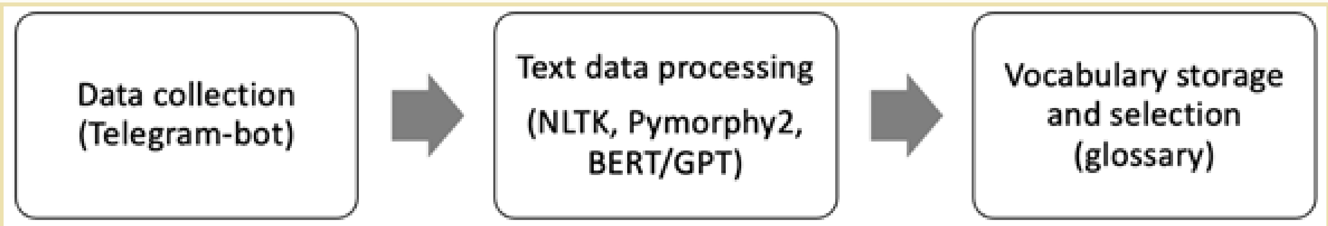

Figure 1 Algorithm for placement test text data processing and vocabulary range selection

After the students got the results of the placement test and automatic processing of their vocabulary range in the native language, they were offered a series of “personal quest” tasks to enlarge their vocabulary following the personalised foreign language learning path. This problem-based task is conducted as part of the classroom curriculum with the assistance of mobile devices.

The objective of the task is to provide students with the opportunity to engage with the material on a given topic for a designated period of time, followed by the completion of reception tasks utilising personally selected and automatically adapted texts. This approach enables students to develop the necessary linguistic proficiency to express their position or interests on subtopics within the general topic. Subsequently, a discussion (roundtable, brainstorming, debates, etc.) is held, during which learners share new information or defend their position using the vocabulary they have acquired.

RESULTS

Let us now examine in greater detail the process of developing and conducting this assignment format, namely the adaptation of texts. Discourse analysis represents the most appropriate and efficient instrument for identifying the boundaries of adaptation of authentic texts. Discourse analysis conceptualises the relationship between language and society as a dialectical process, whereby on the one hand, society exerts an influence on language, and on the other, language in turn shapes and is shaped by society. The process of adapting authentic texts is one that requires a significant expenditure of energy, and the use of automatic tools, such as Oxford Text Checker, can facilitate a more efficient and streamlined approach to this endeavour. Neural networks permit the automatic analysis and determination of the level of the current text by analysing a number of parameters, including the overall level of text complexity according to the Common European Framework of Reference (CEFR), the complexity of syntactic constructions, the complexity of grammatical constructions, the variety of vocabulary, the level of vocabulary complexity according to Cambridge University Press, the British National Corpus, the Corpus of Contemporary American English, the Academic Word List, the frequency of meta-discourse markers and the propositional density. In the process of adapting texts, educators have the option to incorporate or exclude lexical units of elevated complexity, thereby enabling the adjustment of the linguistic level of the materials provided to students.

With regard to the collated quantitative data from the study, the results of the placement and final tests are included for the experimental ($n = 10$) and control ($n = 9$) groups were analysed. In order to ascertain the statistical significance of the presented figures, Student's t-test for placement and final results, and Mann-Whitney U-test for final results in experimental and control groups were conducted. The statistical analysis indicates that the comparison of the final test results in the experimental and control groups reveals the grade growth in the experimental group by on average 10% (statistically significant, according to t-test) compared to 2% (statistically insignificant, according to t-test) in the control group. The statistical significance of the results of the experiment is confirmed by the conducted U-test, where $p<0.05$ (Table 1).

Table 1

The quantitative data obtained from the experiment

	PLACEMENT TEST (mean value)	FINAL TEST (mean value)	Student's t-test	Mann-Whitney U-test
Experimental group	80.7	91.3	t-test = 3.968 (>2.262) $p = 0.003$	U-test = 0 0 ≤ 20 $p < 0,05$
Control group	76.8	78.8	t-test = 1.134 (<2.306) $p = 0.290$	

DISCUSSION

A review of the literature on personalised learning paths revealed a dearth of analysis on the potential of AI and NLP in the automatic selection of personalised learning materials within the formal educational system [15]. The studies concentrate on psychometric parameters that pertain to user behaviour in virtual learning environments (VLEs). Recent literature since 2022 predominantly addresses AI and ChatGPT applications, including data-driven approaches [9], self-regulated learning [16], project-based learning [17], and challenges posed by GPT [18] and AI-detection methods [19]. Nonetheless, these studies also highlight new opportunities arising from GPT-based tasks [20]. For instance, AI systems effectively support factual knowledge acquisition like vocabulary and grammar but may struggle with more complex skills such as persuasive writing or cultural nuances without careful design integration [21]. Learner motivation also affects outcomes, with high-achieving students often benefiting significantly more from AI systems, while others may disengage if systems lack sufficient interactivity or a human-like approach [22]. Thus, sustained effectiveness requires consistent integration and alignment with learner goals. Comparative evaluations reveal that AI systems featuring advanced NLP capabilities, especially conversational agents allowing natural language interactions, yield deeper language acquisition compared to limited interaction platforms [23].

Accordingly, this research was conducted within the framework of NLP tools for the analysis of students' texts in their native language, with the objective of developing and practising their personalised vocabulary range in the foreign language.

In order to achieve the aforementioned personal quest, it is essential to adhere to the following requirements: the provision of a comprehensive and diverse range of educational materials in both written and oral formats, ensuring the authenticity and relevance of the materials; the ability to comprehend the linguistic phenomena that are necessary for the completion of the task; the establishment of an environment conducive to communication in the target language throughout the task [24]; the availability of resources that can facilitate the comprehension of the content and direct the students' attention towards the analysis of language forms. Technical difficulties, such as an unstable internet connection or a low battery charge on a mobile device, may prove to be a significant limitation to the task. It is therefore recommended that the independent format be replaced by working in pairs, whereby students can jointly prepare a message using the same device.

The results of the conducted experiment demonstrate that the formation of a personalised foreign language learning path enables the consideration of individual strengths and weaknesses in the learner's vocabulary range, facilitating the adaptation of the learning process to the personal requirements. Furthermore, the addition of authentic materials, tailored with the assistance of neural networks, allows for the updating of course materials. Nevertheless, the aforementioned limitations of contemporary AI technologies [25] necessitate substantial enhancements and further investigation for the utilisation of available big data and the capabilities of neural networks to facilitate the development of a digital tutor as an electronic assistant for the establishment of a personalised learning path. As NLP is only one segment of the broader field of computational linguistics, it encompasses computational methods in all areas of language and communication: computational semantics; computational syntax (concerning grammatical frameworks); computational pragmatics (focusing on discourse analysis). Our analysis of texts has been conducted within the framework of computational semantics, exclusively in the written register. However, the application of syntax and pragmatics frameworks has the potential to facilitate further analysis in both written and oral registers.

Empirical studies to date generally support the effectiveness of AI-driven personalized learning in improving certain educational outcomes, though results can vary by context and implementation. One consistent finding is that personalized AI systems tend to **increase learner engagement and motivation**. In language learning, for example, Ouyang et al. found [9] that students using an AI-powered app exhibited higher engagement on multiple dimensions (behavioral, cognitive, and affective) than those who did not, attributing this to the app’s adaptive and interactive nature that kept learners challenged but not frustrated. Gamification elements commonly integrated into these systems (points, badges, immediate feedback) also play a role in sustaining motivation, especially in self-study contexts. Another study reported that AI-assisted personalized learning had moderately positive effects on student performance, with improvements in test scores and homework completion rates compared to traditional instruction. These positive effects are often linked to the timely feedback and tailored support that AI systems provide, which can address misunderstandings immediately and prevent learners from developing persistent knowledge gaps.

In terms of **learning outcomes**, personalized learning platforms have shown gains in various language competencies. Some adaptive systems focusing on vocabulary acquisition have led to faster vocabulary growth and better retention, as the AI’s spacing algorithms present review items at optimal intervals for memory consolidation. For grammar learning, AI tutors that analyze free-form student input (like an essay or spoken response) and provide corrections have helped learners improve their accuracy more than generic workbook exercises. Comparative studies of AI-tutored groups vs. control groups frequently show that the AI-tutored groups perform as well as or better than their peers on standardized assessments, despite having more self-paced learning. For instance, students practicing writing with an AI feedback tool might achieve similar improvement in essay scores as those who get human feedback, which is promising for scaling individualized feedback to large classes. In the Duolingo EFL study, beyond engagement, the **willingness to communicate** (a key soft outcome in language learning) was significantly higher in the AI group, indicating that the personalized practice and confidence-building exercises translated into greater readiness to use the language in real contexts.

Moreover, AI-driven personalization can contribute to **efficiency in learning** [26]. Several studies have noted that when instruction is tailored to the individual, students often reach mastery in less time. A literature review [26] on reinforcement learning in educational systems documented cases where AI policies optimized practice schedules, resulting in students completing learning tasks faster while maintaining performance. Similarly, in corporate training scenarios, organizations observed reduced training times – for example, employees could finish required learning modules sooner because the AI skipped content they already knew and focused on areas they needed to learn, without redundant lessons. In one case, an adaptive learning system cut down the time employees spent in training by a significant margin while improving their post-training assessment scores, indicating a more efficient learning process.

However, results are not universally positive; some studies show **mixed outcomes**. The effectiveness of AI personalization can depend on the subject matter, the quality of the AI system, and the context of use. For example, an AI system might effectively improve factual knowledge (like vocabulary or grammar rules) but be less impactful on higher-order skills (like persuasive writing or cultural nuances in communication) without careful design. Additionally, learner outcomes can vary – high-achieving or self-motivated students might take great advantage of a self-paced AI tutor and soar, whereas other students might disengage if they feel the AI lacks the human touch or if technical issues arise. It has been observed that the novelty of an AI tool can boost engagement initially, but sustained usage requires that the tool genuinely addresses learner needs and aligns with their goals. Hence, while many studies celebrate improved test scores or engagement metrics, a few report no significant difference between AI-personalized instruction and conventional methods, particularly if the traditional instruction was already high-quality or if the AI was not fully integrated into the learning process (e.g., used sporadically rather than as a core component).

When comparing different AI-powered tools, research suggests that **effectiveness varies by design features**. Tools that allow more interactivity (such as conversational agents where learners can ask questions in natural language) tend to receive positive feedback for making learning feel more human-like. On the other hand, systems that rely on multiple-choice adaptations without richer feedback may be perceived as less helpful. A comparative analysis of AI tutoring systems in one study found that those employing advanced NLP (able to handle open responses) led to deeper learning gains in language tasks than systems with limited input options

The role of human teachers remains essential in all successful implementations. Blended approaches—where instructors leverage AI analytics to inform teaching strategies and provide personalized encouragement—consistently outperform either AI-only or human-only approaches. Researchers emphasize that AI tools are most effective when thoughtfully integrated into pedagogical frameworks, such as when teachers receive AI-generated insights about student progress and intervene in personalized ways. This human-AI partnership represents the most promising direction for educational technology integration [27].

Overall, recent empirical studies present an optimistic picture of AI-driven personalized learning in language education, with demonstrated improvements in engagement, proficiency, communication skills, and learning efficiency. However, even positive studies caution against viewing AI as a comprehensive solution without appropriate implementation. The discussion consistently returns to the importance of pedagogical alignment and teacher involvement, ensuring that AI serves as a tool to enhance rather than replace effective language teaching. Our research contributes to this growing body of evidence by demonstrating how NLP analysis of native language proficiency can inform personalized foreign language vocabulary acquisition. Future research should explore extending these techniques to other language dimensions while addressing the practical and technical limitations identified in our study.

Challenges and Ethical Considerations

Despite its promise, AI-driven personalized learning comes with significant challenges and ethical considerations. One major concern is **data privacy and security**. Personalized learning platforms rely on collecting vast amounts of learner data – from demographics and learning performance to behavioral logs. Protecting this sensitive data is paramount. There are risks of data breaches or misuse of data if proper safeguards are not in place. For instance, a language learning app may record a student's voice to evaluate pronunciation; storing and processing such data requires compliance with privacy regulations (like GDPR in Europe or FERPA for student data in the U.S.). Ensuring informed consent is another aspect: learners (or their guardians) should know what data is being collected and how it will be used. Ethical deployment of AI in education demands transparent data practices, secure storage (encryption, access controls), and clear data retention and deletion policies. If a learner decides to stop using a platform, they should have the right to have their personal data deleted. Many educational institutions now require AI vendors to undergo privacy audits and to provide options for data anonymization to protect student identities [28].

Another challenge is **algorithmic bias and fairness** in AI systems. AI models learn from data, and if the training data contains biases, the AI may perpetuate or even amplify those biases. In the context of personalized learning, this could mean an AI tutor that unknowingly provides less encouragement or fewer challenging tasks to certain groups of students based on gender, ethnicity, or other factors, because of biased historical data. For example, if an AI system was trained mostly on data from native English-speaking students, it might not adapt as effectively for learners from different language backgrounds, thus disadvantaging them. Bias can also appear in content recommendations: a personalized system might consistently steer underrepresented learners toward "easier" content, creating a self-fulfilling prophecy where those learners don't get the same growth opportunities as others. Addressing this requires careful attention in both development and deployment. Techniques such as **data bias audits** (inspecting training data for skewed representations) and **fairness-aware machine learning** (adding constraints or adjustments to algorithms to ensure fair outcomes) are being advocated. Moreover, involving diverse teams in the design of educational AI can help – when developers and educators from varied backgrounds collaborate, they are more likely to catch biases and ensure the tool serves a broad learner population.

The **transparency of AI systems** is also an ethical concern. Many AI-driven tools, particularly those based on deep learning, operate as "black boxes" that make it hard to interpret why a certain recommendation or score was given. In education, lack of transparency can erode trust. Both learners and teachers might be skeptical of an AI tutor's guidance if they don't understand the rationale. For instance, if an AI system advises a student to review a particular grammar topic, the student and teacher might want to know whether this suggestion is based on the student's recent mistakes, comparison with peers, or some other factor. Explainable AI (XAI) is a growing area focusing on making AI decisions more interpretable [29]. Ethically, providing explanations ("The system suggested this lesson because you struggled with related exercises last week") can help users trust and effectively use AI recommendations, and also allows educators to double-check the AI's appropriateness.

Ethical use of AI in assessment is another consideration. AI models can automatically grade essays or evaluate speaking skills, but if they are biased or inaccurate, they could unfairly affect student grades. There have been cases where automated essay scoring systems were fooled by verbose nonsense or penalized non-native writing styles. Ensuring these models are used as support rather than sole arbiters of assessment is important to maintain fairness.

Moreover, there is the question of **dependency and student agency**. If learners become too dependent on AI guidance, they might not develop self-regulation skills or might lose confidence in learning without AI support. Striking a balance between providing personalized help and encouraging independent problem-solving is a pedagogical and ethical challenge [30]. Some educators worry that over-reliance on AI could diminish the role of social learning and critical thinking if not carefully moderated.

Teacher perspectives and ethical integration also come into play. Teachers may feel challenged or even threatened by AI systems if they are introduced without proper training or clear purpose. It is crucial to involve educators in the deployment of personalized learning tools – both in interpreting AI outputs and in overriding the system when necessary. Research has shown that teachers' acceptance of AI is higher when they see it as aligned with their pedagogical aims and when they retain control (a concept known as human-in-the-loop oversight). Ethically, maintaining human oversight ensures that the AI serves as a tool for human educators rather than replacing their judgment. For example, an AI might flag a student as at-risk of dropping out based on low engagement metrics; a teacher or counselor should then intervene, verify, and decide how to help that student, bringing in empathy and contextual knowledge that an AI lacks.

Lastly, there are **broader equity issues**. While AI personalized learning has the potential to democratize education (by bringing tutoring to those who might not afford human tutors), it also could widen gaps if not accessible to all. AI systems require technology (devices, internet) and produce benefits mainly if used properly. Under-resourced schools might not have the infrastructure, or students at home might not have reliable internet to use an AI language app extensively. If only some students get access to advanced personalized tools, that could increase educational inequity. Ensuring equitable access and designing AI for low-bandwidth or offline use are important ethical considerations to truly use AI for the benefit of all learners.

In summary, the challenges and ethical considerations surrounding AI-driven personalized learning include safeguarding data privacy, eliminating biases to ensure fairness, making AI systems transparent and accountable, maintaining the human element in education, and ensuring equal access. Addressing these issues is essential for the sustainable and responsible adoption of AI in language learning and education in general. Researchers and practitioners are increasingly cognizant of these factors, and recent literature often accompanies reports of AI's benefits with discussions on mitigating risks. By proactively tackling privacy, bias, and transparency concerns, the educational community can work toward AI systems that are not only effective but also trustworthy and just.

CONCLUSION

Thus, in order to address the issue of material selection through the utilisation of artificial intelligence and big data technologies, a set of selection criteria has been established. In accordance with the aforementioned criteria, we have devised a chatbot that is able to gather learner data through the utilisation of automated diagnostic testing and questionnaires. Moreover, we have developed code based on open-source big model languages with the objective of processing the obtained data. Additionally, we have pre-trained our neural network using a pre-built modular corpus. The experiment of personalised vocabulary learning paths based on NLP has proven our hypothesis concerning the effectiveness of the automatic analysis of texts in one's native language with the aim of tailoring their personalised vocabulary range in the foreign language.

AI-driven personalized learning in language education sits at the intersection of pedagogical innovation and technological advancement. The integration of constructivist and self-directed learning principles into AI systems has allowed for learning experiences that adapt to each student, promoting active engagement and autonomy. Breakthroughs in NLP, especially large language models like GPT-4, have endowed educational software with unprecedented language understanding and generation capabilities, enabling interactive and responsive tutoring that can mimic one-on-one human instruction. Case studies from recent years demonstrate tangible benefits of these approaches: higher student motivation, faster learning gains, and improved proficiency, whether in formal classrooms, online platforms, or professional training programs. At the same time, this expanded analysis makes clear that maximizing the potential of AI in personalized learning requires careful consideration of research methodology and ethical safeguards. Robust experimental studies and qualitative inquiries provide evidence of effectiveness while also highlighting where AI tools need refinement.

As we look ahead, the future of AI-driven personalized language learning will likely involve even more seamless collaboration between humans and AI. Teachers and AI tutors will work in tandem – AI handling routine personalization and data analysis, and teachers focusing on complex cognitive skills, empathy, and mentorship. Ongoing research is needed to refine models (ensuring they are fair, explainable, and aligned with educational values) and to explore long-term impacts on learning outcomes. With at least ten recent studies affirming the benefits and analyzing the challenges of AI in education, there is a growing consensus that when implemented thoughtfully, AI-driven personalized learning can significantly enhance language education. However, realizing this promise universally will depend on addressing current limitations and ethically integrating these technologies. By combining insights from pedagogy, computational advances, and empirical evaluations, stakeholders can ensure that AI becomes a powerful ally in fostering effective and inclusive personalized learning experiences.

FUNDING

This work was done with the support of MSU Program of Development, Project No 23-SCH02-26.

REFERENCES

1. Maphalala, M. C., & Ajani, O. A. (2025). Leveraging artificial intelligence as a learning tool in higher education. *Interdisciplinary Journal of Education Research, 7*(1), 1–16. DOI: 10.38140/ijer-2025.vol7.1.01
2. Zhang, L., Basham, J. D., & Yang, S. (2020). Understanding the implementation of personalized learning: A research synthesis. *Educational Research Review, 31*, 100339. https://doi.org/10.1016/j.edurev.2020.100339
3. Sysoev, P. V., & Filatov, E. M. (2023). Method of the development of students' foreign language communication skills based on practice with a chatbot. *Perspektivy nauki i obrazovania = Perspectives of Science and Education, 63* (3), 201–218. https://doi.org/10.32744/pse.2023.3.13
4. Shemshack, A., & Spector, J. M. (2021). A comprehensive analysis of personalized learning components. *Journal of Computers in Education, 8*(4), 485–503. https://doi.org/10.1007/s40692-021-00188-7
5. Sitarov, V. A., Glagolev, S. N., & Shutenko, A. I. (2010). The evolution of the idea of integrated education in pedagogical science and practice. *Znanie. Ponimanie. Umenie,* (1), 152-160. https://zpu-journal.ru/zpu/contents/2010/1/Sitarov~Glagolev~Shutenko/23_2010_1.pdf
6. Sukhomlinsky, V. A. (1979). Selected works: In 5 vols. (T. 2). (D. A. others, Ред.) Kiev: Glad.school.
7. Khutorskoy, A. V. (2005). Methodology of personality-oriented education. How to tech everyone differently? Moscow: Vlados-Press.
8. Vdovina, S. A., & Kungurova, I. M. (2013). The nature and direction of the individual educational trajectory. *Naukovedeniye,* (6), 1-8. Availble at: https://naukovedenie.ru/PDF/40PVN613.pdf
9. Ouyang, Z., Jiang, Y., & Liu, H. (2024). The Effects of Duolingo, an AI-Integrated Technology, on EFL Learners' Willingness to Communicate and Engagement in Online Classes. *The International Review of Research in Open and Distributed Learning, 25*(3): 97–115. https://doi.org/10.19173/irrodl.v25i3.7677
10. Mon, B. F., Wasfi, A., Hayajneh, M., Slim, A., & Abu Ali, N. (2023). Reinforcement Learning in Education: A Literature Review. *Informatics, 10*(3): 74. https://doi.org/10.3390/informatics1003007
11. Liu, S., Guo, X., Hu, X., & Zhao, X. (2023). Advancing Generative Intelligent Tutoring Systems with GPT-4: Design, Evaluation, and a Modular Framework for Future Learning Platforms. *Electronics, 13*(24): 4876. https://doi.org/10.3390/electronics13244876
12. Wang, J., Hussain, Y., & Mao, C. (2025). Artificial Intelligence-Driven Personalized Learning: Psychological Implications and Educational Outcomes. *International Journal of Education Humanities and Social Sciences, 2*(1): 1–10. https://doi.org/10.70088/3wnrs278
13. Katiyar, N. (2021). AI-Driven Personalized Learning Systems: Enhancing Educational Effectiveness. *KUEY, 30*(5): 4961–4972. (Contains discussion of learning theories and case studies)
14. Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education, 16*(1): 39. https://doi.org/10.1186/s41239-019-0171-0
15. Basham, J. D., Hall, T. E., Carter Jr, R. A., & Stahl, W. M. (2016). An operationalized understanding of personalized learning. *Journal of Special Education Technology, 31*(3), 126–136. https://doi.org/10.1177/0162643416660835
16. Liu, D. Y. T., Bartimote-Aufflick, K., Pardo, A., & Bridgeman, A. J. (2017). Data-driven personalization of student learning support in higher education. In *Learning analytics: Fundaments, applications, and trends,* 143–169. https://doi.org10.1007/978-3-319-52977-6_5
17. Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in psychology, 8*, 422. https://doi.org/10.3389/fpsyg.2017.00422
18. Pogorskiy, E. (2015). Using personalisation to improve the effectiveness of global educational projects. *E-Learning and Digital Media, 12* (1), 57–67. https://doi.org/10.1177/2042753014558378
19. Cingillioglu, I. (2023). Detecting AI-generated essays: The ChatGPT challenge. *The International Journal of Information and Learning Technology, 40*(3), 259–268. https://doi.org/10.1108/IJILT-03-2023-0043
20. Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 228–239. https://doi.org/10.1080/14703297.2023.2190148
21. Creely, E. (2024). Exploring the Role of Generative AI in Enhancing Language Learning: Opportunities and Challenges. *International Journal of Changes in Education,* 158–167. https://doi.org/10.47852/

bonviewIJCE42022495
22. Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal,* 537–550. https://doi.org/10.1177/00336882231162868
23. Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies,* 13943–13967. https://doi.org/10.1007/s10639-023-11742-4
24. Fahim, M., & Ebadi, S. (2020). The effect of AI-based instruction on learners' engagement and motivation in language learning. *Journal of Language Teaching and Research, 11*(5), 825–834.
25. Li, Y. (2024). Study on the Role and Influence of AI in Personalized Learning for Second Language Acquisition. *Journal of Education, Humanities and Social Sciences, 38*, 125–131. https://doi.org/10.54097/wfe3gk31
26. Wei, L. (2023). Artificial intelligence in language instruction: impact on English learning achievement, L2 motivation, and self-regulated learning. *Frontiers in Psychology, 14*, Article 1261955. https://doi.org/10.3389/fpsyg.2023.1261955
27. Panwale, S. B., & Vijayakumar, S. (2025). Evaluating AI-Personalized Learning Interventions in Distance Education. *The International Review of Research in Open and Distributed Learning, 26*(1), 157–174. https://doi.org/10.19173/irrodl.v26i1.7813
28. AlAqlobi, O., Alduais, A., Alasmari, M., & Qasem, F. (2024). Artificial Intelligence in Language Acquisition: A Balancing Act of Potential and Challenges. *Forum for Linguistic Studies, 6*(6), 1103–1122. https://doi.org/10.30564/fls.v6i6.7524
29. Divekar, R. R., Drozdal, J., Chabot, S., Zhou, Y., Su, H., Chen, Y., Zhu, H., Hendler, J. A., & Braasch, J. (2022). Foreign language acquisition via artificial intelligence and extended reality: Design and evaluation. *Computer Assisted Language Learning, 35*(9), 2332–2360. https://doi.org/10.1080/09588221.2021.1879162
30. Hung, H. T., Yang, J. C., Hwang, G. J., Chu, H. C., & Wang, C. C. (2018). A scoping review of research on digital game-based language learning. *Computers & Education, 126*, 89–104. https://doi.org/10.1016/j.compedu.2018.07.001
31. Bahman Zohuri (2023). Artificial super intelligence (ASI) the evolution of AI beyond human capacity. *Current trends in engineering science (CTES),* 3 (7). DOI: 10.54026/CTES/1049

Авторы

Авраменко Анна Петровна
(Россия, Москва)
Кандидат педагогических наук, доцент кафедры лингвистики и информационных технологий
МГУ им. М. В. Ломоносова
E-mail: avram4ik@gmail.com
ORCID ID: 0009-0009-3004-2192

Хван Мария Владимировна
(Россия, Москва)
Преподаватель кафедры лингвистики и информационных технологий факультета иностранных языков и регионоведения
МГУ им. М. В. Ломоносова
(Россия, Москва)
E-mail: mkhvan@yandex.ru
ORCID ID: 0009-0008-8327-6788

Терновский Владимир Владимирович
(Россия, Москва)
Кандидат физико-математических наук, доцент, доцент кафедры вычислительных методов факультета вычислительной математики и кибернетики
МГУ им. М. В. Ломоносова
E-mail: vladimir@chatroulette.com
ORCID ID: 0000-0001-5817-0211
Scopus Author ID: 58844603500

Authors

Anna P. Avramenko
(Russia, Moscow)
PhD, Associate Professor
Department of Linguistics and Information Technologies
Faculty of Foreign Languages and Area Studies
Lomonosov Moscow State University
E-mail: avram4ik@gmail.com
ORCID ID: 0009-0009-3004-2192

Mariya V. Khvan
(Russia, Moscow)
Lecturer
Department of Linguistics and Information Technologies
Faculty of Foreign Languages and Area Studies
Lomonosov Moscow State University
E-mail: mkhvan@yandex.ru
ORCID ID: 0009-0008-8327-6788

Vladimir V. Ternovskii
(Russia, Moscow)
Associate professor, PhD
Associate Professor, Department of Computational Methods
Faculty of Computational Mathematics and Cybernetics
Lomonosov Moscow State University
E-mail: vladimir@chatroulette.com
ORCID ID: 0000-0001-5817-0211
Scopus Author ID: 58844603500

Вклад авторов

Авраменко А. П.: концептуализация, методология, создание рукописи и ее редактирование
Хван М. В.: проведение исследования, создание черновика рукописи и ее редактирование
Терновский В. В.: разработка программного обеспечения, формальный анализ

Author's contribution

Anna P. Avramenko: Conceptualization, Methodology, Writing - Review & Editing
Mariya V. Khvan: Investigation, Writing - Original Draft
Vladimir V. Ternovskii: Software, Formal analysis

Авторы заявляют об отсутствии конфликта интересов

The author's declared no conficts of interest