PHD Research Plan

introEduAI – Evaluation Tool to Assess the Impact of Artificial Intelligence in Education

PhD candidate: Pedro Mendonça

PhD supervisor: Filipe Quintal

PhD co-supervisor: Fábio Mendonça

Summary (max:150 words)

We live in a time where artificial intelligence (AI) technologies are proliferating at dizzying speed. We currently find this technology everywhere: companies and industry, finance, education, medicine and health, internet, among others. One of the most popular AI technologies used to support teaching and learning activities is the Chatbot system and are considered good technological innovations that improve student learning interest, acquisition of cognitive skills, and achievement (Lin e Chang 2020; Murad et al. 2019). This project focus on design an AI based tool and analyse the impact of these new technologies, namely the use of chatbots in the educational system.

State of the art *(max:500 words)*

Chatbots are digital systems that can be interacted with entirely through natural language via text or voice interfaces. They are intended to automate conversations by simulating a human conversation partner and can be integrated into software, such as online platforms, digital assistants, or be interfaced through messaging services (Wollny et al. 2021). The use of AI in education is rapidly expanding, particularly the use of chatbots, providing a focused, personalised, and result-oriented online learning environment (Cunningham-Nelson et al. 2019).

Large language models (LMMs), have made significant advancements in natural language processing (NLP) in recent years. These models are trained on massive amounts of text data and are able to generate human-like text, answer questions, and

complete other language-related tasks with high accuracy (Kasneci et al. 2023). LMMs have the potential to revolutionize teaching and assist in teaching processes: personalized learning; lesson planning; language learning; research and writing; assessment and evaluation (Kasneci et al. 2023). On November 30, 2022, ChatGPT, from OpenAI, was launched, a tool that aimed to revolutionize all aspects of society: and it did!

This new reality brings challenges for all educational agents: students, educators and educational institutions. Educators have to consider adjustments to the educational learning goals, learning tasks, and assessments and evaluations to better prepare future citizens (Zhai 2022). It is also a powerful tool for producing comprehensive lesson plans, engaging presentations, and other educational resources (Atlas 2023). The AI systems could adapt the instructional approach to accommodate each student's unique learning style and progress. Additionally, AI systems have seen successful deployments as tutors beyond the traditional academic disciplines, serving as personal mindset coaches (Grassini 2023a).

The use of this technology also brings concerns and problems:

- Plagiarism: during the generation for a new prompt, the answer may contain a full sentence or even a paragraph seen in the training set, leading to copyright and plagiarism issues (Kasneci et al. 2023).
- Cheating: it can be used to complete written assignments and examinations on behalf of students, leading to concerns about AI-assisted cheating (Lo 2023).
- Data privacy and security: concerns about data breaches, unauthorized access to student data, and the use of student data for purposes other than education (Kasneci et al. 2023).
- Bias and fairness: it is important to ensure that the training data or the data used for fine-tuning on down-stream tasks for the model is diverse and representative of different groups of people (Kasneci et al. 2023).
- Fake information: ChatGPT occasionally does hallucinate and spout nonsense, for instance, by inventing references (Rudolph, Tan, e Tan 2023).

After analysing the literature, the following gaps were identified:

- Train educators on how to integrate ChatGPT into their teaching practices, ensuring that they possess a deep understanding of how LLMs like ChatGPT operate to provide appropriate guidance to students (Guo e Lee 2023).

- It's crucial to assess its (NLP Tools) effectiveness in relation to specific learning outcomes, such as knowledge retention, skill development, and other pertinent objectives (Hadi Mogavi et al. 2023).
- Limited effort has been made to include AI in learning tasks in subject domains. Integrating AI in domain learning tasks is critical as it reflects how humans are solving real-world problems (Zhai 2022).

The objective of this PhD thesis is to make a contribution to solving the problems outlined above, creating a multidisciplinary AI tool that serves as a smart virtual assistant for both educators and students.

Objectives (max:300 words)

The key focus of this thesis is to create an AI-driven tool to support educators and students in the integration of LMMs into educational practices, particularly through chatbot integration. The objectives I aim to achieve in this project are:

1) Integrating AI tools into education to provide our students with a more effective and constructive learning process, establishing and solidifying knowledge, and fostering new essential skills crucial for success in a progressively competitive job market;
2) Make use of AI tools to create inclusive lesson plans, captivating presentations, and other educational assets;
3) Using AI tool to generate personalized reports on students' performance in solving formative worksheets;
4) Enhancing the learning setting into a more personalized, organized, engaging, and results-focused space;
5) Simplify the daily routine for educators and students by partially automating administrative tasks (Assisting chatbots).

Detailed description (max:1000 words)

The purpose of this doctoral thesis is to develop an AI tool that simplifies the adoption of generative AI technologies in the teaching-learning process. Chatbots in education can serve multiple purposes: teach content or skills; simplifying the student's everyday life; deal with the student's personal development (Wollny et al. 2021). The most recent

technological advancements in Chatbot development exploit sophisticated AI techniques, including NLP, machine learning (ML), and deep learning (DL). These AI-powered chatbots learn how to respond to user inquiries based on a vast human language dataset (Jiang et al. 2022).

As an initial goal, this project aims to investigate how to incorporate the AI tool into the teaching process to enhance experiences and outcomes in the teaching and learning process. Educators should be upskilling their competencies and practices to meet the new demands of technology, therefore, teachers are required to think about new teaching philosophies, which could rely on to assess their students (Tlili et al. 2023). We'll explore the key advantages of using these AI tools among educators, showcasing their multidisciplinary application.

One of the numerous applications of this technology is, for instance, in the teaching of a second language. AI-powered chatbots can provide linguistic input and daily conversation practice (Huang, Hew, e Fryer 2022), stimulate language learners' interest (Fryer et al. 2020; Gallacher, Thompson, e Howarth 2018; Kohnke 2022) and contribute to their overall growth (Kohnke 2022; Kohnke, Moorhouse, e Zou 2023). Because these chatbots are available 24/7, students can practise their language skills anytime and anywhere (Haristiani 2019; Winkler e Soellner 2018), additionally, can conduct formative assessments and provide immediate feedback (Huang, Hew, e Fryer 2022; Kuhail et al. 2023), which are essential components of successful language acquisition (Kohnke, Moorhouse, e Zou 2023).

Regarding the second objective, AI tools have the capability to produce resources almost instantly: summarizing texts; creating topics for presentations; creating personalized images; producing videos and animations; generating activities for formative and summative assessments. LLMs have the potential to revolutionize teaching from a teacher's perspective by providing them with a wide range of tools and resources that can assist with lesson planning, personalized content creation, differentiation and personalized instruction, assessment, and professional development (Kasneci et al. 2023). Regarding assessment and evaluation, these tools can partially automate the grading of student work by identifying potential strengths and weaknesses in the submitted work (Kasneci et al. 2023). In this context, educators can adapt the reports generated by such a model to deliver beneficial feedback to students, whether in formative or summative assessment scenarios (Grassini 2023b), allowing them to target interventions more effectively (Kasneci et al. 2023).

Regarding the third goal, students, with the AI tool, can: interact via chatbot with the educator's provided content; seek formative assessment with feedback to self-assess; request suggestions or guidance to gain deeper knowledge, among other functionalities.

As the concluding objective, the AI tool can partially assist educators and students in their administrative tasks. Therefore, teachers could potentially lessen their workloads, redirecting their primary focus towards crafting innovative lesson plans, engaging in professional development, and offering personalized coaching and mentorship to each student (Grassini 2023b). Concerning students, chatbots can help in their day-to-day life by offering information like: the Math exam date; the deadline for the Portuguese assignment; the cafeteria menu for the next week; the due date for paying this month's tuition fee. Chatbots are also being employed to counsel students on academic issues, thereby assisting them in making critical decisions regarding their diverse programs or academic activities (Murad et al. 2019; Troussas, Krouska, e Virvou 2017; Lin e Chang 2020).

For the development of this AI tool, the general requirements are:

At first, the platform will be adjusted for browser-based functionality and will interact with LLM models through LangChain (framework for developing applications powered by language models);

Interaction with the chatbot via prompt and voice.

Primary features:

The platform will include a multi-profile system: institution, educator, and student. Exclusive features are incorporated into each specific profile, namely:

Institutional profile:

Within this profile lies the responsibility to add documents related to the institution's management: enrolments, exams, events, contests, fees, and other items.

Educator profile:

It will be the professor's responsibility to oversee their courses, which involves: enrolling students; adding classes and their corresponding content; managing the course

documentation (assessment criteria, evaluation moments, utilized tools, among other aspects). Teacher will have access to a chatbot that aids in creating educational materials and providing individualized learning reports, allowing for more personalized and targeted support for students.

Student profile:

Students will be granted entry to their enrolled disciplines/courses and the educational materials supplied by their teachers. Additionally, they'll have access to a chatbot facilitating self-guided learning through the teacher-shared resources, ensuring immediate feedback. This chatbot will further function as a virtual assistant, delivering details about institution-related affairs and specifics of the enrolled courses, including assignment deadlines and exam dates.

Planned interventions

IntroEduAI framework will consist of multiple modules designed to incorporate LLM models into educational practices, benefiting both educators and students. We plan to evaluate the proposed system using real students enrolled in higher education in Portugal. The system will also be evaluated by these students' educators. The completion of the PhD plan will involve a comprehensive study assessing a holistic system integrating services for all stakeholders involved in the education process. Nevertheless, smaller studies are planned to validate specific modules of the system or even the experience itself:

- For educators

  Integrating LLMs in education facilitates task automation, enabling educators to provide more personalized support to students. To meet these objectives, tools will be added, enabling teachers to: create topics/summaries from documents; generate formative/summative assessment sheets; update the chatbot with information related to the subject's proper functioning; receive feedback from students' formative assessments.

  Challenges:

  - o Educators are expected to assess the quality and relevance of the content generated by the LLM through survey completion;

- o The system is expected to analyse students' responses during formative assessments and then send a report to the educator. This process ensures immediate feedback for the educator, enabling identification, among other things, of students facing greater difficulties and areas where content may need a different approach. To automate the evaluation of this solution, a feedback mechanism by the educator will be integrated to validate the accuracy of the AI analysis. Finally, a statistical study will be carried out with the gathered data.

- <u>For students</u>

  The main focus of this platform is to ensure that our students achieve better results. Accordingly, various services will be developed to aid in achieving this objective, including the creation of virtual assistants. These assistants will aid educators in delivering content and simultaneously evaluate students' performance, emphasizing their primary challenges.

  Challenges:
  - o Besides attending face-to-face classes, students are encouraged to use the virtual assistant for further exploration and review of covered concepts or to tackle formative worksheets. The interaction will employ the question-answer method, culminating in feedback provided by the assistant. The assistant's functionality will be expanded to remember past conversations, with a specified character limit to prevent unforeseen cost escalation. In order to evaluate the effectiveness of this tool, we should: add only specific content to the assistant; allow access to the assistant for selected groups, aiming to compare their performances with groups that do not have access; Evaluate the students' achievements in comparison with prior years; Apply inquiries to students to assess the integration of chatbots into their teaching and learning process. Ultimately, a statistical analysis will be conducted using the collected data.

Evaluation of results

Several methods will be employed to assess the efficacy of the AI tool developed as part of the doctoral thesis:

- Measure the improvement in students' academic performance, such as grades, test scores, or completion rates.
- Retention - An analysis will be conducted on the retention rates of students who have experienced AI-enhanced instruction compared to those who have not.
- The quality and quantity of feedback provided to students by the AI tool will be evaluated. Educators and students have the option to confirm the feedback they've received by assessing it through a quantitative scale, and they may also choose to include a brief comment.
- Accounting for the time saved by students and instructors through the use of AI in various learning tasks, such as assessment, personalized feedback provision, or instructional content delivery.
- Collect feedback from students and instructors regarding their satisfaction with AI-enhanced learning experiences: surveys and/or interviews.
- The data collected by the AI tool (as stated in the previous points) will be analyzed using techniques such as machine learning and data mining. Once the analysis process is finished, reports and forecasts will be produced, providing the educator with all the tools needed to make optimal decisions for enhancing the teaching-learning process.

Timetable

First Year:

In the first year, I will start with the literature review and problem definition, subsequently outlining the research strategies. Next, I'll analyse in detail the various available AI models, gathering their advantages and disadvantages. After determining the models to use, I'll conduct a detailed analysis of their APIs. Finally, I will develop a prototype of a prompt and voice-activated chatbot capable of analysing provided documentation, generating content, and providing responses and learning suggestions.

Second Year:

In the second year, i will design, implement, integrate, and support the use of a platform in the teaching-learning process, followed by the collection and analysis of all relevant data.
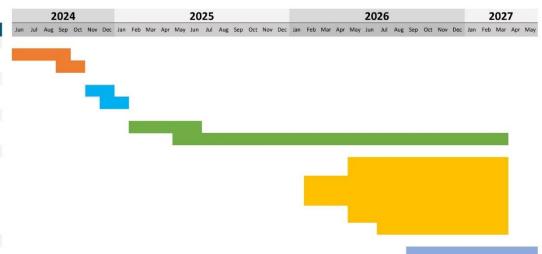
Third Year:

During the third year, I will be working on my doctoral dissertation.

introEduAI – Evaluation Tool to Assess the Impact of Artificial Intelligence in Education

Project start date: Jun-24

Description	Start	End
Literature review		
State of the art	Jun-24	Sep-24
Draft research proposal	Sep-24	Oct-24
AI Models definition		
LLM AI models research and selection	Nov-24	Dec-24
Analysis and testing of Model API	Dec-24	Jan-25
Development		
Develop Prototype chatbot (prompt and voice)	Feb-25	Jun-25
Build, test, and maintain a web-based system incorporating a chatbot	May-25	Jan-26
Data collection and analysis		
Conducting surveys with educators to evaluate the quality and relevance of the chatbot-generated content	May-26	Mar-27
Assessment validation by the educator regarding the report provided by the chatbot on formative evaluations	Feb-26	Mar-27
Collect data related to student assessments using introEduAI	Feb-26	Mar-27
Administer surveys to students to assess the use of chatbots in their teaching and learning process	May-26	Mar-27
Data analysis	Jul-26	Mar-27
Development		
Drafting the doctoral thesis	Sep-26	May-27

References

Atlas, Stephen. 2023. «ChatGPT for Higher Education and Professional Development: A Guide to Conversational AI». *College of Business Faculty Publications*, janeiro. https://digitalcommons.uri.edu/cba_facpubs/548.

Cunningham-Nelson, Sam, Wageeh Boles, Luke Trouton, e Emily Margerison. 2019. «A Review of Chatbots in Education: Practical Steps Forward». Em *30th Annual Conference for the Australasian Association for Engineering Education (AAEE 2019): Educators Becoming Agents of Change: Innovate, Integrate, Motivate*, 299–306. Australia: Engineers Australia. https://search.informit.com.au/documentSummary;dn=068364390172788;res=I ELENG.

Fryer, Luke, David Coniam, Rollo Carpenter, e Diana Lăpușneanu. 2020. «Bots for Language Learning Now: Current and Future Directions», junho. http://hdl.handle.net/10125/44719.

Gallacher, Andrew, Andrew Thompson, e Mark Howarth. 2018. «"My Robot Is an Idiot!" – Students' Perceptions of AI in the L2 Classroom». Em *Future-Proof CALL: Language Learning as Exploration and Encounters – Short Papers from EUROCALL 2018*, por Peppi Taalas, Juha Jalkanen, Linda Bradley, e Sylvie Thouësny, 70–76. Research-publishing.net. https://doi.org/10.14705/rpnet.2018.26.815.

Grassini, Simone. 2023a. «Kasneci». *Education Sciences* 13 (7): 692. https://doi.org/10.3390/educsci13070692.

———. 2023b. «Shaping the Future of Education: Exploring the Potential and Consequences of AI and ChatGPT in Educational Settings». *Education Sciences* 13 (7): 692. https://doi.org/10.3390/educsci13070692.

Guo, Ying, e Daniel Lee. 2023. «Leveraging ChatGPT for Enhancing Critical Thinking Skills». *Journal of Chemical Education*, novembro. https://doi.org/10.1021/acs.jchemed.3c00505.

Hadi Mogavi, Reza, Chao Deng, Justin Juho Kim, Pengyuan Zhou, Young D. Kwon, Ahmed Hosny Saleh Metwally, Ahmed Tlili, et al. 2023. «Wollny A qualitative study exploring early adopters' utilization and perceptions». *Computers in Human Behavior: Artificial Humans*, novembro, 100027. https://doi.org/10.1016/j.chbah.2023.100027.

Haristiani, Nuria. 2019. «Artificial Intelligence (AI) Chatbot as Language Learning Medium: An Inquiry». *Journal of Physics: Conference Series* 1387 (1): 012020. https://doi.org/10.1088/1742-6596/1387/1/012020.

Huang, Weijiao, Khe Foon Hew, e Luke K. Fryer. 2022. «Chatbots for Language Learning—Are They Really Useful? A Systematic Review of Chatbot-Supported Language Learning». *Journal of Computer Assisted Learning* 38 (1): 237–57. https://doi.org/10.1111/jcal.12610.

Jiang, Hua, Yang Cheng, Jeongwon Yang, e Shanbing Gao. 2022. «AI-powered chatbot communication with customers: Dialogic interactions, satisfaction, engagement, and customer behavior». *Computers in Human Behavior* 134 (setembro): 107329. https://doi.org/10.1016/j.chb.2022.107329.

Kasneci, Enkelejda, Kathrin Sessler, Stefan Küchemann, Maria Bannert, Daryna Dementieva, Frank Fischer, Urs Gasser, et al. 2023. «ChatGPT for good? On opportunities and challenges of large language models for education». *Learning and Individual Differences* 103 (abril): 102274. https://doi.org/10.1016/j.lindif.2023.102274.

Kohnke, Lucas. 2022. «A qualitative exploration of student perspectives of chatbot use during emergency remote teaching». *International Journal of Mobile Learning and Organisation* 16 (4): 475–88. https://doi.org/10.1504/IJMLO.2022.125966.

Kohnke, Lucas, Benjamin Luke Moorhouse, e Di Zou. 2023. «ChatGPT for Language Teaching and Learning». *RELC Journal* 54 (2): 537–50. https://doi.org/10.1177/00336882231162868.

Kuhail, Mohammad Amin, Nazik Alturki, Salwa Alramlawi, e Kholood Alhejori. 2023. «Interacting with Educational Chatbots: A Systematic Review». *Education and Information Technologies* 28 (1): 973–1018. https://doi.org/10.1007/s10639-022-11177-3.

Lin, Michael Pin-Chuan, e Daniel Chang. 2020. «Enhancing Post-Secondary Writers' Writing Skills with a Chatbot: A Mixed-Method Classroom Study». *Educational Technology & Society* 23 (1): 78–92.

Lo, Chung Kwan. 2023. «What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature». *Education Sciences* 13 (4): 410. https://doi.org/10.3390/educsci13040410.

Murad, Dina Fitria, Muhamad Irsan, Pas Mahyu Akhirianto, Erick Fernando, Silvia Ayunda Murad, e Michael Hendra Wijaya. 2019. «Learning Support System using Chatbot in "Kejar C Package" Homeschooling Program». Em *2019 International Conference on Information and Communications Technology (ICOIACT)*, 32–37. https://doi.org/10.1109/ICOIACT46704.2019.8938479.

Rudolph, Jürgen, Samson Tan, e Shannon Tan. 2023. «ChatGPT: Bullshit Spewer or the End of Traditional Assessments in Higher Education?» *Journal of Applied Learning and Teaching* 6 (1): 342–63. https://doi.org/10.37074/jalt.2023.6.1.9.

Tlili, Ahmed, Boulus Shehata, Michael Agyemang Adarkwah, Aras Bozkurt, Daniel T. Hickey, Ronghuai Huang, e Brighter Agyemang. 2023. «What If the Devil Is My Guardian Angel: ChatGPT as a Case Study of Using Chatbots in Education». *Smart Learning Environments* 10 (1): 15. https://doi.org/10.1186/s40561-023-00237-x.

Troussas, Christos, Akrivi Krouska, e Maria Virvou. 2017. «Integrating an Adjusted Conversational Agent into a Mobile-Assisted Language Learning Application». Em *2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI)*, 1153–57. https://doi.org/10.1109/ICTAI.2017.00176.

Winkler, Rainer, e Matthias Soellner. 2018. «Huang». *Academy of Management Proceedings* 2018 (1): 15903. https://doi.org/10.5465/AMBPP.2018.15903abstract.

Wollny, Sebastian, Jan Schneider, Daniele Di Mitri, Joshua Weidlich, Marc Rittberger, e Hendrik Drachsler. 2021. «Are We There Yet? - A Systematic Literature Review on Chatbots in Education». *Frontiers in Artificial Intelligence* 4. https://www.frontiersin.org/articles/10.3389/frai.2021.654924.

Zhai, Xiaoming. 2022. «ChatGPT User Experience: Implications for Education». SSRN Scholarly Paper. Rochester, NY. https://doi.org/10.2139/ssrn.4312418.