A Digital Assistant for the Vocational Guidance of Peruvian Students Using the LLaMA

Abel Cierto
Software Engineering Program, Universidad Peruana de Ciencias Aplicadas, Lima, Peru
u20201b561@upc.edu.pe

Branco Villegas
Software Engineering Program, Universidad Peruana de Ciencias Aplicadas, Lima, Peru
u20201c082@upc.edu.pe

Lenis Wong
Software Engineering Program, Universidad Peruana de Ciencias Aplicadas, Lima, Peru
pcsilewo@upc.edu.pe (corresponding author)

Received: 18 July 2025 | Revised: 8 September 2025 and 19 September 2025 | Accepted: 20 September 2025

 DOI: https://doi.org/10.48084/etasr.13482

ABSTRACT

Vocational guidance is very useful in helping students make informed academic and professional choices worldwide. However, in Peru, many young people do not have access to this type of specialized support, contributing to issues such as school dropout and poor career decision-making. To help address this gap, we developed a digital assistant software that provides vocational guidance in Spanish using the Large Language Model Meta Artificial Intelligence (LLaMA) 3.2 3B. The development process followed a four-phase methodology. First, Holland's test was selected as the psychometric tool. Second, we trained and optimized LLaMA 3.2 3B using specialized vocational guidance datasets, enabling the system to correctly interpret test responses. Third, we designed and implemented a mobile application that allows students to interact with a digital assistant via voice or text messages. Finally, we conducted a usability and effectiveness evaluation of the system with 40 students from a public high school. By comparing the assistant's recommendations to those provided by an expert psychologist, we obtained a concordance rate of 75.83%, while 80% of participating students were able to use the system without external assistance. These findings indicate that the proposed digital assistant has strong potential to serve as an effective and accessible tool for vocational guidance in Peru.

Keywords-english; vocational guidance; digital assistant; LLaMA; fine-tunning; Holland's test

I. INTRODUCTION

Vocational guidance is a determining factor in students' academic and professional development, supporting them as they navigate crucial decisions regarding their educational and career trajectories. In Peru, however, access to higher education options has declined from 36.6% in 2019 to 30.9% in 2022 [1], while 42% of students dropped out due to a lack of adequate guidance [2]. This issue disproportionately affects low-income youth, among whom 74.9% attribute their university dropout to inadequate career choices due to the absence of specialized support [3].

Concurrently, Artificial Intelligence (AI) has emerged as a powerful driver of innovative educational solutions. Various studies have applied AI techniques for vocational guidance, including chatbots on social platforms [4], web-based data analysis systems [5], conversational virtual robots [6], and Natural Language Processing (NLP)-based assistants [7]. However, most of these AI programs are limited to basic interfaces, without exploiting the advanced capabilities of current Large Language Models (LLMs), which enable natural, contextualized, and adaptive conversational interactions.

To address these gaps, recent research in automated vocational guidance has made adjustments in four main areas by integrating: (i) validated psychometric instruments, (ii) NLP techniques, (iii) LLM, and (iv) developing tools for digital assistants. Among existing psychometric models, Holland's Realistic, Investigative, Artistic, Social, Enterprising, and Conventional (RIASEC) test offers a robust foundation, identifying six vocational personality types proven effective at predicting job satisfaction. Additionally, this vocational values scale [9] enables career alignment with personal values, while cross-cultural studies have validated the model's universal applicability [10].

In the domain of NLP, the Transformer architecture has revolutionized automatic language understanding through its attention mechanism [11]. Models such as Bidirectional Encoder Representations from Transformers (BERT) achieve over 96% accuracy in text classification tasks [12], making them highly effective for interpreting user responses in conversational environments, while their optimization capabilities allow efficient implementation without compromising performance [13]. Additionally, LLMs like Large Language Model Meta AI (LLaMA) stand out for their balance of efficiency and accuracy. LLaMA 2-Chat, for instance, attains 91.3% accuracy in conversational tasks, surpassing BERT's 90.1% [14], while LLaMA-2 13B even outperforms GPT-3 175B despite its significantly smaller size [15]. Techniques such as Quantized Low Rank Adapters (QLoRA) further enhance these models by reducing memory consumption and improving Bilingual Evaluation Understudy (BLEU) and Recall-Oriented Understudy for Gisting Evaluation (ROUGE-L) metrics [16].

Building on these advancements, this study proposes the development of a digital vocational guidance assistant based on the LLaMA language model, specifically adapted to the Peruvian educational context. The system integrates Holland's test as a validated psychometric instrument and enables students to interact via voice or text to receive personalized career recommendations through a dedicated mobile application.

II. MATERIALS AND METHODS

A. Vocational Test Selection

Holland's test [17] was selected as the psychometric instrument in this study due to its proven reliability and widespread use in vocational guidance studies. This test identifies six RIASEC dimensions (realistic, investigative, artistic, social, enterprising, and conventional), rated on a 1-5 Likert scale with the final score reflecting the participant's vocational interest profile, indicating alignment with specific career fields based on the three highest-scoring dimensions.

B. LLM Selection

Three open-source models were evaluated: LLaMA 3.2 3B (M1), Gemma 2 2B IT (M2), and Phi-3.5-mini IT (M3). The selection of these models was based on balancing between capability and hardware requirements, without the need to provide expensive computing resources. The evaluation used five criteria weighted through a confrontation matrix [18], as shown in Table I. The criteria included:

- C1: Instruction comprehension (IFEval), which measures adherence to complex and multi-step instructions.
- C2: Response clarity (TLDR9), which evaluates concise and coherent summarization.
- C3: Long-conversation consistency (NIH/Multi-needle), which assesses contextual coherence across extended dialogues.
- C4: Multi-step guidance capability (BFCL V2), which evaluates the model's ability to decompose and explain multi-stage procedures.
- C5: Logical reasoning (HellaSwag), which measures performance on context-rich inference tasks requiring plausible continuation selection.

Each criterion's impact was normalized over the 13 positive relationships identified in the matrix.

TABLE I. LLM COMPARISON BY IDENTIFIED CRITERIA

Criteria	**C1**	**C2**	**C3**	**C4**	**C5**	**Total**	**Impact**
C1	-	1	1	1	1	4	31%
C2	0	-	1	1	1	3	24%
C3	0	0	-	1	1	2	15%
C4	0	0	1	-	1	2	15%
C5	0	0	1	1	-	2	15%
Total						13	100%

The evaluation scores (S, based on a 1-5 scale) and averages (A) based on the impact values of each criterion listed in Table I for each of the three LLMs are presented in Table II.

LLaMA 3.2 3B achieved the highest weighted score of 4.85, excelling particularly in instruction comprehension and response clarity; therefore, it was selected as the base model for the assistant developed.

TABLE II. WEIGHTED SCORING OF LLM MODELS

Criteria	**Impact**	**M1**		**M2**		**M3**	
		S	***A***	***S***	***A***	***S***	***A***
C1	31%	5	1.55	4	1.24	3	0.93
C2	24%	5	1.20	4	0.96	3	0.72
C3	15%	5	0.75	3	0.45	3	0.45
C4	15%	5	0.75	2	0.30	4	0.60
C5	15%	4	0.60	3	0.45	5	0.75
Total	100%		4.85		3.40		3.45

C. LLaMA 3.2 3B Model Fine-Tuning

A Kaggle dataset [20] containing 145,828 Holland test responses of diverse RIASEC profiles (2015-2018) was utilized. Table III summarizes the dimensions of RIASEC, the characteristics of each dimension, and the associated career paths.

From the original 93 variables, the 48 items corresponding to the standardized RIASEC questionnaire (eight per vocational dimension) were retained, while attributes such as age, race, religion, and others were excluded to prevent potential bias. Furthermore, the dataset responses were converted into 144 natural-language conversational sequences, each representing a complete Holland test interaction between user and assistant. A total of 120 were used for training and 24 for validation.

Fine-tuning was conducted using PyTorch and the Unsloth library on an NVIDIA A10G Graphics Processing Unit (GPU) for 60 training steps. Figure 1 illustrates the full fine-tuning workflow, including dataset preparation, conversation simulation, and iterative prompt refinement. The finalized model was deployed on Hugging Face.

TABLE III. HOLLAND PERSONALITY TYPES

Holland Test	Description
Realistic (R)	Prefers practical and hands-on activities. Suggested careers: engineering, construction, mechanics, and related fields.
Investigative (I)	Interested in research and analysis. Suggested careers: science, technology, medicine, and related fields.
Artistic (A)	Enjoys creative expression. Suggested careers: art, design, music, architecture, and related fields.
Social (S)	Likes helping and working with people. Suggested careers: teaching, social work, psychology, and related fields.
Enterprising (E)	Oriented toward leadership and persuasion. Suggested careers: sales, management, law, and related fields.
Conventional (C)	Prefers organized and detailed tasks. Suggested careers: accounting, banking, logistics, and related fields.

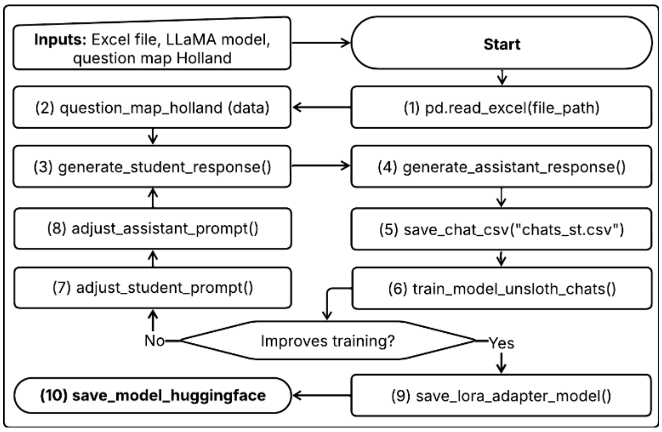


Fig. 1. Fine-tuning flowchart of synthetic Holland test data.

Three metrics were used to evaluate model training performance: training loss (1), which measures prediction error during training, validation loss (2), which evaluates generalization to unseen data, and perplexity (3), which assesses prediction confidence.

$$\text{Loss} = -\sum_{i=1}^{N} y_i \, log(p_i) \quad (1)$$

$$\text{Validation Loss} = -\frac{1}{M}\sum_{j=1}^{M}\sum_{t=1}^{T_j}\sum_{i=1}^{N} y_{j,t,i} \log\left(p_{j,t,i}\right) \quad (2)$$

$$\text{Perplexity} = e^{L} \quad (3)$$

where N is the total number of possible words in the model's vocabulary, y_i is the true value for word i (1 if correct, 0 otherwise), p_i is the probability assigned to the word i by the selected model, M is the total number of examples used for validation, T_j is the number of words (or tokens) in the example j, $y_{j,t,i}$ is the true value for word i, at position t of example j (1 if correct, 0 otherwise), and $p_{j,t,i}$ is the probability assigned by the model to the word i, at position t of example j.

D. Model Training

Figure 2 shows the training loss evolution of the LLaMA 3.2 3B model, where steps are used instead of epochs (15 steps per epoch) to visualize the rapid convergence pattern in detail. The initial training loss of 1.60 dropped to 1.25 within the first ten steps and stabilized at 0.75 around step 45, indicating rapid convergence without evidence of memorization. Figure 3 presents the validation loss and perplexity of the proposed model.

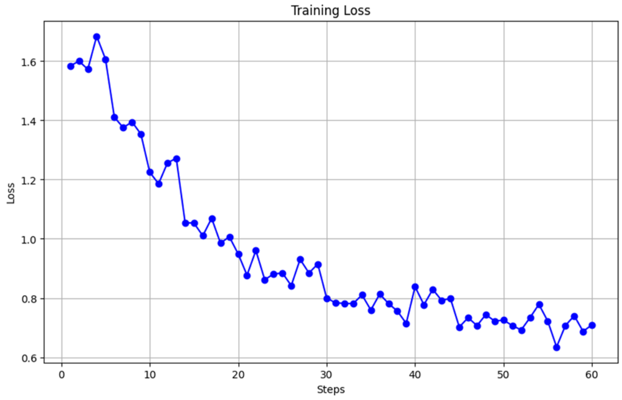
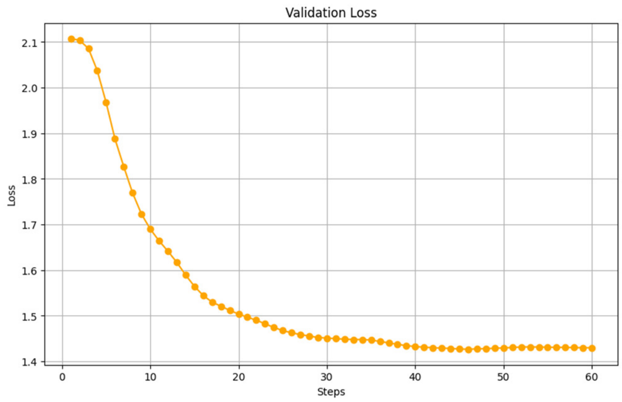
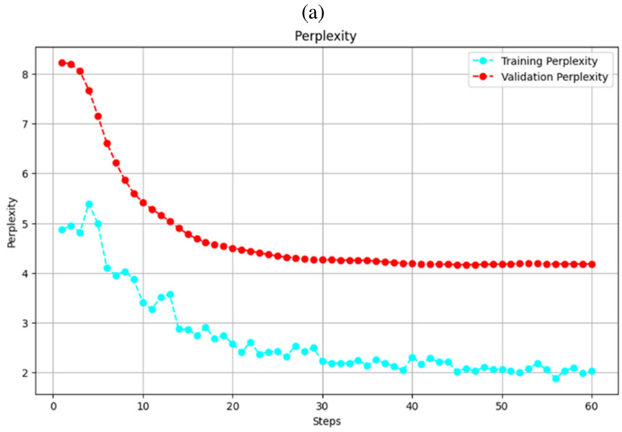


Fig. 2. Training loss over steps of LLaMA 3.2 3B.



(a)



(b)

Fig. 3. (a) Validation loss and (b) perplexity of LLaMA 3.2 3B.

Figure 3(a) illustrates that the initial validation loss of 2.11 plateaued to approximately 1.45 after 25 steps, demonstrating the model's ability to respond well to previously unseen questions. Additionally, Figure 3(b) shows that perplexity improved from 5.0 to 2.0 during training and from 8.2 to 4.2 during validation after 60 steps, confirming that the model successfully captured relevant patterns from the training data and can interpret vocational guidance results without overfitting.

E. Mobile Application Development

The mobile application was developed using Flutter to ensure full cross-platform compatibility, following a microservices-based architecture deployed on Amazon Web Services (AWS). The system consists of two primary components: (i) a front-end module that supports both voice and text input and includes an integrated Portable Document Format (PDF) report generator, and (ii) a back-end module composed of Flask-based microservices that manage application logic, with PostgreSQL used for user profile storage and MongoDB for conversation data management. Figure 4 illustrates the physical architecture of the system.

All microservices operate on Amazon EC2 instances within a Virtual Private Cloud (VPC). PostgreSQL is hosted on Amazon Relational Database Service (RDS), while MongoDB is deployed on a dedicated EC2 instance. For real-time inference, a Network Load Balancer (NLB) distributes WebSocket requests across two GPU-enabled EC2 instances running the Application Programming Interface (API) responsible for serving the fine-tuned LLaMA 3.2 3B model.

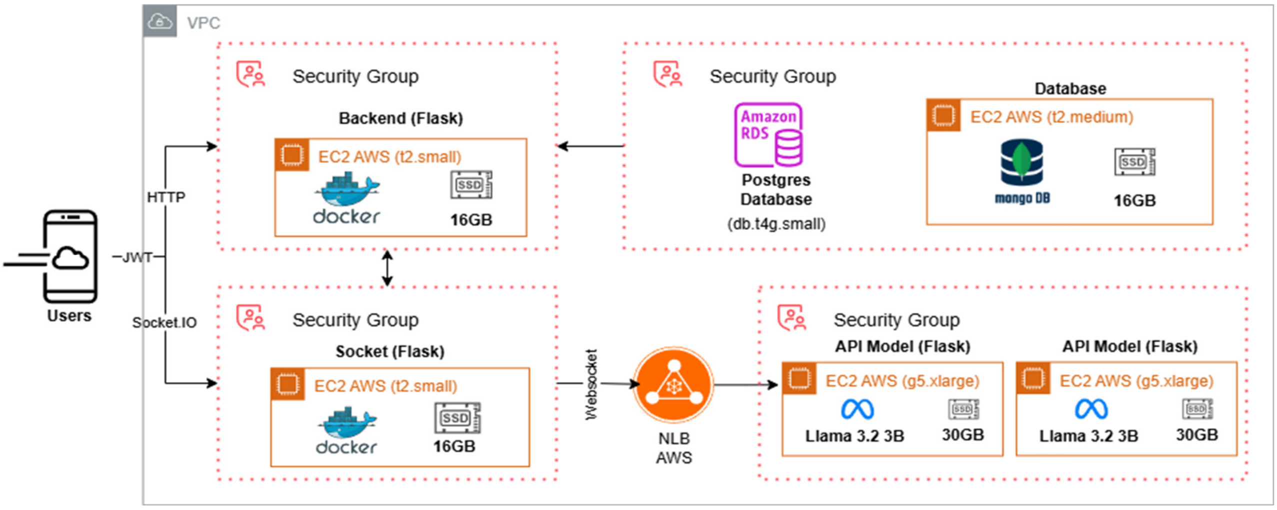


Fig. 4. Mobile application architecture.

III. EXPERIMENTATION

A. Experimental Design

To validate the AI-based vocational guidance system, a comparative experiment was conducted between the traditional method and the digital assistant method. The experiment involved 40 high school students from a public school in Ventanilla, Peru. All 40 students completed both formats; however, the second session was administered two weeks after the traditional method, following literature recommendations [21]. This approach reduced the likelihood that students would recall specific item-level responses, thus affecting their performance during the digital assistant method testing. The two experimental conditions implemented are summarized in Table IV.

TABLE I. EXPERIMENTAL DESIGN AND SCENARIOS

Experiment	Instrument	Participants	Metrics
Traditional method	Holland test in printed format	40 students 1 psychologist	ART, TST, PCI
Digital assistant	Holland test integrated into a mobile application	40 students 2 facilitators	ART, TST, PCI

Average Resolution Time (ART), Task Success Time (TST), Partial Concordance Index (PCI)

1) Experiment 1: Traditional Method

The first phase involved administering the Holland test in printed form under the supervision of psychologist Sung Jing Ferreira, in which each student had to complete the questionnaire manually. Before beginning, the psychologist provided detailed instructions on how to complete the test. During the session, the instances when students requested help and the time each took to complete the test were recorded. Upon completion, the psychologist scored the forms and delivered the results to participants. Figure 5 shows the complete sequence of activities, from participant selection to result delivery, allowing the psychologist to closely accompany each stage of the process.

2) Experiment 2: Digital Assistant

In the second phase, the same 40 students used the mobile application on pre-configured devices. Before starting, an explanation of the application usage was provided, and a brief demonstration of registration and test completion was conducted. During the session, students interacted with the assistant through chat or voice messages. The system automatically recorded responses, tracked completion times, and generated personalized recommendations. Two facilitators documented interventions for students who requested additional guidance. Figure 6 presents the complete sequence of activities.

3) Model Evaluation

Three main metrics were defined to evaluate both methods: i) the Average Resolution Time (ATR), which measures the average time needed to complete the vocational test, ii) the Task Success Time (TST), which measures the percentage of students who completed the test without external assistance, and iii) the Partial Concordance Index (PCI), which measures the rate of concordance between traditional method and digital assistant results. The formulas to calculate these metrics are:

$$\text{ART} = \frac{1}{n}\ \sum_{i=1}^{n} t_i \tag{4}$$

$$\text{TST} = \left(\frac{n_e}{n}\right) \times 100 \tag{5}$$

$$\text{PCI} = \left(\frac{1}{n \times 3}\ \sum_{i=1}^{n} m_i\right)\ \times 100 \tag{6}$$

where t_i is the time student i needed to complete the test, n is the total number of students, n_e is the number of students who correctly completed all items of the vocational test without requiring help, m_i is the number of matches (1, 2, or 3) obtained from the Holland test by applying both the traditional method and the software application for student i, and $n \times 3$ is the maximum possible number of matches (if all options matched).

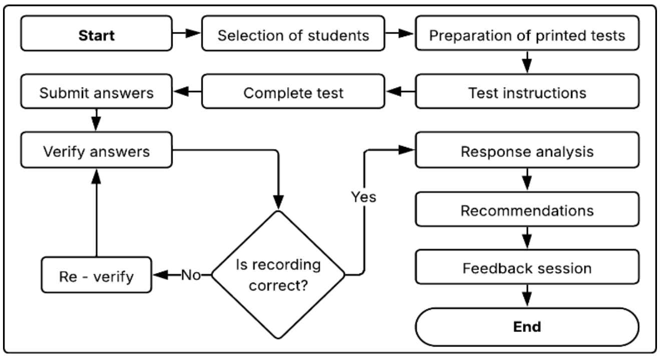


Fig. 5. Process flow of Holland's test with the traditional method.

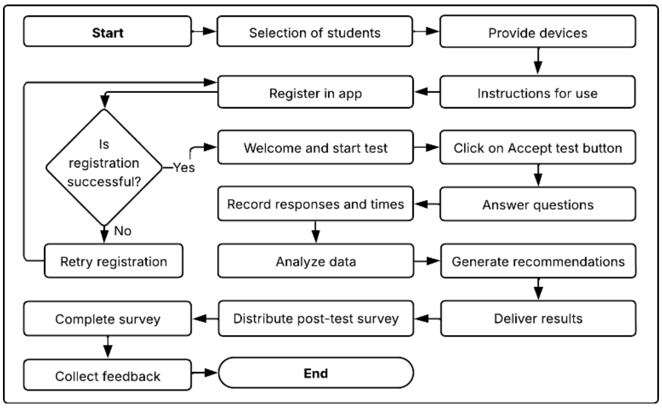


Fig. 6. Process flow of Holland's Test with the digital assistant.

A PCI score of above 70% is interpreted as good concordance, a PCI score between 50% and 69% is considered a moderate concordance, while a PCI score below 50% corresponds to the model significantly deviating from the expert judgment.

4) Satisfaction Evaluation

Lastly, a satisfaction survey based on the ISO/IEC 25010 standard [21] was administered. The survey contained 10 items across two dimensions: functionality and usability, rated on a 5-point Likert scale (1 = Very poor, 2 = Poor, 3 = Acceptable, 4 = Good, 5 = Very good), as summarized in Table V.

TABLE II. EXPERT SURVEY QUESTIONS

	Question
	Attributes: Functionality
Q1	Do you consider that the voice assistant improves fluency and comfort while taking the vocational test?
Q2	Does the text chat function allow a clear and effective interaction during the vocational test development?
Q3	Does the test result-generated PDF file present information in a comprehensible and useful way that allows you to make vocational decisions?
Q4	Do the functionalities offered by the application comprehensively cover all necessary steps for effective vocational guidance?
Q5	Do you consider that the recommendations generated by the system are consistent with your interests, abilities, or vocational profile?
	Attributes: Usability
Q6	Do you consider the mobile application to be easy to understand from its first use?
Q7	Did you find it simple to learn how to use all the functions of the digital assistant?
Q8	Did the application interface allow you to navigate clearly between questions?
Q9	Are you satisfied with the vocational recommendation generated by the system?
Q10	Would you recommend this application to a classmate?

IV. RESULTS AND DISCUSSION

A. *Experiment 1: Traditional Method*

Figure 7 illustrates the distribution of time (min) needed by the 40 students to complete the vocational test using the traditional method. The fastest completion time was 13.85 min (S27), and the slowest was 32.87 min (S18), while, as shown in Figure 8, 23 students needed assistance to complete the test.

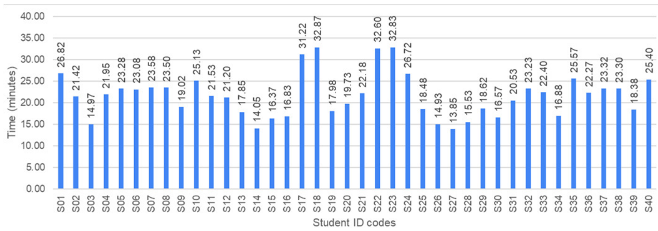


Fig. 7. Holland test completion time per student (Experiment 1).

Additionally, Figures 9 and 10 present the vocational profiles and consolidated RIASEC results, respectively. The top three dimensions per student are highlighted in Figure 9, showing both common and individual preferences. Overall, the most frequently appearing dimensions were Artistic, Social, and Enterprising, reflecting students' inclination toward creative activities, social welfare, and leadership.

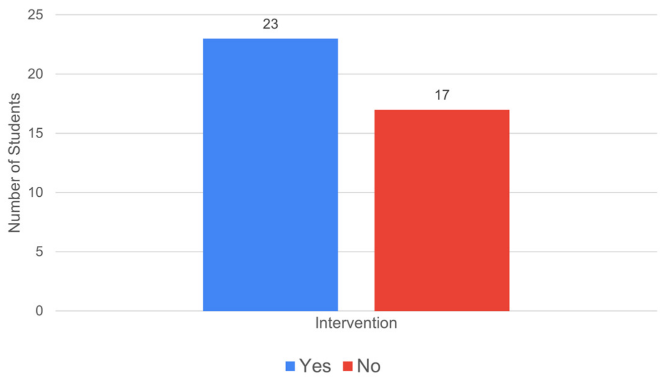


Fig. 8. Student intervention during the Holland test (Experiment 1).

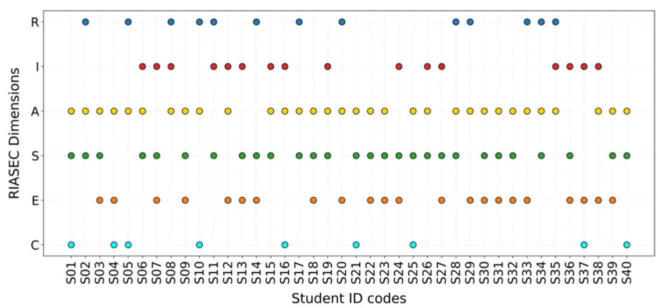


Fig. 9. RIASEC profile per student in the Holland test (Experiment 1).

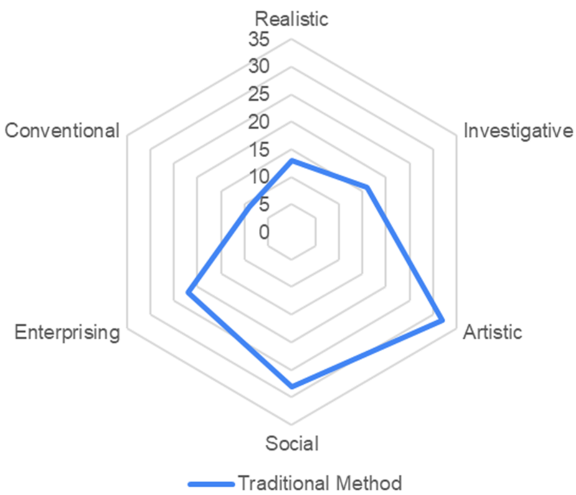


Fig. 10. Distribution of RIASEC dimension matches (Experiment 1).

B. *Experiment 2: Digital Assistant*

Figure 11 represents various examples of the uses and interfaces of the mobile application during the interaction of students with the assistant. Specifically, the Figure displays the use of the voice and text chat functionality, as well as the vocational report generated, along with the additional feature of an interpretive report with clear explanations of prominent and possible occupations in the Peruvian labor context.

After completing the vocational test with the digital assistant, each student completed a printed survey to evaluate the mobile application. Figure 12(a) shows a student reading the survey, Figure 12(b) shows a student marking their responses, and Figure 12(c) depicts the student completing the questionnaire.

Fig. 11. Digital assistant use and interfaces: (a) use of the voice functionality, (b) response to the the vocie message, (c) using of the text chat, (d) example of vocational report generated in PDF format, (e) example of interpretive report, and (f) option to share the report through digital media.

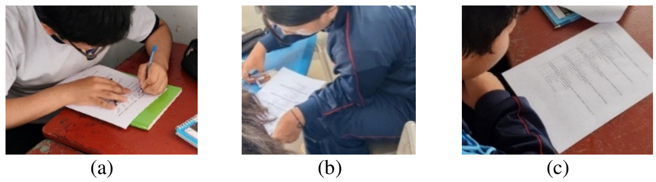


Fig. 12. Mobile application evaluation survey.

Figure 13 displays the distribution of completion times, with a minimum of 14.82 min (S04) and a maximum of 53.40 min (S20).

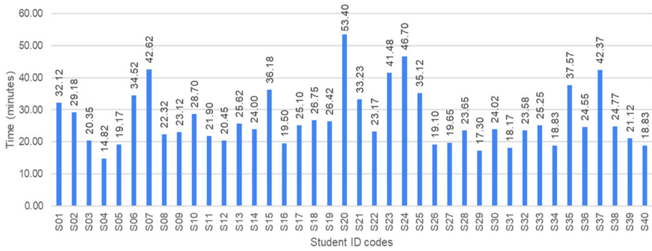


Fig. 13. Holland test completion time per student (Experiment 2).

Figure 14 shows that 32 students completed the test independently, while 8 requested assistance from the 2 facilitators. Figures 15 and 16 depict individual and consolidated RIASEC profiles, respectively, with the most frequent dimensions matching those from experiment 1: Artistic, Social, and Enterprising.

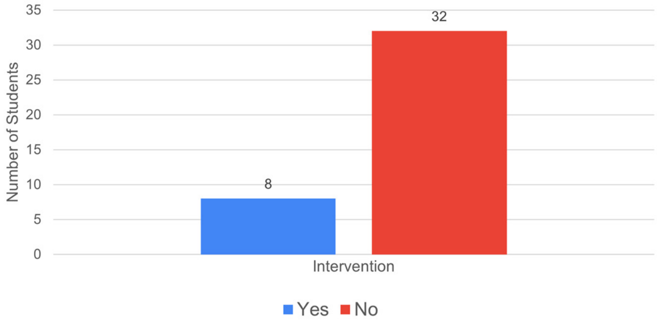


Fig. 14. Student intervention during the Holland test (Experiment 2).

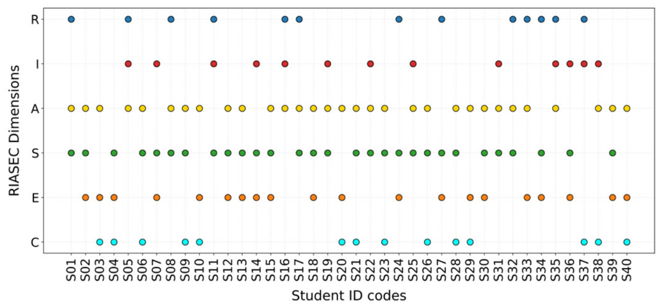


Fig. 15. RIASEC profile per student in the Holland test (Experiment 2).

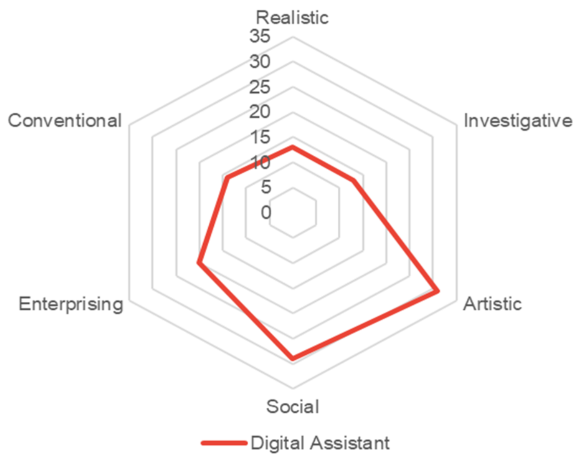


Fig. 16. Distribution of RIASEC dimension matches (Experiment 2).

C. Comparative Analysis

After collecting data from both experiments, the ART, TST, and PCI metrics were calculated using equations (4), (5), and (6), respectively. The traditional method obtained an ART of 21.65 min, while the digital assistant recorded 27.12 min. The increased ART with the digital assistant can be explained by students alternating between writing and speaking with the assistant, which seems to have taken more time than direct questions to the psychologist.

In contrast, the TST during the 2nd experiment was 80.00% while during the 1st experiment it was 42.50%, reflecting that almost double the students felt more comfortable using the application and were able to progress on their own without depending on external help.

Additionally, the PCI score was 75.83%, exceeding the "good concordance" threshold. Notably, this 24.17% difference from psychologist evaluations is an expected gap since digital interactions allow students to express interests through natural language rather than selecting predefined options. Thus, this difference reflects communication richness, not system error, as students with similar interests naturally articulate them differently in conversation.

Table VI and Figure 17 show the matches by RIASEC dimension for both experiments, with high concordance observed for Artistic, Social, and Enterprising, while Conventional showed an increase in matches with the digital assistant. Overall, the digital assistant demonstrated reliable performance, with minor differences reflecting richer, conversational expression rather than system error.

TABLE III. MATCHES BY DIMENSION

Dimensions	Traditional Method	Digital Assistant
Realistic	13	13
Investigative	16	13
Artistic	32	31
Social	28	29
Enterprising	22	20
Conventional	9	14

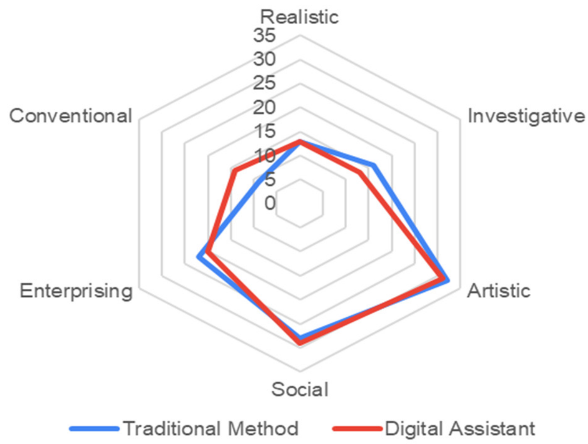


Fig. 17. Partial matches in RIASEC dimensions comparing both methods.

D. Student Satisfaction Evaluation

The results of the satisfaction survey conducted are summarized in Figure 18. As shown, the "Functionality" and "Usability" attributes obtained averages of 4.19/5 and 4.29/5, respectively, both falling within the "good" to "very good" range on the evaluation scale. These results reflect that features such as voice/text interaction, PDF report generation, and personalized career recommendations contributed to high functionality scores, while intuitive navigation and ease of learning enhanced usability. These results indicate that the digital assistant met both operational and user experience expectations.

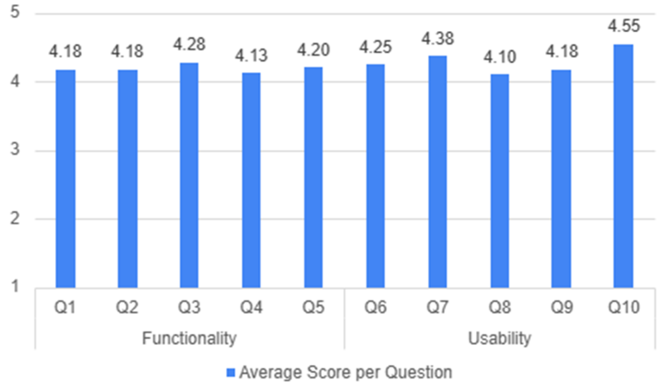


Fig. 18. Student digital assistant satisfaction questionnaire results.

V. CONCLUSION AND FUTURE WORK

This study developed and validated a digital assistant for vocational guidance tailored to Peruvian students. The mobile application, powered by the fine-tuned Large Language Model Meta Artificial Intelligence (LLaMA) 3.2 3B model, accurately interpreted Holland's test responses and generated personalized profiles through voice or text without requiring a specialist.

To experimentally validate the mobile application functionality and ease of use, a comparison experiment was conducted with 40 student participants completing the Holland's test with the traditional method and two weeks later with the digital applications. The results showed that the mobile application achieved 75.83% concordance with psychologist-generated profiles, and the most frequent vocational dimensions aligned across both methods, confirming the system's reliability. Additionally, student feedback highlighted the application's intuitive design, ease of use, and clarity of recommendations, while 80% of participants completed the test without assistance, underscoring the platform's accessibility and user-friendly interaction.

Future work will expand the training dataset to include a broader range of vocational profiles, enhancing the model's ability to capture diverse student interests.

ACKNOWLEDGMENT

The authors gratefully acknowledge the support provided through the UPC-Expost-2025-2 incentive program by the Research Department of Universidad Peruana de Ciencias Aplicadas.

REFERENCES

[1] Senaju National Youth Secretariat. "International Education Day: Only 30.9% of young Peruvians managed to transition to higher education." Senaju National Youth Secretariat. [Online]. Available: https://juventud.gob.pe/2024/01/dia-internacional-de-la-educacion-solo-el-30-9-de-jovenes-peruanos-logro-transitar-a-la-educacion-superior.

[2] Andina News Agency. "Four out of ten university students drop out due to lack of vocational guidance." Peruvian News Agency Andina. https://andina.pe/agencia/noticia-cuatro-cada-10-universitarios-deja-su-carrera-falta-orientacion-vocacional-927027.aspx.

[3] J. I. E. López, C. J. M. Valderrama, and A. V. Muñoz, "University dropout: an unresolved problem in Peru," *Hacedor - AIAPÆC*, vol. 7, no. 1, pp. 60–72, June 2023, https://doi.org/10.26495/rch.v7i1.2421.

[4] N. Suresh, N. Mukabe, V. Hashiyana, A. Limbo, and A. Hauwanga, "Career Counseling Chatbot on Facebook Messenger using AI," in *Proceedings of the International Conference on Data Science, Machine Learning and Artificial Intelligence*, Windhoek Namibia, Aug. 2021, pp. 65–73, https://doi.org/10.1145/3484824.3484875.

[5] A. F. Cruz, L. Orozco, and C. Gonzales, "Intelligent Web Platform for Vocational Guidance," in *2019 International Conference on Virtual Reality and Visualization (ICVRV)*, Hong Kong, China, Nov. 2019, pp. 205–207, https://doi.org/10.1109/ICVRV47840.2019.00049.

[6] A. C. R. Tade and I. G. García, "Virtual Robot in Vocational Guidance," *Revista Iberoamericana de Producción Académica y Gestión Educativa*, vol. 1, no. 2, Dec. 2014.

[7] G. D'Silva, M. Jani, V. Jadhav, A. Bhoir, and P. Amin, "Career Counselling Chatbot Using Cognitive Science and Artificial Intelligence," in *Advanced Computing Technologies and Applications*, H. Vasudevan, A. Michalas, N. Shekokar, and M. Narvekar, Eds. Singapore: Springer Singapore, 2020, pp. 1–9.

[8] J. S. Batista and S. M. G. Gondim, "Personality and Person-Work Environment Fit: A Study Based on the RIASEC Model," *International Journal of Environmental Research and Public Health*, vol. 20, no. 1, Dec. 2022, Art. no. 719, https://doi.org/10.3390/ijerph20010719.

[9] K. A. Atitsogbe and J.-L. Bernaud, "Vocational values scale: initial development and testing of the student form (VVS–S)," *International Journal for Educational and Vocational Guidance*, vol. 24, no. 2, pp. 501–524, Aug. 2024, https://doi.org/10.1007/s10775-022-09561-z.

[10] K. B. MacDonald, A. Benson, J. K. Sakaluk, and J. A. Schermer, "Pre-Occupation: A Meta-Analysis and Meta-Regression of Gender Differences in Adolescent Vocational Interests," *Journal of Career Assessment*, vol. 31, no. 4, pp. 715–738, Nov. 2023, https://doi.org/10.1177/10690727221148717.

[11] K. T. Chitty-Venkata, M. Emani, V. Vishwanath, and A. K. Somani, "Neural Architecture Search for Transformers: A Survey," *IEEE Access*, vol. 10, pp. 108374–108412, 2022, https://doi.org/10.1109/ACCESS.2022.3212767.

[12] S. R *et al.*, "Analyzing Sentiments Regarding ChatGPT Using Novel BERT: A Machine Learning Approach," *Information*, vol. 14, no. 9, Aug. 2023, Art. no. 474, https://doi.org/10.3390/info14090474.

[13] X. Gong, W. Ying, S. Zhong, and S. Gong, "Text Sentiment Analysis Based on Transformer and Augmentation," *Frontiers in Psychology*, vol. 13, May 2022, Art. no. 906061, https://doi.org/10.3389/fpsyg.2022.906061.

[14] J. Insuasti, F. Roa, and C. M. Zapata-Jaramillo, "Computers' Interpretations of Knowledge Representation Using Pre-Conceptual Schemas: An Approach Based on the BERT and Llama 2-Chat Models," *Big Data and Cognitive Computing*, vol. 7, no. 4, Dec. 2023, Art. no. 182, https://doi.org/10.3390/bdcc7040182.

[15] A. Kumar, R. Sharma, and P. Bedi, "Towards Optimal NLP Solutions: Analyzing GPT and LLaMA-2 Models Across Model Scale, Dataset Size, and Task Diversity," *Engineering, Technology & Applied Science Research*, vol. 14, no. 3, pp. 14219–14224, June 2024, https://doi.org/10.48084/etasr.7200.

[16] J. Goswami, K. K. Prajapati, A. Saha, and A. K. Saha, "Parameter-efficient fine-tuning large language model approach for hospital discharge paper summarization," *Applied Soft Computing*, vol. 157, May 2024, Art. no. 111531, https://doi.org/10.1016/j.asoc.2024.111531.

[17] P. Koolnaphadol, P. Inang, and J. Dudsdeemaytha "Online Career Intelligence Test: Self-Assessment for Students' Career and Abilities," *International Journal of Instruction*, vol. 15, no. 2, pp. 1075–1086, Apr. 2022, https://doi.org/10.29333/iji.2022.15259a.

[18] L. Huaroto, L. Wong, and V. Alvarado, "Mobile Application: For Anxiety and Cardiovascular Depression Monitoring Using a Smartwatch Based on Cognitive Behavioral Therapy," in *2022 32nd Conference of*

Open Innovations Association (FRUCT), Tampere, Finland, Nov. 2022, pp. 112–120, https://doi.org/10.23919/FRUCT56874.2022.9953848.

[19] Meta AI. "Llama 3.2: Revolutionizing edge AI and vision with open, customizable models." Meta AI Blog. [Online]. Available: https://ai.meta.com/blog/llama-3-2-connect-2024-vision-edge-mobile-devices.

[20] *Holland Code (RIASEC) Test Responses*. (2020), L. Greenwell. [Online]. Available: https://www.kaggle.com/datasets/lucasgreenwell/holland-code-riasec-test-responses.

[21] D. Vallejos, K. Villalobos, J. L. Castillo-Sequera, and L. Wong, "Intelligent System Based on Round Robin and Genetic Algorithm for Managing Nurse Schedules in Health Centres in Peru," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 20, no. 14, pp. 116–138, Nov. 2024, https://doi.org/10.3991/ijoe.v20i14.50571.