JAICOB: A Data Science Chatbot

DANIEL CARLANDER-REUTERFELT[ID], **ÁLVARO CARRERA**[ID], **CARLOS A. IGLESIAS**[ID], **ÓSCAR ARAQUE**[ID], **JUAN FERNANDO SÁNCHEZ RADA**[ID], **AND SERGIO MUÑOZ**[ID]

Grupo de Sistemas Inteligentes, Universidad Politécnica de Madrid, 28040 Madrid, Spain

Corresponding author: Álvaro Carrera (a.carrera@upm.es)

ABSTRACT The application of natural language to improve students' interaction with information systems is demonstrated to be beneficial. In particular, advances in cognitive computing enable a new way of interaction that accelerates insight from existing information sources, thereby contributing to the process of learning. This work aims at researching the application of cognitive computing in blended learning environments. We propose a modular cognitive agent architecture for pedagogical question answering, featuring social dialogue (small talk), improved for a specific knowledge domain. This system has been implemented as a personal agent to assist students in learning Data Science and Machine Learning techniques. Its implementation includes the training of machine learning models and natural language understanding algorithms in a human-like interface. The effectiveness of the system has been validated through an experiment.

INDEX TERMS Cognitive informatics, educational technology, human-computer interaction, machine learning, natural language processing.

I. INTRODUCTION

Cognitive computing has grown in the last few years, increasing the research and commercial interest in the topic [1]. Conversational agents have evolved from simple pattern-based programs into rather complex systems, including Natural Language Understanding and Machine Learning Techniques, which have allowed them to be more flexible in maintaining a conversation. Every day more businesses include chatbots as a way to interact with consumers to answer requests and FAQs. Natural Language Interface (NLI) increases user satisfaction and can help to find the information needed in a more comfortable way than other less sophisticated and time-consuming search interfaces [2].

Like humans, cognitive systems can use their knowledge to deduce data meaning based on context [3]. By having the advantage of computational power, a system like this can be even more successful than a human in this kind of task. Though they do not understand the meaning as humans do, the insights these systems provide can be beneficial. As they grow in time, it is expected that they gain abilities such as sensing and awareness [4].

Some of the benefits of the application of cognitive computing in the development of learning applications

The associate editor coordinating the review of this manuscript and approving it for publication was Tony Thomas.

are: (1) They can actively enhance students' performances [5], especially in computer science classes [6]; (2) studying cognitive computing behavior can lead to significant results in educational applications, especially in AI-related studies [6]; (3) *using a cognitive computing layer for digital interactions with students can enhance their performances and ease the teachers' job in managing classes and learning materials [6];* and (4) chatbots are excellent analysis tools, as students feel more inclined to send more messages to chatbots than real people [2].

Compared to other traditional e-learning training, chatbots generate a more positive response from the users [7]. Moreover, there are advantages in this type of learning, such as interaction, active learning, and sociability [8].

Despite these reasons, these technologies have not been widely adopted yet in education, and the ones that have are usually very rule-based and, therefore, less practical and functional. This article presents a modular architecture chatbot named Jaicob, adapted to the learning of Data Science techniques that aims to take advantage of all the benefits for education previously described. It is designed in a modular way that allows its adaptation to other areas of knowledge. It includes a flexible conversation workflow and is easy to maintain. This contribution has been evaluated with real users for a specific use case in a Data Science class.

The rest of the paper is organized as follows. Section II analyses related works about chatbots and the techniques applied in their development. Section III describes the different modules of the architecture and how they are interconnected. Section IV describes the evaluation process and results. Finally, Section V summarizes the learnings of this article with conclusions and defining future works.

II. RELATED WORK

A comprehensive systematic review of the use of chatbots in education is provided in this recent survey [9]. The authors identify several perspectives for analyzing current research following the theoretical model of Technology-Mediated Learning (TML) [10]: structure (input), learning process (process), and learning outcome (output). Regarding the input perspective, several dimensions have been identified [9]: student profile, educational settings, and chatbot technology. Learning outputs depend on individual student characteristics such as personality traits, technological skills as well as educational and social background [9].

Some research works claim that chatbot technology is so disruptive that *it will eliminate the need for websites and apps* [11]. Chatbots have been used in different educational settings, such as language learning [12], health-related coaching agents [13], chatbots designed to provide feedback to students [14], programming language learning [15], administrative support [16] or increase students' motivation [17]. These are examples if we don't take into account open-domain solutions such as Amazon or Google's [18], which aim to answer any kind of question, instead of a specific area of knowledge. While these types of chatbots are astoundingly ambitious and function with a near-human precision, sometimes, they come at a very high cost. Closed-domain question answering systems benefit from the ability to respond with more profound and specific knowledge [19], and also can achieve high quality at a lower complexity cost.

Design aspects of chatbots can influence the learning process. Flow-based chatbots, like [20]–[22], also called rule-based, can require an extensive database of questions and answers and need to have a clear flow of conversation that, if the user decides not to follow, can result in a bad experience. A study on chatbots of this type [23] concludes that they are quite limited to human direction and control. These can be built with frameworks like *Landbot.ai*,[1] or with simple coding abilities, but require great sophistication to work correctly. There is where its limitations lie. An extension to this kind of bots is button-based, like *HelloFreshus*,[2] that avoids the possibility of exiting the pre-planned flow. These can work well but can be very limited in scope and depth.

On the other hand, artificial intelligent based chatbots can better understand student intents. Even the most simple non-rule-based natural language understanding methods

[1]https://landbot.io
[2]https://chatfuel.com/bot/HelloFreshus

significantly outperform the most carefully crafted rule-based systems [24]. The reason is that they can achieve a more profound understanding of the intent and the requested information, thanks to machine learning techniques [25]. The most usual and effective approach [26], which is explained in greater detail in Section III, is based on intent-entity and Knowledge Base (KB).

Another aspect to take into account is if they are text or voice-based. Users tend to use longer sentences with voice-based chatbots and prefer reading expanded answers in a text-like manner. However, there is no significant difference in perceived effectiveness, learnability, and humanness between text-based and voice-based chatbots [27].

III. PROPOSED ARCHITECTURE FOR THE COGNITIVE BOT

The first step to design the proposed architecture was to identify the way students learn and the types of questions. Different types of requirements for different types of learning (inductive and deductive) [28] were identified due to the nature of students' curiosity, and the specifics of the topic. The following **pedagogical solutions** were identified:

1) A **definition** of a concept is a consequence of the usual teaching style, which is deductive, starting from the main concepts and developing towards the applications. It is part of the process of learning, but cannot be the whole process. In the Oliver model [29], definitions provide learning content.
2) As stated in [30], the learning of programming techniques can be enhanced by using **examples** of code using analogy [31] and induction. Also, learning is significantly facilitated by examples in initial coding attempts. Furthermore, surveys suggest that engineering students usually view themselves as inductive learners [28]. In the Oliver model [29], examples can provide learner support.
3) Lastly, the human need for **small-talk**, such as joking and asking for the weather, must be satisfied to provide a more significant communication source [2].

With that in mind, the architecture was designed, having identified the pedagogical needs of the student. There are several steps involved in the process and are explained below and represented in Figure 1.

A Knowledge Base (KB) was populated with pertinent information regarding the topic at hand, to satisfy the requests for definitions and examples. The Question Answering (QA) module is designed to extract meaning from all the data with the pedagogical requirements in mind to make sense of that information.

To analyze the students' question, we use the Speech Act Classifier. It selects the module where the question must be delegated. The way it works will be explained in greater detail in Section III-B. If small talk is detected, it is passed onto the Small Talk module or into the QA Module if a question regarding Data Science is detected.
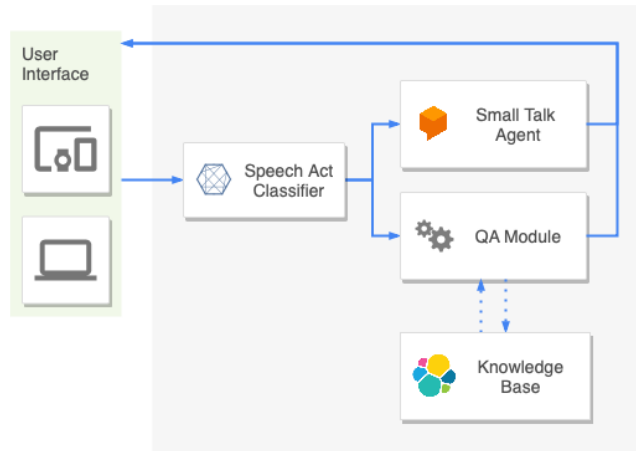
FIGURE 1. System architecture.

Afterward, the modules generate an answer to satisfy the student request. The answer is sent back to the student, and feedback is collected to evaluate and improve the model.

A. KNOWLEDGE BASE

The KB is the place where all the information used by the chatbot is stored. It has been populated from several online academic resources. Its selection was based on the previously identified student interactions with the chatbot. Glossaries and Frequently Asked Questions (FAQs) of the topic have been mined using web scraping techniques to provide concept definitions. Regarding searching code snippets, technical documentation has also been mined.

This approach benefits from providing students a curated list of pedagogical sources that are credible and useful. According to some studies [32], undergraduates tend to use Google for searching for information, and the usage of academic resources is low. Thus, our system increases the use of curated academic resources since the bot can enhance their familiarity [33].

The adaptation of the bot to other domains could be made replicating the same approach.

According to the categories previously described, the sites that fit the necessities of the definition answering are: (1) **Big Data glossary**[3] with a list of terms regarding big data. (2) **Machine Learning glossary**[4] with a complete glossary of machine learning and statistics terms and definitions.

The documentation sites used to populate the KB for answering with examples are *Pandas Documentation*. The use of the Python Pandas library is widely used when developing machine learning models. It is beneficial to have examples available for standard implementations of data handling. This documentation is structured with brief descriptions with code examples; and *Scikit-Learn Documentation*, being the library used widely for Machine Learning purposes, Scikit

---

[3]Big Data glossary: https://bigdata-madesimple.com/big-data-a-to-zz-a-glossary-of-big-data-terminology/
[4]Machine Learning glossary: https://www.analyticsvidhya.com/glossary-of-common-statistics-and-machine-learning-terms/

TABLE 1. Post classification examples.

Classification	Example	Next Module
Wh-Question	What is an elasticsearch database?	QA
Yes Answer	yes I do, lol	Small Talk
Yes/No Question	is overfitting a real problem	QA
Emotion	lol	Small Talk

examples of implementations is an obvious use case for the chatbot, and therefore an nearly important part of the KB.

For more complex questions, the use of FAQs solves the problem. The *Machine Learning Mastery* site[5] used for this purpose is structured as a list of questions with the answers associated. It was selected because of the rich and adequate answer for the project.

B. SPEECH ACT CLASSIFIER

The speech act classification task involves classifying a specific sentence into a set of predefined speech act categories. This classification is relevant to the project because it is indispensable to know the student's intention [34] and answer accordingly.

The dataset [35] used to train the classifier consists of 10567 posts from five different age-oriented chat rooms at an internet chat site. It is sanitized to protect user privacy. The posts were tagged using 15 post categories (*Accept, Bye, Clarify, Continuer, Emphasis, Greet, No Answer, Other, Reject, Statement, System, Wh-question, Yes Answer, Yes/No question, Emotion*). Examples of these classes are shown in Table 1.

Since this is a chatbot system that requires a fast response time, the preprocessing has been simplified to improve the model's time complexity while not sacrificing relevant performance. Each phrase is processed into machine-understandable information using a raw pipeline. The overall process is (1) Simple tokenization because n-grams did not present a significant improvement in accuracy, (2) Stemming, and (3) Feature extraction by vectorization.

By training and evaluating some of the most popular classification algorithms, the best one is selected based on the score achieved by a K-Fold. This process can be automated through a grid search that finds the best parameters optimally.

The scores following are calculated, saving a fourth of the dataset for testing afterward and using the rest to get these results.

The scores shown in Table 2 are obtained by performing a 5-Fold and calculating the mean of the scores. Support Vector Machines, such as the SVC algorithm, achieves higher performance.

Using as the training data 3/4 of the dataset and the rest as testing data, we obtain with the SVC a final accuracy score of 0.799.
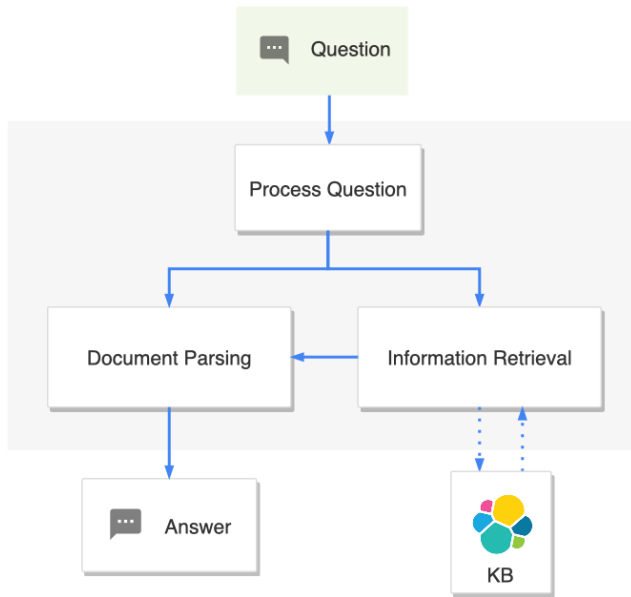
---

[5]Site for complex questions: https://machinelearningmastery.com/faq/

FIGURE 2. QA architecture.

TABLE 2. Evaluation scores.

Model	Accuracy	Precision	Recall	F1
Multinomial NB	0.685 (+/- 0.012)	0.75	0.71	0.66
Decision Tree	0.711 (+/- 0.005)	0.76	0.76	0.75
Random Forest	0.745 (+/- 0.005)	0.78	0.77	0.76
SVC	0.768 (+/- 0.007)	0.79	0.8	0.77

C. QUESTION ANSWERING MODULE

The Question Answering module comes into place when the user asks for a specific piece of information. These can range from doubt, a consultation, or documentation clarifications. It must be able to understand what the user is asking for to retrieve the information effectively.

Using natural language processing techniques, it answers the question in near real-time. This general-purpose model is enhanced to attend specific cases to the task at hand, such as code examples.

The general view of the architecture is defined in Figure 2. The modules involved in the process are the following.

The **Process Question** module extracts the relevant information and intention of the question. The output contains a type of question, a type of answer, and a vector with the relevant ideas.

The **Information Retrieval** This module receives the `question vector` and the `answer type` from the question processor as an input. The question vector is, in essence, a list of keywords ordered by importance. An ElasticSearch query is generated to retrieve relevant documents and pieces of information that match the keywords, using this valuable information.

The **Document Parsing** module receives and parses the retrieved information, so it matches the questions intented to generate an answer.
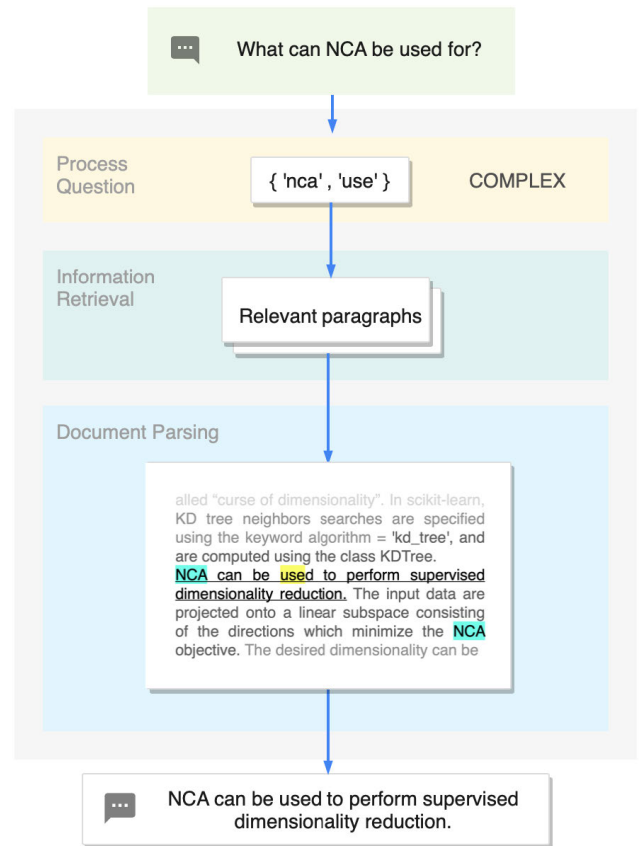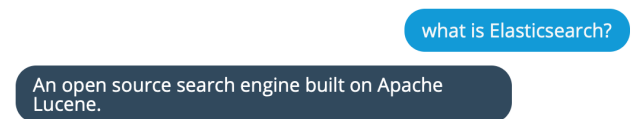

FIGURE 3. QA example.

FIGURE 4. Definition intent use case.

1) DEFINITION ANSWERING

When the `Answer Type` is of the *definition* type, the module searches in the Knowledge Base's Glossary index. It searches for a match with the terms in the index. When a match is found, the corresponding definition is sent as an answer. Common questions of this type are:

- What is a neural network?
- Can you give me a definition of overfitting?

This module is implemented as a DialogFlow agent, with an intent to recognize that the user wants a definition. The intent is trained with multiple training phrases that can be used to ask for a definition. It extracts a term as the slot. These slots are recognized thanks to an entity[6] defined as all the terms available in the Knowledge Base. An example can be seen in Figure 4.

---

[6]An entity is the definition of a type of slot. It can be defined as a list with all the possible terms for that slot.

how do you implement a svc in scikit?

```
>>> from sklearn import svm
>>> X = [[0, 0], [1, 1]]
>>> y = [0, 1]
>>> clf = svm.SVC(gamma='scale')
>>> clf.fit(X, y)
SVC(C=1.0, cache_size=200,
class_weight=None, coef0=0.0,
    decision_function_shape='ovr',
degree=3, gamma='scale', kernel='rbf',
    max_iter=-1, probability=False,
random_state=None, shrinking=True,
    tol=0.001, verbose=False)
```

FIGURE 5. Example intent use case.

2) EXAMPLE ANSWERING

When the `Answer Type` is of the *example* type, we need a more complex type of search. There is a search across the documentation text to match the keywords of the query. When a match is found, the corresponding code snippet is sent to respond with the appropriate format. Examples of these type of questions are:

- How is a dataframe defined in Pandas?
- How can I implement a k-fold using scikit?

This module is implemented as a DialogFlow agent with and intent trained to detect example queries. The slot, in this case, is more open, so there is no Entity defined. The example can be of any kind. The result can be seen in Figure 5.

D. SMALL TALK MODULE

According to [36], the users' satisfaction with a certain chatbot is influenced by various factors. By testing which of these factors were more influential, the results revealed that the bot's human-likeness was significantly correlated with the users' satisfaction.

It was stated [2] that people were *inclined to send more than twice as many messages to chatbots with a human-like interaction compared to other people, contrary to our expectations and disconfirming the notion that people feel less confident or comfortable communicating with chatbots.*

Including a module to handle small-talk improves the bot's human-likeness and makes it more fun and engaging. Instead of answering with the fallback answer, if the question is not about the topic, it triggers the small-talk module to simulate human interaction and cleverness. Some examples of the behavior that the bot can answer are collected in Figure 6.

1) IMPLEMENTATION

This module is implemented with Google's DialogFlow technology. There is a specially trained agent to provide the desired output. This agent can detect more than 100 different intents.

Among these intents are some of the provided with the default Small Talk module and some custom ones. The intents are defined to fit the purpose of this project. For example, when asked what it can do, it responds with directions to ask questions about Data Science.

Hi!

Good day!

What is your purpose?

I am here to answer your questions while trying not to take over humanity, so don't annoy me.

How old are you?

Age is just a number. You're only as old as you feel.

You are bad

I can improve with continuous feedback. My training is ongoing.

FIGURE 6. Small talk examples.

TABLE 3. Intent distribution.

Intent	Queries Rate
Example Intent	36.5%
Definition Intent	25.7%
Complex	21.1%
Small Talk	26.7%

E. GRAPHICAL USER INTERFACE

The bot needs an identifier to generate a more personal relationship [37]. Being a Cognitive bot and an intelligent one, it was decided to be called *Just an Artificial Intelligence Cognitive Bot* (JAICOB).

In contrast with Jaicob, a general-purpose bot would gain quality from a text-to-speech transformer, giving it a more human appearance. This feature is not the case of Jaicob because it is centered on answering documentation and programming related questions. The frequent use of acronyms and code examples in the answers would not make for a pleasant listening experience. Instead, the use of text is the best option in this case.

IV. EVALUATION

A. EVALUATION TECHNIQUES

The evaluation method for Jaicob chatbot is a Partial Least Squares (PLS) analysis. A detailed example [38] is followed to perform PLS methods. The tool being used is Smart PLS.[7]

The method is based on a questionnaire and requires the definition of latent variables to be evaluated, which are abstract variables that are connected to directly measurable variables. These variables' values are scored by the

[7]https://www.smartpls.com/documentation/pls-sem-compared-with-cb-sem

TABLE 4. Outer loadings.

Constructs and Items	Outer loading	CR	AVE
Social Handling (original scale, 4 items)		0.9132	0.7246
SH1: I felt a sense of human contact in the agent.	0.8608		
SH2: I felt a sense of personality in the agent.	0.8902		
SH3: I felt a sense of human warmth in the agent.	0.8327		
SH4: I felt a sense of sociability in the agent.	0.8196		
Behavioral intentions (original scale, 3 items)		0.9551	0.8764
R1: I would be willing to talk positively about JAICOB to my friends.	0.9319		
R2: I would be willing to recommend JAICOB to my friends.	0.9313		
R3: When visiting forums, I would be willing to say positive things about JAICOB to other people.	0.9451		
Satisfaction (original scale, 4 items)		0.9287	0.7655
S1: After visiting the website, I am satisfied.	0.8103		
S2: After visiting the website, I am happy.	0.8977		
S3: I found the visit on the website pleasant.	0.8614		
S4: Compared to other things I could have done, the time spent talking with JAICOB was enjoyable.	0.9261		
Utilitarian value (original scale, 3 items)		0.8891	0.7285
U1: Using JAICOB would make my life easier.	0.7812		
U2: I think of this Web site as an expert.	0.8815		
U3: If I want to solve doubts, the information JAICOB is able to provide would be what I would look for.	0.8933		
Answer Accuracy (original scale, 4 items)		0.9147	0.7288
A1: JAICOB responded effectively to my questions regarding definitions of terms.	0.8569		
A2: JAICOB was able to answer my doubts about code implementations correctly.	0.7866		
A3: The answer JAICOB gave me was what I was looking for.	0.9116		
A4: When he did not have the answer he told me so in a touchful way.	0.855		

responses of the questionnaires. These latent variables can also have relations, and these can be hypothesized, as shown in Section IV-C.

Being a conversational interface, the way to test it is with real users who answer the questionnaire after using the chatbot. The number of observations (number of questionnaires answered by users) should be at least ten times the number of relations between latent variables.

B. PARTICIPANTS
The experiment was done with 50 participants, all of them with technical backgrounds. All of them were unaware of the inner workings of Jaicob. They were asked to use the chatbot as a tool to answer any questions or doubts that may arise in understanding Data Science related topics or writing the corresponding code.

The median of the ages of the participants is 22 years. A 51% of them were studying a Telecommunication Engineering Grade and the rest a Master or superior studies.

About their technological background, 54% of the participants had developed and implemented some machine learning programs. The rest had some basic knowledge in the field.

C. EXPERIMENT DESIGN
As explained in Section III-D, small talk is an essential part of the architecture of the chatbot. Therefore, before making the measurements, it is taken into account.

Five latent variables were defined to evaluate the conversational bot:

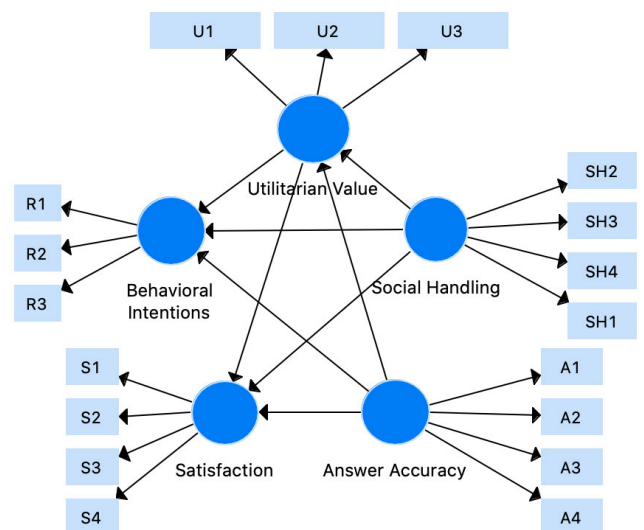- **Social Handling (SH)** refers to the personality and human-likeness of the bot.

FIGURE 7. Structural paths in the applied PLS model.

- **Behavioral intentions (BI)** refers to the recommendation of users to others to use the bot.
- **Satisfaction (SS)** refers to the feeling after using the bot.
- **Utilitarian value (UV)** refers to the value it provides to the task you are looking to complete.
- **Answer Accuracy (AA)** refers to the performance in the task it was programmed to do.

These latent variables are not independent, as represented in Figure 7. They present relations between latent variables, which are hypothesized and tested. Moreover, the relations between latent variables and questions, summarized in Table 4, are shown in the structural paths of the applied PLS Model.

Research [39] suggests that the quality of information on an e-commerce website has a positive impact on perceived value. Reference [40] suggests that accurate information can help users make better decisions, thus improving both utilitarian values. According to [41], the utilitarian values increases when the interaction with the process improves. These hypotheses are proposed:

H1. Perceptions of a better answer accuracy improve utilitarian value.

User satisfaction is influenced by the human-likeness of the chatbot [36]. Also, [2] state that people are more inclined to send messages to a chatbot that handles this type of small-talk well. A website's social dimension is another important antecedent of perceived value [42]. Research [43], [44] reveals that there is a direct link between perceived sociability and satisfaction.

H2. The social handling of the bot improves the overall satisfaction of the user.

H3. The social handling of the bot improves the utilitarian value a user perceives.

H4. Good social handling improves the behavioral intentions of users after using the bot.

Utilitarian value is central to user satisfaction and behavioral intentions. If the perceived value is low, the user probably switches to other sources [39].

H5. A higher perceived answer accuracy value increases positive behavioral intentions.

H6. A higher perceived utilitarian value increases positive behavioral intentions.

Perceived utilitarian value also enhances satisfaction [40]. Research [45] demonstrates that utilitarian value can improve the final user satisfaction:

H7. Perceived utilitarian value has a positive effect on user satisfaction.

H8. Perceived answer accuracy has a positive effect on user satisfaction.

D. RESULTS

The testers made an average of 15.86 queries per session. The intent that matched most of the queries was related to code example requests, which means that users used the bot for what it was intended. After that, there is the Definition intent and then the complex intent. Also, 26.7% of the queries resulted in small talk handling. The distribution can be seen in Table 3.

The results extracted from the PLS modeling, having used *SmartPLS 3.0* [46] meet the requirements, being the sample size ten times the largest number of structural paths directed at a particular construct in the structural model. There are three paths directed to Behavioral Intentions and Satisfaction in this model, so the minimum sample size should be 30, and the sample size is above this minimum.

To test the experiment's internal coherence, and therefore, reliability, we look at the outer loadings. These coefficients need to meet a threshold for every measure that points to the latent variables. All the measures met this reliability index,

TABLE 5. Discriminant validity.

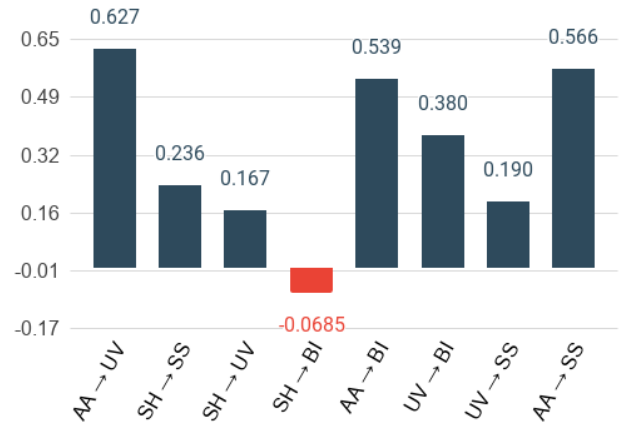|	AA	BI	SS	SH	UV
Answer Accuracy	**0.8537**				
Behavioral Intentions	0.7746	**0.9361**			
Satisfaction	0.8276	0.8009	**0.8749**		
Social Handling	0.5328	0.4089	0.6326	**0.8513**	
Utilitarian Value	0.7152	0.7315	0.7126	0.5002	**0.8535**

FIGURE 8. Path coefficients.

as shown in Table 4. The PLS analysis also provides us with the Composite Reliability of each latent variable. This index surpassed the minimum acceptable value of .70 in all variables, being all over .85.

The average variance extracted (AVE) for each variable must surpass a threshold of .50 [43], [47], and provide a square root that is much larger than the correlation of the specific construct with any other construct in the model. All the latent variables surpass a .70 AVE, as shown in Table 4. Table 5 shows that the square roots of the AVE (on the diagonal) are higher than any other values, in support of the discriminant validity of the measurement scales [38].

Then, discriminant validity is tested, which indicates the extent to which a given construct (variable) differs from other latent constructs. The validity of these variables requires that each measurement item correlates weakly with all constructs, except for which it is theoretically associated. The results in Table 5 support the validity of the measurement scales.

All the direct hypotheses received support, except for H4, as shown in Figure 8. From these results, we can extract some insights, such as the impact that Answering Accuracy has on all the other variables. Therefore, the quality of the system and its ability to respond effectively is what makes the difference for overall user Satisfaction, Utilitarian Value, and Behavioral Intentions (H1, H5, H8). Also, the perceived Utilitarian Value has a positive effect on Behavioral Intentions and Satisfaction (H6, H7). Surprisingly, Social Handling was not significant in positive behavioral intentions (H4), contrasting with the Utilitarian Value and Satisfaction (H2, H3).

V. CONCLUSION

The use of chatbots has become prevalent in the last years in shopping, customer support, general assistance, and, though less developed, education. The use of chatbots as a form of e-learning brings lots of opportunities.

This article identified the advantages of cognitive assistants in education and the corresponding challenges in implementation. A result is a tool for students with a comfortable and usable interface and a human experience. It can provide insights and solve doubts about Data Science. The main contribution is the adaptation of students' real pedagogic needs to the design of the architecture and being flexible in maintaining a conversation.

Teachers can also use it as a tool to identify gaps in the knowledge of their students. They can also outsource to Jaicob the answering of all the questions. The pedagogue is also an excellent asset to select the most valuable sources of information from which Jaicob feeds from, thus providing a curated source of information instead of a regular Google Search.

The project was evaluated with a sample of students, achieving very favorable results in usability and originality. The experiment confirms that the system can answer effectively, that the answer accuracy affects the satisfaction, utilitarian value, and behavioral intentions of the user, and that proper social handling is significant in satisfaction and utilitarian value but not in behavioral intentions.

As these technologies evolve, more and more people will study these subjects. Therefore, the future impact of the project is promising, and the affected groups will increase. In future work, to achieve a broader reach in the areas of knowledge, it is straightforward to place additional information in the Knowledge Base and the corresponding Dialogflow intents.

REFERENCES

[1] A. Soffer, D. Konopnicki, and H. Roitman, "When watson went to work: Leveraging cognitive computing in the real world," in *Proc. 39th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr. (SIGIR)*, 2016, pp. 455–456.

[2] J. Hill, W. Randolph Ford, and I. G. Farreras, "Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations," *Comput. Hum. Behav.*, vol. 49, p. 245–250, Aug. 2015.

[3] Y. Chen, J. E. Argentinis, and G. Weber, "IBM Watson: How cognitive computing can be applied to big data challenges in life sciences research," *Clin. Therapeutics*, vol. 38, no. 4, p. 688–701, Apr. 2016.

[4] R. High, "The era of cognitive systems: An inside look at IBM Watson and how it works," IBM Corp., Redbooks, North Castle, NY, USA, Tech. Rep., 2012.

[5] M. Coccoli, A. Guercio, P. Maresca, and L. Stanganelli, "Smarter universities: A vision for the fast changing digital era," *J. Vis. Lang. Comput.*, vol. 25, no. 6, pp. 1003–1011, Dec. 2014.

[6] M. Coccoli, P. Maresca, and L. Stanganelli, "Cognitive computing in education," *Big Data*, vol. 12, no. 2, p. 15, 2016.

[7] S. Kowalski, K. Pavlovska, and M. Goldstein, "Two case studies in using chatbots for security training," in *Proc. IFIP World Conf. Inf. Secur. Educ.* Cham, Switzerland: Springer, 2009, pp. 265–272.

[8] P. Bii, "Chatbot technology: A possible means of unlocking student potential to learn how to learn," *Educ. Res.*, vol. 4, no. 2, pp. 218–221, 2013.

[9] R. Winkler and M. Söllner, "Unleashing the potential of chatbots in education: A state-of-the-art analysis," in *Proc. Acad. Manage. Annu. Meeting*, 2018, pp. 1–40. [Online]. Available: https://www.alexandria.unisg.ch/254848/

[10] R. P. Bostrom, "Technology-mediated learning: A comprehensive theoretical model," *J. Assoc. Inf. Syst.*, vol. 10, no. 9, pp. 686–714, Sep. 2009.

[11] R. Dale, "The return of the chatbots," *Natural Lang. Eng.*, vol. 22, no. 5, pp. 811–817, Sep. 2016.

[12] X. L. Pham, T. Pham, Q. M. Nguyen, T. H. Nguyen, and T. T. H. Cao, "Chatbot as an intelligent personal assistant for mobile language learning," in *Proc. 2nd Int. Conf. Educ. E-Learn. (ICEEL)*, 2018, pp. 16–21.

[13] J. Beaudry, A. Consigli, C. Clark, and K. J. Robinson, "Getting ready for adult healthcare: Designing a chatbot to coach adolescents with special health needs through the transitions of care," *J. Pediatric Nursing*, vol. 49, pp. 85–91, Nov. 2019.

[14] M. Awais Hassan, U. Habiba, H. Khalid, M. Shoaib, and S. Arshad, "An adaptive feedback system to improve student performance based on collaborative behavior," *IEEE Access*, vol. 7, pp. 107171–107178, 2019.

[15] M. Coronado, C. A. Iglesias, Á. Carrera, and A. Mardomingo, "A cognitive assistant for learning java featuring social dialogue," *Int. J. Hum.-Comput. Stud.*, vol. 117, pp. 55–67, Sep. 2018.

[16] H. T. Hien, P.-N. Cuong, L. N. H. Nam, H. L. T. K. Nhung, and L. D. Thang, "Intelligent assistants in higher-education environments: The FIT-EBot, a chatbot for administrative and learning support," in *Proc. 9th Int. Symp. Inf. Commun. Technol. (SoICT)*, New York, NY, USA, 2018, pp. 69–76, doi: 10.1145/3287921.3287937.

[17] L. Benotti, M. C. Martínez, and F. Schapachnik, "Engaging high school students using chatbots," in *Proc. Conf. Innov. Technol. Comput. Sci. Educ. (ITiCSE)*, 2014, pp. 63–68.

[18] D. Adiwardana, M.-T. Luong, D. R. So, J. Hall, N. Fiedel, R. Thoppilan, Z. Yang, A. Kulshreshtha, G. Nemade, Y. Lu, and Q. V. Le, "Towards a human-like open-domain chatbot," 2020, *arXiv:2001.09977*. [Online]. Available: http://arxiv.org/abs/2001.09977

[19] H. Doan-Nguyen and L. Kosseim, "Improving the precision of a closed-domain question-answering system with semantic information," in *Proc. Coupling Approaches, Coupling Media Coupling Lang. Inf. Retr.*, Le Centre De Hautes Etudes Internationales D'Informatique Documentaire, 2004, pp. 850–859.

[20] B. Heller, M. Proctor, D. Mah, L. Jewell, and B. Cheung, "Freudbot: An investigation of chatbot technology in distance education," in *Proc. EdMedia+ Innovate Learn. Assoc. Adv. Comput. Educ. (AACE)*, 2005, pp. 3913–3918.

[21] S. Hussain and G. Athula, "Extending a conventional chatbot knowledge base to external knowledge source and introducing user based sessions for diabetes education," in *Proc. 32nd Int. Conf. Adv. Inf. Netw. Appl. Workshops (WAINA)*, May 2018, pp. 698–703.

[22] J. Jia and W. Chen, "Motivate the learners to practice English through playing with Chatbot CSIEC," in *Proc. Int. Conf. Technol. E-Learn. Digit. Entertainment*. Cham, Switzerland: Springer, 2008, pp. 180–191.

[23] S. Roos, "Chatbots in education: A passing trend or a valuable pedagogical tool?" M.S. thesis, Dept. Inform. Media, Social Sci., Disciplinary Domain Humanities Social Sci., Uppsala Univ., Uppsala, Sweden, 2018, pp. 1–58.

[24] D. Suendermann, K. Evanini, J. Liscombe, P. Hunter, K. Dayanidhi, and R. Pieraccini, "From rule-based to statistical grammars: Continuous improvement of large-scale spoken dialog systems," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Apr. 2009, pp. 4713–4716.

[25] S. Paliwal, V. Bharti, and A. K. Mishra, "Ai chatbots: Transforming the digital world," in *Recent Trends and Advances in Artificial Intelligence and Internet of Things*. Cham, Switzerland: Springer, 2020, pp. 455–482.

[26] H. T. Hien, P.-N. Cuong, L. N. H. Nam, H. L. T. K. Nhung, and L. D. Thang, "Intelligent assistants in higher-education environments: The FIT-EBot, a chatbot for administrative and learning support," in *Proc. 9th Int. Symp. Inf. Commun. Technol. (SoICT)*, 2018, pp. 69–76.

[27] O. Mangkang and J. Y. Yun, "A comparative study of self-diagnosis user interfaces for depression: Focusing on usability of paper-based, text-based and voice-based conversational interface," in *Proc. Korean HCI Conf.*, 2019, pp. 262–267.

[28] R. M. Felder and L. K. Silverman, "Learning and teaching styles in engineering education," *Eng. Educ.*, vol. 78, no. 7, pp. 674–681, 1988.

[29] R. Oliver, "Exploring strategies for online teaching and learning," *Distance Educ.*, vol. 20, no. 2, pp. 240–254, Jan. 1999.

[30] P. Brusilovsky, "Webex: Learning from examples in a programming course," in *Proc. WebNet*, vol. 1, 2001, pp. 124–129.

[31] P. L. Pirolli and J. R. Anderson, "The role of learning from examples in the acquisition of recursive programming skills," *Can. J. Psychol./Revue Canadienne de Psychologie*, vol. 39, no. 2, p. 240, 1985.

[32] J. R. Griffiths and P. Brophy, "Student searching behavior and the Web: Use of academic resources and Google," *Trends*, vol. 53, no. 4, p. 539, 2005.

[33] J. Y. Lee, W. Paik, and S. Joo, "Information resource selection of under-graduate students in academic search tasks," *Inf. Res., Int. Electron. J.*, vol. 17, no. 1, p. n1, 2012.

[34] B. Samei, H. Li, F. Keshtkar, V. Rus, and A. C. Graesser, "Context-based speech act classification in intelligent tutoring systems," in *Proc. Int. Conf. Intell. Tutoring Syst.* Cham, Switzerland: Springer, 2014, pp. 236–241.

[35] E. N. Forsythand and C. H. Martell, "Lexical and discourse analysis of online chat dialog," in *Proc. Int. Conf. Semantic Comput. (ICSC)*, Sep. 2007, p. 19.

[36] M. Xuetao, F. Bouchet, and J.-P. Sansonnet, "Impact of agent's answers variability on its believability and human-likeness and consequent chatbot improvements," in *Proc. AISB*, 2009, pp. 31–36.

[37] J. Finch, "Naming names: Kinship, individuality and personal names," *Sociology*, vol. 42, no. 4, pp. 709–725, Aug. 2008.

[38] M. S. Ben Mimoun and I. Poncin, "A valued agent: How ECAs affect website customers' satisfaction and behaviors," *J. Retailing Consum. Services*, vol. 26, pp. 70–82, Sep. 2015.

[39] M. Magni, M. Susan Taylor, and V. Venkatesh, "'To play or not to play': A cross-temporal investigation using hedonic and instrumental perspectives to explain user intentions to explore a technology," *Int. J. Hum.-Comput. Stud.*, vol. 68, no. 9, pp. 572–588, Sep. 2010.

[40] C. Kim, R. D. Galliers, N. Shin, J.-H. Ryoo, and J. Kim, "Factors influencing Internet shopping value and customer repurchase intention," *Electron. Commerce Res. Appl.*, vol. 11, no. 4, pp. 374–387, Jul. 2012.

[41] T. Ahn, S. Ryu, and I. Han, "The impact of Web quality and playfulness on user acceptance of online retailing," *Inf. Manage.*, vol. 44, no. 3, pp. 263–275, Apr. 2007.

[42] J. Lawler, P. Vandepeutte, and A. Joseph, "An exploratory study of apparel dress model technology on European Web sites," *J. Inf., Inf. Technol., Organizations*, vol. 2, no. 244, pp. 31–46, Nov. 2007.

[43] S. S. Al-Gahtani, G. S. Hubona, and J. Wang, "Information technology (IT) in Saudi Arabia: Culture and the acceptance and use of IT," *Inf. Manage.*, vol. 44, no. 8, pp. 681–691, Dec. 2007.

[44] O. Turel, A. Serenko, and N. Bontis, "User acceptance of hedonic digital artifacts: A theory of consumption values perspective," *Inf. Manage.*, vol. 47, no. 1, pp. 53–59, Jan. 2010.

[45] T. W. Traphagan, Y.-H.-V. Chiang, H. M. Chang, B. Wattanawaha, H. Lee, M. C. Mayrath, J. Woo, H.-J. Yoon, M. J. Jee, and P. E. Resta, "Cognitive, social and teaching presence in a virtual world and a text chat," *Comput. Educ.*, vol. 55, no. 3, pp. 923–936, Nov. 2010.

[46] C. Ringle, S. Wende, and J.-M. Becker, "Smartpls 3," SmartPLS GmbH, Bönningstedt, Germany, Tech. Rep., Jan. 2015.

[47] C. Fornell and D. F. Larcker, "Structural equation models with unobservable variables and measurement error: Algebra and statistics," *J. Marketing Res.*, vol. 18, no. 3, pp. 382–388, 1981, doi: 10.1177/002224378101800313.

DANIEL CARLANDER-REUTERFELT is currently pursuing the master's degree in telecommunications engineering with the Universidad Politécnica de Madrid (UPM). He was awarded an honorary mention for the bachelor's thesis (Development of a Cognitive Bot for Data Science Tutoring based on a Big Data Natural Language Analytics Platform). He has been a part of the Intelligent Systems Group, UPM, since 2019. His research interests include intelligent agents, natural language processing and understanding, and time series prediction.

ÁLVARO CARRERA received the Telecommunication Engineering degree and the Ph.D. degree in telecommunications engineering from the Universidad Politécnica de Madrid (UPM), Spain, in 2010 and 2016, respectively. He is currently an Assistant Professor with the School of Telecommunications Engineering, UPM. His research interests include applying intelligent agents in telecommunication networks for critical tasks, such as fault diagnosis or intrusion detection in real-time, and in the application of simulation techniques for complex systems in dynamic environments, such as telecommunication networks and social networks.

CARLOS A. IGLESIAS received the Telecommunications Engineering degree and the Ph.D. degree in telecommunications from the Universidad Politécnica de Madrid (UPM), Spain, in 1993 and 1998, respectively. He is currently an Associate Professor with the Telecommunications Engineering School, UPM. He has been leading the Intelligent Systems Group, UPM, since 2014. He has been the Principal Investigator on numerous research grants and contracts in the field of advanced social and the IoT systems, funded by the regional, national, and European bodies. His main research interests include social computing, multiagent systems, information retrieval, sentiment and emotion analysis, linked data, and web engineering.

ÓSCAR ARAQUE received the graduate and master's degrees in telecommunication engineering from the Technical University of Madrid (Universidad Politécnica de Madrid), Spain, in 2014 and 2016, respectively, where he is currently pursuing the Ph.D. degree. He is currently a Teaching Assistant with the Universidad Politécnica de Madrid. His research interest includes the application of machine learning techniques for natural language processing. The main topic of his thesis is the introduction of specific domain knowledge into machine learning systems in order to enhance sentiment and emotion analysis techniques.

JUAN FERNANDO SÁNCHEZ RADA received the Ph.D. degree from the Universidad Politécnica de Madrid (UPM), Spain, in 2020. He is currently a Researcher with the Intelligent Systems Group, Universidad Politécnica de Madrid. His research interests include natural language processing (sentiment and emotion analysis), social network analysis (social context and graph embedding), the web and distributed systems (interoperability and federation), and semantic technologies (linked data, ontologies, and knowledge graphs).
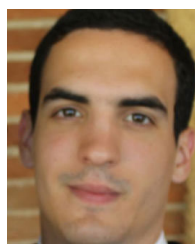
SERGIO MUÑOZ received the graduate and master's degrees in telecommunication engineering from the Technical University of Madrid (Universidad Politécnica de Madrid), Spain, in 2016 and 2017, respectively, where he is currently pursuing the Ph.D. degree. He is currently a Teaching Assistant with the Technical University of Madrid. His research interests include ambient intelligence and agent-based simulation. The main topic of his thesis is the adaptation of smart environments to users' emotions, to enhance well-being and performance.

● ● ●