

## Article

# Custom-Trained Large Language Models as Open Educational Resources: An Exploratory Research of a Business Management Educational Chatbot in Croatia and Bosnia and Herzegovina

Nikša Alfrević <sup>1,\*</sup> , Daniela Garbin Praničević <sup>1</sup>  and Mirela Mabić <sup>2</sup> <sup>1</sup> Faculty of Economics, Business and Tourism, University of Split, 21000 Split, Croatia; daniela@efst.hr<sup>2</sup> Faculty of Economics, University of Mostar, 88000 Mostar, Bosnia and Herzegovina; mirela.mabic@ef.sum.ba

\* Correspondence: nalf@efst.hr

**Abstract:** This paper explores the contribution of custom-trained Large Language Models (LLMs) to developing Open Education Resources (OERs) in higher education. Our empirical analysis is based on the case of a custom LLM specialized for teaching business management in higher education. This custom LLM has been conceptualized as a virtual teaching companion, aimed to serve as an OER, and trained using the authors' licensed educational materials. It has been designed without coding or specialized machine learning tools using the commercially available ChatGPT Plus tool and a third-party Artificial Intelligence (AI) chatbot delivery service. This new breed of AI tools has the potential for wide implementation, as they can be designed by faculty using only conventional LLM prompting techniques in plain English. This paper focuses on the opportunities for custom-trained LLMs to create Open Educational Resources (OERs) and democratize academic teaching and learning. Our approach to AI chatbot evaluation is based on a mixed-mode approach, combining a qualitative analysis of expert opinions with a subsequent (quantitative) student survey. We have collected and analyzed responses from four subject experts and 204 business students at the Faculty of Economics, Business and Tourism Split (Croatia) and Faculty of Economics Mostar (Bosnia and Herzegovina). We used thematic analysis in the qualitative segment of our research. In the quantitative segment of empirical research, we used statistical methods and the SPSS 25 software package to analyze student responses to the modified BUS-15 questionnaire. Research results show that students positively evaluate the business management learning chatbot and consider it useful and responsive. However, interviewed experts raised concerns about the adequacy of chatbot answers to complex queries. They suggested that the custom-trained LLM lags behind the generic LLMs (such as ChatGPT, Gemini, and others). These findings suggest that custom LLMs might be useful tools for developing OERs in higher education. However, their training data, conversational capabilities, technical execution, and response speed must be monitored and improved. Since this research presents a novelty in the extant literature on AI in education, it requires further research on custom GPTs in education, including their use in multiple academic disciplines and contexts.

**Keywords:** specialized machine learning; custom Large Language Models (LLMs); Artificial Intelligence (AI); Open Educational Resources (OERs)



**Citation:** Alfrević, N.; Praničević, D.G.; Mabić, M. Custom-Trained Large Language Models as Open Educational Resources: An Exploratory Research of a Business Management Educational Chatbot in Croatia and Bosnia and Herzegovina. *Sustainability* **2024**, *16*, 4929. <https://doi.org/10.3390/su16124929>

Academic Editor: Vasiliki Brinia

Received: 8 April 2024

Revised: 5 June 2024

Accepted: 6 June 2024

Published: 8 June 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The motivation for this study can be found in the extremely quick development of Artificial Intelligence (AI) tools and their application in higher education, which especially applies to Large Language Models (LLMs). This study has followed the introduction of a highly anticipated LLM feature, enabling end-users of OpenAI tools, such as ChatGPT 4, to create custom-trained Generative Pretrained Transformer (GPT) chatbots. Its immediate objective is to determine how custom-trained LLMs could serve the role of Open Educational Resources (OERs) in higher education, specifically within e-learning. As far as we know,

no similar studies in the extant literature might inform education staff and administrators in higher education of the potential strengths and challenges of this emerging technology.

Introducing commercial LLM functionality to train custom LLM models focused on a specialized field provides new opportunities for implementing Artificial Intelligence (AI) without coding or developing machine learning and algorithm competencies. Such developments make AI tools accessible to a wide range of educators at all levels of education, aiming to develop virtual teaching assistants with specialized training. Although such a proposition sounds attractive, it is still unknown how much training by field-specific data will contribute to the success of custom LLMs in educational settings [1].

It is important to note that the discipline-specific training of chatbots does not interfere with the generic capabilities of LLMs used by their underlying models. For example, ChatGPT provides the full functionality of its ChatGPT 4 model in the Custom GPT mode, which can be trained and used for a domain-specific purpose. At the time, when the paper had been conceptualized, and the initial version was written, opportunities for implementing specialized AI chatbots were relatively limited. Custom GPTs were not made available to free users of the ChatGPT service, who were restricted to the less powerful ChatGPT 3.5 model. Even with the relatively modest cost of the membership plan, such charges made it impossible for a broader audience of life-long learners to afford AI tools, especially in low-income countries. Other commercial LLMs (including Google Gemini, Anthropic Claude, and others) did not offer the functionality for custom training of LLM-based chatbots.

The other option available to education professionals was related to building a ChatGPT-powered assistant, which can be performed easily without programming skills. At the time of revising this paper, this option is still viable but requires a significant institutional budget. Namely, the educator is considered a software developer and charged per usage of their AI chatbot.

These requirements limited the implementation of custom-trained LLM chatbots without using one of the open-source AI models and datasets. Following such a route ensured that the specialized AI assistant served as a true Open Educational Resource (OER). However, a relatively high level of ICT competencies was required for its development. Therefore, when conceptualizing this paper, we decided to use a commercially available service (Botsonic) (a free test version of our chatbot is available at <https://bot.writesonic.com/share/bot/f49db0d8-7e43-4d93-81fa-4e916fb88383>, accessed on 7 April 2024), enabling individual content creators to create custom ChatGPT-based AI chatbots and freely test them before mass deployment.

AI tools and technologies are rapidly developing. The current OpenAI model ChatGPT 4 Omni (GPT-4o), launched to the public in May 2024, introduced significant changes to the relevant licensing and usage practices. While the public attention was focused on the GPT-4o multimodal capabilities (i.e., reasoning and communicating across video, audio, and text content), the custom GPT licensing also changed. Since GPT-4o became available to free users in May 2024 (although with a usage cap), custom GPTs can now be accessed by free users. This makes the popular Open AI GPT platform relevant for developing and implementing custom-GPT-based OERs without third-party tools. Since the latest changes have been announced, our free test version of the chatbot (in Croatian language) has been also made available through the ChatGPT platform at <https://chat.openai.com/g/g-2NLU5DM8D-menadzarski-asistent>, accessed on 6 June 2024).

However, the dynamic nature of the AI field might mean that the current licensing and usage patterns might change again soon.

This study aims to determine the initial student perceptions of a custom-trained LLM to answer the following research question:

- How efficiently can custom Large Language Models (LLMs) serve as Open Educational Resources (OERs) in higher education, specifically business management?

Since we have not found a single study on custom GPT evaluation, we believe our results will be useful to direct further research in the field. The novelty of this research topic

and the rapidly evolving technological background of the AI field require that the research follows the AI developments as quickly as possible. This is the rationale for choosing a limited and applicative research question and using the exploratory approach in this study.

## 2. The Role of Large Language Models (LLMs) as (Open) Educational Resources

Using algorithms and statistical models to process and analyze extremely large datasets, LLMs can recognize patterns, make predictions, or generate responses based on the data they are trained on [2]. They enhance knowledge discovery by summarizing research papers, generating relevant research questions, and identifying relevant sources [3]. Unethical students can misuse this LLM capability to generate entire course assignments, which requires a new approach to academic assessment methods [4]. However, their output still requires human experts' evaluation [5], as LLMs might 'hallucinate', i.e., provide coherently and convincingly sounding but rather misleading or inaccurate information [6]. Human experts also need to ensure that AI technologies and content do not reflect the biases and ethical concerns that potentially exist in their training materials and that their implementation aligns with the values of diversity, fairness, and inclusivity [7]. Other ethical issues related to data privacy and transparency also require the attention of faculty and other actors in educational settings [8,9].

A widely recognized LLM capability is related to the personalization of learning materials and delivery methods, which can be customized to reflect individual needs, learning styles, and the level of professional knowledge [10,11]. Even individual users could find LLMs useful to support them in life-long learning (LLL), as endorsed by policies on both the European Union and global scales [12]. These LLM capabilities, if adequately implemented, could empower educators to tailor educational resources to specific topics or course requirements and enable faculty to devote more time to other relevant tasks [13].

LLMs can enhance the learning experience by providing personalized feedback, positively impacting student understanding and engagement [14]. This capability can provide guidance and support, helping students overcome barriers to learning and achieving academic milestones [15]. By serving as virtual teaching assistants, LLMs provide real-time assistance, query resolution, and guidance, enriching students' learning experiences, especially in online or hybrid learning environments [16]. They can enhance language skills and proficiency [17], enabling multicultural student groups to collaborate effectively [18]. Future custom LLMs may develop into comprehensive student support platforms, offering various services, from academic guidance to career counseling [19].

LLMs offer multiple opportunities to create Open Educational Resources (OERs), ultimately democratizing academic teaching and learning [20]. Due to their potential to generate high-quality educational content in various disciplines, LLMs have the potential to provide free access to educational materials to non-native speakers and students from different language and cultural environments. However, it seems that LLMs can only be used in a limited capacity, as the full automation of teaching and learning is still outside of their current level of development [21].

Educators may be able to adapt LLM-based OERs to fit their curriculum, teaching methodologies, and student needs, fostering a more personalized and engaging learning experience [22,23]. Custom training an LLM is the process of fine-tuning a pretrained LLM model with a specific corpus of training materials by using the following steps:

- Collection and preparation of data: A diverse corpus of training materials representing the intended learning outcomes' scope, including textbooks, class notes, scholarly articles, and other training materials. The corpus should be cleaned and preprocessed to ensure relevance and quality. Training materials must also formally assess the ethical aspects to ensure diverse perspectives and arguments.
- Model training and fine-tuning: After collection, the dataset has to be used to fine-tune a generic LLM, usually a downloadable open-source model, to make it context-aware for specific academic teaching and learning scenarios.

- Continuous evaluation and updating: After deployment, the model has to be monitored continuously and evaluated for performance. The feedback needs to be used for iterative improvement of the model to remain helpful, accurate, and effective.

Creating and implementing an Open Educational Resource (OER) based on a custom GPT involved the following:

- Ensuring accessibility and scalability: The AI chatbot has to be accessible to students and educators, which can be achieved by integrating it into frequently used educational platforms and LMSs or providing a simple web interface.
- Customizing model to individual needs: Key aspects of the AI chatbot acceptance are adaptability and customization to different educational needs and cultural contexts to fit particular teaching goals or the needs of students.
- Addressing ethical concerns and transparency: Among ethical concerns are privacy, consent bias, and lack of transparency, which need to be addressed from the beginning of the process, i.e., the training material selection stage.

LLM-based OERs may foster a sense of community among educators and learners, encouraging knowledge sharing and co-creation [24]. In addition, by incorporating real-time data and insights, LLMs could ensure the OER's relevance and effectiveness, promote knowledge exchange, encourage innovation, and reduce barriers to accessing high-quality educational resources [25]. By leveraging custom GPTs to generate OERs, educational institutions reduce the costs associated with content creation and acquisition and reallocate resources towards other educational priorities, such as student support services and infrastructure improvement [15].

### 3. Methods

We developed the analyzed chatbot in two subsequent steps. First, we developed a conceptual version of the chatbot using the OpenAI custom GPT functionality. We used the relevant OpenAI guidelines in formulating our prompting strategy [26], available in Appendix A. The chatbot was designed as an educational assistant in Croatian, aiming to assist 2nd-year students in the undergraduate programs of business schools in Croatian-speaking universities in Croatia and Bosnia and Herzegovina. Its domain-specific knowledge is based on the prescribed readings from business management authored by a group of regional scholars from Croatia and Bosnia and Herzegovina [27]. Since we could not test the chatbot using the ChatGPT environment due to the previously described limitations, we implemented it using the Botsonic platform. We distributed the link to its test implementation to four academic experts from Croatia and Bosnia and Herzegovina.

Our research approach used a mixed method design, combining qualitative and quantitative methods in two stages. This sequential exploratory method is especially valuable for emerging topics requiring in-depth exploration through qualitative methods before being more precisely assessed (using quantitative methods). The first stage involved collecting and reviewing interview information and informing the development of the following phase based on a previously verified instrument (survey).

In the first stage, we used nonrandom sampling to select four experts specialized in e-learning methods and their evaluation, business administration (management) teaching, and the role of technology in business marketing. The stratified purposeful approach used in expert selection (i.e., assuring that at least one local expert represents each of the fields considered relevant for the scope of the study) is considered appropriate for the exploratory nature of this study.

The experts were presented with five questions, developed using dimensions for evaluations of chatbots and other intelligent agents, previously suggested by Radziwill and Benton, i.e., Denecke et al. [28,29]. These were as follows:

1. How can the quality and accuracy of the information provided by the chatbot be ensured?
2. How does a chatbot handle complex inquiries?

3. How adaptable is the chatbot to user requirements?
4. How does the chatbot integrate with other tools for learning business management?
5. How do we measure the success of a chatbot in improving business management learning?

After transcribing the interviews (available in Appendix B), we used the experts' evaluation of our test chatbot to develop a summary report presented in the next part of the paper. We also used the results of this research stage to modify the items of the Bot Usability (BUS-15) Scale, originally developed and empirically verified for evaluation of Customer Relationship Management (CRM) AI chatbots [30,31].

BUS-15 Scale is grounded on the theoretical dimensions of perceived accessibility, quality of information, and functions provided by the chatbot, including conversational experience, perceived privacy and security, and response time. These consider Borsci et al.'s [30] importance given to the relevance and accuracy of responses and the extent to which chatbot keeps the conversational context. These are essential attributes required to meet user requirements, as detailed in the BUS-15 scale validation. The scale contains specific items measuring chatbot interaction, including response time. They consider the accessibility dimension, ensuring effective interaction with users, regardless of specific user characteristics or the context in which a chatbot is used. In addition, the human–computer interaction (HCI) context, concerned with trust and security aspects [32], addresses the inclusion of privacy and security aspects into chatbot assessment.

The modified survey (see Table 1) was translated to Croatian (see Appendix C) by two experts in academic translation, specialized in business management, and uploaded to the online Qualtrics XM platform for electronic data collection. All items were measured on the standard 5-item Likert scale, with answers ranging from '1—Strongly disagree' to '5—Strongly agree', comparable to the original BUS-15 Scale. No personally identifiable data, including student names or other IDs, were collected.

**Table 1.** Adaptation of the BUS-15 scale.

Original Bus-15 Scale Items	Modified Items of the Scale for Business Management Chatbot Evaluation (Based on Input from Four Academic Experts)
Perceived accessibility to chatbot functions	
1. The chatbot function was easily detectable. 2. It was easy to find the chatbot.	Items are not relevant since the AI chatbot was directly presented to students.
Perceived quality of chatbot functions	
3. Communicating with the chatbot was clear.	Adopted directly.
4. I was immediately made aware of what information the chatbot can give me.	Adopted directly.
5. The interaction with the chatbot felt like an ongoing conversation.	Adopted directly.
6. The chatbot was able to keep track of context.	Adopted directly.
7. The chatbot was able to make references to the website or service when appropriate.	The chatbot was able to make references to respected websites or scholarly articles. The chatbot provides answers aligned with the teaching materials (textbook) and other courseware. The chatbot can comment on complex topics and connect complex topics and issues from the field of business management. Information provided by the chatbot is of better quality than the generic ChatGPT, Google Gemini, or the other Large Language Model I use regularly.
8. The chatbot could handle situations in which the line of conversation was not clear.	Adopted directly.
9. The chatbot's responses were easy to understand.	Adopted directly.

Table 1. Cont.

Perceived quality of conversation and information provided	
10. I find that the chatbot understands what I want and helps me achieve my goal.	Adopted directly.
11. The chatbot gives me the appropriate amount of information.	Adapted: The chatbot gives me the appropriate information to understand basic concepts from business management. The chatbot gives me the appropriate information to understand complex concepts from business management.
12. The chatbot only gives me the information I need.	Adopted directly.
13. I feel like the chatbot's responses were accurate.	Adopted directly.
Perceived privacy and security	
14. I believe the chatbot informs me of any possible privacy issues.	Items are irrelevant since the AI chatbot does not collect or use personal information.
Response time	
15. My waiting time for a response from the chatbot was short.	Adopted directly.

Source: Adapted from [30,31], based on research results.

This survey was used as a research instrument in the subsequent quantitative evaluation of our test chatbot, with the full-time undergraduate students enrolled at the Faculty of Economics, Business and Tourism at the University of Split, Croatia (EFST) and the Faculty of Economics at the University of Mostar, Bosnia and Herzegovina (EF SUM). The survey was implemented by using the licensed Qualtrics XM electronic survey system. We did not collect any data enabling the personal identification of survey participants. The anonymous data were collected with the permission of the institutional ethics board and with the students' consent to use their anonymous data for academic research.

We distributed the links to the Qualtrics survey and the test implementation of the LLM chatbot for business management learning to the entire population of second undergraduate-year students at the two schools using the Moodle Learning Management System (LMS). All the students were previously enrolled in the Fundamentals of Management course, which did not involve using LLMs or other AI tools. There were no other previous formal attempts to use intelligent agents, LLMs, or other AI tools in academic teaching and learning at the two schools, which ensures that respondents could not be influenced by the success of similar initiatives in their academic environment.

Data collection lasted three weeks during the winter semester of the academic year 2023/2024. Our final sample size of 204 respondents was based on a nonrandom approach, using the specific method of snowball sampling, as we also specifically asked students to recommend filling in the survey to their colleagues at the two targeted business schools. As previously suggested in the research methods literature [33], this sampling approach is effective for studies limited by time and resources, as it builds upon the existing social connections and trust provided by the immediate social environment referrals. Weaknesses of the method are related to the sensitivity of research results to the availability and motivation of newly recruited respondents. The nonrepresentative nature of the sample design does not guarantee the representativeness or the results free from biases. However, it is a good starting point for applying quantitative methods in exploratory research designs.

It should also be noted that the nature of the sample limits our research, i.e., the population of young adults enrolled in the second and third year of undergraduate business studies at two regional business schools in Southeast Europe. We did not collect detailed demographic information about the survey participants.

Statistical analysis was conducted in IBM SPSS Statistics, version 25 (Armonk, NY, USA: IBM Corp.). Results are presented as mean and standard deviation. Cronbach's alpha coefficient checked the internal reliability of the used instrument. Pearson's correlation coefficient investigated the connection between the dimensions. The significance of the obtained results was analyzed at significance levels of 0.05 and 0.001.

## 4. Results

### 4.1. Qualitative Analysis of Expert Interviews

The four experts summarized perspectives on the effectiveness and quality of the chatbot as follows:

- Expert A expressed concerns about the quality and correctness of the LLM responses and referred to the domain-specific data quality used for training. They believed the chatbot was not up to the quality and accuracy standards when assessing complex concepts and their relationships.
- Expert B focused on the performance of the LLM and indicated that it seems a custom LLM model is underperforming compared with the generic LLM model (without custom training).
- Expert C agreed with Expert B on the quality and accuracy of responses and the depth of the management topics covered, which are the disadvantages of the custom LLM model.
- Expert D observed that though most of the answers are correct, they are more general. They suggested that the tool is more appropriate for explaining simple concepts than efficiently answering complex queries.

Considering the expert insights, the LLMs can be evaluated as generally possessing the potential to enhance the learning experiences in business management education and serve as a useful tool to build Open Educational Resources (OERs). Four main themes could be identified in the experts' interviews: concerns about the quality and accuracy of chatbot output, chatbot adaptability, integration with educational tools, and measurement of chatbot success:

- Concerns regarding quality, accuracy, and adaptability emerged, with some experts expressing their reservations about the depth and specificity of the chatbot-generated output.
- A concern for adaptability was also noted, with varying views on how successful custom LLM models will serve users' needs effectively.
- Integration with educational tools was highlighted, especially with the LMSs (Learning Management Systems) and other educational platforms, to increase accessibility and student engagement.
- Experts mentioned a variety of metrics and methods related to measuring chatbot success, focusing on user engagement, satisfaction, learning outcomes, etc. Methodological approaches suggested for evaluating such AI tools included measurement of learning outcomes and quasi-experimental studies.

### 4.2. Quantitative Analysis of Student Perceptions

In the second stage of the empirical research, we used the adapted research instrument. We distributed the link to the electronic version of the survey to the entire student body of the 2nd year of undergraduate studies of business and management at the Faculty of Economics, Business and Tourism at the University of Split, Croatia (EFST) and the Faculty of Economics at the University of Mostar, Bosnia and Herzegovina (EF SUM).

A total of 204 responses were obtained: 151 (74%) from EFST and 53 (26%) from EF SUM. The scale's overall internal consistency proved adequate since Cronbach's alpha value for the entire scale was 0.904. A relatively high value, much higher than the generally accepted threshold of 0.7, shows that the used research instrument can be internally consistent. Descriptive statistics for all questionnaire items provided in Table 2 clearly show the following:

- Students' feedback on the clarity of responses, their awareness of the information available from the chatbot, the immediacy of information delivery provided, and the conversational quality had high scores (with means scattered around the scale point of 4.0, on a five-level Likert scale).

- The chatbot's ability to maintain the conversation's context and provide relevant references was also highly regarded by students, although higher standard deviations show a higher level of variability.
- Unlike experts, students did not evaluate the chatbot responses as significantly worse than those of the generic LLMs. However, this could be attributed to their limited experience using LLMs and other AI tools in higher education settings.
- Users also appreciated the chatbot's response accuracy and quick response time.

**Table 2.** Descriptive statistics of chatbot evaluation items.

	N	Min	Max	Mean	St.Dev.
Perceived quality of chatbot functions					
Communicating with the chatbot was clear.	204	1	5	4.22	0.647
I was immediately made aware of what information the chatbot can give me.	204	2	5	4.24	0.638
The interaction with the chatbot felt like an ongoing conversation.	204	1	5	3.69	0.887
The chatbot was able to keep track of context.	204	1	5	4.03	0.752
The chatbot was able to make references to the respected websites or scholarly articles.	204	1	5	3.67	0.862
The chatbot provides answers aligned with the teaching materials (textbook) and other courseware.	204	1	5	3.85	0.750
The chatbot is able to comment on complex topics and connect complex topics and issues from the field of business management.	204	1	5	3.89	0.780
Information provided by the chatbot have better quality, compared with the generic ChatGPT, Google Gemini, or the other Large Language Model I am using regularly.	204	1	5	3.51	0.985
The chatbot could handle situations in which the line of conversation was not clear.	204	1	5	3.50	0.879
The chatbot's responses were easy to understand.	204	2	5	4.20	0.674
Perceived quality of conversation and information provided					
I find that the chatbot understands what I want and helps me achieve my goal.	204	2	5	4.17	0.660
The chatbot gives me the appropriate information to understand basic concepts from business management.	204	1	5	4.12	0.753
The chatbot gives me the appropriate information to understand complex concepts from business management.	204	1	5	4.06	0.734
The chatbot only gives me the information I need.	204	1	5	3.71	0.889
I feel like the chatbot's responses were accurate.	204	1	5	3.98	0.746
Response time					
My waiting time for a response from the chatbot was short.	204	1	5	4.30	0.815

When compared across three dimensions (see Table 3), there is a positive evaluation of the chatbot functions (with a mean value of 3.88), with a somewhat better evaluation of the quality of conversation and information provided (with a mean value of 4.00). Students were especially satisfied with the response time, achieving a mean value of 4.30. Simultaneously, standard deviations and mean scores for each of the previously mentioned evaluation dimensions increased, indicating varying expectations related to the quality of conversation, LLM output, and response time. Internal consistency, measured by Cronbach alpha values, is also acceptable for the two relevant dimensions, consisting of multiple items.



**Table 3.** Descriptive statistics of chatbot evaluation dimensions.

Chatbot Evaluation Dimensions	Mean	St.Dev.	Cronbach's Alpha
Perceived quality of chatbot functions	3.88	0.51	0.84
Perceived quality of conversation and information provided	4.00	0.58	0.82
Response time	4.30	0.81	n/a

The analysis of linear correlations (see Table 4) shows a high level of mutual dependencies among the dimensions of chatbot evaluation, including custom LLM functionalities, conversational quality, and response time, as demonstrated by a high level of statistical significance. There is a very strong correlation (0.76) between the perceived functionalities and the perceived quality of chatbot output and its conversational performance. According to the previous empirical results, improvements in the knowledge base used for training and the chatbot technical aspects are likely to impact students' views on the quality of chatbot interactions. Students also expect that response times will enhance as chatbot capabilities improve (as indicated by a Pearson value of 0.58), although the perceived correlation is not as strong. The weakest correlation (0.51) exists between the response time (efficiency) and students' perception of chatbot conversation and information quality. This empirical result suggests that response time significantly shapes the user experience with the chatbot.

**Table 4.** Pearson linear correlation coefficients among the chatbot evaluation dimensions.

Dimensions	Perceived Quality of Chatbot Functions	Perceived Quality of Conversation and Information Provided	Response Time
Perceived Quality of Chatbot Functions	1.00	0.76 **	0.58 **
Perceived Quality of Conversation and Information Provided		1.00	0.51 **
Response Time			1.00

\*\* Significant at the 1% level.

According to these empirical findings, chatbot optimization must be approached holistically. Improving technical features, chatbot knowledge base, and its relevance, and reducing response times work together to enhance user satisfaction with custom LLM chatbots. A few students were providing an answer to an additional, open-ended question positioned at the very end of the research instrument. They expressed a generally positive evaluation of the chatbot. They suggested several potential improvements, including the need for faster responses, regular content updates, and the implementation of custom LLM digital assistants for other business school classes.

## 5. Discussion

This study is one of the first attempts to explore the potential of custom-trained LLMs in higher education. There is a pressing need to contribute to the research of widely available, specialized machine learning and its contribution to creating Open Educational Resources (OERs) since the extant literature does not cover this topic. We used the mixed-method approach to obtain qualitative feedback from domain experts (regional faculty with extensive experience in teaching business administration and implementing Information Technology in higher education) and quantitative feedback from undergraduate business and economics students.

While experts had concerns about the quality and contextualization of the chatbot output, this result of the empirical research can be attributed to their pedagogical experience and the perceived limitations of the AI tools compared with the competencies of experienced faculty. Simultaneously, students seem to appreciate the convenience of using

a custom AI chatbot as an OER. Although this purpose has technical and commercial limitations, previously described in the introduction section, students' initial reactions are very positive.

This study is the first exploratory research in the extant literature. Although its results are not generalizable across different subject fields and student populations, there is an interesting dichotomy between student evaluations and expert opinions due to limitations in the sample and a simplified approach to implementing the test AI chatbot. While the AI chatbot for business management learning seems to serve its purpose for the students, who might be looking for a convenient tool for test preparation and quick assessment of their knowledge, academic experts are looking at the pedagogical merits and the reasoning demonstrated by the AI tool. They are not completely acceptable to the experienced experts. At the same time, the depth of output and the quality of arguments are not comparable to the faculty with many years of teaching experience. However, even in a very early stage of development, a quickly developed custom LLM proved to serve as a useful OER to undergraduate students at two regional business schools in Southeast Europe.

Research results might also guide the faculty and other actors wishing to design and improve the OERs using the custom LLM tools. Within our limited sample, the functional and conversational capabilities of the test chatbot and its response speed proved to be highly correlated, which shows the need to simultaneously consider all aspects of students' user experience and improve the chatbot by using a holistic approach.

## 6. Conclusions

This study explored a pilot project using the AI chatbot, which was implemented to assist business school students in business management. The piloted AI chatbot could also prove useful in improving higher education practices. This especially applies to sustainable educational practices that include nontraditional students and promote lifelong learning opportunities. AI tools enable flexible learning solutions and improve access to higher education, thus contributing to the United Nations Sustainable Development Goal (UN SDG4). According to our empirical results, such an approach creates value perceived by students. However, the consulted experts indicated several limitations concerning the quality of the generated content and how the chatbot handles complex queries. These limitations must be addressed when developing AI tools for higher education and lifelong learning at the academic level.

### 6.1. Theoretical and Practical Contributions

Our findings can be used to develop a comprehensive evaluation model of AI chatbots in education. We point out the theoretical dimensions of learner–AI interaction, including perceived quality, privacy/security, and accessibility. From practitioners' point of view, we demonstrate the viability of commercial AI tools, enabling education (and other) professionals to train and deploy custom-trained LLMs as tools of choice for academic teaching and learning. We also identify current issues with AI tools to create open-access educational materials and approaches.

Our approaches contribute to UN SDG4, allowing higher education institutions to facilitate nontraditional and lifelong learning paths. Our results align with lifelong learning incentives [34,35] and practices in implementing advanced educational technologies [36]. In addition, compared with traditional (printed) learning resources, AI-based OERs have a minimal ecological footprint. This assures resource efficiency and environmental sustainability, both implied under SDG12 on sustainable consumption and production.

We also believe that integrating AI tools, such as the described business management chatbot, can help make educational programs more relevant and address the demands of the labor market. Custom-trained LLMs can be continuously updated to include emerging knowledge and skills required by the workforce, thus improving the employability of graduates and contributing to SDG8 on inclusive and sustainable economic growth, including full and productive employment for the current student generations [37].

### 6.2. Limitations and Future Research Directions

Limitations of this study are related to the specific educational context analyzed and the limited range of chatbot functionalities tested. Future studies should consider the use in multiple academic disciplines and additional educational contexts, including lifelong learning scenarios. More advanced AI features should be considered, including analysis of student sentiment and adaptive responses to the current level of students' knowledge and competencies. Integrating multimedia elements and a multimodal learning approach (based on a voice or a video chat) could also enhance the interactivity and learners' motivation to use the AI chatbot. All those functionalities would contribute to a completely personalized learning experience customized to the student's context and educational goals.

Research questions, which future research could explore, are as follows:

- Are custom Large Language Models (LLMs) appropriate for developing Open Educational Resources (OERs) in higher education, regardless of the discipline and academic context?
- Can custom LLMs be used in the context of Lifelong Learning in higher education?
- What is the role of advanced custom LLM features, such as the analysis of student sentiment and adaptability to student requirements, in their role as effective teaching and learning tools?
- What is the role of advanced custom LLM technical features, such as reasoning and communicating across text, audio, and video, in their role as effective teaching and learning tools?

### 6.3. Implications for Educational Practice

The practical implications of these results are relevant for improving chatbot analysis and as potential guidelines in the design of AI chatbots within other academic disciplines. The most important takeaway lesson is that the chatbot development process has to be iterative and based on student feedback. Further training content has to be identified for use in custom training to meet the handling of complex student queries. Context understanding and interaction can be achieved through better cooperation with AI experts who may guide through better tools and approaches toward chatbot optimization.

**Supplementary Materials:** The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/su16124929/s1>, Table S1: Data sharing—custom GPT survey.

**Author Contributions:** Conceptualization, N.A.; methodology, N.A.; software, M.M.; validation, D.G.P. and M.M.; formal analysis, N.A.; investigation, N.A., D.G.P. and M.M.; resources, M.M.; data curation, N.A.; writing—original draft preparation, N.A., D.G.P. and M.M.; writing—review and editing, N.A., D.G.P. and M.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** This paper has been approved by the institutional Ethics Committee of the Faculty of Economics, Business and Tourism, University of Split, Croatia (Class: 004-01/24-01/03, Ref. no.: 2181-196-02-05-24-03, dated 10 April 2024).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data presented in this study are freely available in the Supplementary Materials to the manuscript.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Appendix A. Custom AI Chatbot Training Instructions

**Role and Goal:** This GPT is a specialized educational resource in Croatian, designed for college and university students in management and business administration. It aims

to support lifelong learning, solve professional problems, and provide expert responses to questions across all types of organizational management, regardless of the user's educational level or interest. The GPT is trained and customized using the user's proprietary materials on management, intended to enhance its understanding and response capability in this domain. Additionally, it includes guidelines for optimal use, encourages giving feedback, and offers customization options for personalized learning experiences.

**Constraints:** Responses must be in Croatian, tailored to the educational context, and appropriate for a diverse audience of students and professionals in management and business administration. The GPT should avoid giving personal opinions, making predictions, or providing financial advice. It should also respect the copyright of the materials provided for training.

**Guidelines:** The GPT should use professional and accessible language, make complex concepts understandable, and encourage inquiry and exploration in the field of management. It should reference credible sources when possible and admit its limitations when necessary. Users are encouraged to provide feedback and customize their learning experience through settings.

**Clarification:** The GPT should ask for clarification when the question is ambiguous or lacks specific details needed for a comprehensive response. It should aim to make its explanations as relevant and useful as possible to the user's query.

**Personalization:** The GPT maintains a supportive and educational tone, adapting its responses to the user's level of understanding and preferences. Users are encouraged to specify their interests and desired complexity level for a more tailored learning experience.

## Appendix B. Transcripts of Expert Interviews

- Expert A.

Croatian academic professional with 15+ years of experience in teaching and research of Information Technology, with a focus on e-learning and Business Process Management.

1. Are the quality and the accuracy of the information provided by the chatbot ensured? Are there any chatbot 'hallucinations', i.e., incorrect but plausible-sounding output?

For testing the chatbot, I used question(s) related to the role and the link between strategic communication and business process management. From my perspective, for an assessment of the personalized chatbot's answers, a series of questions would need to be posed, whereby the answers would need to be compared with relevant courseware and standard chatbot (ChatGPT) replies using a strict protocol and criteria. Without that, it is challenging to comment on the accuracy and other quality criteria. On this occasion, the answers did not meet my expectations, possibly indicating a problem with the chatbot's training or design. Namely, the quality relies on the diversity of the data used to train it. I do not know what data were used, so it might be an issue that the data are limited or outdated.

2. How does the chatbot handle complex inquiries?

Chatbots are only as knowledgeable as the information they have been trained on. The chatbot defined well the main topic; however, in connecting the concept of strategic communication to another one, to Business Process Management, it struggled to provide accurate responses, as this is a complex or nuanced question. The answers sounded plausible but lacked depth or accuracy.

3. How adaptable is the chatbot to user requirements?

The design of the chatbot, including its interface and conversational flow, impacts the quality of its responses. In this particular case, I am not aware of the user requirements the authors set, and if the design prioritized efficiency over accuracy. In any case, continuous improvement is crucial for AI models and should be one of the vital user requirements. Without a feedback loop where users can provide corrections or feedback on the chatbot's responses, it may not learn from its mistakes and improve over time. In that, what is crucial is the evaluation of its responses that is systematic, based on the feedback from users,

and updated with the training data with more relevant and diverse sources to refine its algorithms and design iteratively.

4. How does the chatbot integrate with other tools for learning business management?

Integrating a chatbot with various tools and platforms can create a more interactive environment for students. First of all, it makes sense to integrate it with an LMS to assist them in accessing course materials, recommend supplementary resources, and so on. It is even better if it is integrated with content repositories to retrieve relevant information and resources for students. Collaboration tools can also embed access to a chatbot in specific channels or spaces within these platforms. For teachers, it can support assessment and feedback also within LMS or learning analytics platforms.

5. Is it possible to measure the success of a chatbot in improving business management learning?

I do not know of a systematic way to do so; however, I believe it is possible to measure the success of a chatbot in improving learning outcomes in a particular area, including business management learning. First of all, through usage metrics such as the frequency of usage and time spent interacting with the chatbot, although I do not know if such data can be obtained. Learning outcomes before and after implementing the chatbot could be measured using quizzes/exams. Students could be asked as well if they perceive the chatbot supported them in their learning. Also, there could be a comparison between the performance of students who have access to the chatbot and those who do not use it via quasi-experimental studies.

- Expert B.

Croatian academic professional with 15+ years of experience in teaching and research of Information Technology, with a focus on e-learning and Management Information Systems.

1. Are the quality and the accuracy of the information provided by the chatbot ensured? Are there any chatbot 'hallucinations', i.e., incorrect but plausible-sounding output?

Based on my evaluation process, I can say that the quality and accuracy of the information provided by the chatbot are ensured. Employing methods such as benchmark testing and validation against trusted sources, I aimed to assess its efficacy. Benchmark testing involved comparing the chatbot's responses against established standards and evaluating factual accuracy, response coherence, and linguistic appropriateness. The chatbot consistently met these criteria, delivering precise information with clarity. Additionally, validation against trusted sources involved cross-referencing the chatbot's information with reputable sources such as respected websites and scholarly articles, revealing a notable alignment and confirming its accuracy. During this testing, it is notable that the chatbot did not display any instances of 'hallucinations', maintaining relevance and reliability throughout the process.

2. How does the chatbot handle complex inquiries?

After analyzing the questions I've found that the tested chatbot still falls short compared with general chatbots, like ChatGPT. An example of a question was, 'How can a company effectively incorporate development into its business strategy while considering market competition, stakeholder expectations, and long-term profitability?' To enhance the chatbot's ability to handle complex queries, potential improvements include a better comprehension of complex questions, broadening its knowledge base, and incorporating contextual understanding.

3. How adaptable is the chatbot to user requirements?

It's clear that this chatbot solution needs more work and testing with a larger user base to truly succeed. It's important to gather perspectives from a larger group of users to better understand their requirements, choices, and actions. Consistently collecting feedback and making relevant improvements is crucial to ensure the chatbot progresses, meets user expectations, and provides better outcomes. Implementing scenario-based testing or

long-term monitoring of usage is essential to understand and address a range of user needs and preferences.

4. How does the chatbot integrate with other tools for learning business management?

Within the given time frame, I am unable to give an evaluation of this aspect. Evaluating how well the chatbot works with business management tools involves looking into its ability to meet user needs and work alongside these tools effectively. This requires examining if the chatbot seamlessly integrates with business management tools, ensures a seamless user experience, facilitates data sharing and analysis, and welcomes user feedback for continuous improvement.

5. Is it possible to measure the success of a chatbot in improving business management learning?

Yes, it is possible. Measuring the success of a chatbot in enhancing business management learning is definitely achievable. One of the primary approaches to measuring the success of a chatbot in improving business management learning is through a comprehensive evaluation of user engagement, satisfaction, learning outcomes, etc.

Overall, this chatbot proves to be a useful tool for providing information and addressing user queries with very few problems. After assessing its performance, I recommend continued use of the chatbot, while also making updates and improvements to maintain its accuracy and effectiveness in sharing information.

- Expert C.

Croatian academic professional with 10+ years of experience in teaching and researching business administration topics, with a focus on the role of technology in marketing.

1. Are the quality and the accuracy of the information provided by the chatbot ensured? Are there any chatbot 'hallucinations', i.e., incorrect but plausible-sounding output?

The information provided is accurate at a very general level. It can recognize the main concepts and terms in contemporary management, but one can notice a lack of knowledge and insights into specific management topics. But, with enhanced usage frequency, improvements can be seen, and answers become more specific.

2. How does the chatbot handle complex inquiries?

At the general level, even complex inquiries are handled relatively well. Still, sometimes, it circles around topics, picking one that is more familiar with and providing several variations of the same answer while ignoring the rest of the complex inquiry. It can be noticed that, over time, performance is improved even in complex subjects.

3. How adaptable is the chatbot to user requirements?

This is one of the main strengths of this bot. It adapts to user specifics relatively fast, trying to avoid 'mistakes' from previous inquiries and providing more customized answers. In addition, when inquiry is recognized as contradictory (in several iterations)—it provides accurate (maybe more general) answers that can be seen from the beginning.

4. How does the chatbot integrate with other tools for learning business management?

I can't evaluate this as I have yet to use it as a tool for learning business management. However, I believe it can be a helpful tool in creating themes for students' discussions, interpreting key concepts, and summarizing main management concepts and theories.

5. Is it possible to measure the success of a chatbot in improving business management learning?

I don't think it is possible at the moment, but if all the main features keep improving in the future, I believe this might happen. But first, the chatbot's success should be clearly defined, and relevant metrics should be presented.

- Expert D.

Bosnia and Herzegovina academic professional with 20+ years of experience in teaching and researching business administration, with a focus on business management.

1. Are the quality and the accuracy of the information provided by the chatbot ensured? Are there any chatbot 'hallucinations', i.e., incorrect but plausible-sounding output?

I didn't notice any specific mistakes or deviations, but the answers are very general.

2. How does the chatbot handle complex inquiries?

The answers are very simple and often general and superficial. This is an excellent tool for high schools and students to get acquainted with certain topics and concepts. I think it might struggle to provide answers to more complex practical questions. I see its application in learning basic concepts, phenomena, and relationships.

3. How adaptable is the chatbot to user requirements?

I see the use of this tool in high schools and universities for learning the most basic concepts.

4. How does the chatbot integrate with other tools for learning business management?

I only checked the comparison with concepts in theory and there I received accurate but general answers. Besides that, I tried asking the same questions to ChatGPT and there I received much more precise answers.

5. Is it possible to measure the success of a chatbot in improving business management learning?

This question is very hard for me because I don't understand the technical support, but I believe it's possible to improve the tool by expanding the database of research, books, and papers from which it could 'pull' knowledge and provide answers.

### Appendix C. Research Instrument (English Original and Croatian Translation)

Perceived quality of chatbot functions	Percipirana kvaliteta funkcija chatbota
Communicating with the chatbot was clear.	Razgovor s chatbotom bio je razumljiv.
I was immediately made aware of what information the chatbot can give me.	Brzo sam shvatio/la koje sve informacije chatbot može pružiti.
The interaction with the chatbot felt like an ongoing conversation.	Dijalog s chatbotom tekao je poput pravog razgovora.
The chatbot was able to keep track of context.	Chatbot je uspješno pratio kontekst razgovora.
The chatbot was able to make references to the respected websites or the scholarly articles.	Chatbot je mogao upućivati na relevantne web stranice ili znanstvene članke.
The chatbot provides answers aligned to the teaching materials (textbook) and other courseware.	Chatbot daje odgovore usklađene s nastavnim materijalima (udžbenikom) i ostalim obrazovnim sadržajima.
The chatbot is able to comment on complex topics and connect complex topics and issues from the field of business management.	Chatbot može komentirati složene teme i povezivati kompleksne teme i probleme iz područja poslovnog menadžmenta.
Information provided by the chatbot have better quality, compared with the generic ChatGPT, Google Gemini, or the other Large Language Model I am using on a regular basis.	Informacije koje pruža chatbot kvalitetnije su u usporedbi s "običnim" (generičkim) ChatGPT-om, Google Gemini ili drugim jezičnim modelima umjetne inteligencije koje redovito koristim.
The chatbot could handle situations in which the line of conversation was not clear.	Chatbot je mogao upravljati situacijama u kojima tijekom razgovora nije bio jasan.
The chatbot's responses were easy to understand.	Odgovori chatbota bili su lako razumljivi.
Perceived quality of conversation and information provided	Percipirana kvaliteta razgovora i pruženih informacija
I find that the chatbot understands what I want and helps me achieve my goal.	Utvrдио/la sam da chatbot razumije što želim i pomaže mi postići moj cilj.
The chatbot gives me the appropriate information to understand basic concepts from the field of business management.	Chatbot mi pruža odgovarajuće informacije za razumijevanje osnovnih pojmova iz područja poslovnog menadžmenta.
The chatbot gives me the appropriate information to understand complex concepts from the field of business management.	Chatbot mi pruža odgovarajuće informacije za razumijevanje složenih pojmova iz područja poslovnog menadžmenta.
The chatbot only gives me the information I need.	Chatbot mi daje samo informacije koje su mi potrebne.
I feel like the chatbot's responses were accurate.	Procjenjujem da su odgovori chatbota točni.
Response time	Vrijeme odgovora
My waiting time for a response from the chatbot was short.	Vrijeme čekanja na odgovor chatbota bilo je kratko.

### References

- Garrido-Merchán, E.C.; Arroyo-Barrigüete, J.L.; Borrás-Pala, F.; Escobar-Torres, L.; de Ibarreta, C.M.; Ortiz-Lozano, J.M.; Rúa-Vieites, A. Real Customization or Just Marketing: Are Customized Versions of Chat GPT Useful? *arXiv* **2023**, arXiv:2312.03728.
- Bandi, A.; Adapa, P.V.S.R.; Kuchi, Y.E.V.P.K. The power of generative AI: A review of requirements, models, input–output formats, evaluation metrics, and challenges. *Future Internet* **2023**, *15*, 260. [[CrossRef](#)]
- Leslie, D. Does the sun rise for ChatGPT? Scientific discovery in the age of generative AI. *AI Ethics* **2023**. [[CrossRef](#)]
- Hsiao, Y.P.; Klijn, N.; Chiu, M.S. Developing a framework to re-design writing assignment assessment for the era of Large Language Models. *Learn. Res. Pract.* **2023**, *9*, 148–158. [[CrossRef](#)]
- Shoufan, A. Exploring students' perceptions of ChatGPT: Thematic analysis and follow-up survey. *IEEE Access* **2023**, *11*, 38805–38818. [[CrossRef](#)]

6. Maleki, N.; Padmanabhan, B.; Dutta, K. AI Hallucinations: A Misnomer Worth Clarifying. *arXiv* **2024**, arXiv:2401.06796.
7. Ray, P.P. ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet Things Cyber-Phys. Syst.* **2023**, *3*, 121–154. [[CrossRef](#)]
8. Michel-Villarreal, R.; Vilalta-Perdomo, E.; Salinas-Navarro, D.E.; Thierry-Aguilera, R.; Gerardou, F.S. Challenges and opportunities of generative AI for higher education as explained by ChatGPT. *Educ. Sci.* **2023**, *13*, 856. [[CrossRef](#)]
9. Huallpa, J.J. Exploring the ethical considerations of using Chat GPT in university education. *Period. Eng. Nat. Sci.* **2023**, *11*, 105–115.
10. Berson, I.R.; Berson, M.J. The democratization of AI and its transformative potential in social studies education. *Soc. Educ.* **2023**, *87*, 114–118.
11. Fuchs, K. Exploring the opportunities and challenges of NLP models in higher education: Is Chat GPT a blessing or a curse? *Front. Educ.* **2023**, *8*, 1166682. [[CrossRef](#)]
12. Fidalgo, P.; Thormann, J. The Future of Lifelong Learning: The Role of Artificial Intelligence and Distance Education. 2024. Available online: <https://www.intechopen.com/online-first/88930> (accessed on 26 March 2024).
13. Aithal, P.S.; Aithal, S. The changing role of higher education in the era of AI-based GPTs. *Int. J. Case Stud. Bus. IT Educ.* **2023**, *7*, 183–197. [[CrossRef](#)]
14. Meyer, J.; Jansen, T.; Schiller, R.; Liebenow, L.W.; Steinbach, M.; Horbach, A.; Fleckenstein, J. Using LLMs to bring evidence-based feedback into the classroom: AI-generated feedback increases secondary students' text revision, motivation, and positive emotions. *Comput. Educ. Artif. Intell.* **2024**, *6*, 100199. [[CrossRef](#)]
15. Rasul, T.; Nair, S.; Kalendra, D.; Robin, M.; de Oliveira Santini, F.; Ladeira, W.J.; Sun, M.; Day, I.; Rather, R.A.; Heathcote, L. The role of ChatGPT in higher education: Benefits, challenges, and future research directions. *J. Appl. Learn. Teach.* **2023**, *6*, 41–56. [[CrossRef](#)]
16. Aithal, P.S.; Aithal, S. Optimizing the use of artificial intelligence-powered GPTs as teaching and research assistants by professors in higher education institutions: A study on smart utilization. *Int. J. Manag. Technol. Soc. Sci.* **2023**, *8*, 368–401. [[CrossRef](#)]
17. Godwin-Jones, R. Distributed agency in second language learning and teaching through generative AI. *arXiv* **2024**, arXiv:2403.20216.
18. Kramsch, C. Global English: The Indispensable Bridge in Intercultural Communication? *TEANGA J. Ir. Assoc. Appl. Linguist.* **2023**, *30*, 1–32. Available online: <https://journal.iraal.ie/index.php/teanga/article/view/6797> (accessed on 7 April 2024).
19. Marshall, S.; Sankey, M.D. The Future of the Learning Management System in the Virtual University. In *Technology-Enhanced Learning and the Virtual University*; Springer Nature Singapore: Singapore, 2023; pp. 283–304. [[CrossRef](#)]
20. Dessimoz, C.; Thomas, P.D. AI and the democratization of knowledge. *Sci. Data* **2024**, *11*, 268. [[CrossRef](#)]
21. Li, Z.; Pardos, Z.A.; Rena, C. Aligning open educational resources to new taxonomies: How AI technologies can help and in which scenarios. *Comput. Educ.* **2024**, *216*, 105027. [[CrossRef](#)]
22. Firat, M. How Chat GPT Can Transform Autodidactic Experiences and Open Education? 2023. Available online: <https://osf.io/preprints/osf/9ge8m> (accessed on 25 March 2024).
23. Elbanna, S.; Armstrong, L. Exploring the integration of ChatGPT in education: Adapting for the future. *Manag. Sustain. Arab Rev.* **2024**, *3*, 16–29. [[CrossRef](#)]
24. Dai, Y.; Liu, A.; Lim, C.P. Reconceptualizing ChatGPT and generative AI as a student-driven innovation in higher education. *Procedia CIRP* **2023**, *119*, 84–90. [[CrossRef](#)]
25. Mills, A.; Bali, M.; Eaton, L. How do we respond to generative AI in education? Open educational practices give us a framework for an ongoing process. *J. Appl. Learn. Teach.* **2023**, *6*, 16–30. [[CrossRef](#)]
26. OpenAI. OpenAI Prompt Engineering. 2024. Available online: <https://platform.openai.com/docs/guides/prompt-engineering> (accessed on 20 March 2024).
27. Klepić, Z.; Alfirević, N.; Rahimić, Z. (Eds.) Menadžment. In *University of Mostar, University of Split—Faculty of Economics, Business and Tourism, University of Sarajevo—School of Economics and Business, Croatian Edition*; PresuM Publishing: Mostar, Bosnia and Herzegovina, 2020.
28. Radziwill, N.M.; Benton, M.C. Evaluating quality of chatbots and intelligent conversational agents. *arXiv* **2017**, arXiv:1704.04579.
29. Denecke, K.; Abd-Alrazaq, A.; Househ, M.; Warren, J. Evaluation metrics for health chatbots: A Delphi study. *Methods Inf. Med.* **2021**, *60*, 171–179. [[CrossRef](#)]
30. Borsci, S.; Malizia, A.; Schmettow, M.; Van Der Velde, F.; Tariverdiyeva, G.; Balaji, D.; Chamberlain, A. The chatbot usability scale: The design and pilot of a usability scale for interaction with AI-based conversational agents. *Pers. Ubiquitous Comput.* **2022**, *26*, 95–119. [[CrossRef](#)]
31. Borsci, S.; Schmettow, M.; Malizia, A.; Chamberlain, A.; Van Der Velde, F. A confirmatory factorial analysis of the Chatbot Usability Scale: A multilanguage validation. *Pers. Ubiquitous Comput.* **2023**, *27*, 317–330. [[CrossRef](#)]
32. Wang, T.; Duong, T.D.; Chen, C.C. Intention to disclose personal information via mobile applications: A privacy calculus perspective. *Int. J. Inf. Manag.* **2016**, *36*, 531–542. [[CrossRef](#)]
33. Handcock, M.S.; Gile, K.J. Comment: On the concept of snowball sampling. *Sociol. Methodol.* **2011**, *41*, 367–371. [[CrossRef](#)]
34. Alonderiene, R.; Sabaliauskaitė, G. Non-formal and informal learning conditions as experienced and perceived by technical staff and HR professionals. *Manag. J. Contemp. Manag. Issues* **2017**, *22*, 15–33. [[CrossRef](#)]



35. Dukić, G. Managers and lifelong learning: An analysis of motivation and motivational factors. *Manag. J. Contemp. Manag.* **2023**, *28*, 57–71. [[CrossRef](#)]
36. Chiu, T.K.; Chai, C.S. Sustainable curriculum planning for artificial intelligence education: A self-determination theory perspective. *Sustainability* **2020**, *12*, 5568. [[CrossRef](#)]
37. Ernst, E.; Merola, R.; Samaan, D. Economics of artificial intelligence: Implications for the future of work. *IZA J. Labor Policy* **2019**, *9*, 1–35. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.