

Empathetic Intelligent Virtual Assistant to Support Distance Learning: Systematic Review

Claudia De Armas*

Universidade de São Paulo, Brazil

Email: claudiadearmas@usp.br

ORCID iD: <https://orcid.org/0000-0002-7830-1247>

*Corresponding Author

Anderson A. A. Silva

Universidade de São Paulo, Brazil

Instituto de Pesquisas Tecnológicas (IPT), Brazil

Universidade Paulista (UNIP), Brazil

Centro Universitário SENAC, Brazil

Faculdade de Informática e Administração Paulista (FIAP)

Email: anderson@uol.com.br

ORCID iD: <https://orcid.org/0000-0001-5426-6478>

Fabio S. Oliveira

Instituto de Pesquisas Tecnológicas (IPT), Brazil

Email: fabiod.oliveira@uol.com.br

ORCID iD: <https://orcid.org/0000-0003-2442-9736>

Sergio T. Kofuji

Universidade de São Paulo, Brazil

Email: kofuji@usp.br

ORCID iD: <https://orcid.org/0000-0002-8477-4091>

Romero Tori

Universidade de São Paulo, Brazil

Email: tori@usp.br

ORCID iD: <https://orcid.org/0000-0001-9381-9565>

Adilson E. Guelfi

Universidade do Interior Paulista (UNOESTE), Brazil

Email: guelfi@unoeste.br

ORCID iD: <https://orcid.org/0000-0002-4636-1878>

Received: 27 October, 2024; Revised: 13 January, 2025; Accepted: 09 March, 2025; Published: 08 April, 2025

Abstract: Distance learning as a modality has been growing for some time, however it received a significant boost from 2020. As teachers' activities have increased, Intelligent Virtual Assistants (IVAs) have helped to cope with the high workload and volume of requests. However, IVAs that include modules on empathy and teaching personalization are scarce. In the current work, we intend to map, through a systematic literature review, the level of maturity of the IVAs and how they can include empathy and personalization to improve results in conversations. The study considers a systematic review methodology that analyzes a series of works involving the use of IVAs and empathetic modules, and the platforms, resources, and functionalities available. We demonstrate the relevance of the topic in the scientific area, the diversity of countries involved, and the limitations and challenges that still need to be discussed.

Index Terms: Cognitive Assistants, Intelligent Virtual Assistant, Natural Language Processing, Distance education and online learning, Artificial Intelligence

1. Introduction

Education is a service area which experiences constant evolution and change, with the aim of improving the learning and management process. From the year 2020 a migration from face-to-face to virtual teaching was necessary. Although this modality has existed for a long time, some institutions had not previously considered having to implement it and were then obliged to do so in order for students to continue their studies. This brought a re-adaptation of teaching methods and planning in these institutions, for both teachers and students [1]. Although migration has been a difficult task, it was made possible with the good practices and technological advances available. However, engaging with virtual platforms is still a challenge for many teachers and students [2]. In addition, the effectiveness of the learning is questionable, and student retention is low, especially in fully asynchronous modalities [3]. In these cases, it is common to have long response times, low student engagement, and high dropout rates. These are some of the problems that need to be addressed in order to ensure quality teaching and an increase in the number of graduating students.

In this context, assistants that can support both the student and the teacher can be useful within the context of smart cities and education 4.0. Advances in the field of Artificial Intelligence (AI) and cloud computing have enabled the creation of Intelligent Virtual Assistants (IVAs) or Advanced Chatbots, which represent an innovative form of user service. An IVA is a conversational software system designed to emulate the communication capabilities of a human to interact with a user automatically [4]. Systems of this type can handle a high volume of requests with uninterrupted availability, respond immediately, interact with customers, and be integrated into different systems. Chatbots can be applied in education as a smart tutor, in self-learning, in collaborative learning, in mediation of learning, and in the use of Frequently Asked Questions (FAQs) [5].

In this way, it is possible to envisage solutions to some of the problems mentioned above, as IVAs can provide technical support with full availability and immediate response times and serve as a helpful guide in the use of platforms. In addition, they can establish a more objective student-teacher connection, supporting the management of activities, schedules, and timetables.

For this, some works focus on the inclusion of empathetic modules. These modules are associated with sentiment analysis which can be performed in many ways. Most research focuses on sentiment analysis through text, while little research analyzes facial expressions, voice, or multimodal data. As feelings and emotions directly influence the level of commitment and motivation to carry out an activity, the teacher's mediation is essential to boost and enhance positive feelings and mitigate negative emotions [6]. Gathering this information with technology to personalize the treatment of students and course planning is still a challenge.

However, not all approaches to empathetic IVA development yield equivalent outcomes. For example, text-based sentiment analysis has shown utility in asynchronous communication contexts but struggles with real-time adaptation [7]. In contrast, multimodal systems that integrate facial and vocal emotion recognition provide richer emotional context, enhancing empathetic responses [8, 9]. Yet, such systems are often resource-intensive, requiring specialized hardware and extensive data for training, posing challenges for scalability in broader educational settings.

The objective of the current paper is to group and analyze the studies that developed IVA with empathetic modules and personalization. Understanding the state of the art, level of maturity, and relevance are some of the aims of this research, developed using detailed planning of a systematic literature review.

The article is divided into sections and subsections: the current study, materials and methods, analysis and results, and finally, conclusions.

2. Materials and Methods

In this section, the research questions that motivate this research are presented, as well as the databases and the string used to survey the articles. The conduction and procedures established in the systematic review are detailed, as well as the process flowchart and the initial surveys.

Research questions

In order to study IVAs with empathetic modules in a deeper way, we elaborated the following research questions:

Q1: Is it a current topic and researched worldwide?

Q2: Are there empathetic academic-focused IVAs to support online education?

Q3: What platforms or resources are used to create IVAs and what functionalities do they have?

Q4: What are the challenges or limitations of IVAs in the area of online education?

Databases

We conducted the search in the Web of Science (WoS), IEEE, and Scopus databases using the periods, search fields, and search terms defined in the (Table 1).

Table 1. References, Databases, periods, search fields, and string

	Web of Science (WoS)	IEEE	Scopus
Period	1900-2023	1884-2023	1900-2023
Search field	Topic	Data field	Title, abstract and keywords
String	("virtual assistant*" OR chatbot* OR "conversational bot" OR "conversational agent*" OR "educational bot" OR "advanced chatbot" OR chatterbot) AND ("online education" OR "distance learning" OR "distance education" OR e-learning OR "online learning" OR "MOOC" OR education) AND ("affective computing" OR "emotion recognition" OR emotion* OR "speech analysis" OR "Sentiment analysis" OR "facial expression" OR "affective learning")		

To manage the articles collected, facilitate the identification and exclusion of duplicate articles, and centralize all the information obtained, we used the software StArt¹ (State of the Art through Systematic Review) - Version 2.3.4.2 [10]. We stored the result in a bibtex file and used Start to import databases with relevant metadata for the initial filtering: authors, number of citations, year, title, keywords, and abstract.

3. Procedure

The search was initially performed on 01/23/2022 and updated on 03/23/2023. The accepted languages were Spanish, Portuguese, and English. We also defined some criteria for the exclusion of works resulting from the search:

(E1) Abstracts, incomplete works, or those at a very early stage that do not allow the research questions to be answered;

(E2) Works that implement an IVA, or related modality, without using AI to capture emotions or that do not use affective computing;

(E3) Works that carry out applications outside the area of education;

(E4) Works that do not develop or apply an IVA or related modality.

We carried out a step-by-step process until we reached the final number of articles accepted for analysis. Initially, we identified 199 articles using the defined search terms, in the following distribution: 55 articles from the IEEE base, 101 from Scopus, and 43 from WoS.

We then applied the first search filter, derived from the application of exclusion criteria in the title and abstract of each work, which led to the exclusion of a total of 154 articles, 32 of which were duplicates and 122 that met the predefined exclusion criteria. Regarding criterion E1, we excluded articles presented in conference proceedings, in the very early stages, incomplete or with only a proposal, that is, without development and without results. Considering E2, we excluded articles that developed simple IVAs, only with the exchange of messages without including sentiment analysis through any intelligent capture method, that is, articles that did not consider aspects of AI in relation to Affective Computing. As for E3, despite incorporating keywords related to education in the search term, the search result identified articles from the medical field and automation. Finally, regarding E4, we excluded articles that focused only on a specific topic within the researched universe. For example, articles focused on chatbot evaluation, studies centered on the analysis of feelings in assistants, as well as works focused on identifying offensive responses when interacting with chatbots or creating censorship systems. Finally, we excluded articles that did not develop or apply an IVA. The details of the quantities and reasons for exclusion are detailed in the flowchart represented in Fig. 1. The most frequent exclusion criteria during the survey were E3 and E4. After this verification process, a total of 45 articles remained: 14 (IEEE), 24 (Scopus), and 7 (WoS).

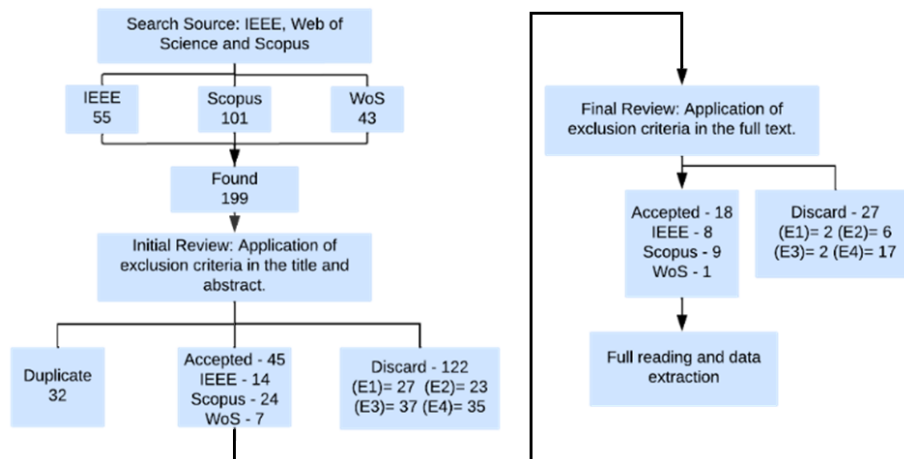


Fig. 1. Systematic review flowchart

¹ http://lapes.dc.ufscar.br/tools/start_tool

To select the final works to be analyzed, we performed the last filtering step. The purpose of this step is to apply the criteria to the full text of each article. Finally, we separated 18 articles, consolidated in Table 2, and rejected another 27, based on the distribution by criteria presented in the Fig. 1.

The main exclusion criterion was E4, since most of the rejected works do not present application development with IVA, contrary to what was expected in the initial filtering, where we analyzed only the abstract and title. Many of these excluded works implement 3D conversational agents within a subject to support a specific content, developed as an instructional agent and, in most cases, students are only able to interact with predefined responses within a set of multiple responses. Other works also address the issue of responses and affective feedback, but starting from the emotional condition that the student declares through a questionnaire. That is, there is no intelligent capture in which the system itself identifies the emotional state of students. The objective of the current research goes beyond these points, as it proposes an intelligent system that captures students' emotions in a passive way, that is, without student intervention.

Table 2. Summary of works accepted for analysis

Reference	Country	Aim	Data Base
[18]	Tunisia	To propose a multi-agent system with agents that manage the cognitive and affective model of the learner, which are able to express emotions through Emotional Embodied Conversation Agents (EECA) and recognize the learner's facial expression through emotional agents in peer interaction.	Scopus
[19]	China	To introduce an interactive care agent that uses facial expressions, hand gestures, body movements, and social dialogue to show empathy and active listening to interact with students in a learning system.	IEEE
[30]	UK	To investigate the non-verbal behavior of a tutor and propose a model for Embodied Conversation Agents (ECAs) acting as virtual tutors, in addition to building an ECA with conversational skills, episodic memory, emotions and expressive behavior based on the result of the empirical study.	Scopus
[9]	Greece	To examine the impact of emotional facial expressions and tone of voice of ECAs, combined with empathic verbal behavior when displayed as feedback to students' emotions of fear, sadness, and happiness in the context of an interview.	IEEE
[17]	Brazil	To develop a Virtual Learning Companion, called Buti, to provide healthy lifestyle information to children and tweens, introducing Affective Computing; applying personality and emotion in order to obtain a better relationship between system and user.	Scopus
[20]	Netherlands	To create a prototype that provides cognitive, affective, and social learning support and compare it to a prototype that only provides cognitive support in an in- and between-subjects experiment.	Scopus
[21]	Belgium	To develop a 3D chatbot integrated into the learning management system using elements of cognitive therapy to help with procrastination problems, lack of study planning and communication problems (with colleagues and team).	WoS
[22]	Greece	To propose a framework for building pedagogy-oriented conversational agents based on Reinforcement Learning combined with Sentiment Analysis, also inspired by the theory of pedagogical learning of Basic Cognitive Skills.	Scopus
[15]	France & China	To propose a Collaborative Learning Environment using technology-supported tools organized as a system, with different learning approaches combining pedagogy and a neuroscience-based view of emotions and emotional approaches.	Scopus
[8]	China	To develop a Deep Neural Network (DNN) based on an emotionally aware campus virtual assistant.	IEEE
[25]	France	To present how an emotional chatbot that considers the emotional state of the student can offer an adapted response to the concerns that arise during their learning processes.	Scopus
[7]	Malaysia & Indonesia	To develop an intelligent chatbot using Neural Language Processing (NLP) and Telegram API	IEEE
[29]	Sri Lanka	To develop a chatbot using Natural Language and Dense Neural Network processing techniques.	IEEE
[24]	Russia	Application of Chatbot Technology in the studio of the Quality Management discipline for high quality learning and communication composed of artificial intelligence, an analysis component and an intelligent assistant based on RPA (Robotic Process Automation) systems.	Scopus
[23]	Spain	To present a virtual assistant to improve communication between the presenter and the audience during educational or general presentations to enhance interaction during presentations. The main objective is the design of an interaction framework to increase the audience's attention level.	Scopus
[28]	China	To combine the technology acceptance model (TAM) and emotional intelligence (EI) theory to evaluate the elements influencing the adoption of chatbot applications by implementing the proposed theoretical model.	IEEE
[26]	Republic of Mauritius	To determine whether a virtual tutor can effectively deal with student queries in online distance learning environments	Scopus
[16]	China & Turkey & USA & Australia	To examine ChatGPT in education, among early adopters, through a qualitative instrumental case study conducted in three stages: (1) public discourse in social media; (2) educational transformation, response quality, usefulness, personality and emotion, and ethics; and (3) user experiences through ten educational scenarios.	IEEE

4. Scope

As the subject of this paper is diverse and the applications are many, there are several directions in which a survey like this can point. Precisely for this reason, it is extremely important to establish a scope to limit the coverage of the research and ensure that accurate and credible results are obtained. Therefore, the following comments are made on some topics that, although important, are not addressed in this research.

Pedagogical and psychological effects of empathic IVA in education: most studies suggest that, when well applied, empathic IVA has positive effects on education, mainly in improving student engagement, acting as a motivator and fostering more immersive, interactive and adaptive learning experiences [11]. Some studies also suggest that personalization of learning was more effective when empathic IVA features were put into practice [12, 13]. On the other hand, there are also criticisms about harmful pedagogical effects, such as reducing students' cognitive load due to easy access to information [14].

Performance analysis of IVA: it is not within the scope of this research to analyze the performance of the IVA studied. The main reason is the difficulty in replicating the tools, environments and results used in the researched references. Since, configurations or implementations different from those used in related works may generate false results in relation to the performance obtained, we do not perform performance comparisons.

5. Results and Analysis

From the surveyed works, we carried out complete reading and analysis of the texts in order to answer the research questions and survey the state of the art. Next, we divided each research question into subsections and answered each one based on the 18 included papers.

Is it a current topic and researched worldwide?

First, we tried to assess the relevance of this work proposal in the scientific area. For this, we analyzed the years (Fig. 2a) of the research and the territorial distribution of the works (Fig. 2b), in order to verify if there is a global and recent interest in the theme.

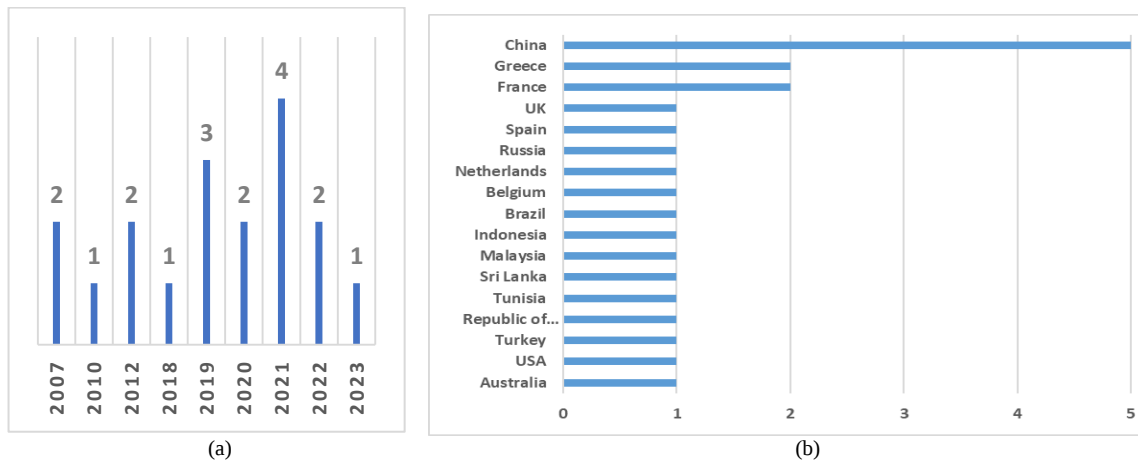


Fig. 2. (a) Distribution of research years (b) Distribution of research countries

By only considering the initial number of works identified (199) with the keywords in the area, we can state that it is a topic of interest and widely researched with similar objectives, but in different scales and contexts of application. The evaluation of the 18 accepted articles shows that IVA as a topic has been strongly returned to in recent years, but there is little research that addresses and deals with the subject. Observing the years of research, one can intuit that the impact of COVID-19 has driven the resumption of research. Finally, it is also acceptable to conclude that several countries are interested in the topics addressed. Furthermore, it is noteworthy that in three works [7, 15, 16] international cooperation between institutions was implemented. In Eurasia, wide distribution is visible with 90% of the total works, and only 5% for Latin America represented by one work from Brazil [17].

Are there empathetic academic-focused IVAs to support online education?

Our research shows that yes, some IVAs are developed only for the academic area, but we did not find any product being marketed for this purpose. Under the research context, we included a total of 18 works developing IVAs to support remote learning empathetically. Each author uses a different term for the created assistants, which we decided to maintain.

The first work is from 2007 [18]. The authors propose a multi-agent system that manages the student's cognitive and affective model and is capable of expressing emotions through Emotional Embodied Conversational Agents (EECA)

from the emotions captured from the students themselves. In their work, the authors create a system called an Emotional Multi-Agents System for Peer-to-peer E-Learning (EMASPEL), based on a multi-agent architecture that makes it possible to recognize the emotional state during learning in a Peer-to-Peer (P2P) network. Despite the work not including a development or validation experiment, we accepted it in the systematic review because it is one of the first works with a proposal similar to our own.

In a very similar context, experiments are created with an empathetic agent in the form of a humanoid avatar that modifies facial expressions and tone of voice in relation to the student's state of mind through intelligent capture (e.g., [9, 19]). Through three experiment scenarios, it is demonstrated that empathy is a characteristic that, when attributed to virtual agents, can help transform the student's emotional state [9]. Preliminary results reveal that the caregiver agent can not only increase the interaction between students and the learning system, but also have positive effects on students' emotions to make them more engaged in learning [19].

In 2012, a chatbot called Buti was developed, based on a research project. The aim of the chatbot is to encourage children aged between 7 and 10 to reflect on their own eating habits and sedentary lifestyle, teaching them how to improve these behaviors [17]. In 2018 the authors created the Integrated Conversation Agent called BUQUE, which provides cognitive, affective, and social learning support for people with low literacy levels [20]. In 2019, a proposal was developed for a 3D chatbot aimed at higher education. The virtual avatar integrates with the learner management system and uses cognitive therapy elements to help overcome typical problems that affect students, such as procrastination, lack of study planning, and communication problems [21].

In the same year, a collaborative learning environment supported by chatbot technologies was created, with the aim of promoting active learning in two ways: (1) individual; and (2) collaborative. The authors use a combination of pedagogy and neuroscience to reach emotions. The chatbot has the function of helping the teacher with the management of students individually and collectively. With the built-in cameras, it is possible to notify the teacher about the student's behavior [15]. In this same context, an agent is created based on the pedagogical concept of reinforcement learning, combined with sentiment analysis, in addition to the pedagogical learning theory of Basic Cognitive Skills [22].

In 2020, an emotional chatbot was proposed for the analysis of a discussion forum on a Distance Education (EAD) platform. The authors proposed reading the students' personality through the Five Factors Model to personalize learning on the platform. For this, they use different text analysis and data mining techniques [25].

In 2021, a survey presented the structure of an automated system with intelligent robots to be used within the Quality Management discipline. The intention of the article is to use a chatbot, connected with the institution's Enterprise Resource Planning (ERP) and with the BigData system, to compile information about students and disciplines, in addition to creating statistics. Affective computing is only used during the conversation with the bot through the semantic analysis of the text [24].

In the same year, an intelligent chatbot was developed using Natural Language Processing (NLP) and the Telegram API. For this, the authors use text processing with Python and recognize emotions in already recorded chats. Finally, they send the appropriate response to the user. The results show that the chatbot can interact with users without any problems, identifying their emotions [7].

On the other hand, a virtual assistant is also available to improve communication between the presenter and the audience. The assistant can be used in conferences to improve interaction during presentations in large auditoriums (e.g., 200 people). The authors design an interaction architecture to increase the audience's level of attention on key aspects of the presentation. The intelligent system needs to detect different patterns of public behavior, such as diverting/distracting attention, sleeping people, positive, negative, and confused participation [23].

For university students, some authors created a chatbot using NLP and Dense Neural Network (DNN) techniques. The article focuses mainly on the development of the bot, on which tests are carried out with very high levels of accuracy [29].

Finally, in the year 2023, an experiment was carried out with the aim of examining concerns about the use of chatbots in education based on ChatGPT. The results obtained reveal that, although ChatGPT is a powerful tool in education, it still needs to be used with caution, since some guidelines on safe use in education must be previously established [6].

In Table 3 it is possible to observe the evolution of the articles focused on online education in relation to trends in development and research based on the works raised.

Table 3. Research trend in Empathetic IVA Development (2015-2023)

Year	Focus Areas	Key Gaps Identified	Key Outcomes Identified	Methodologies Used	Notes on Early Work
2015-2020	Sentiment analysis, text-based	Lack of multimodal approaches	Basic emotion capture achieved	Text-based sentiment analysis, keyword tracking	Early multimodal attempts, e.g., Moridis & Economides (2012), were foundational but not widely adopted yet.
2021-2023	Multimodal sentiment systems	Limited scalability, insufficient real-time adaptation	Enhanced emotional context in interactions	Integration of facial, vocal, and multimodal AI	Multimodal methods became prominent, with broader adoption and advancements in AI technologies.

What platforms, resources and tools are used to create IVAs and what features do they have?

For the development of the IVA some works use a popular strategy, the use of Questions and Answers (QA). These are models that are quite easy to implement, especially when standard short answers are implemented. The scope of the model varies depending on the built-in intelligence, from the most basic with just chat to the most advanced with the inclusion of emotion analysis. To do this, the developers use a Deep Reinforcement Learning (DRL) model powered by QA systems to produce an IVA that can learn from the inputs it receives [22].

Commercial IVAs that are available on the market, such as Alexa², Siri³, and Google Assistant⁴, among others, do not present 3D representations, however, several works use efforts to create characters or avatars that use the Unity game engine for development (e.g., [10, 21]). Some works create female avatars [9] and others use the Poser 6 design and animation tool [19]. In all these works the common factor is the creation of an Embedded Conversational Agent, a vital differentiator in relation to traditional chatbots that use only text. Still within this context, we also see the development of avatars with a webcam design [23], that is, not all developed characters have a humanoid aspect. Another approach is to make virtual characters available for users to choose from [8]. Additionally, as the interaction method used is the smartphone, the authors also implement Augmented Reality (AR) to project the character in the student's environment, as shown in example of Fig. 3. In order to provide cross-platform features and an AR experience to the user, the System front-end interface was also implemented in Unity.



Fig. 3. Example of user interface and screen.

For data management, the authors use the Vertica system, which is an analytical database management software [24]. To raise common problems or concerns of students, a text mining technique is applied [25].

In the context of Affective Computing, we survey all the technologies used. iAIDL is a mechanism to treat intentionality in chatbots based on Conversational Analysis Theory (e.g., [17, 21]) - an OCC theory widely used by computer scientists is that proposed by Ortony, Clore, and Collins (OCC henceforth). It is one of the most popular when it comes to implementing computational models due to the large taxonomy of triggers and resulting emotions⁵.

The Facial Action Coding System (FACS) verifies whether basic emotions match the analyzed facial models (e.g., [9, 19, 20, 26]). That is, it consists of a deep study of the facial muscles and the representation of their contractions in a set of values called Action Units (AUs) [27]. Fig. 4 presents some representations of AUs categorized by upper and lower face action units.

In Fig. 4, for example, an AU could be the inner brow lift (AU1) or the cheek lift (AU6). Expressions can be formed by joining two or more AUs.

² <https://alexa.amazon.com.br/>

³ <https://www.apple.com/br/siri/>

⁴ <https://assistant.google.com/>

⁵ emotions related to the consequences of events (example: pity and joy), actions of agents (example: pride and disapproval), and appearance of objects (example, love and hate).

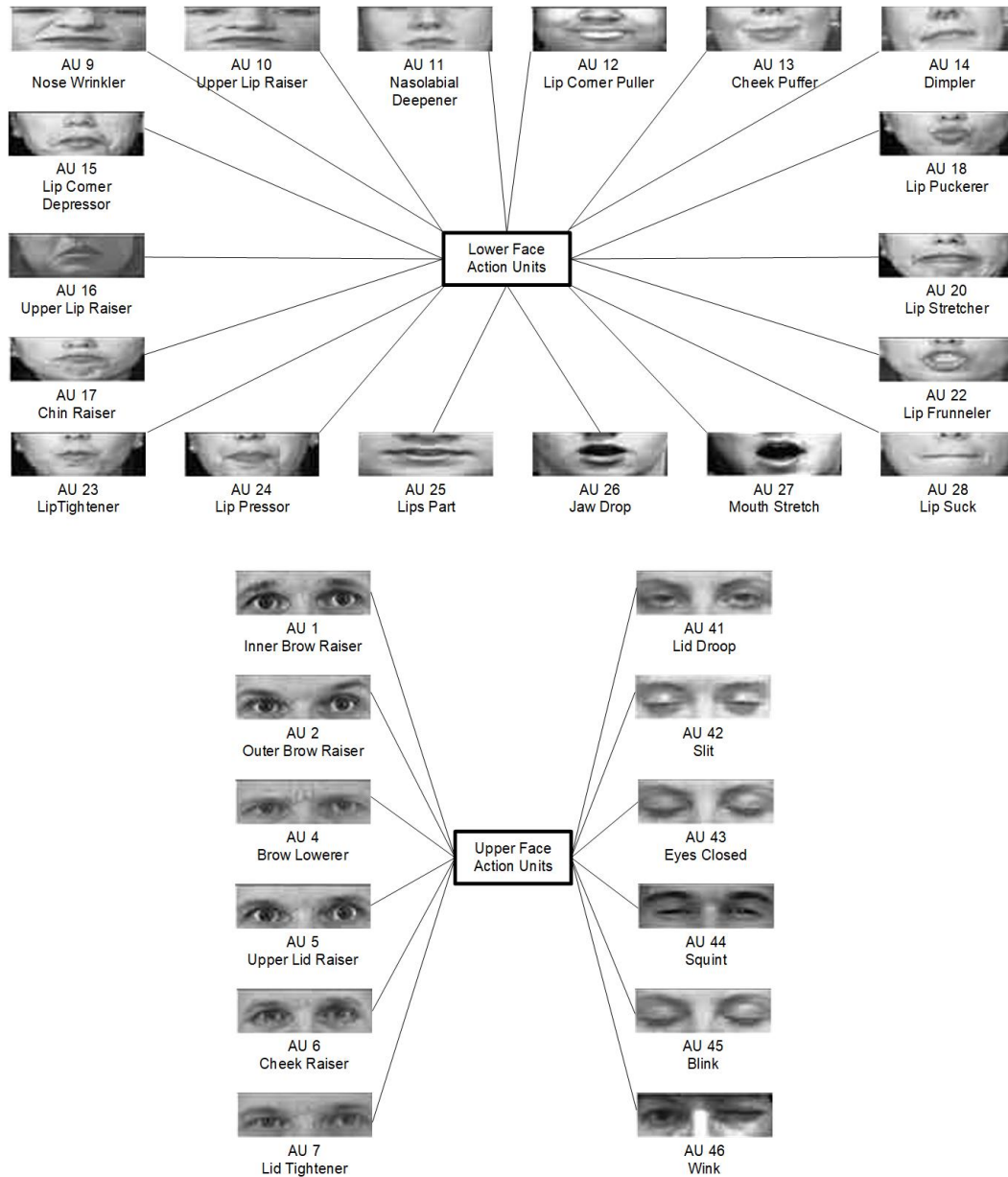


Fig. 4. Face Action Units for emotion analysis [27]

To determine students' emotions, passive methods are used through image capture and processing using a web camera or cell phone (e.g., [8, 15, 19-21]). Other authors use a commercial device called faceLAB⁶ from the company Seeing Machines⁷ which is used to track gaze, head, and facial features to elicit non-verbal behavior [30].

Another passive method is text analysis of data that the student inputs in the chat or forum. This processing is performed using NLP techniques (e.g. [7, 15, 19, 21, 22, 24, 25, 29]). In the mentioned works, all the tools used in both approaches are raised (image or text processing): Facereader⁸, Google NLP⁹, Aipia¹⁰, Textobserver¹¹, Parallel dots¹², Q Emotion¹³, Aylien¹⁴, and OntoSenticNet¹⁵. Table 4 presents the limitations, tools, and technologies used in each job, while, in a complementary way, Fig. 5 and 6 group the targets and results achieved.

⁶ <https://www.eyecomtec.com/3132-faceLAB>

⁷ <https://seeingmachines.com/>

⁸ <https://www.noldus.com/facereader>

⁹ <https://cloud.google.com/natural-language>

¹⁰ <https://www.aipia.info/>

¹¹ <http://explorationdecampus.corpusecrits.huma-num.fr/textobserver/>

¹² <https://www.paralldots.com/technology>

¹³ <https://www.qemotion.com/>

¹⁴ <https://aylien.com/>

¹⁵ <https://www.sentic.net/downloads/>

Table 4. Survey of related works

Reference	Term or name used	Technologies	Limitation
[18]	Emotional Embodied Conversational Agents (EECA)	MadKit Platform	No experimental development or validation
[19]	Embodied Conversational Agents (ECA) – Maggie	Avatar 3d, Text Analysis (NLP), Image capture and processing via Webcam	Limited scalability for large-scale deployment
[30]	Expressive Virtual Tutor	FaceLAB, Microsoft Speech Technology	Requires specialized hardware
[9]	Empathetic Agent	Avatar 3D, FaceReader, MySQL	High resource requirements
[17]	Chatterbot – Buti	iAIML, computational model of emotion	Focused on a specific demographic group
[20]	Embodied Conversational Agents (ECA) – Buque	Avatar 3D, Unity, Shimmer sensor pack, Facereader, Wizard of Oz Method	Low literacy-level target audience only
[21]	Chatbot 3D	Avatar 3D, Unity, AIML, Microsoft LUIS, Google Cloud Service	Lacks robust empirical evaluation
[22]	Pedagogical Conversational Agents	Deep Learning Model, QA System, Text Analysis	Theoretical proposal, lacks implementation
[15]	Virtual Assistant Chatbot	Image capture and processing, Alexa Skills Kit (ASK)	Limited adaptation to different contexts
[8]	Emotionally Aware Virtual Assistant	Speech recognition and analysis, Natural Language Processing NLP, Augmented Reality, Digits Nvidia, AlexNet, GoogLeNet	High computational demands
[25]	Emotional chatbot	Text mining, WordNet Affect, Google NLP, TextObserver, Parallel dots, Google Dialogflow	Limited to discussion forums
[7]	Intelligent Telegram Chatbot	API Telegram, Google Dialogflow, Aipia – NLP, SVM & Naïve Bayes	Lacks multimodal emotion recognition
[29]	Mind Relaxation Chatbot	Tensorflow, NLTK, numPY, TFLear, Pycharm IDE, Anaconda Navigator, NLP, AIML	No personalization capabilities
[24]	Chat-Bot	NativeChat, Vertica (Data Base), Morphological, syntactic, and semantic analysis	Limited emotional recognition
[23]	Virtual Assistant	Audio Analysis (Python-‘Pliers’ NLP), Image capture and processing, OntoSenticNet, Reinforcement Learning Models	Limited use case applications
[28]	Chatbot	Technology Acceptance Model (TAM), Simple Logistic Classifier, Iterative Classifier Optimizer, Logistic Model Tree (LMT), Weka tool	Focused on acceptance metrics
[26]	Virtual Tutor	Recurrent Neural Network (RNN), Convolutional Neural Network (CNN), T-distributed stochastic neighbor embedding (t-SNE)	Limited scalability in real-time applications
[16]	ChatGPT	OpenIA	Ethical and contextual concerns

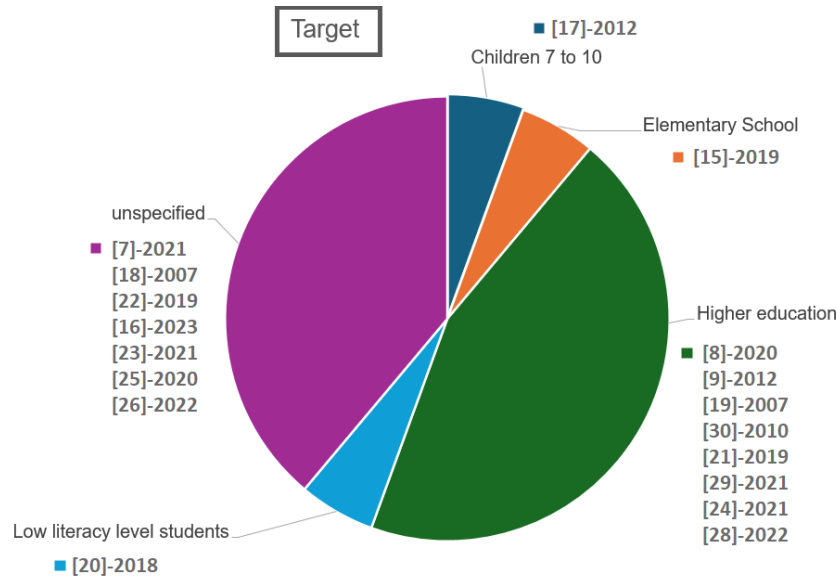


Fig. 5. Targets of related works

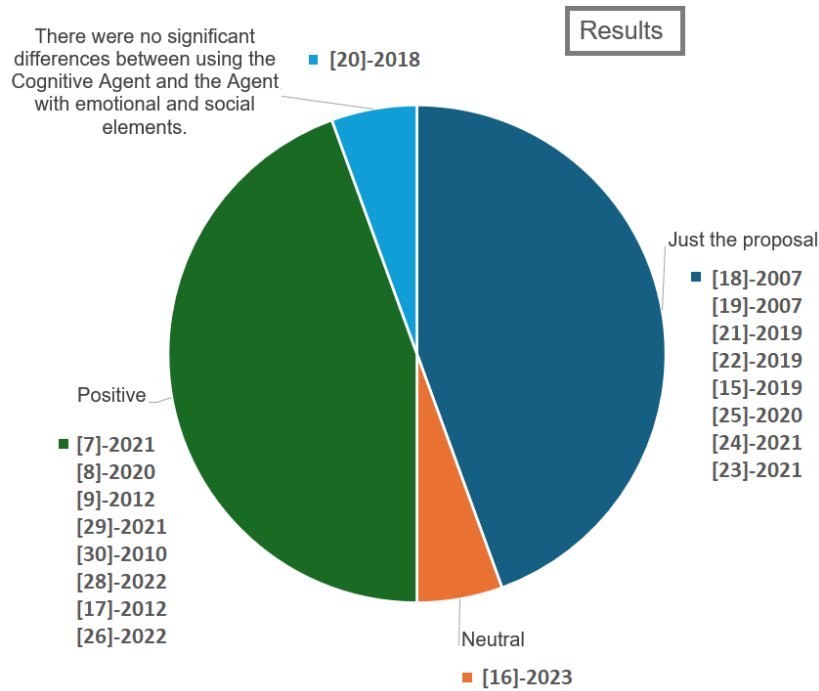


Fig. 6. Results of related works

Finally, six works use the Assistant's built-in speech recognition: Microsoft Speech Technology [30], Nvidia DIGITS¹⁶ [8]; Python open source solutions like 'Pliers' [23], which provides an extensible framework to automatically implement semantic feature extraction from audiovisual resources; Google Cloud Speech third-party services¹⁷ [21]; the Wizard-of-Oz method [20]; and the use of the Alexa Skills Kit (ASK) tool [15]. In articles that do not contain this technology, it is possible to see the interest in adding this feature to IVAs in future works and challenges [29].

Active methods are also raised through feedback from the students themselves, who use emojis to classify the emotions they are experiencing in the learning process [15], or through questionnaires [19, 20].

It is also possible to detect emotions through physiological measurements [20]. For this, the authors used the Shimmer¹⁸ sensor package to measure student arousal using a photoplethysmogram (PPG) attached to the earlobe. These sensors can measure heartbeat by measuring changes in light absorption that result from blood flow.

Finally, we also found works that mix different active and passive methods (e.g., [15, 19, 20]).

¹⁶ <https://developer.nvidia.com/digits>

¹⁷ <https://cloud.google.com/speech-to-text>

¹⁸ <https://shimmersensing.com/>

A common component in almost all works is the use of Machine Learning, specifically classification systems, both for text and images. Among the models used, we can mention: Simple Logistic Classifier, Iterative Classifier Optimizer Logistic Model Tree (LMT) [28], Reinforcement Learning Models [23], SVM & Naïve Bayes [7], and Deep Learning Model [22], among others.

An interesting element presented in the work that evaluates chatGPT is the feeling that the bot is understanding the conversation, since it uses the same words used in the question. This type of element or metric adds humanization and empathy to the task [16].

In the implementation part, some authors not only use the web, but also Telegram Bot to integrate with the cell phone, which, compared to other platforms such as MS Bot Framework, Chatfuel and Amazon Lex, has several advantages that the authors summarize in Table 5 [7]. Fundamentally Telegram bot has good documentation, is simple to use, is free, and supports several programming languages, such as PHP, Python, and Java.

Table 5. Comparison of different chatbot platforms [7]

Platform	Documentation	Ease of Use	Pricing	Program
MS Bot Framework			\$0.50 per 1000 messages	.Net, Javascript
Chatfuel			Free	Pyhton, PHP, Javascript
Amazon Lex			\$0.00075 per 1000 messages	Pyhton, PHP
Telegram Bot			Free	Pyhton, PHP, Java

Most studies focus on facial expression recognition, text-based sentiment analysis, or multimodal emotion detection as primary methods to incorporate empathy into Intelligent Virtual Assistants (IVAs). However, broader contextual elements such as course progress, student grades, and learning behavior metrics are notably absent. These elements could significantly enhance the ability of IVAs to offer more personalized and empathetic interactions.

This gap in research underscores the need for future IVAs to move beyond isolated emotional analyses. By incorporating broader educational metrics, IVAs can create more context-aware interactions that not only respond to immediate emotions but also anticipate and adapt to long-term learning needs. The limitations highlighted here represent a critical opportunity for advancing the empathetic capabilities of IVAs in education.

What are the challenges or limitations of IVAs in the area of online education?

While some authors use NativeChat to organize dialogue [24], others explain the importance of determining the target audience of the application [17] since these systems need to follow human development to adapt to the user's environment. That is, adaptations are necessary, as it is not always possible to adopt the same system for adults, young people, and children. The adequacy of interaction arises because children are in the process of human development, especially with regard to language, reflection, and logic. Another adaptation is derived from the interests and motivations of the audience. For example, for children's use, graphic components and figures represent important aspects that must be added to the systems, while in the case of adults these resources are not as essential.

Regarding the placement of avatars, it is very important to take into account the result of the experiment presented in one of the works studied [19] In the study, several of the students included in the experiment argued that they did not like Maggie (Virtual Agent) because they considered that she was not a real person. This point is interesting, because instead of helping in the learning process, this avatar becomes a character that does not convey confidence and seriousness about the subject, generating a perception more focused on games and fun than on education. Gamification can be a controversial subject when you consider that there is an audience that doesn't like games or animated characters.

Among other limitations raised is the fact that students still prefer to communicate with the teacher and not with the robot [24]. In addition, during communication, the student can make mistakes that the robot can misinterpret, generating a bad situation for the student. In this context, some authors attest that one of the biggest limitations of this type of research is getting the chatbot to think like a human being, with wisdom and empathy [29]. It helps a lot if the chatbot knows how to deal with hypernyms, synonyms, and hyponyms, instead of just dealing with specific terms. In this sense, we can foresee the next challenge, which is the need for large amounts of data for effective training and learning of assistants [22]. Response quality is vital for the success and effective adoption of chatbots for school operations [16].

Although many studies are focused on the level of acceptance of chatbots by students, little research points to acceptance by teachers [28]. It is important to highlight that the development of a competent chatbot system requires a

multidisciplinary team, and the same importance should be attributed to technical development as to the content of the system [21]. In this last point, the participation of teachers is more present, as in the reforms and adaptations of contents and teaching methods for the use of chatbots. Instructions on how they should be used are also relevant for good use and enjoyment of technology [16].

6. Conclusions

To carry out this work, 4 research questions were designed, which were answered following a strict systematic review methodology. During the research we were able to understand that the topic studied is of current and worldwide interest. The works form part of the research of 17 different countries, and the fruits of international cooperation were concentrated in the last 5 years. The challenge of creating a broad-ranging IVA that is able to understand the student's needs, that manifests empathy, and that is able to personalize learning in an instructional way still exists. The need for this type of IVA is growing with the rise in distance courses. Maintaining the quality of education and supporting teachers is a current need in educational institutions.

Despite being a relevant topic, with initial survey results of 199 studies, the use of empathetic modules through intelligent capture for teaching personalization is a scarce topic in literature. Most of the works excluded from the review do not feature intelligent capture of emotions or focus more on giving instructions than talking. In addition, instructional agents are more frequent than conversational agents, and within the context of this paper, the ideal would be a mix between instructional and conversational. One of the conclusions, therefore, is that instructional agents do not have enough intelligence to go beyond the scope of the initial conversation. They are programmed to pass on information in a closed domain and their intelligence does not need to be as sophisticated as a mixed agent. Developing an agent that can truly understand students' concerns and transform them into instructions to support and mediate student-teacher interaction is still a challenge.

In discussions about the techniques and methods used, it is possible to verify that there are several technologies, tools, and resources that can support the development of these IVAs using NLP techniques, AI, and specifically Deep Learning, since cloud computing technologies that support the creation of conversational assistants with AI techniques are becoming increasingly improved. Research in recent years on the topic is increasing and the introduction of empathetic elements is gaining strength. The direction points to the humanization of the agents and the incorporation of elements related to feelings to achieve not only effective interactions, but also affective ones.

The rapid advancement of artificial intelligence has created unprecedented opportunities to bridge emotional intelligence with educational technology. This study highlights the scarcity of empathetic IVAs in education and identifies critical gaps in their design and implementation. While research has made strides in integrating text-based and multimodal sentiment analysis, the practical application remains hindered by resource limitations, scalability challenges, and insufficient adaptation to real-time educational needs.

To address these gaps, we propose a comprehensive framework encompassing technical innovation, pedagogical alignment, and cultural adaptation. For technical innovation, the focus should be on modular, resource-efficient designs capable of scaling across diverse educational contexts. Multimodal approaches must leverage advancements in machine learning and cloud technologies to optimize real-time responsiveness and emotional understanding. In terms of pedagogical alignment, collaboration with educators is essential to co-design IVAs that align with curriculum goals and foster student-teacher synergy. A user-centered design approach ensures the integration of meaningful, context-aware interactions. Finally, cultural adaptation involves tailoring IVAs to reflect linguistic, cultural, and regional nuances, ensuring their relevance and inclusivity in global educational scenarios.

The future of empathetic IVAs lies in fostering interdisciplinary collaboration among technologists, educators, and psychologists. By integrating emotional intelligence into educational tools, these systems can enhance student engagement, retention, and satisfaction. However, achieving this vision requires a shift from descriptive implementations to proactive, context-sensitive, and scalable solutions.

As limitations, the work was directed towards research with experiments focused on online education, that is, works developed in face-to-face environments were disregarded. Works that did not include affective computing in the form of an empathetic module for tutor development were also excluded. Finally, the work is limited to research contained in the databases studied IEEE, Scopus, and Web of Science.

References

- [1] Rocha, E.M., & Lima, J.D. Impactos e desafios do ensino on-line decorrentes da pandemia COVID-19. *Revista Ibero-Americana de Estudos em Educação*, 377-390, 2021. <https://doi.org/10.21723/riaee.v16i2.14526>.
- [2] Dube, S. The 21st Century Students' Educational Ict Preferences. *International Journal of Robotics and Automation*, 3. <https://doi.org/10.15406/iratj.2017.03.00069>.
- [3] Goel, A. K., & Polepeddi, L. Jill Watson: A Virtual Teaching Assistant for Online Education. *Georgia Tech Library*, 1–21. 2017. <https://doi.org/10.4324/97813511861937>.
- [4] Nuruzzaman, M., & Hussain, O.K. A Survey on Chatbot Implementation in Customer Service Industry through Deep Neural Networks. 2018 IEEE 15th International Conference on e-Business Engineering (ICEBE), 54-61, 2018. <https://doi.org/10.1109/ICEBE.2018.00019>.

- [5] Larisane K., N., Antunes, C. A., Jo ão de Barros Vanzin, V., Luis Tavares da Silva, J., Loureiro Krassmann, A., & Margarida Rockenbach Tarouco, L. Chatbots na educa ção: uma Revis ão Sistem ática da Literatura. *Renote*, 123–132, 2018. <https://doi.org/10.22456/1679-1916.86019>.
- [6] Silva, R. F. d. As emo ções e sentimentos na rela ção professor-aluno e sua import ância para o processo de ensino e aprendizagem: Contribui ções da teoria de Henri Wallon. 2017. <http://hdl.handle.net/11449/150708>
- [7] Gunawan, T.S., Falemlula Babiker, A.B., Ismail, N., & Effendi, M.R. (2021). Development of Intelligent Telegram Chatbot Using Natural Language Processing. 2021 7th International Conference on Wireless and Telematics (ICWT), 1-5. <https://doi.org/10.1109/icwt52862.2021.9678471>
- [8] Chiu, P., Chang, J., Lee, M., Chen, C.H., & Lee, D. Enabling Intelligent Environment by the Design of Emotionally Aware Virtual Assistant: A Case of Smart Campus. *IEEE Access*, 8, 62032-62041, 2020. <https://doi.org/10.1109/ACCESS.2020.2984383>
- [9] Moridis, C.N., & Economides, A. A. Affective Learning: Empathetic Agents with Emotional Facial and Tone of Voice Expressions. *IEEE Transactions on Affective Computing*, 3, 260-272, 2012. <https://doi.org/10.1109/T-AFFC.2012.6>
- [10] Zamboni, A. B., Thommazo, A. D., Hernandez, E. C. M., & Fabbri, S. C. P. F. StArt – Uma ferramenta para apoio ao processo de revis ão sistem ática. *Brazilian Conference on Software: Theory and Practice - Tools Session*, 2010. https://nemo.inf.ufes.br/wp-content/papercite-data/pdf/revisis_uma_ferramenta_de_apoio_a_realizacao_de_estudos_baseados_em_revisao_sistemica_da_literatura_2012.pdf.
- [11] Grivokostopoulou, F. et al. The Effectiveness of Embodied Pedagogical Agents and Their Impact on Students Learning in Virtual Worlds. *Applied Sciences*. 2020. <https://doi.org/10.3390/app10051739>.
- [12] Katsarou, E. et al. A Systematic Review of Voice-based Intelligent Virtual Agents in EFL Education. *Int. J. Emerg. Technol. Learn.*, 18, 2023, 65-85. <https://doi.org/10.3991/ijet.v18i10.37723>.
- [13] Nikolopoulou, K. et al. Generative Artificial Intelligence in Higher Education: Exploring Ways of Harnessing Pedagogical Practices with the Assistance of ChatGPT. *International Journal of Changes in Education*. 2024. <https://doi.org/10.47852/bonviewijce42022489>.
- [14] Sajja, R. et al. Artificial Intelligence-Enabled Intelligent Assistant for Personalized and Adaptive Learning in Higher Education. *ArXiv*, abs/2309.10892. 2023. <https://doi.org/10.48550/arXiv.2309.10892>.
- [15] David, B., Chalon, R., Zhang, B., & Yin, C. Design of a Collaborative Learning Environment integrating Emotions and Virtual Assistants (Chatbots). 2019 IEEE 23rd International Conference on Computer Supported Cooperative Work in Design (CSCWD), 51-56, 2019. <https://doi.org/10.1109/CSCWD.2019.8791893>
- [16] Tlili, A., Shehata, B., Adarkwah, M.A., Bozkurt, A., Hickey, D.T., Huang, R., & Agyemang, B. What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10, 1-24, 2023. <https://doi.org/10.1186/s40561-023-00237-x>.
- [17] Jacob, A., Mata, E., Santana, A., Franc ãs, C., Costa, J., Hirata, C., & Barros, F. Adaptation of interaction with chatterbots for use in the education of children. *IADIS Multi conference on Computer Science and Information Systems* 2012. <https://doi.org/10.4018/978-1-4666-4623-0.ch021>
- [18] Ammar, M. B., Neji, M., & Alimi, A. M. Emotional multiagents system for peer to peer e-learning (EMASPEL). In *Proceedings of the 5th WSEAS International Conference on Distance Learning and Web Engineering table of contents*, 164-170, 2005. https://www.researchgate.net/publication/242080562_Emoional_Multi-Agents_System_for_Peer_to_peer_E-Learning_EMASPEL.
- [19] Lee, T., Chang, C., & Chen, G. Building an Interactive Caring Agent for Students in Computer-based Learning Environments. *Seventh IEEE International Conference on Advanced Learning Technologies (ICALT 2007)*, 300-304, 2007. <https://doi.org/10.1109/ICALT.2007.87>.
- [20] Schouten, D.G., Venneker, F., Bosse, T., Neerincx, M.A., & Cremers, A.H. A Digital Coach That Provides Affective and Social Learning Support to Low-Literate Learners. *IEEE Transactions on Learning Technologies*, 11, 67-80, 2018. <https://doi.org/10.1109/TLT.2017.2698471>
- [21] Samyn, K. 3D Chatbot in higher education, helping students with procrastination and study planning problems. *EDULEARN19 Proceedings*, 2019. <https://doi.org/10.21125/edulearn.2019.2336>.
- [22] Feidakis, M., Kasnesis, P., Giatraki, E., Giannousis, C., Patrikakis, C.Z., & Monachelis, P. Building Pedagogical Conversational Agents, Affectively Correct. *International Conference on Computer Supported Education*, 100 -107, 2019. <https://doi.org/10.5220/0007771001000107>
- [23] Cobos-Guzman, S., Nuere, S., Miguel, L.D., & König, C. Design of a Virtual Assistant to Improve Interaction Between the Audience and the Presenter. *Int. J. Interact. Multim. Artif. Intell.*, 7, 232. 2021. <https://doi.org/10.9781/ijimai.2021.08.017>
- [24] Tsukanov, A. V., & Rusina, N. A. Application of Chatbot Technology in the study of the discipline quality Management. *CEUR Workshop Proceedings*, 2914, 273–283, 2021. <https://ceur-ws.org/Vol-2914/paper23.pdf>.
- [25] Fraoua, K., Leblanc, J. M., & David, A. Use of an emotional chatbot for the analysis of a discussion forum for the improvement of an e-learning platform. *Interacci ón*, 25–35, 2020. https://doi.org/10.1007/978-3-030-50506-6_3
- [26] Suddul, G., Lillmond, C., & Armoogum, S. A Smart Virtual Tutor with Facial Emotion Recognition for Online Learning. 2022 IEEE Zooming Innovation in Consumer Technologies Conference (ZINC), 67-72, 2022. <https://doi.org/10.1109/ZINC55034.2022.9840683>.
- [27] Friesen, E., & Ekman, P. Facial action coding system: a technique for the measurement of facial movement. *Palo Alto*, 3(2), 5, 1978. <https://www.scirp.org/reference/referencespapers?referenceid=3274392>.
- [28] Rokaya, A., Md Touhidul Islam, S., Zhang, H., Sun, L., Zhu, M., & Zhao, L. Acceptance of Chatbot based on Emotional Intelligence through Machine Learning Algorithm. 2022 *2nd International Conference on Frontiers of Electronics, Information and Computation Technologies (ICFEICT)*, 610-616, 2022. <https://doi.org/10.1109/ICFEICT57213.2022.00111>.
- [29] Bopage, H., & Weerakoon, C. Mind Relaxation Chatbot for University Students by Using Dense Neural Network. 2021 5th SLAAI International Conference on Artificial Intelligence (SLAAI-ICAI), 1-6, 2021. <https://doi.org/10.1109/SLAAI-ICAI54477.2021.9664678>.

- [30] Moussa, M.B., Kasap, Z., Magnenat-Thalmann, N., Chandramouli, K., Mirza, S.N., Zhang, Q., Izquierdo, E., Biperis, I., & Daras, P. Towards an expressive virtual tutor: an implementation of a virtual tutor based on an empirical study of non-verbal behaviour. SMVC '10, 39-44, 2010. <https://doi.org/10.1145/1878083.1878096>.
- [31] Soysal, Ö.M., Shirzad, S., & Sekeroglu, K. Facial Action Unit Recognition Using Data Mining Integrated Deep Learning. 2017 International Conference on Computational Science and Computational Intelligence (CSCI), 437-443, 2017. <https://doi.org/10.1109/CSCI.2017.74>

Authors' Profiles



Claudia De Armas earned her B.S. degree in Industrial Engineering, specializing in statistical applications, from the University of Technologies José Antonio Echevarría in Cuba, in 2015. She obtained her M.S. degree in Computer Engineering from University of São Paulo (USP), Brazil, in 2020, specializing in the application of virtual reality and augmented reality in education and training. She is currently a Ph.D. researcher focusing on the use of chatbots in Education 4.0 at USP.



Anderson A. A. Silva holds a PhD in Computer Engineering (USP 2016) and has completed 2 post-docs in engineering. He worked for 23 years in the private sector and is currently a researcher and university professor. His line of research is information security.



Fabio S. Oliveira has achieved his MSc. in Applied Computer Science (IPT-2023). He has a MBA in Strategic Technology Management (IPT-2007) and holds a Bachelor's Degree in Computer Science (Universidade Paulista - 2005). He currently works as an Information Security Officer in the private sector. His main line of research includes information security, industrial networks, Cloud computing and IT governance.



Sergio T. Kofuji holds a PhD in Electrical Engineering from Poli-USP. He is currently Professor at the Polytechnic School of the University of São Paulo.



Romero Tori is an engineer, PhD and associate professor at the University of São Paulo (USP). He is currently Associate Professor Level 3 at the Polytechnic School of the University of São Paulo.



Adilson E. Guelfi holds a PhD in Electrical Engineering from Poli-USP. Adilson served as technical manager of LSI-TECH and currently holds the position of Dean of Research and Graduate Studies at UNOESTE (SP).

How to cite this paper: Claudia De Armas, Anderson A. A. Silva, Fabio S. Oliveira, Sergio T. Kofuji, Romero Tori, Adilson E. Guelfi, "Empathetic Intelligent Virtual Assistant to Support Distance Learning: Systematic Review", International Journal of Education and Management Engineering (IJEME), Vol.15, No.2, pp. 29-43, 2025. DOI:10.5815/ijeme.2025.02.04