Journal of the Brazilian Computer Society, 2026, 32:1, doi: 10.5753/jbcs.2026.7924
This work is licensed under a Creative Commons Attribution 4.0 International License.

AMIVE: A Computational Infrastructure to Identify and Chat to College Students with Possible Depressive Profiles

Paula Maia de Souza [**Federal University of Lavras** | ***paulasouza@ufla.br***]
Franco Eusébio Garcia [**Federal University of São Carlos** | ***francogarcia@protonmail.com***]
Conrado Saud [**Federal University of São Carlos** | ***conradosaud@estudante.ufscar.br***]
Amanda Basso de Oliveira [**Federal University of São Carlos** | ***amanda.oliveira@estudante.ufscar.br***]
Evandro Yudi Alves Ribeiro [**Federal University of São Carlos** | ***evandro@estudante.ufscar.br***]
Isabella da Costa Pires [**Federal University of São Carlos** | ***isabellacosta@estudante.ufscar.br***]
Vanessa de Cássia Alves [**Federal University of São Carlos** | ***vanessa.alves92@estudante.ufscar.br***]
Grasiane Cristina da Silva [**Federal University of São Carlos** | ***grasiane.silva@estudante.ufscar.br***]
Taís Bleicher [**Federal University of São Carlos** | ***tbleicher@ufscar.br***]
Hélio Crestana Guardia [**Federal University of São Carlos** | ***helio.guardia@ufscar.br***]
Vânia Paula de Almeida Neris [**Federal University of São Carlos** | ***vania.neris@ufscar.br***]

Institute of Science, Technology and Innovation, Federal University of Lavras (ICTIN/UFLA), Rua Antônio Carlos Pinheiro Alcântara, nº 855 - Jardim Mediterranée, São Sebastião do Paraíso, MG, 37950-000, Brazil.

Received: 25 March 2026 • **Accepted:** 27 October 2026 • **Published:** 25 November 2026

Abstract. The university context has demands that can increase the risk of depression and other mental disorders in students. Conversational User Interfaces (CUI) have been explored to support the prevention, diagnosis and therapy of people with depression. However, the design and construction of these CUI, focused on college students, has not been deeply investigated. Based on this context, the aim of this paper is to detail the design process for building a conversational solution to support college students with depression. The research method used was Design Science Research (DSR). The results suggest that the chatbot and the AMIVE infrastructure led to enjoyable CUI in this domain, with promising expectations for the mental health of users. The lessons learned point to discussions to be considered by other researchers in the same domain. The main contributions of this paper are: i) the AMIVE infrastructure, in which chatbot was incorporated; ii) the results of the evaluation with the target audience in a real-world context; and iii) a discussion of the lessons learned throughout the research development.

Keywords: Chatbot design, CUI infrastructure, Design Science Research, Student well-being, Mental health.

1 Introduction

The university context involves a series of changes that have the potential to influence students' mental health. It is estimated that between 15% and 25% of university students will experience some mental disorder during their academic period [Arino and Bardagi, 2018]. A study by the World Health Organization (WHO) with university students from 19 colleges spread across 8 countries (Australia, Belgium, Germany, Mexico, Northern Ireland, South Africa, Spain, and the United States) found that Major Depressive Disorder (MDD) was the most common mental health disorder among university students, with 18.5% of students in the first 12 months of college being affected by MDD [Auerbach *et al*., 2018].

Estimates worsened during the pandemic caused by COVID-19. Post-COVID-19 research indicates a general prevalence of mental disorders of 72.9% among undergraduate students [Silva, 2021]. Among the causes of damage to the mental health of university students, adaptations related to the transition from high school to higher education, moving to another city, and distancing from family and friends stand out, as well as pressure regarding academic obligations [Cardoso, 2022]. Another important related cause is the fact that during university life, around 80% of students experience daily stress. High levels of stress are associated with the development of depression and other mental disorders [Fragelli and Fragelli, 2021].

Depression is considered one of the main causes of disability in the world. Depression is characterized by depressed mood (prostration, irritability, feelings of emptiness) and/or loss of energy and pleasure or interest in day-to-day activities. These symptoms should predominate on most days for at least two weeks. Several related symptoms may be present, such as poor concentration, feelings of low self-esteem, guilt, hopeless thoughts about the future, ideas of death, and disturbances in the biological rhythms of sleep, appetite and libido. It is quite understandable that these symptoms imply difficulties in interpersonal relationships and socializing, as well as in work relationships [American Psychiatric Association, 2022].

In addition to conventional treatments, other approaches can also be used to prevent or reduce the damage caused by depression [Whitaker, 2017]. Computational solutions have been developed to support the prevention [Li and Luo, 2021], diagnosis [Gupta and Sharma, 2021] and therapy [Serrano-Ripoll *et al*., 2022] of people with depression. Conversational User Interfaces (CUI) (e.g. chatbots) are among computational solutions developed [Souza *et al*., 2022]. The term

"chatbot" stems from the words "chat", which is related to conversation, and "bot", which is short for robot. Chatbots are automated programs that execute instructions based on specific rules to simulate conversations with humans, in general, using natural language structures [Galvão *et al*., 2019; Kumar, 2021].

Although there are studies in the literature involving chatbots to support people with MDD (see Section 2), in general, the studies do not detail how the design decisions were made. The design of CUI aimed at college students with MDD has also not been deeply investigated.

In this context, we formalize a codesign approach for CUI to support college students with depression. The term codesign has been used as a way of designing a solution for a community with that community being engaged in the design process, emphasizing the aspects of Human-Computer Interaction (HCI) [Baranauskas *et al*., 2013]. We built the proposed approach, instantiated it in a real context and evaluated it with college students with depression [Souza *et al*., 2025]. This paper, in turn, aims to present an extended and revised version of the paper from Souza *et al*. [2025].

Both papers are part of the AMIVE [1] project. The AMIVE project is developed by a multidisciplinary team comprising researchers from multiple institutions, fostering diverse expertise and collaborative innovation. The central objective of the AMIVE project is to design and implement a computational infrastructure capable of identifying users with a Potential Depressive Profile (PDP), as well as supporting appropriate intervention strategies, all autonomously and in real time.

The research method used to develop the solution was Design Science Research (DSR) [Dresch *et al*., 2015] (Section 4). This extended paper includes new contributions such as: i) more details on related work (Section 2), also considering issues regarding the implementation of infrastructure for CUI; ii) a background on the AMIVE project (Section 3), summarizing the main research findings related to the construction of the project's infrastructure; iii) a table summarizing the themes and descriptions of 10 conversations created by mental health specialists (Section 5); iv) details on the implementation of the AMIVE infrastructure (Section 6); v) results unpublished from the evaluation with the target audience (Section 7), including the application of the Technology Acceptance Model Questionnaire (TAM); vi) new lessons learned (Section 8), also considering the infrastructure development process; and vii) in-depth conclusion (Section 9), considering the infrastructure implementation process of the AMIVE project. The results suggest that AMIVE has the potential to contribute to the target audience by delivering a specialized intervention within the domain.

2 Related works

The authors are conducting a Systematic Literature Review (SLR) on chatbot personalization for mental health, currently in the writing and submission phase. To synthesize the works related to this paper, a filter was applied to the studies analyzed in the SLR. The filter considered studies that targeted university students or young adults, and a total of 10 studies were returned. An ad-hoc research on chatbot design approaches for mental health was conducted, and 7 more studies within the domain were included. Thus, a total of 17 papers were analyzed to identify related works.

After reading and analyzing the 17 papers, 8 were excluded due to the following criteria: i) the artifact addressed was not a conversational interface; ii) the work was not related to the target audience; iii) the work was not related to mental disorders; and iv) it was not possible to access the work to read the full text. Therefore, this section presents a summary of the 9 selected studies, ordered by year of publication and author's name.

Fitzpatrick *et al*. [2017] (S1) determined the feasibility, acceptability, and preliminary efficacy of a fully automated conversational agent (Woebot) to deliver a self-help program to college students who identify as having symptoms of anxiety and depression. The authors did not report how the chatbot was developed, whether they used any existing design approaches, and they also did not report the profile of the project team. However, they describe that the chatbot utilized a hybrid computational approach, primarily structured as a decision tree with predefined response options while also supporting natural language input. Targeted natural language processing techniques were embedded at key points to guide user input toward appropriate conversational paths. Its interaction style was designed to reflect human clinical reasoning and social communication patterns. Additionally, the psychoeducational components were based on self-help principles derived from Cognitive Behavioral Therapy (CBT). The authors conducted a 2-group randomized clinical trial with 70 college students. The intervention group interacted with the chatbot and the control group received an e-book on depression. According to the authors, the results suggest that the group that used Woebot showed a significant reduction in symptoms of depression and anxiety, as measured by the Patient Health Questionnaire (PHQ-9) ($F=6.47$; $P=.01$) and General Anxiety Disorder-7 (GAD-7) ($F_{1,54}= 9.24$; $P=.004$), compared to the control group. Although both groups showed a decrease in anxiety symptoms, there was no statistically significant difference between them. The group that interacted with the chatbot also showed a high level of engagement with the tool.

Gabrielli *et al*. [2021] (S2) conducted a proof-of-concept evaluation to measure the engagement and effectiveness of the psychoeducational chatbot Atena, which supports healthy coping with stress and anxiety. The target audience consisted of university students. The authors did not report how the chatbot was developed. The authors indicated that psychologists and User Experience (UX) specialists were part of the project team. A proof-of-concept evaluation was conducted with a sample of 71 volunteer university students recruited by convenience. The evaluation was performed using the 10-item Perceived Stress Scale (PSS-10), GAD-7, and Five-Facet Mindfulness Questionnaire (FFMQ) questionnaires and free-text responses about the user experience with the chatbot. The volunteers were also asked to complete a short online survey to report what they liked and disliked most about their experience with the chatbot. The authors used statistical analyses to evaluate the effectiveness of the chatbot intervention.

[1] https://www.amive.ufscar.br/

Participants' responses to the open-ended questions in the online final survey were analyzed using thematic analysis. The main results suggest that chatbot use is associated with a significant reduction in anxiety and stress symptoms among participants.

He *et al*. [2022] (S3) tested the clinical efficacy and nonclinical performance of a mental health chatbot (XiaoE) based on CBT for young adults with depressive symptoms. The authors did not report how the chatbot was developed, whether they used any existing design approaches, and they also did not report the profile of the project team. The authors conducted a randomized clinical trial with university students to evaluate the effectiveness of the mental health chatbot compared to an e-book intervention and a general chatbot. They used the PHQ-9, Working Alliance Questionnaire (WAQ), Usability Metric for User Experience-LITE (UMUX-LITE) and the Acceptability Scale (AS) as evaluation metrics. The results indicate that participants who used the mental health chatbot showed a significant reduction in depressive symptoms compared to the control groups. Adherence to the intervention was higher in the chatbot group compared to the control groups. They also reported high satisfaction with the use of the mental health chatbot, highlighting the ease of use and the usefulness of the interactions. The majority of users indicated that they would recommend the chatbot to other young people experiencing similar symptoms. The chatbot intervention group showed high engagement rates, with many participants using the tool regularly throughout the study period.

Koulouri *et al*. [2022] (S4) explored whether chatbots can be an acceptable solution in mental health for young adults. The paper presents a high-fidelity prototype of a conversational agent for screening and intervening in mental health of young adults in a university context. The authors point out that they used a user-centered design approach. The authors also report that they involved the target audience only in the data collection phase. An evaluation with the target audience of university students was not carried out. The evaluation described in the paper was carried out with 3 mental health specialists. Semi-structured interviews and a demonstration of the functional prototype were carried out for the evaluation. The results indicate that the specialists considered the chatbot a tool that can help students identify mental health problems and seek help. The authors also point out as a benefit the possibility of 24/7 access, unlike traditional care that has restricted hours. As limitations, the authors highlighted concerns about confidentiality and security; limitations in expressing empathy, especially in serious cases. The chatbot was highlighted as a complement and not a substitute for human professionals.

Kuhlmeier *et al*. [2022] (S5) proposed a theory-based design for personalized conversational agents to treat depression in youth and young adults. They proposed two principles to guide the design. The authors report the participation of young people with depression in the problem identification phase and a multidisciplinary team throughout the project. They built prototypes and for evaluation, interviews were conducted with five healthcare specialists and potential non-diagnosed users.

Souza *et al*. [2022] (S6) formalized design recommendations for chatbots to support students with depression. They proposed a set of 24 design recommendations. The authors reported that mental health professionals and HCI specialists worked on this proposal. The paper has much to contribute to the design process in this context, but it does not present an approach with steps and activities to be carried out. The proposed recommendations were also not used to create a chatbot, and an evaluation with the target audience was not carried out.

Tagarot [2022] (S7) presents a chatbot designed to help college students facing depression. The author does not describe in the paper which approach was used to design the chatbot or the profile of the team that participated in the project. It used the Rasa Open Source platform (Rasa). The author reported that having a basic structure in place that handled the Natural Language Processing (NLP) and dialogue management aspects freed up extra time to work on the training data. The author reported an evaluation with the target audience, in which five volunteers interacted with the chatbot and answered a feedback questionnaire. The results indicate that the volunteers felt better after using the chatbot.

Mustafa *et al*. [2024] (S8) developed a chatbot prototype to detect and treat anxiety in university students. The aim of the paper was to present the evaluation of the level of acceptance of the prototype. The authors do not report in the article which approach was used for the design of the chatbot. The authors also do not report the profile of the team that participated in the project. For the evaluation, semi-structured interviews were conducted with the target audience.

Oghenekaro and Okoro [2024] (S9) implemented an AI-based chatbot to provide personalized and accessible mental health support to college students. The authors point out the use of a user-centered design approach, but do not describe the profile of the team that developed the chatbot. The proposed mental health support system was developed using Visual Studio Code as the primary IDE. The frontend was implemented with ReactJS to provide an interactive user interface, while Vite improved development efficiency and optimized performance. SCSS (Sassy CSS) was used for structured and maintainable styling. Dido supported chatbot training and integration through efficient dataset access. CSS keyframes enabled smooth visual integration of the chatbot into the interface. The authors report that the evaluation of the chatbot was conducted using a mixed-methods approach (engagement metrics, PHQ-9 and GAD-7). Although it is reported in the paper that the evaluation was conducted and that these metrics collectively assessed the effectiveness of the chatbot in promoting mental well-being, the authors do not describe the evaluation process and do not present the results.

Table 1 presents a summary of related work and compares them with the proposal of this paper. The sequence in which the studies are presented in the table follows the same order as that presented in the text. Where N/A appears in the table, it reads "Not applicable".

The works presented indicate that chatbots tend to be a good tool to provide mental health interventions. Chatbots have the potential to be accessed more actively and frequently, unlike traditional care. Chatbots also enhance the engagement of university students. However, little has been reported on design approaches for CUI in this context. The studies also say little about the participation of multidisciplinary teams and

Table 1. Comparison table of related works. Adapted from Souza *et al*. [2025].

ID	Proposed, implemented and/or use an approach to design?	Developed a CUI?	Describe the implementation of the CUI infrastructure?	What mental disorder is covered?	What target audience?	Was there an evaluation with the target audience?
S1	No	Yes	No, it only reports the use of decision trees, natural language and CBT.	Depression and Anxiety	College students	Yes
S2	No	No	N/A	Stress and Anxiety	College students	Yes
S3	No	No	N/A	Depression	College students	Yes
S4	Use the User-Centered Design approach	Yes	No	Mental health	Young adults and counselors in a university context.	No
S5	Proposed and implemented a new approach	Yes	Functional prototype	Depression	Youth and young adults	Yes
S6	Proposed a new approach	No	N/A	Depression	People with depression	No
S7	No	Yes	It used a ready-made infrastructure, the Rasa Open Source platform (Rasa).	Depression	College students	Yes
S8	No	Yes	No	Anxiety	College students	Yes
S9	Use the User-Centered Design approach	Yes	Yes	Mental health	College students	No
This paper	Proposed and implemented a new approach	Yes	Yes	Depression	College students	Yes

end users in the construction of chatbots. Details regarding the implementation of chatbot support infrastructure are also poorly explored in the works found in the literature. Little is also said about how dialogues should be created to ensure the safety of the target audience. Therefore, this paper seeks to minimize this gap in the literature.

3 AMIVE Background

Regarding the development of the AMIVE infrastructure modules to support the identification of PDP, a systematic mapping study of related works was conducted in Saud *et al*. [2024]. The study revealed several relevant points and research gaps that the AMIVE Project could address to provide adequate support for PDP identification among college students. One of the highlighted findings was the prevalence of Machine Learning (ML) and classification algorithms used to analyze mobile sensor data for depression classification. These approaches often attempt to identify behavioral patterns associated with depressive symptoms defined by the American Psychiatric Association [2022]. The mapping study also identified frequently used ML algorithms in related works, as well as limitations and challenges related to the quality of collected data and the recruitment of participants for mental health studies. Furthermore, the mapping study revealed that relatively few works involved multidisciplinary teams composed of computer science researchers and mental health specialists. The absence of this kind of collaboration may contribute to the lack of appropriate protocols to address sensitive mental health issues and emergency cases. Additionally, the study identified a gap in the literature regarding the adoption of Human-in-the-Loop (HITL) approaches in studies related to ML application in autonomous classification of mental health conditions.

In the context of the identification of PDP within the AMIVE solution, one of the identification scopes involved detecting depressive symptoms in participants' social network posts. In this research context, a data collection of Facebook posts was performed using Crowd Tangle to gather public posts related to the university student context. In this dataset, mental health researchers conducted an annotation process aiming to identify correlations between the expressed thoughts in the posts and depression. The study presented in Reche *et al*. [2024] revealed relationships between occupational changes, declining academic performance, and depression. Additionally, the annotated dataset was also used to support the development of a NLP model capable of automatically identifying depressive symptoms in social network posts, as presented in Mendes and Caseli [2024].

To address the lack of studies involving HITL approaches,

the work presented in Alves *et al*. [2023] proposed a college-students-in-the-loop approach to deal with outliers and inaccuracies in mobile sensor data collected through smartphones. In the proposed loop, the end users of the AMIVE application can evaluate and approve or disapprove the collected data before it is sent to the platform. At this stage, participants assess whether the data collected by the sensors is consistent with their daily behavior. In cases where the data does not accurately represent the participant's real-life experience, it is marked as unsuitable for training the identification models. The HITL approach employed in AMIVE integrates participants into the ML loop and also enables the triggering of events related to the chatbot (detailed in Subsection 6.2) within the infrastructure.

Beyond the mobile sensor and social network data collected by the AMIVE application, sociodemographic data also represent a potential source of information for studies related to depressive profiles. In this context, a clustering study conducted in Bayazidi *et al*. [2025] analyzed AMIVE data composed of mobile sensor data, PHQ-9 responses, and sociodemographic informations. The clustering analysis revealed strong associations between depression and behavioral patterns such as physical exercise, as well as factors related to family educational background and the amount of time students dedicate to studying. Despite these findings, the sociodemographic data collected by AMIVE remain relatively unexplored in both the identification and intervention scopes of the project.

4 Methodological path

The research method used in this study was DSR. The DSR grew out of efforts to formalize Design Science in Information Systems research. The method provides researchers with a framework to produce and present high-quality, rigorous, and publishable research in scientific journals [Peffers *et al*., 2007; Dresch *et al*., 2015]. The DSR method was used to propose, instantiate in a real context and evaluate the design generated using the proposed approach by Souza *et al*. [2025]. Grounding our research in DSR provided a validated research framework. We adhered to DSR best practices to ensure that our research was developed systematically and with scientific rigor.

The DSR process [Dresch *et al*., 2015] includes 11 steps (Figure 1), being them: 1) Problem identification; 2) Problem awareness; 3) Artifact identification and problem class configuration; 4) Proposal of artifacts to solve the specific problem; 5) Project of the selected artifact; 6) Artifact development; 7) Artifact evaluation; 8) Explanation of learning; 9) Conclusions; 10) Generalization to a class of problems; and 11) Communication of results.

The introduction (Section 1), the related works (Section 2) and the AMIVE background (Section 3) contains the results to support the identification and awareness of the problem (steps 1 and 2 of the DSR). The phases of codesign (Section 5) and implementation of the AMIVE infrastructure (Section 6) encompass the identification, proposal, project and development of the artifact (steps 3, 4, 5 and 6 of the DSR). The evaluation of the AMIVE solution (Section 7) encompass the artifact evaluation phase (step 7 of the DSR). The lessons learned (Section 8) and conclusions (Section 9) encompass the results of the phases of explanation of learning, conclusions, generalization, and communication of results (steps 8, 9, 10, and 11 of the DSR).

5 Codesign of the AMIVE chatbot

As reported in the introduction, this paper is inserted in the context of the AMIVE project, that intends to build a computational infrastructure (see Section 6) for identification and intervention autonomous of real-time of users with a MDD. Amive's action aims to identify a PDP and, subsequently, offer care strategies for psychological distress suffering and mental illness by delivering autonomous, specialized and personalized interventions. In this context, we sought to insert a CUI into this infrastructure that would deliver the intervention through specialized dialogue with users.

To support this project, an initial research [Souza *et al*., 2022] was conducted to substantiate the problem and the motivation. As reported in Sections 1 and 2, there is a lack of approaches in the literature to support the codesign of CUI in the context of college students with depression.

After identifying the problem and motivation, we defined the objectives and artifacts to solve the specific problem. A multidisciplinary work team has been formed. This work team included computer researchers, designers, psychologists, professionals who work with the target audience and representatives from the target audience (graduate students in computing and health). We drew from our previous experience with therapeutic solutions and incorporated relevant literature on computing and health to support our process. The specialists discussed the initial steps, necessary artifacts, and information required to develop the CUI, taking into account the specific demands of the project.

This section presents the instantiation of the approach proposed in Souza *et al*. [2025] for the design of the AMIVE chatbot to support college students with depression. Figure 2 illustrates the workflow of the approach. The results of the clarification and design stages are described in the Subsections 5.1 and 5.2. The results of the evaluation stage are described in the Section 7.

5.1 Clarification

Firstly, to initiate the implementation of the approach, the work team was defined. The work team was composed of 3 Computer specialists, 2 Healthcare specialists and 4 Representatives of the target audience (university students). Starting the Clarification stage, the literature study activity was carried out. The work team applied the benchmark method to evaluate the scenario of offering chatbots to people with depression. Seven chatbots were evaluated, and none of the chatbots found were focused on college students. Among some of the benchmark findings, we highlight the need to be careful with the responses that the chatbot delivers to users, since, considering the sensitive context, the chatbot cannot generate a response that stimulates risky thoughts [Souza *et al*., 2022].

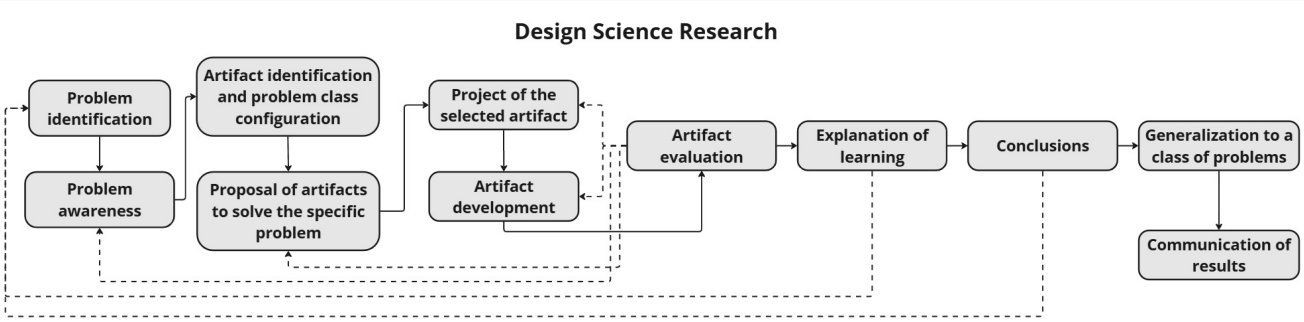

Figure 1. Design Science Research. Adapted from Dresch *et al*. [2015].

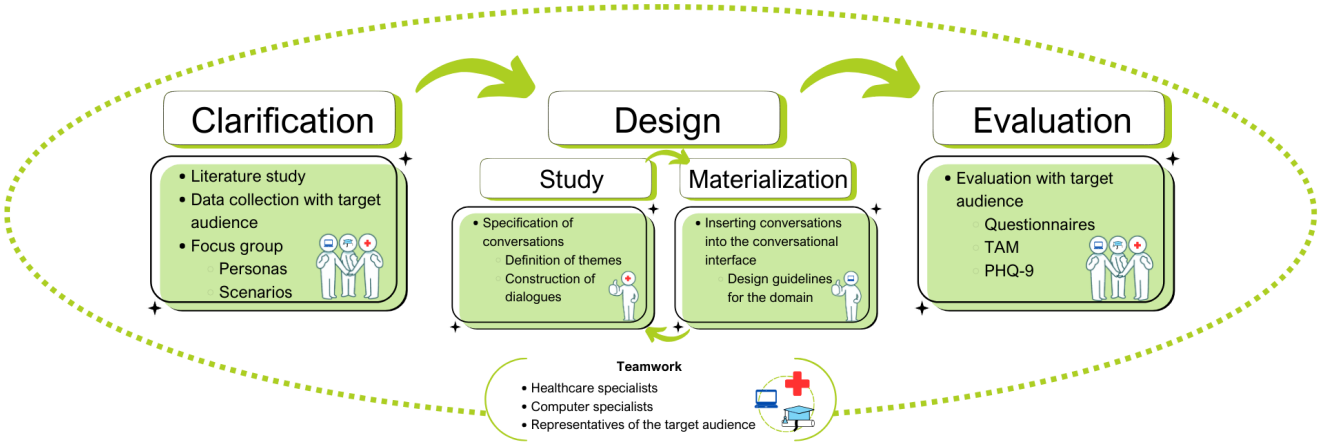

Figure 2. Codesign Approach for CUIs to Support College Students with Depression [Souza *et al*., 2025].

To collect data about the target audience profile, the work team consulted a survey about the socioeconomic and cultural profile of undergraduates at federal universities in the country [Andifes, 2019]. The work team also consulted a survey about the perceptions of college students about ethics in computer-based data collection and interventions for mental health [Menegasso, 2021].

With the results of the literature study and the collection of target audience profiles, the work team carried out a focus group. The objective of the focus group was to use qualitative data to discuss the presentation of the solution to the end user, content parameters and frequency of interventions, and chatbot behavior strategies. All 9 work team members of this project participated in the focus group. The focus group conducted a total of eight meetings to define the chatbot's scope, requirements, personas, and scenarios. These meetings were recorded, and various artifacts were produced by participants as a result. During the meetings, a moderator facilitated the discussions, welcoming diverse perspectives, clarifying terms and concepts as needed, and fostering a shared understanding among participants. No significant conflicts arose among the participants; any differences of opinion stemmed primarily from the team members' varied backgrounds and experiences. The structured use of Persona [Cooper, 2004; Pruitt and Adlin, 2006] and Scenario [Carroll, 1997; Carrol, 1999; Carroll, 2000] techniques helped organize the discussions and align divergent ideas, as decisions were reached collectively based on the perspectives presented and the steps outlined in the adopted techniques. Among the results obtained, the definition of the stages of interaction between the chatbot and the user, the proposition of the chatbot's communication flow with the user (see more in Subsection 6.2), and a set of 6 personas was defined. Figure 3 shows one of the personas developed.

5.2 Design

After completing the clarification stage proposed by the approach, the working group began the design stage. This design stage is divided into two sub-stages, namely: Study and Materialization.

Study. A design decision made by the work team was that the chatbot would not create dialogues autonomously. This decision was made because of the sensitive context in which the chatbot is inserted. Therefore, the team decided that the chatbot would have personalized content developed by the team's healthcare specialists.

The text of the dialogues could be written in non-technical language, facilitating collaboration with healthcare specialists. Later, in the materialization stage, the computer specialists were tasked with transcribing the dialogues into JSON files to be interpreted by the chatbot.

Given the sensitive nature of mental health, all decisions regarding themes, dialogues, presentation methods, and response options were made by psychologists and psychology professionals who work with the target audience. They defined the themes and constructed the conversations considering the defined communication flow and the profiles of

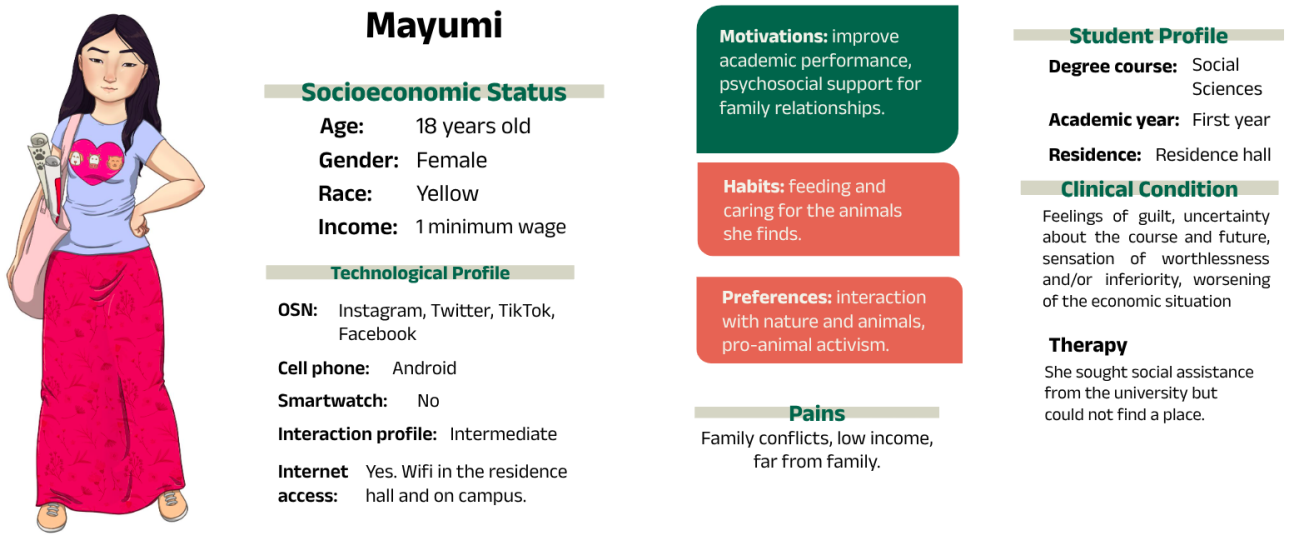

Figure 3. Example of one of the personas created [Souza *et al*., 2025].

the personas created. When writing the content of the dialogues, the specialists based themselves on their experiences and scientific literature on depression and mental health.

In total, 13 themes were defined, namely: Autonomy, Depression in university students, Distance from origin city, First use, First year at University, Gender, Income, Initial dialogue, Insomnia, Poor academic performance, Risk of death, Sleeping deprivation and Support network. Subsequently, the themes "Insomnia" and "Sleep deprivation" were combined into a single dialogue ("Sleep deprivation"). The themes "First use" and "Initial dialogue" were used solely to initiate the interaction between the bot and the user; since no therapeutic intervention took place, no specialized dialogues were constructed for these topics. Based on the other themes, 10 specialized conversations were written. Table 2 provides a brief description of each of the conversations.

Figure 4 shows an excerpt from a dialogue on the topic of "support network". In the dialogue, the health expert indicated a symptom to which the question is linked (e.g. does not find a support network, impaired spiritual support) and the classification of the answers according to severity (light, medium or severe). The user will only see the questions and answer options; the data on symptoms and severity are for internal use by the team to suggest intervention.

Materialization. The dialogues created in the Design/Study stage were inserted into the AMIVE infrastructure through the chatbot dialogue editing functionality. Figure 5a shows a screenshot of the application's home screen, where the development team (system administrator) can enter the dialog editor. When entering the dialog editor, the system administrator is redirected to the second screen (Figure 5b), where he can import a dialog already created in a file in JSON format, or create a new dialog directly in the application. When clicking on "create new" the system administrator is redirected to the screen shown in Figure 5c. The system administrator can then enter data about a specific dialog and link that dialog with symptoms identified in the user's interactions with the application. This linking of the dialog to the symptoms (Figure 5d) is used to suggest personalized dialog themes to users based on their interactions with the application. More details about the AMIVE infrastructure will be presented in the Section 6.

Considering the approach, before inserting the dialogues into the application's chatbot, one of the computer specialists used design guidelines for Chatbots to Support People with Depression [Souza *et al*., 2022] and a manual on writing, behavior and communication journeys for chatbots [Souza *et al*., 2026]. The manual was developed within the context of the AMIVE project, with the purpose of recording the modeling process of the project's conversational agent, validating and recording the design choices made, and guiding further research on iterations of chatbot development projects.

The manual developed for the project is composed of three parts that help in creating the chatbot: (1) writing manual, section 1 of the manual; (2) behavior manual, section 2 of the manual; and (3) communication journey manual, section 3 of the manual. The writing manual aims to be a guide for creating and formatting content for the conversational agent. It deals especially with the language used to create the bot's content, proposes a definition of symbology, including graphic elements that the bot can use during dialogues, and proposes the construction of a glossary with flag words and other important terms that can trigger the bot. The behavior manual aims to describe guidelines and recommendations on how the created content can meet the user in a way that is appropriate to the project's values. Among these behaviors, we can mention the bot's encouragement of the user for self-care and the individual care and attention that the bot should have with each user. This section also includes a catalog of interventions from the bot to the user. The purpose of the communication journey manual is to describe the path that the chatbot should take during the bot-user and user-bot

Table 2. Themes and brief descriptions of each of the written conversations.

Theme	Description
Autonomy	Addresses the development of autonomy among university students, providing support and encouragement for adapting to new responsibilities through personal strategies, as well as social and institutional support.
Depression in university	Seeks to support university students in identifying feelings and difficulties (such as loneliness, sadness, and academic or family-related problems), while also providing emotional support, guidance, and referral information for psychological assistance.
Distance from origin city	Addresses the emotional and practical challenges faced by students living away from their city of origin, particularly those related to adaptation, autonomy, and feelings of loneliness. It also provides support, guidance, and information on social support networks and health services, aiming to facilitate adaptation to university life.
First year at university	Addresses the challenges of the first year at university, including academic adaptation, study difficulties, social integration, and uncertainties about the chosen course. It also provides guidance and information on available institutional, social, and psychological support to facilitate student adjustment.
Gender	Addresses gender-related issues in the university context, exploring whether the user experiences impacts or violence related to this topic, while offering emotional support, guidance, and pathways for assistance.
Income	Addresses the financial difficulties faced by university students and their impact on retention and well-being. It also presents guidance on financial aid, scholarships, and income-generating strategies aimed at supporting student persistence, in addition to information on institutional support services.
Poor academic performance	Addresses academic performance difficulties by exploring their potential causes, such as emotional factors, organization, interpersonal relationships, and the broader university context. It provides support, guidance, and information on institutional, social, pedagogical, and psychological support services.
Risk of death	Seeks to identify and support students experiencing feelings of hopelessness or risk of death by exploring their emotions, the intensity of such thoughts, and their level of safety. It encourages users to seek immediate and specialized help, as well as support from close contacts and emotional support services.
Sleeping deprivation	Addresses factors affecting sleep among university students and suggests care strategies, including behavioral changes, institutional support, and, when necessary, professional follow-up.
Support network	Addresses the importance of support networks in university life, exploring social relationships, feelings of loneliness, and strategies for seeking emotional and institutional support. It also guides students in identifying, strengthening, and expanding their social connections, as well as accessing available support services.

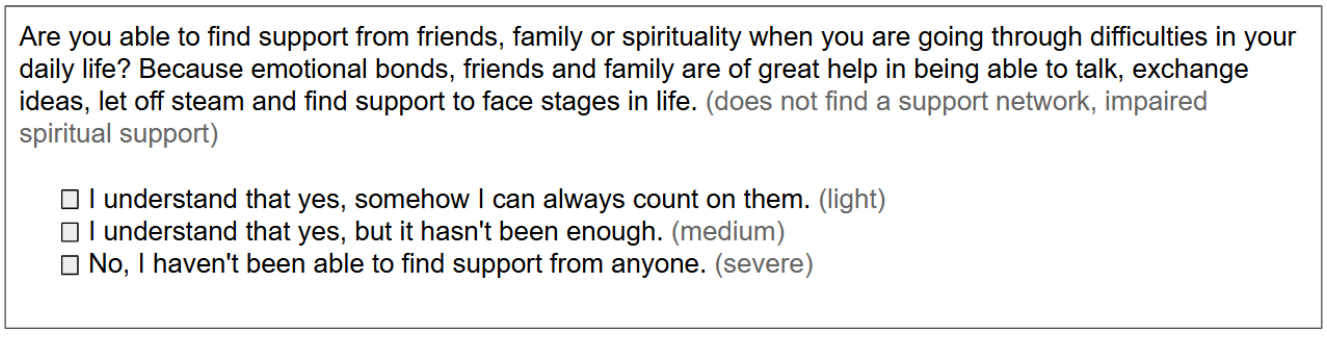

Figure 4. Example of an excerpt from a dialogue [Souza *et al*., 2025].

interaction, seeking a comfortable dialogue flow. The journey proposed by the project is composed of four phases, from the initial greeting to the end of the dialogue [Souza *et al*., 2026].

The guidelines, guidance, and suggestions available in the manual were used to create the content available in the project's chatbot. Figure 6 shows an example of how the manual was applied in a dialogue inserted in the chatbot. The bot starts the dialogue session in phase 01 of the communication journey described in section 3 of the manual, subsection 3.1 (first and second lines of the table in Figure 6). This phase is characterized by having little friction and a definition of the user's current situation. In the fifth and sixth lines of the table, the bot moves on to phase 02 of subsection 3.1 of the manual, this phase is characterized by there being a more in-depth

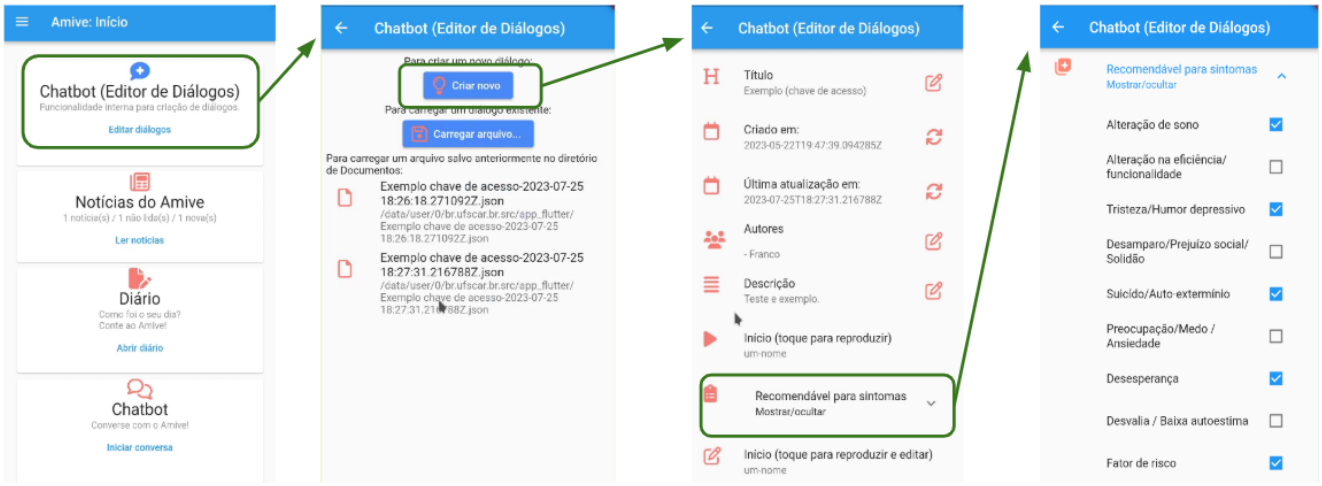

Figure 5. Screenshots of the flow for editing chatbot dialogues in the AMIVE application. Adapted from Souza *et al*. [2025].

analysis of the user's current situation by the bot. The next lines of the table show the transition from phase 02 to phase 03 of subsection 3.1 of the manual. Phase 03 describes the need for the bot to welcome the user, along with the provision of cautionary signs.

Throughout the conversation presented in Figure 6, it is possible to note the presence of personalization, non-textual elements, and empathy of the bot with the end user, following the guidelines present in section 1 of the manual. The personalization subsection of the manual recommends using the user's name or nickname and writing the dialogue directed at the user, for example, in the sentence "Good Morning, User!", in this case, "User" is a fictitious name of a user, indicating the personalization of the content. Subsection 1.3 of the manual (reducing context ambiguity) describes how to write the content to reduce the context ambiguity of the dialogue, creating more specific content; In the image excerpt, the text content converges on problems that may be causing insomnia and what actions can help in the treatment. Regarding non-textual elements, subsection 1.4 of the manual (non-textual elements) describes the importance of inserting emojis and emoticons to bring the bot closer to the user; emojis are present throughout the excerpt, such as the sleeping emoji in the bot's first message. Also in section 1 of the manual, guidelines are given for the content to be written to demonstrate the bot's empathy with the user (subsection 1.6 of the manual - empathy). This characteristic is also present in the bot's first message, specifically in the excerpt "College life can be quite challenging at times, right? There are so many things to do, and so many demands to meet.".

Also present in the excerpt of the dialogue presented is the application of the guidelines on the bot's behavior, present in section 2 of the manual. The guidelines applied in this section are individual care and attention (subsection 2.3 of the manual) and mapping of determinants of suffering (subsection 2.5 of the manual). Subsection 2.3 of the manual describes the care that the bot must have with the user, trying to clearly understand the situation of the individual and the environment in which he or she is inserted. Subsection 2.5 of the manual guides the survey of determinants of suffering so that they can be used in the provision of care actions that the bot offers to the user. In Figure 6, the white lines refer to the bot's messages, while the gray lines refer to the options available to the user. The choice made by the user is marked in bold and underlined.

After applying the manual, directives, and guidelines, the computer specialists transformed each of the dialogues written in non-technical language by the project's researchers into the JSON file format and they added the dialogues to the application's chatbot. To make the chatbot available for the evaluation described in Section 7, a total of 10 dialogues were created, each covering a different topic about the mental health of university students, including dialogues about insomnia, income, and poor academic performance.

6 AMIVE Infrastructure

In order to offer identification and support for college students with PDP, the AMIVE project developed a mobile application where the students could submit their data to track depression and interact with a specialized chatbot to receive an initial mental health care.

The computational infrastructure of the AMIVE project consists of the mobile application for end-users and alongside with systems running on a dedicated server to collect and process the student's data. The application contains the follow functionalities: (1) collect student's digital phenotyping data (mobile sensor and textual data); (2) a journaling feature that allow user's to write about their feelings; (3) application of periodic questionnaires about subjective feeling and mental health assessments; and (4) a chatbot interaction.

The AMIVE platform infrastructure comprises front-end and back-end computational systems. The front-end includes the AMIVE mobile application and web pages, with controlled access, available to project and mental health researches. These web interfaces enable the creation and editing of the chatbot dialogues, as well as the visualization of student's data. As illustrated in Figure 7, the platform integrates services and systems to support the overall solution. All the communication between front-end applications and back-end

	Dialogue Session	Manual Application
BOT:	Good Morning, User! I'm here to help you! 🤗 College life can be quite challenging at times, right? There are so many things to do, and so many demands to meet. 😔 At times we may feel distressed or lost with so many responsibilities, after all, balancing several activities can be tiring and stressful. When we enter college, we may end up facing some concerns, including financial, permanence, and survival. 😔 In this back and forth of college life we may also eventually start to have problems with sleep. 😴 Did you know that you may be unwell and therefore have insomnia or have insomnia and therefore be unwell? Both possibilities are true. 🤔 I'm here to talk whenever you want. If at any time you no longer want to talk, just close the conversation. 🤗 Knowing this, what would you like to talk about insomnia? 🧐	The bot starts the session with low friction, as indicated in phase 01 of 3.1, in addition to applying guidelines 1.1 (personalization) and 1.4 (non-textual elements). In this section, the bot also applies guidelines 1.6 (empathy), and 2.3 (care and individualized attention).
USER:	• I'm feeling unwell, and that's why I have insomnia. I want to know what it can cause me. • I have insomnia, and that's why I'm feeling unwell. I want to know what it can cause me. • **I have insomnia and I want to know what it can cause me.**	
BOT:	Do you want to know what sleep deprivation can cause in our lives? 🧐	In this section, the bot prepares to transact between phase 01 and phase 02 of 3.1. It also applies guidelines 1.3 (reducing context ambiguity), and 1.4.
USER:	• Yes, I want to understand a little more. • **I don't want to understand this right now.**	
BOT:	To properly address sleep problems, let's first try to understand why you're suffering from insomnia. Could it be? 😄	The bot continues the session by understanding the user's context, as indicated in phase 02 of 3.1, in addition to applying guidelines 1.3, 1.4, and 2.5 (mapping of determinants of suffering).
USER:	• No, I know why I have insomnia. • I don't want to understand that now. • **Yeah, let's go.**	
BOT:	Do you know what is depriving you of sleep? 🤔	The bot prepares to transact between phases 02 and 03 of 3.1. Guidelines 1.3, 1.4, 2.3, and 2.5 still apply here.
USER:	• I am overwhelmed with work. • **I need to meet deadlines.** • I need to study for tests. • None of these.	

Figure 6. Example of dialogue session and application of manual items [Souza *et al*., 2025].

services occur through a RESTful API.

In relation to the AMIVE front-end, the computational application was developed using the Dart programming language with the Flutter framework, enabling the development of the application for both Android and iOS mobile operating systems. In the context of the project, the front-end system represents the main interaction interface with students, providing access to the identification and intervention solutions for PDP. The application has access to two different databases: a local database implemented in SQLite to store the student's personal journal and configuration settings related to the specific instantiation of the application on the student's smartphone. Furthermore, the front-end is also responsible for enabling the data collection flow of the infrastructure and for supporting communication with the chatbot. Additionally, the dialogue editor for the chatbot solution is also part of the front-end of the application; however, it is only accessible to privileged users (researchers) and is implemented in Dart within the mobile application and in HTML pages with JavaScript to allow access through a web viewer.

The back-end of the infrastructure is responsible for handling the stored data and executing routine scripts related to the event-oriented communication between the intervention module and the identification module. In the back-end infrastructure, there are two different managed-access databases, one implemented in SQLite and another implemented in Post-

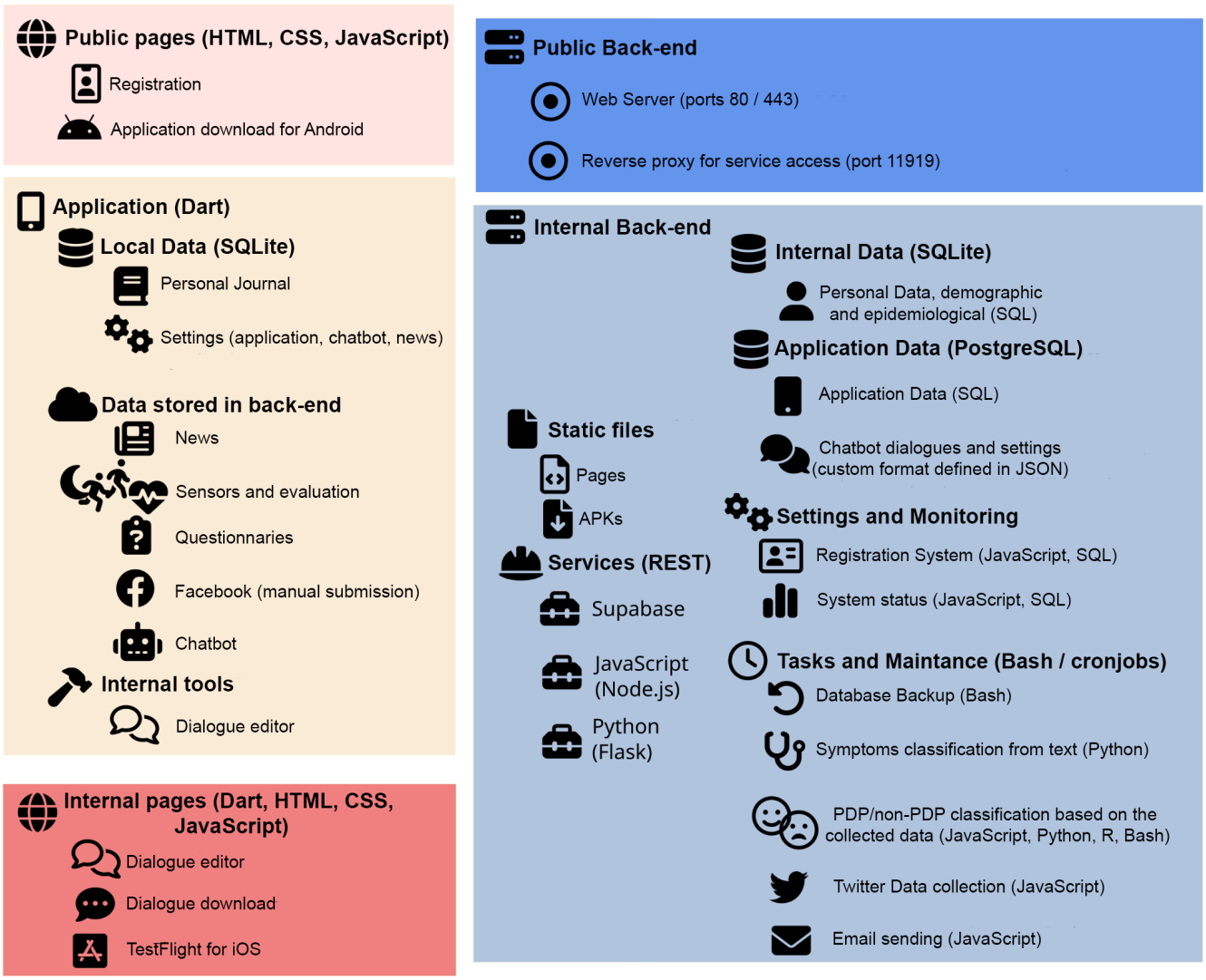

Figure 7. AMIVE infrastructure schematic, including system components and services. Source: The authors.

greSQL. The SQLite[2] database stores sensitive student data (demographic and identification information). This design choice was made considering that this sensitive database is not accessed frequently, allowing the use of a simpler and self-contained database. Furthermore, it is not directly accessible by the application, which was a design decision based on the need to ensure the security of students' data.

The second database, implemented in PostgreSQL[3] to support multiple instantiations and simultaneous access, serves as the main application database. It is designed to store anonymized data from mobile sensors, mental health questionnaires, online social networks (OSN), and the dialogues generated during chatbot interactions. Additionally, within the back-end scope, there are routine scripts implemented in Bash and Python to perform database backups, symptom classification using the NLP model, and PDP identification based on the collected sensor data, as well as the automatic collection of X data from the OSN's API. The outputs of the symptom and PDP classification processes are communicated through JavaScript routines in JavaScript Object Notation (JSON)[4] format files to be presented in the front-end. The communication between the front-end and the back-end is performed through the AMIVE RESTful API (detailed in Subsection 6.1), hosted on the Supabase[5] platform.

The computational solution comprises two main scopes: identification and intervention for PDP. In the identification scope, students' digital phenotyping data are collected and processed by two ML models. The first model performs binary classification of mobile sensor data (collected from smartwatch and smartphone devices) to classify students as PDP or non-PDP. The second model processes the students' textual data (extracted from online social networks and the AMIVE journal) to identify depressive symptoms through an NLP model.

In relation to the ML modules in the identification scope, as cited, there are currently two modules, one for depressive symptom identification and another for PDP classification based on students' digital phenotyping data. The PDP classification model was developed in an experimental study aimed at identifying the best combination of mobile sensor datasets based on the data collected by AMIVE and traditional ML

[2] https://sqlite.org
[3] https://www.postgresql.org
[4] https://www.json.org
[5] https://supabase.com

algorithms. In the study conducted in Ribeiro *et al*. [2026], a series of experiments were carried out regarding feature extraction from the students' sensor data, as well as variations of the dataset designed to minimize the need for artificial imputation of missing data in order to train ML algorithms. The ML approach employed in the experimental study was supervised learning, in which the models require some form of ground truth label associated with each data entry in order to evaluate classification performance during training and tests. In this case, the labels for the PDP and non-PDP groups were derived from the outcomes of the PHQ-9 questionnaire.

The algorithms evaluated in the experiments included Logistic Regression, Decision Trees, Support Vector Machines, Random Forest, K-Nearest Neighbors, and Naive Bayes. In the performance evaluation of the trained models, the best-performing model was the Random Forest classifier, achieving 77% accuracy, Area Under the ROC Curve (AUC) of 0.77, and Matthews Correlation Coefficient (MCC) of 0.53. The accuracy achieved by the model can indicate a decent performance, considering that this metric is calculated as the ratio of correct predictions to the total number of predictions made. At 77%, the performance clearly surpasses the majority class baseline on the AMIVE dataset (55% for non-PDP entries), indicating that the model is not only predicting the dominant class. However, this metric alone may not be sufficient given the slight class imbalance present in the dataset, which could still bias the result toward the majority class's correct classification. In this case, the model's reliability is further reinforced by the AUC, which represents the ability of the model to correctly classify the two labeled classes [Faceli *et al*., 2021; Fawcett, 2006], and the MCC, a robust metric to measure classification quality in imbalanced classes [Chicco *et al*., 2021], both of which ensure that the accuracy result does not stem from a bias toward the majority class. Beyond these metrics, cross-validation was applied during training to further reinforce the robustness of the results. Although the results of the PDP classification model are not perfect, they represent a decent performance considering the complexity of the task.

Regarding the NLP model developed to identify depressive symptoms in students' textual data, as proposed in Mendes and Caseli [2024], the model was trained using a pre-trained BERT model adapted for the Brazilian Portuguese language and fine-tuned for the depressive symptoms classification task. It was trained with annotated datasets of OSN posts from public Facebook pages related to the university student context. The dataset used for training the BERT model was annotated for depressive symptoms with the support of mental health researchers involved in the AMIVE project. The NLP model is capable of identifying depressive symptoms within a list of 18 specific symptoms, as well as a neutral symptom flag indicating the absence of depressive symptoms. Among the depressive symptoms, the ones with the highest identification precision were sleep disorders, worry/fear/anxiety, sadness/depressed mood, and suicide/self-extermination. The two modules perform the PDP and symptoms identification tasks separately, and the output of the PDP model is used by the chatbot to send an invite to the students for conversation, while the output of the NLP model is used to suggest conversation topics. Additionally, the results of both models can be viewed by students within the application.

In the intervention scope, the AMIVE chatbot is designed as a personalized interaction channel to support students' mental health care. The chatbot module provides the following features: (1) Dialogue files in JSON-based format; (2) functionalities for incorporating interactive media; (3) support for text formatting using MARKDOWN[6](sections headers, mono-spaced fonts, bold, italic and emojis); (4) support for figures, animations and sounds; (5) support for tags, dialogue suggestions based on the text classification (application journal) and X (formerly Twitter); (6) support for usage data collection (identified symptoms, suggested dialogues, accessed dialogues, feedback dialogues and questionnaires); (7) support for specific user parameters (name, campus, age, etc); (8) string templates for personalizing dialogue content for specific campuses, such as health services and contact information; (9) support for loading dialogues from external files; (10) integration with a music generation system; and (11) support for enabling or disabling functionalities to allow the use of either the full solution or a placebo version in clinical studies.

To assist the creation and publication of dialogues, an editing tool was implemented to enable dialogue development without requiring modifications on the application's source code. The dialogue tool allows the content creation for all types of media supported by the chatbot, including text, images, sounds, questionnaires, forms and dialogue transitions. Subsection 6.1 presents the architecture of the AMIVE. Subsection 6.2 presents the conversational flow and interaction behavior of the AMIVE chatbot.

6.1 AMIVE architecture

The AMIVE's systems components communicates through an API. The developed API acts as the central layer of the architecture, responsible for receiving, validating and provide data for the services and modules of the infrastructure. This architecture choice was made to ensure that the database remains inaccessible, by direct means, to the mobile application and the other services. It also allows for scalability, since the components are independent.

Between the API and the processing modules, an event-oriented message queue was implemented. The message queue is responsible for temporarily storing the data received by the application and enabling communication between the identification and intervention modules. This behaviour allows the processing modules (identification models) to consume the students' data to perform PDP classification and symptoms identification tasks independently of the API and to provide the outputs as event triggers for other modules. Additionally, the implementation of a message queue allows the processing modules to asynchronously perform their tasks without impacting the API layer. In this way, the API is able to provide the data to the message queue and, while the ML models consume and process the data according to their own capacity, still remain available for other services. Considering that ML models demand significant time and processing resources to perform their tasks, this mechanism improves

[6]https://daringfireball.net/projects/markdown/

the system's responsiveness and scalability.

Within the data processing and storage cycle, the AMIVE chatbot module operates as an interceptor in the architecture. The chatbot consumes the classification outputs produced by the processing modules and, when a student is identified as PDP, it initiates an intervention process. The chatbot then invite the student to start an interaction, providing a personalized and structured conversational flow specifically designed to support the PDP. The chatbot also send data related to the current conversation to the message queue, to be used by other services or to be stored in the database. This architecture design allows the chatbot to remain an independent module in the infrastructure while still begin dynamically triggered by the identification module.

6.2 Conversational flow and interaction behaviour of the AMIVE chatbot

As stated previously, the AMIVE chatbot does not provide automatic or artificially generated dialogue options. This decision was made to ensure control over the dialogue content and to guarantee security of the research participants and end-users. In this context, the chatbot was developed using a rule-based approach, in which dialogue options are organized as nodes within a graph structure. The dialogues are presented as content statements based on the current context, and dialogue flow state, followed by a predefined set of response options from which participants can choose as their responses.

The dialogue statements are not passive presentations of content, each dialogue instance has a sentence that instigates the user's response, aiming to keep the interaction flow. Each response option represents a transition decision, where the selected choice leads to another dialogue instance that is contextually related to the previous exchange and coherent with the current conversation topic. Additionally, fixed options are also available to allow users to end the interaction or return to the beginning of the conversational flow.

Beyond the chatbot's basic structure, it was necessary to define a behavioural and interaction baseline to enable the chatbot to provide mental health support. Through the focus group, it was defined that the chatbot's interactions with the user would follow the following steps: 1) Greeting: in order to hail the user and begin to understand the context and problem; 2) Analysis: with the aim of understanding the user's demand at the moment and seeking the next most appropriate courses of action; 3) Intervention: to reaffirm the user's welcome, guide and offer intervention for the current demand; 4) Feedback: with the aim of understanding the effectiveness of the previous stage through user feedback; and 5) Conclusion: with the aim of farewell.

After defining the interaction steps between the chatbot and the user, the communication flow between them was defined (Figure 8). The proposed interaction flow begins with an initial greeting, during which the chatbot welcomes the user, explains the triggering factor that initiated the conversation, and inquires about the user's immediate state. Following this, the chatbot conducts a preliminary exploration of the user's context to establish a foundational understanding. Subsequently, the interaction progresses to a more detailed contextual specification, wherein the chatbot deepens its inquiries to gain a comprehensive understanding of the user's situation. Upon reaching this stage, the chatbot delivers an intervention, directing the user toward an appropriate course of action for care. To assess the efficacy of the intervention, the chatbot then solicits user feedback. Throughout the process, the chatbot also functions as a data collector, leveraging both initial and prior user inputs to identify previously attempted protective measures. The interaction concludes with an appropriate closing message, ensuring a structured and coherent conversational experience.

The interaction between the user and the chatbot can be initiated in two distinct ways: (1) through the user's voluntary decision to engage with the conversational feature; or (2) through an application-triggered interaction based on the results of the depressive profile classification.

As stated, the AMIVE infrastructure comprises two main scopes: (1) a PDP identification module, which uses ML models classification based on the user's digital phenotyping data; and (2) a mental health support module, implemented as the chatbot. There is an interaction point between these two modules made by the message queue. Within the AMIVE identification scope, a classification model categorizes the user's data as PDP or non-PDP, and a NLP model is used to identify depressive symptoms in the user's textual data, such as content from online social networks or the AMIVE journaling feature. The outputs of these ML models are used to raise the user's self-awareness regarding their potential mental health state, and to support the chatbot's functionality. The classification result, through the message queue, may prompt the application to encourage the user to initiate a conversation with the chatbot, if the results shows true for PDP, while the symptoms identified in textual data can be used to define the conversation topic, acting as a filter for the offered dialogue options.

The interaction between the modules is part of the HITL interaction cycle implemented within the AMIVE infrastructure to provide personalized intervention and mental state monitoring for college students. The AMIVE HITL approach was originally proposed in [Alves *et al*., 2023]. Figure 9 illustrates the interaction flow between the student and the AMIVE modules, highlighting the cyclical exchange of information that supports adaptive chatbot dialogues. Additionally, in accordance with the HITL approach, participants were also able to evaluate their experience with the chatbot after each interaction.

7 Evaluation of the AMIVE solution

This section presents the protocol (Subsection 7.1) for recruiting students for the clinical study to evaluate the AMIVE chatbot, the profile of the students (Subsection 7.2) participating in the study and the results (Subsection 7.3) obtained from the evaluation. This study was approved by the Research Ethics Committee of the Federal University of São Carlos (UFSCar) - São Paulo - Brazil.

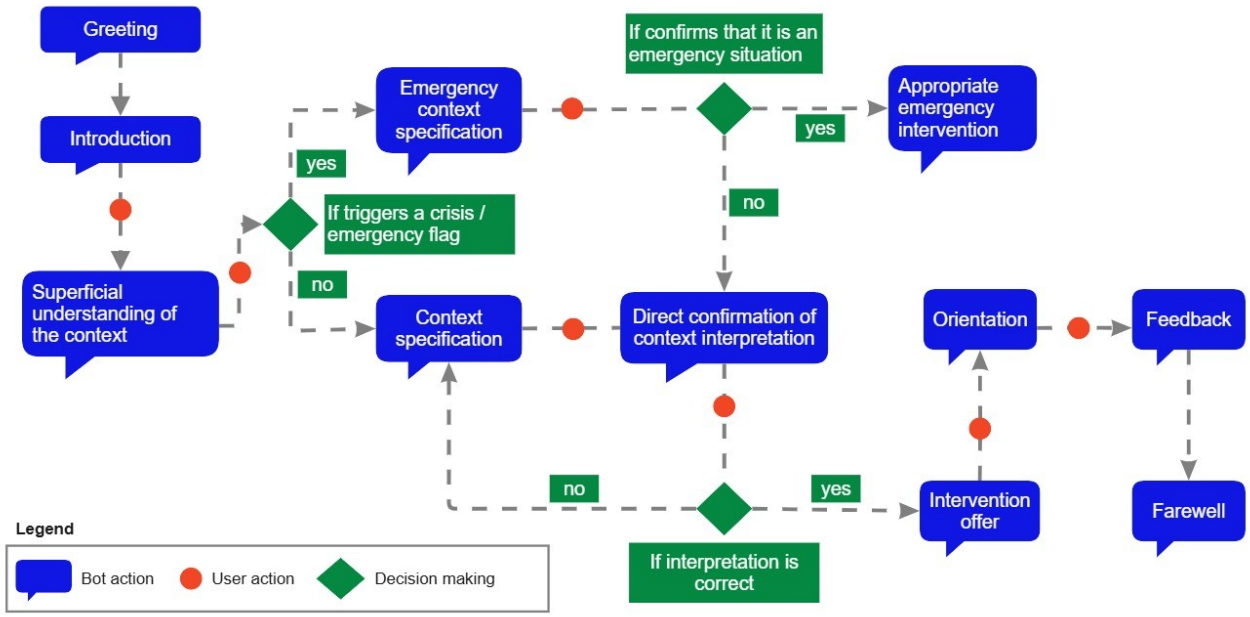

Figure 8. Communication flow between the chatbot and the user [Souza *et al*., 2025].

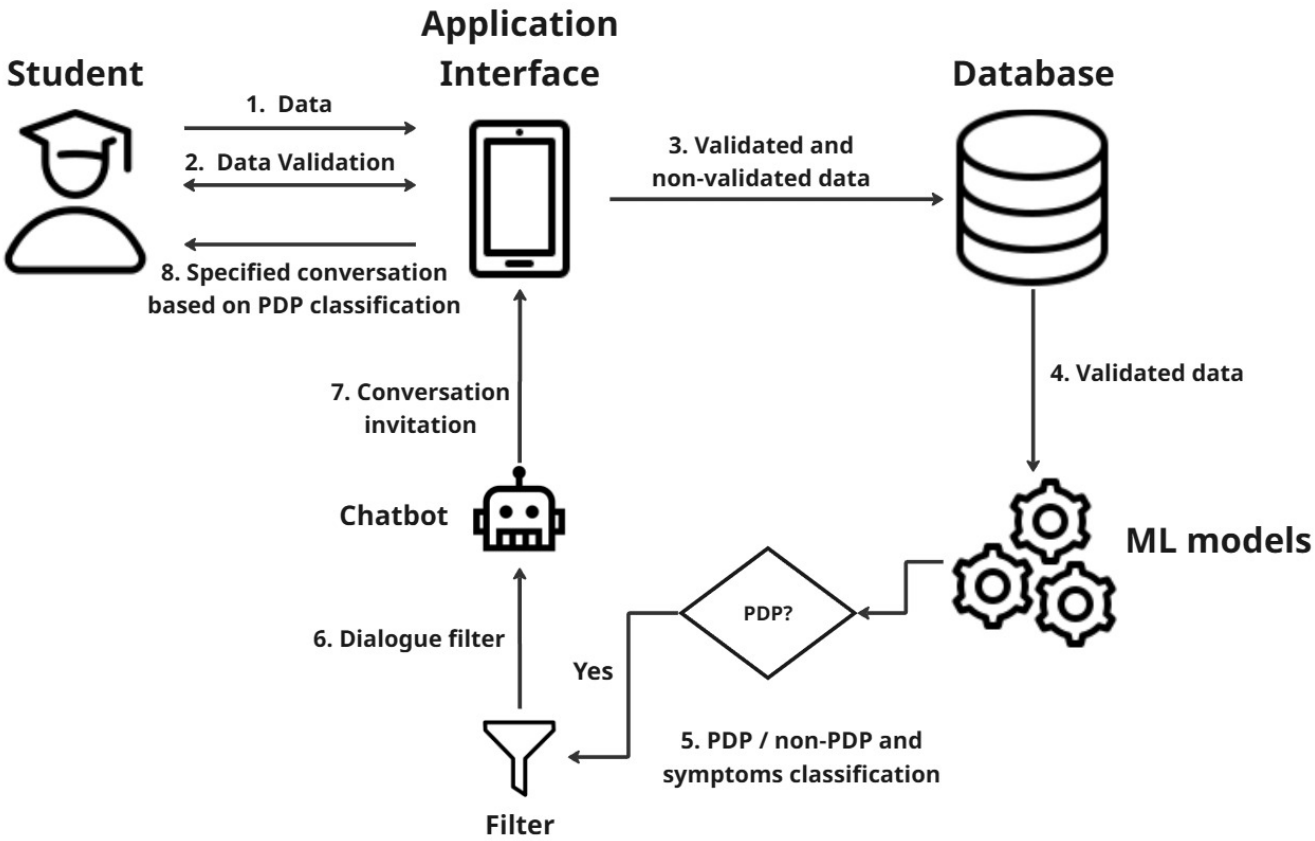

Figure 9. Interaction between Student and the AMIVE modules. Adapted from Alves *et al*. [2023].

7.1 Recruitment protocol

The invitation to participate in the clinical study was sent to student email lists and was published on all official channels of the UFSCar. A total of 222 students expressed interest in the study and responded to a sociodemographic survey. Students interested in the study also answered the PHQ-9 questionnaire [Santos *et al*., 2013].

Students with moderately severe symptoms (PHQ-9 score between 15 and 19), severe symptoms (PHQ-9 score between 20 and 27) or active suicidal ideation (response greater than 0 on item 9 of the PHQ-9) were not recruited for the second phase of the study. These students were offered support and listening by a health professional, with the intention of psychoeducation and referral to a health service. Students who presented symptoms compatible with mild depression (PHQ-9 score between 5 and 9) or moderate depression (PHQ-9 score

between 10 and 14) were invited to participate in the second phase of the study. In total, of the 222 interested students, 44 were identified as having mild or moderate depression and were invited to participate in the second phase of the study.

The second phase of the study evaluated different functionalities of the AMIVE application. Study participants received a link to download and install the AMIVE application on their personal mobile devices. 44 students who were invited to participate in the second phase of the study and were divided into two groups: treatment and control. To ensure sample homogeneity and control for confounding variables, participants were assigned to groups through a simultaneous stratification process carried out by the project's mental health professionals. This procedure took into account sociodemographic, academic, and clinical criteria, resulting in a balanced distribution regarding field of study, gender identity, age (mean of 25 years in both groups), and severity of depressive symptoms (mean PHQ-9 score of 9). This rigorous matching ensured that both groups exhibited comparable profiles, thereby minimizing bias and enhancing the reliability of the intervention results. The control group had access to a placebo solution, while the treatment group had access to the complete AMIVE, including the chatbot and, consequently, had access to the dialogues.

7.2 Students profiles

The profiles of the 44 students participating in the study was identified considering the applied sociodemographic research. The research data indicate that the students participating in the study, for the most part, are in undergraduate studies (79.5%), do not have a partner (50.0%), do not have children (88.6%), use or have used psychiatric medication (47.7%) and are being monitored by a mental health professional (52.3%). 45.5% of students are from the Exact and Technology area, 13.6% are from the Humanities and Management area and 40.9% are from the Biological and Health area. 47.7% of students reported being in the initial years of their degree, with less than 50% of the course completed.

In relation to the distance from their hometown and, consequently, from their family, 38.7% of students study less than 100km away from their hometown. 13.6% study between 100km and 200km away from their hometown. 31.8% study between 200km and 300km away from their hometown. And 15.9% study more than 300km away from their hometown. In relation to monthly family income, 27.3% of students are part of a family with a monthly income of up to one minimum wage per person. 18.2% are part of a family with a monthly income of between 1 and 2 minimum wages per person. 27.3% are part of a family with a monthly income of between 2 and 3 minimum wages per person. 13.6% are part of a family with a monthly income of between 3 and 4 minimum wages per person. And 13.6% are part of a family with a monthly income greater than 4 minimum wages per person.

7.3 Evaluation results

The chatbot was made available to volunteer students and data was collected over a period of four weeks. For the analysis included in this article, the results obtained through the following were considered: i) TAM [Davis, 1989]; and ii) the responses given by users in questionnaires evaluating the use of the chatbot. The chatbot evaluation questionnaire and the TAM were applied to the treatment group, which used the full AMIVE application and had access to the dialogues.

Figure 10 shows the graph with the number of accesses by dialogue theme. As can be seen in the graph, the most accessed dialogues were about the themes "Depression in university students" and "Support Network". Of the available dialogues, the only one that was not accessed once was about the theme "First year at University". Although a qualitative analysis was not conducted with the volunteer students regarding their reasons for accessing the dialogues, the analysis of access frequencies opens up room for discussion about which of these topics have been of greater interest to the participants.

At the end of each dialogue, students were invited to answer two questions to evaluate that specific dialogue. Not every dialogue access resulted in responses to the evaluation questions, which accounts for the difference between the number of accesses and the number of responses. The questions asked and the possible answers were:

1. How do you evaluate the experience of chatting with chatbot? Possible answers: i) I didn't like talking to chatbot; ii) I'm neutral; and iii) I enjoyed talking to chatbot.
2. How do you evaluate the impact of conversation? Possible answers: i) It didn't help me; ii) I'm neutral; and iii) It helped me.

Figure 11 shows the graph with the results of the responses to the first question for each theme (n = 21 responses). The themes "first year at University", "initial dialogue" and "First use" are not included in the graph because these themes did not receive any evaluation from the students. As can be seen in the graph, among the 21 responses, 16 indicated an enjoyable experience when interacting with the chatbot, four were neutral, and one indicated that the interaction was not liked. The two most accessed themes, Depression in university students" and Support Network", received only responses indicating an enjoyable interaction, with four and three responses, respectively. Neutral responses were also recorded for Autonomy", Distance from origin city", and Gender", while the only response indicating that the interaction was not liked was associated with "Poor academic performance".

Figure 12 shows the graph with the results of the responses to the second question for each theme (n = 21 responses). Of these, 9 indicated that the conversation was helpful, 10 were neutral, and 2 indicated that the conversation was not helpful. Responses indicating that the conversation was helpful were recorded for seven of the nine evaluated themes, while responses indicating that it was not helpful were recorded only for Distance from origin city" and Poor academic performance".

At the end of the study, participants were asked to answer 6 questions associated with a questionnaire based on the TAM. This questionnaire helps researchers evaluate the acceptability of technology solutions considering the perception of usefulness and ease of use [Davis, 1989]. The questionnaire was applied within the Amive application itself, as one of the chatbot conversations, which was only available at the end of the

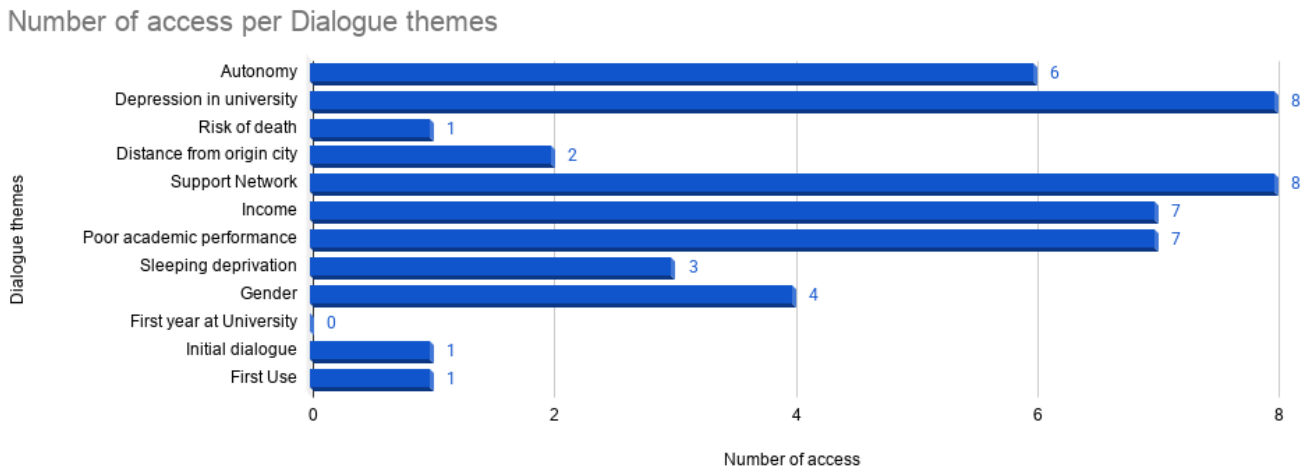

Figure 10. Graph with the number of accesses to the dialogue for each theme. Adapted from Souza *et al*. [2025].

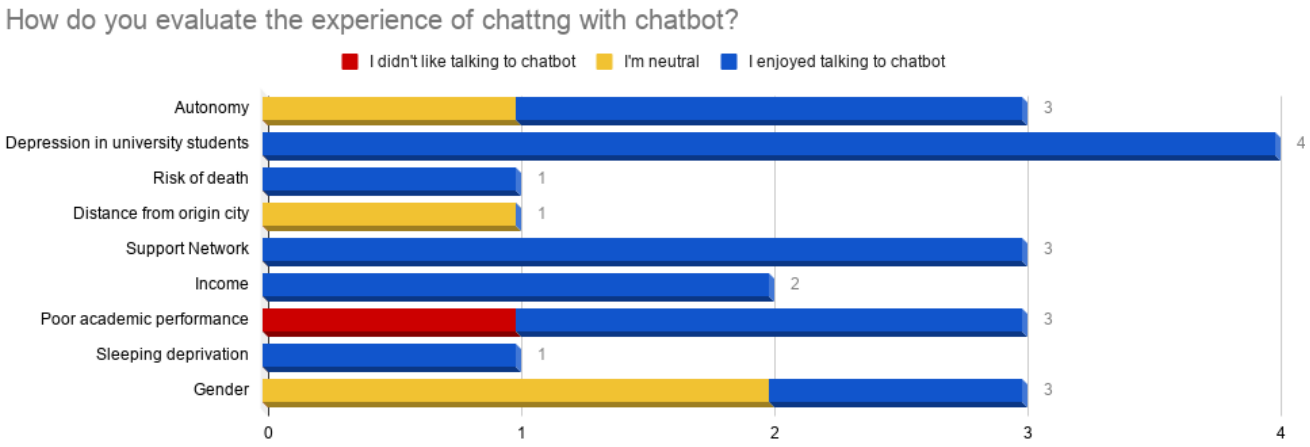

Figure 11. Graph with the results of the answers to the first evaluation question. Adapted from Souza *et al*. [2025].

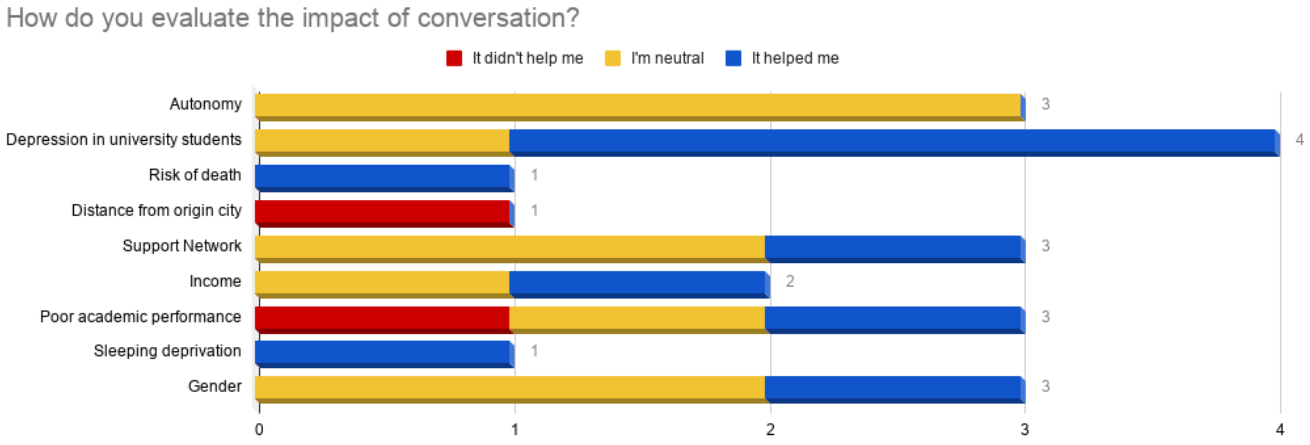

Figure 12. Graph with the results of the answers to the second evaluation question. Adapted from Souza *et al*. [2025].

study. Although it was made available to all 22 participants in the treatment group, only 8 participants responded to the TAM. Figure 13 shows a graph with the responses obtained.

As can be seen in the graph, 7 out of 8 participants who answered the questionnaire reported being satisfied with the application, and the same number considered and the same number considered the application easy to use. Six respondents considered the application useful in their lives, and the same number reported that they would recommend it to friends. In addition, all eight respondents found it easy to learn how to use the application's interface.

Overall, the TAM responses indicate favorable perceptions

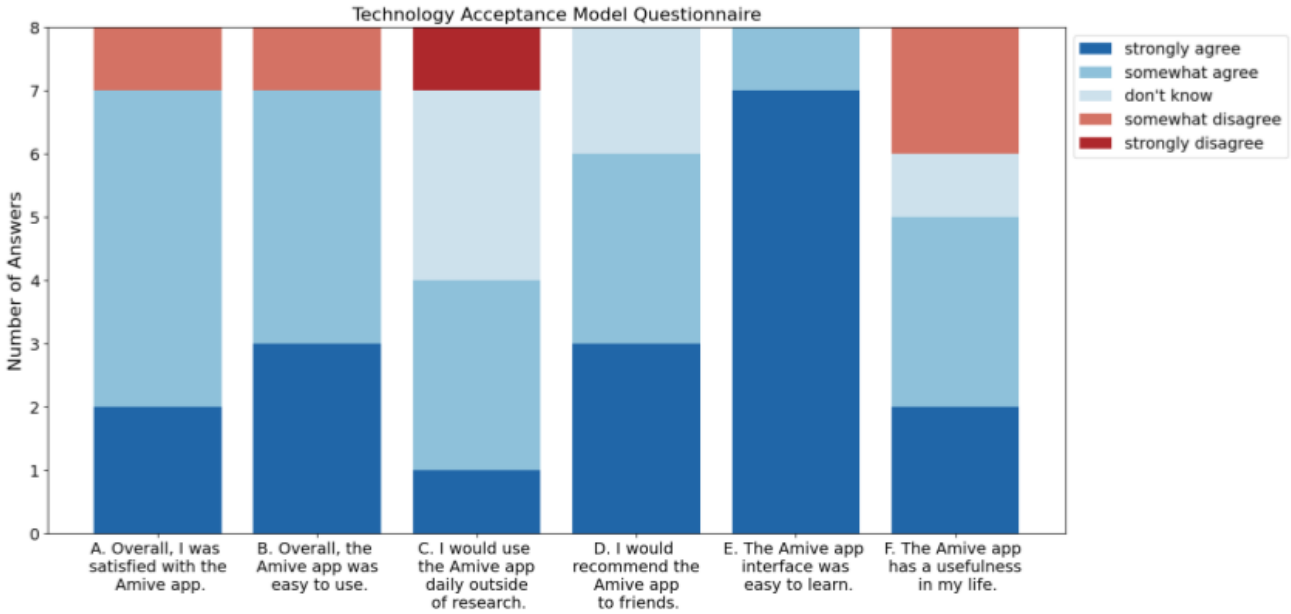

Figure 13. Graph showing the results of applying the Technology Acceptance Model Questionnaire (TAM). Source: The authors.

of the application's ease of use, usefulness, and interaction experience among the eight respondents. However, only half reported that they would use the application daily outside the research context. Given the small number of respondents, these findings are descriptive and should be interpreted with caution. These results also point to the need for further qualitative investigation to better understand the motivations of participants and their reasons for using or not using the application beyond the research context.

8 Lessons learned

This research involved a co-design process for the AMIVE chatbot, presented the AMIVE infrastructure, and reported the results of an evaluation of AMIVE with the target audience in a real-world usage context. Some lessons learned from this research include:

The need for stakeholders from different areas and expertise. CUI for mental health require designers and healthcare professionals on the team. Collecting data for design, applying techniques and methods that inform design decisions, interaction design, and designing user interfaces, among others, requires specialized design knowledge. Furthermore, it is extremely important that health experts take responsibility for the content of conversations, which aim to improve the health situation. Finally, the students, as final users, may have a voice and be as involved as possible.

The need to apply specialized knowledge to design and evaluate. Specialized knowledge was used at different moments in the process and by different actors. Designers used design guidelines that supported the decision-making and interface design. Health professionals used specific databases about the target audience to select conversational topics and provide scientific highlights in the texts. Evaluations may use specialized instruments depending on which aspect will be considered (e.g. TAM or PHQ-9).

Be guided and collaborate. The collaborative approach employed in the design process of the AMIVE application integrated qualitative data analysis, development requirements, and the conceptualization of tools and artifacts. This methodology facilitated a cyclical process of production and research, wherein the outcomes of each phase were systematically assessed and validated in subsequent stages. Consequently, an iterative progression was established throughout the project's development. While the primary focus of the team was to delineate the chatbot's development scope, additional factors—such as the communication framework and user interaction strategies—were also systematically examined. Moreover, previously established design elements were refined through engagement with specialized professionals who contributed to discussion groups. The inclusion of diverse perspectives during these collaborative stages fostered a multidisciplinary validation process, ensuring the rigorous evaluation of key design components. Thus, the project's multidisciplinary and cooperative nature played a pivotal role in consolidating and validating its outcomes. The resulting discussions reinforced the chatbot's design elements, ultimately ensuring a coherent and effective development process tailored to the needs of the target audience.

Choose tools wisely. Different technologies can be applied in the design process as well as in implementing the desired design solution. In this research, a conversation editor was created to support multidisciplinary work, since specific programming languages were not common to all. Also, conversational agent technology can influence many design decisions (e.g. ruled-based bots vs large language models). The technologies that will be applied must be discussed as soon as possible in the project.

Evaluate not only with specialists but with students. Students should be involved at various stages of the design process, as they experience unique mental health situations and may have particular interaction demands. Furthermore, the evaluation of the CUI is a fundamental aspect of the approach and must involve the target audience. Longitudinal

studies are best suited for MDD, given that symptoms are observed over time and provision of interventions must be continuous.

Treat personal information as a central architectural concern. Since the system processed sensitive data, security principles were embedded from the initial design phase. Confidentiality, integrity, authentication, and authorization mechanisms guided decisions regarding data storage, access control, communication between services, and message handling, reducing exposure to risks.

Design for scalability. Software systems are inherently dynamic, and the initial infrastructure was designed with that assumption. The adoption of loosely coupled microservices communicating via RESTful APIs and consuming from a message queue facilitated incremental evolution over time. The current architecture evolved from its earlier versions, and design for scalability reduced the impact of adaptations.

Align with Data Protection Laws. Regulatory alignment was treated as a design requirement, therefore compliance with the Brazilian General Data Protection Law (LGPD) influenced architectural and operational decisions related to data processing and access.

9 Conclusion

The motivation for this research came from observations at the UFSCar regarding a rise in mental health issues among students. Universities often struggle to provide adequate support for these students, and technology could play a significant role in addressing this challenge.

While this audience typically engages with technology systems, the significance of this paper lies in application of a design approach that supports addressing their specific interaction demands related to mental health that have not been adequately explored in existing literature. For instance, aspects such as trustworthiness, empathy and specific conversation topics (e.g. academic performance, distance from family) emerged with the approach proposed.

This paper presented an approach to codesigning CUI for colege students with depression. While the literature emphasizes a participatory and multidisciplinary approach, one of the key differences highlighted in this paper is our focus on specific methods to foster collaboration and participation from both the target audience and specialists. While user-centered design acknowledges the importance of this collaboration, this paper provides practical strategies for implementing it within this unique and relatively unexplored context. For example, a design specialist applied guidelines for CUI on dialogues created by health specialists, a form of collaboration that remains underexplored in this context. The primary contribution of the paper is to present the approach utilized in the research, detailing its application and the results of its evaluation with the target audience. This approach serves as a reusable resource that other researchers or developers can adopt for similar projects.

The suggested steps, methods, techniques, and artifacts were reported. The interfaces created through the instantiation of this approach were evaluated by students, in a longitudinal clinical study and in their daily lives. The results suggest that the approach led to enjoyable user interfaces as desired.

Furthermore, this article also details the AMIVE infrastructure, built to identify and support university students with PDP. Regarding the computational infrastructure of AMIVE, the main focus was on supporting data collection, real-time PDP identification, and the chatbot intervention module, as well as defining how these modules communicate with each other. Additionally, the platform was designed to support the collaboration of mental health researchers. In this context, several measures and functionalities were implemented beyond what is available to the end user. In this regard, the platform serves not only as a tool for PDP identification and intervention for college students, but also as a research platform to support studies in the field of mental health. One example is the existence of internal pages that allow researchers to add new chatbot dialogues and to enable or disable specific functionalities, facilitating the execution of studies in a case-control scenario.

Moreover, beyond the main objective of the AMIVE project, the platform also has the potential to support mental health monitoring even for students who are not classified within a PDP group. As noted by Burgess *et al*. [2025], pervasive and ubiquitous technologies play an important role for individuals with mental health issues, often acting as tools for continuous mental health monitoring and self-regulation. This context aligns with the AMIVE project, as it provides a computational solution for PDP identification and intervention for depressive users in an automatic, private, and autonomous manner, delivered through a smartphone application. Although seeking mental health care may still be associated with Saud *et al*. [2024], technology-mediated interaction may make individuals feel more comfortable engaging with the system's solutions.

One limitation of this research is that the approach was instantiated in only one project. Another limitation is the sample size, with 44 students participating in the evaluation. However, there are significant challenges in recruiting participants for mental health research. Due to the stigma associated with participating in such studies, obtaining large sample sizes is often difficult [Iflaifel *et al*., 2024].

Regarding the limitations related to the AMIVE infrastructure, an important concern involves the scalability of the platform, as well as the ethical and privacy implications associated with the dissemination of the computational solution. The AMIVE infrastructure was initially developed to support the collection of digital phenotyping data from university students and was later used to test the proposed solution for PDP support. In this context, the studies conducted with the AMIVE platform represent only one instantiation of the project. As a relatively small-scale deployment with a limited number of participants, it was possible to provide mental health support through professional assistance in severe and critical cases, considering that one of the pillars in the AMIVE project is the Latin phrase of "primum non nocere" (first do no harm), this condition is important to assure that the solution would not worsen the student's mental health state. Also, in this regard, the chatbot for mental health support implemented in the AMIVE, currently only supports pre-written dialogues organized in a graph-based dialogue structure. This design decision was made to guarantee the quality and content of the

information provided to students. However, this approach may limit the range of conversational variations and the chatbot's behavior, as students may frequently reach the same response node even when initiating different interactions.

In the context of the limitations of the AMIVE infrastructure, future work includes improving the replicability and adaptability of the computational solution for other contexts, research projects, and universities. In this way, the same computational infrastructure design could be instantiated and applied across different universities, campuses, and studies. As a tool capable of collecting digital phenotyping data, administering periodic mental health questionnaires, interacting with a chatbot, and performing automatic PDP identification, the platform has the potential to support a wide range of mental health research.

Furthermore, exploring the use of Large Language Model (LLMs) agents as an alternative to the graph-based chatbot could make user interactions more natural and increase the variability of dialogue. However, this approach would still require careful validation studies, as well as strict ethical and privacy considerations.

Future work also include, as cited, investigating this approach in other cultural contexts and universities and also for other mental health conditions. We also highlight as future work the investigation of how Artificial Intelligence can be further explored and used to improve dialogues within the AMIVE application, without compromising the security and reliability of the application in a context as sensitive as mental health.

9.1 Threats to Validity

As with any empirical study, the results of this research are subject to threats to validity that must be considered when interpreting the findings. Following the classification by Wohlin *et al*. [2000], threats to conclusion, internal, construct, and external validity must be considered.

Conclusion validity - The limited sample size and the absence of inferential statistical analyses restrict the statistical power and robustness of the findings. Therefore, the results should be interpreted as descriptive and exploratory, but it is important to highlight the nature of the research involving sensitive contexts, which sometimes makes access to the research's target audience difficult.

Internal validity - Only 8 of the 22 participants in the treatment group responded to the TAM questionnaire, which may have introduced response bias. Furthermore, the lack of a comparison with the control group limits causal interpretations, as the results could have been influenced by contextual factors or prior experiences.The four-week study period may not have been sufficient to rule out a potential novelty effect. Future studies should include comparison groups and longer-term longitudinal assessments.

Construct validity - Although TAM provides a solid foundation for assessing perceived usefulness and ease of use, the limited number of responses and the reliance on self-reported data may limit the assessment of these constructs.

External validity - The study was conducted at a single Brazilian public university, which limits the generalizability of the findings to other institutions and contexts.

Taken together, these threats define the scope of the conclusions without invalidating the findings within the context investigated. The results should be interpreted in light of the specific context in which the research was conducted and the limitations presented. Future studies involving larger samples, different contexts, comparison groups, and longitudinal assessments could evaluate the consistency and generalizability of the results.

10 Ethical Issues

To ensure ethical standards during the study, services psychologist were made available to support the research participants at any time. The study, conducted with university students, received approval from the Research Ethics Committee of the Federal University of São Carlos (SP-BR) under opinion number 67249823.5.0000.5504. Participants signed an informed consent form, which outlined the risks and benefits of the research. They were informed that their participation was voluntary, that they could withdraw their consent and discontinue their participation at any time without any negative consequences, and that the data collected would be anonymized. Furthermore, the data was anonymized even among team members, meaning that not all team members had access to all participants' data. The focus group was carried out with members of the research project.

Regarding the ethical issues related to the AMIVE infrastructure, the concerns about research ethics and privacy primarily lies on the data collected from users and its management protocol. The project's data management plan was part of the study approved by the Research Ethics Committee of the Federal University of São Carlos and is guided by the following policies: (1) all data used or published in researches will be anonymized and made available to the scientific community; (2) data storage will follow the preservation policy of the University Institutional Repository for a minimum period of five years; (3) the main application infrastructure does not have direct access to the database, only to the anonymized data, ensuring that the data used by the modules does not include student identification, thereby preventing potential data security breaches.

Furthermore, the AMIVE project has guidelines aligned with the Brazilian Data Protection Law (LGPD), such as: (1) specifying to users which data is collected, the sensitivity of the data, when it is collected, and the purpose of the collection; (2) informing users about data sharing with third parties, data retention periods, and rights over their own data (such as correction, access, and exclusion); (3) notifying users about changes in the privacy policy; (4) providing information about the data management plan and security measures taken; and (5) highlighting that data not initially intended may be collected, as long as it is collected with the user's consent.

Declarations

Authors' Contributions

All authors read and approved the final manuscript. Among the main contributions of the authors of this paper, the following stand out:

- Souza, P. M. is the main contributor and writer of this manuscript.
- Conceptualization - Souza, P. M.; Neris, V. P. A.
- Data curation - Garcia, F. E.; Alves, V. C.
- Formal analysis - Alves, V. C.; Souza, P. M.
- Investigation - Alves, V. C.; Pires, I. C.; Bleicher, T.
- Project administration and supervision - Neris, V. P. A.
- Software - Garcia, F. E.; Saud, C.; Oliveira, A. B.; Ribeiro, E. Y. A.; Guardia, H. C.
- Writing – original draft - Souza, P. M.; Silva, G. C.; Ribeiro, E. Y. A.
- Writing – review & editing - Souza, P. M.; Ribeiro, E. Y. A.; Neris, V. P. A.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

We would like to thank all the volunteer research participants, the entire AMIVE project team, and all the students and researchers involved.

Funding

This study was financed in part by the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES)* - Finance Code 001. Grants #2020/05157-9 and #2024/12772-2, São Paulo Research Foundation (FAPESP).

Availability of data and materials

The research instruments and non-sensitive materials associated with this study are available in the UFSCar repository[7] as a dataset titled "AMIVE public data." These materials include a file illustrating the structure of the JSON files, an excerpt from one of the conversations, and a file containing the questions from the TAM questionnaire used in the study. Due to ethical considerations and the sensitive nature of the research involving mental health, materials containing anonymized participant data will not be made publicly available. These materials may be made available upon reasonable request, subject to the signing of a data use responsibility agreement by the requester.

References

Alves, V. d. C., Garcia, F. E., Saud, C., Mendes, A., Medeiros Caseli, H., Genaro Motti, V., de Oliveira Neris, L., Blecher, T., and Almeida Neris, V. P. (2023). College students-in-the-loop for their mental health: a case of ai and humans working together to support well-being. *Interaction Design and Architecture(s)*, (59):79–94. DOI: 10.55612/s-5002-059-003.

American Psychiatric Association (2022). *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Association Publishing, Washington, DC, 5th, text revision edition. DOI:10.1176/appi.books.9780890425596.

Andifes (2019). Associação nacional dos dirigentes das instituições federais de ensino superior: V Pesquisa Nacional de Perfil Socioeconômico e Cultural dos (as) Graduandos (as) das IFES – 2018. Available at:`https://www.andifes.org.br/2019/05/27/v-pesquisa-nacional-de-perfil-socioeconomico-e-cultur`

Arino, D. O. and Bardagi, M. P. (2018). Relação entre fatores acadêmicos e a saúde mental de estudantes universitários. *Psicologia em Pesquisa*, 12:44 – 52. DOI: 10.24879/2018001200300544.

Auerbach, R. P., Mortier, P., Bruffaerts, R., Alonso, J., Benjet, C., Cuijpers, P., Demyttenaere, K., Ebert, D. D., Green, J. G., Hasking, P., Murray, E., Nock, M. K., Pinder-Amaker, S., Sampson, N. A., Stein, D. J., Vilagut, G., Zaslavsky, A. M., Kessler, R. C., and WHO WMH-ICS Collaborators (2018). WHO World Mental Health Surveys International College Student Project: Prevalence and distribution of mental disorders. *Journal of Abnormal Psychology*, 127(7):623–638. DOI: 10.1037/abn0000362.

Baranauskas, M. C. C., Martins, M. C., and Valente, J. A. (2013). *Codesign de Redes Digitais: Tecnologia e Educação a Serviço da Inclusão Social*. Penso, 1ª edição edition. Available at:`https://minhabiblioteca.com.br/catalogo/livro/74130/codesign-de-redes-digitais/`.

Bayazidi, A. S., Romero, N., Ribeiro, E. Y. A., Alves, V. d. C., Valejo, A. D. B., Saud, C. D. S. A., Garcia, F. E., Neris, V. P. d. A., and Bento, M. (2025). Unsupervised clustering of depression profiles in university students using behavioral and sociodemographic features. In *2025 21st International Symposium on Biomedical Image Processing and Analysis (SIPAIM)*, pages 1–5. DOI: 10.1109/SIPAIM67325.2025.11283203.

Burgess, E. R., Munson, S. A., Mohr, D. C., and Reddy, M. C. (2025). What's in your kit? mental health technology kits for depression self-management. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI '25, New York, NY, USA. Association for Computing Machinery. DOI: 10.1145/3706598.3713585.

Cardoso, P. d. F. (2022). Prevalência de ansiedade e depressão nos estudantes universitários do Brasil: uma revisão integrativa. Available at:`http://pantheon.ufrj.br/handle/11422/16801`.

Carrol, J. (1999). Five reasons for scenario-based design. In *Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences. 1999. HICSS-32. Abstracts and CD-ROM of Full Papers*, volume Track3, pages 11 pp.–. DOI: 10.1109/HICSS.1999.772890.

Carroll, J. (2000). Five reasons for scenario-based design. *Interacting with Computers*, 13(1):43–60. DOI: 10.1016/S0953-5438(00)00023-0.

Carroll, J. M. (1997). Chapter 17 - Scenario-Based Design. In Helander, M. G., Landauer, T. K., and Prabhu, P. V., editors, *Handbook of Human-Computer Interaction (Second Edition)*, pages 383–406. North-Holland, Amsterdam. DOI: 10.1016/B978-044481862-1.50083-2.

Chicco, D., Tötsch, N., and Jurman, G. (2021). The matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness

[7] https://repositorio.ufscar.br/home

in two-class confusion matrix evaluation. *BioData Mining*, 14(13). DOI: 10.1186/s13040-021-00244-z.

Cooper, A. (2004). *The Inmates Are Running the Asylum: Why High Tech Products Drive Us Crazy and How to Restore the Sanity (2nd Edition)*. Pearson Higher Education. Available at:`https://dl.acm.org/doi/10.5555/984201`.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3):319–340. Publisher: Management Information Systems Research Center, University of Minnesota. DOI: 10.2307/249008.

Dresch, A., Lacerda, D., and Júnior, J. (2015). *Design Science Research: Método de Pesquisa para Avanço da Ciência e Tecnologia*. Bookman Editora. DOI:10.2753/MIS0742-1222240302.

Faceli, K., Lorena, A. C., Gama, J., Almeida, T. A. d., and Carvalho, A. C. P. d. L. F. d. (2021). *Inteligência artificial: uma abordagem de aprendizado de máquina*. LTC. Available at:`https://repositorio.usp.br/item/003128493`.

Fawcett, T. (2006). An introduction to roc analysis. *Pattern Recognition Letters*, 27(8):861–874. ROC Analysis in Pattern Recognition. DOI: 10.1016/j.patrec.2005.10.010.

Fitzpatrick, K. K., Darcy, A., and Vierhile, M. (2017). Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial. *JMIR Mental Health*, 4(2):e7785. Company: JMIR Mental Health Distributor: JMIR Mental Health Institution: JMIR Mental Health Label: JMIR Mental Health Publisher: JMIR Publications Inc., Toronto, Canada. DOI: 10.2196/mental.7785.

Fragelli, T. B. O. and Fragelli, R. R. (2021). Por que estudantes universitários apresentam estresse, ansiedade e depressão? uma rapid review de estudos longitudinais. *Revista Docência do Ensino Superior*, 11:1–21. DOI: 10.35699/2237-5864.2021.29593.

Gabrielli, S., Rizzi, S., Bassi, G., Carbone, S., Maimone, R., Marchesoni, M., and Forti, S. (2021). Engagement and Effectiveness of a Healthy-Coping Intervention via Chatbot for University Students During the COVID-19 Pandemic: Mixed Methods Proof-of-Concept Study. *JMIR mHealth and uHealth*, 9(5):e27965. DOI: 10.2196/27965.

Galvão, V. F., Maciel, C., and Garcia, A. C. B. (2019). Creating chatbots to talk with humans: HCI evaluations and perspectives. In *Proceedings of the 18th Brazilian Symposium on Human Factors in Computing Systems*, IHC '19, pages 1–11, New York, NY, USA. Association for Computing Machinery. DOI: 10.1145/3357155.3358460.

Gupta, G. K. and Sharma, D. K. (2021). Depression detection on social media with the aid of machine learning platform: A comprehensive survey. In *2021 8th International Conference on Computing for Sustainable Global Development (INDIACom)*, pages 658–662. DOI: 10.1109/ACCESS.2025.3530862.

He, Y., Yang, L., Zhu, X., Wu, B., Zhang, S., Qian, C., and Tian, T. (2022). Mental Health Chatbot for Young Adults With Depressive Symptoms During the COVID-19 Pandemic: Single-Blind, Three-Arm Randomized Controlled Trial. *Journal of Medical Internet Research*, 24(11):e40719. Company: Journal of Medical Internet Research Distributor: Journal of Medical Internet Research Institution: Journal of Medical Internet Research Label: Journal of Medical Internet Research Publisher: JMIR Publications Inc., Toronto, Canada. DOI: 10.2196/40719.

Iflaifel, M., Hall, C. L., Green, H. R., Willis, A., Rennick-Egglestone, S., Juszczak, E., Townsend, M., Martin, J., and Sprange, K. (2024). Strategies to improve recruitment in mental health clinical trials: a scoping review (RE-MIND study). *Trials*, 25:832. DOI: 10.1186/s13063-024-08665-x.

Koulouri, T., Macredie, R. D., and Olakitan, D. (2022). Chatbots to support young adults' mental health: An exploratory study of acceptability. *ACM Trans. Interact. Intell. Syst.*, 12(2). DOI: 10.1145/3485874.

Kuhlmeier, F. O., Gnewuch, U., Lüttke, S., Brakemeier, E.-L., and Mädche, A. (2022). A personalized conversational agent to treat depression in youth and young adults – a transdisciplinary design science research project. In *The Transdisciplinary Reach of Design Science Research: 17th International Conference on Design Science Research in Information Systems and Technology, DESRIST 2022, St Petersburg, FL, USA, June 1–3, 2022, Proceedings*, page 30–41, Berlin, Heidelberg. Springer-Verlag. DOI: 10.1007/978-3-031-06516-3_3.

Kumar, J. A. (2021). Educational chatbots for project-based learning: investigating learning outcomes for a team-based design course. *International Journal of Educational Technology in Higher Education*, 18(1):65. DOI: 10.1186/s41239-021-00302-w.

Li, Y. J. and Luo, H. I. (2021). Depression prevention by mutual empathy training: Using virtual reality as a tool. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pages 60–63. DOI: 10.1109/VRW52623.2021.00017.

Mendes, A. R. and Caseli, H. (2024). Identifying fine-grained depression signs in social media posts. In Calzolari, N., Kan, M.-Y., Hoste, V., Lenci, A., Sakti, S., and Xue, N., editors, *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 8594–8604, Torino, Italia. ELRA and ICCL. DOI: 10.63317/29ethqs8vpdg.

Menegasso, P. V. G. (2021). Perceptions of Brazilian college students about ethics in computer-based data collection and interventions for mental health. Available at:`https://repositorio.ufscar.br/handle/20.500.14289/15542`.

Mustafa, S., Mohammed, E., Salih, A., Palani, K., Albushra, M., Makkawi, S., and Mustafa, A. (2024). A Qualitative Exploration of Acceptance of a Conversational Chatbot as a Tool for Mental Health Support among University Students in Sudan. *International Journal of Medical Sciences and Nursing Research*, 4:16–23. DOI: 10.55349/ijmsnr.2024411623.

Oghenekaro, L. U. and Okoro, C. O. (2024). Artificial Intelligence-Based Chatbot for Student Mental Health Support. *Open Access Library Journal*, 11(5):1–14. Num-

ber: 5 Publisher: Scientific Research Publishing. DOI: 10.4236/oalib.1111511.

Peffers, K., Tuunanen, T., Rothenberger, M., and Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, 24:45–77. DOI: 10.2753/MIS0742-1222240302.

Pruitt, J. and Adlin, T. (2006). *The Persona Lifecycle: Keeping People in Mind Throughout Product Design*. Morgan Kaufmann, Amsterdam Boston. Available at:`https://learn.uni-med.net/pluginfile.php/17456/mod_resource/content/2/Constantine_Users-Roles-and-Personas_2006.pdf`.

Reche, S. A. M., Barbosa Neto, J. B., Caseli, H. d. M., and Frizzo, H. C. F. (2024). Significant changes in occupations among college students: A study carried out on a correlated online social network profile. *RECIIS – Revista Eletrônica de Comunicação, Informação & Inovação em Saúde*, 18(4):898–911. DOI: 10.29397/reciis.v18i4.4187.

Ribeiro, E. Y. A., Garcia, F. E., Saud, C. d. S. A., Caseli, H. d. M., Motti, V. G., Bleicher, T., Neto, J. B., Frizzo, H. C. F., Martini, L. C., Neris, L. d. O., Ara, A., Valejo, A. D. B., and Neris, V. P. d. A. (2026). A machine learning classification model for identifying college students with depression based on digital phenotyping. *Journal of the Brazilian Computer Society*, 32(1):766–782. DOI: 10.5753/jbcs.2026.5939.

Santos, I. S., Tavares, B. F., Munhoz, T. N., Almeida, L. S. P. d., Silva, N. T. B. d., Tams, B. D., Patella, A. M., and Matijasevich, A. (2013). Sensibilidade e especificidade do Patient Health Questionnaire-9 (PHQ-9) entre adultos da população geral. *Cadernos de Saúde Pública*, 29:1533–1543. Publisher: Escola Nacional de Saúde Pública Sergio Arouca, Fundação Oswaldo Cruz. DOI: 10.1590/0102-311X00144612.

Saud, C. S., Motti, V., Neris, L., Bleicher, T., and Neris, V. P. d. A. (2024). Identification of possible depressive profile through mobile sensor data in university students. In *Proceedings of the XXIII Brazilian Symposium on Human Factors in Computing Systems*, IHC '24, New York, NY, USA. Association for Computing Machinery. DOI: 10.1145/3702038.3702096.

Serrano-Ripoll, M. J., Zamanillo-Campos, R., Fiol-DeRoque, M. A., Castro, A., and Ricci-Cabello, I. (2022). Impact of Smartphone App-Based Psychological Interventions for Reducing Depressive Symptoms in People With Depression: Systematic Literature Review and Meta-analysis of Randomized Controlled Trials. *JMIR mHealth and uHealth*, 10(1):e29621. DOI: 10.2196/29621.

Silva, E. P. d. (2021). Determinação social da saúde e sofrimento psíquico na universidade: uma pesquisa com estudantes do Instituto de Ciências Matemáticas e de Computação (ICMC) do campus da USP de São Carlos. Available at:`https://repositorio.ufscar.br/handle/20.500.14289/14389`.

Souza, P. M., Alves, V., Pires, I., Bleicher, T., Garcia, F., da Silva, G., and Neris, V. (2025). A codesign approach for conversational user interfaces to support college students with depression. In *Anais do XXIV Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais*, pages 1093–1115, Porto Alegre, RS, Brasil. SBC. DOI: 10.5753/ihc.2025.10757.

Souza, P. M., Pires, I. C., Bleicher, T., Alves, V. C., Passador, R. V. P., Monteiro, M. S., Frizzo, H. C. F., Oliveira, R., Silva, A., Conejo, S. R., Caseli, H. M., and Almeida, V. P. A. (2026). Mental health communication through chatbots: A design experience report. *XXV Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais (IHC-2026)*. Available at:`https://ihc.sbc.org.br/2026/`.

Souza, P. M. d., Pires, I. d. C., Motti, V. G., Caseli, H. M., Neto, J. B., Martini, L. C., and de Almeida Neris, V. P. (2022). Design recommendations for chatbots to support people with depression. In *Proceedings of the 21st Brazilian Symposium on Human Factors in Computing Systems*, IHC '22, New York, NY, USA. Association for Computing Machinery. DOI: 10.1145/3554364.3559119.

Tagarot, S. (2022). All The Feels: Building a Chatbot to Assist Students Struggling with Depression. *Computer Science Honors Theses*. Available at:`https://digitalcommons.trinity.edu/compsci_honors/67`.

Whitaker, R. (2017). *Anatomia de uma Epidemia: pílulas mágicas, drogas psiquiátricas e o aumento assombroso da doença mental*. Fiocruz, Rio de Janeiro. DOI:10.7476/9786557082324.

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., and Wesslén, A. (2000). *Experimentation in Software Engineering*, volume 6 of *The Kluwer International Series in Software Engineering*. Springer US, Boston, MA. DOI:10.1007/978-1-4615-4625-2.