

## C951 – Introduction to Artificial Intelligence

# C.A.L.I.C.O. – The Career Artificial Intelligence ChatbOt

Maxwell Brown

9-24-2020

Version 1.0



**WESTERN GOVERNORS UNIVERSITY®**

## CONTENTS

A.	Part One.....	3
A.1.	Chatbot Description .....	3
A.1.1.	Information System.....	3
A.1.2.	Recommendation System .....	3
A.2.	Previous Literature.....	4
A.3.	Career Recommendations .....	5
B.	Training.....	5
C.	Optimization .....	6
D.	Installation.....	7
E.	Bot Effectiveness .....	7
F.	Challenges.....	8
G.	Development Environment Strengths and Weaknesses .....	8
H.	Sources .....	8



## A. PART ONE

The enrollment of computer science students at WGU University has been steadily increasing over the past several years. Because of this, the number of computer science students seeking guidance in selecting a career path has exceeded the bandwidth of available career advisors within the Student Advisory Department at WGU University. To mitigate the demand for career advisors at WGU University, the Student Advisory Department has decided to develop CALICO (Career Artificial Intelligence ChatBot) to interact with computer science students at WGU University and provide initial career recommendations. Students can then bring the information and recommendations they receive from CALICO to a meeting with a career advisor with expertise on the career recommendation provided by CALICO. This will allow for the Student Advisory Department to have more targeted and productive conversations with its students.

### A.1. CHATBOT DESCRIPTION

CALICO has two primary functions. The first is to provide computer science students with information about the top 5 career paths taken by computer science students following graduation. The second is to provide a recommendation for which career path CALICO believes is most suited to the student in question. To generate this recommendation, CALICO asks a series of probing questions to gain insight into the interests and qualifications of the student. Based upon this information, CALICO will provide its recommendation as to which career path it believes it is the best fit for the student as well as the information for the most appropriate career advisor to contact for further information. CALICO will also display a help menu for students to choose an option from if unrecognized information is entered into the system.

#### A.1.1. Information System

The information system built into CALICO is designed to provide students with information about the current top five career paths taken by computer science students following their graduation from WGU University. CALICO will also link out to additional resources so that students can obtain additional information about these careers and what they entail. This information, coupled with the recommendation system outlined below, arms the student with information with the goal of facilitating a much more targeted and productive conversation between the student and the recommended career advisor.

#### A.1.2. Recommendation System

To both facilitate the entry of structured responses and enhance the speed with which students can get their recommendation, the recommendation system built into CALICO functions primarily as a structured questionnaire. CALICO poses questions along with a list of preset responses in the form of buttons. Students can either select a preset response by pressing the corresponding button or can type a response into the chat window. Upon completion of the questionnaire, CALICO will use the entered information to display its recommendation to the student. In future versions, CALICO will also provide a link to the most appropriate career advisor's appointment calendar so that the student can have a targeted discussion with a career advisor about the recommendation (should they choose to).



## A.2. PREVIOUS LITERATURE

The inspiration for linking to external resources from the information system built into CALICO came from a publication by Ranoliya et al. (2017) which describes the development of an AIML-based interactive chatbot for the purpose of answering frequently asked questions about Manipal University in India. The authors describe a system in which users interact with their chatbot and are either provided answers to their queries or are directed to external resources based on areas of expressed interest.

The inspiration for implementing the CALICO recommendation system in the form of a questionnaire came from a publication by Reshmi et al. (Reshmi, 2016) which discusses the impact of the quality of a user's query on the response received from a chatbot. The authors postulate that users generally do not provide enough information in their query for a chatbot to respond with a reasonable answer. In such cases, the chatbot system must probe the user for additional information before it can appropriately respond to a user's query. To address this, the authors integrated a complex knowledge base engine into the existing ALICE AIML-based chatbot which, given an arbitrary query from a user, will continue prompting the user for information that it needs to resolve that query. Based on this information, the decision was made to implement CALICO in such a way that the chatbot itself could direct the conversation, as opposed to the user.

The recommendation system itself was derived from the review of existing chatbot technology provided in the publication from Nahdatul et al. (Nahdatul, 2019). One of the existing chatbot implementations reviewed by the authors is the use of multiple option buttons to help attract a user to use the chatbot as well as help to direct the flow of conversation between the chatbot and the user. This directly influenced the implementation of the recommendation system for CALICO as a questionnaire in which users are presented with several preset responses in the form of buttons which can be clicked to answer the questions.

The training approach taken to enhance user interaction with CALICO was developed after reading the publication from Cuayahuitl et al. (Cuayahuitl, 2019). Deep Reinforcement Learning for Chatbots Using Clustered Actions and Human-Likeness Rewards. 2019 International Joint Conference on Neural Networks (IJCNN). doi:10.1109/ijcnn.2019.8852376 about using an unsupervised, deep reinforcement approach to train chatbots. The authors describe their implementation of a simple reward function which rewards their deep reinforcement machine learning model for "good" dialogues and penalizes the model for "bad" dialogues. The reward function itself was derived from human-to-human dialogue, in which a "good" dialogue correlated strongly with coherent, human-like responses and a "bad" dialogue correlated with incoherent responses. Due to both time and resource limitations, training of the CALICO chatbot was not able to be performed using deep reinforcement machine learning models. Instead, we adopted a modified approach to reinforcement learning in which we iteratively garnered feedback from test users on the responses they received while interacting with CALICO. Based on this iterative feedback approach, we made subsequent adjustments to the underlying AIML model and continued testing.



Finally, the optimization approach taken to improve the quality of CALICO was inspired by the chatbot deployed by Penn Medicine (in collaboration with Verily and Google) which leveraged machine learning and natural language processing to provide users with accurate information about the novel coronavirus (COVID-19) during the height of the COVID-19 pandemic (Herriman, 2020). Initial development of the COVID-19 chatbot was based upon identification of 97 frequently asked questions, which were then answered in collaboration with expert clinicians and incorporated into the chatbot database. The authors describe their process for ongoing optimization as an iterative process that involved both identification and implementation of additional frequently asked questions extracted from conversations with the chatbot, as well as continuous re-evaluation and updating of the accuracy of the information provided by the chatbot. Optimization of CALICO continues to proceed in a similar fashion. The chatbot logs are periodically reviewed for questions asked that resulted in the user receiving the default response from the chatbot, and the most frequent of these questions are answered and incorporated into the chatbot's AIML architecture.

### A.3. CAREER RECOMMENDATIONS

The CALICO chatbot can provide computer science students with information about the top 5 career paths for computer science students following graduation as ranked by the U.S. News & World Report. These careers are as follows:

1. Software Developer
2. Information Technology Manager
3. Web Developer
4. Database Administrator
5. Information Security Analyst

All source files detailing interactions that a user can have with the CALICO bot can be found in the companion AIML files. Each job type listed above has its own AIML file containing category declarations which detail the questions that a user can ask as well as the responses that a user will receive.

## B. TRAINING

The initial training of the CALICO bot was conducted with a sample of 30 computer science seniors from WGU University who volunteered to be a part of the CALICO beta. Students were divided into three groups of 10 students each, with each group of students participating in subsequent phases of the CALICO beta. To facilitate easy interaction with the chatbot from home, CALICO was deployed to a staging environment and unique URLs were provided to each student to allow access to the CALICO bot.

Each student was instructed to interact with CALICO on 5 separate occasions over the course of a three-day period. All interactions with CALICO were recorded and logged to identify frequently asked questions that resulted in the student receiving the universal default response from CALICO. These cases were targeted for implementation first and were quickly incorporated into CALICO's AIML knowledge base as additional categories of interaction. This process was repeated with each of the three groups of students to ensure that CALICO had a broad range of interactions with its target users prior to being



deployed into a production environment. An example of this process based upon beta user interaction with CALICO can be found below:

*Student A is a computer science senior at WGU University participating in the beta phase of CALICO implementation. Student A really enjoys web development and thinks that they may want to pursue a career as a full stack developer following graduation. However, Student A also enjoys many other aspects of computer science. Below is an excerpt of their conversation with CALICO.*

**STUDENT A:** Hello

**CALICO:** Hello, I am CALICO, WGU University's virtual career advisor for computer science students!

I can provide you with information about the top careers for computer science students after graduation.

I can also recommend careers for you based upon your interests! Would you like to:

[Learn About Computer Science Careers]

[Get a Career Recommendation]

**STUDENT A:** \*Clicks option two\*

**CALICO:** Do you enjoy managing and coordinating various information systems as well as navigating the interplay between various IT professionals?

[Yes] [No]

**STUDENT A:** I really like to use HTML, CSS, and JavaScript to make beautiful sites.

**CALICO:** I am sorry, I did not understand that. Would you like to:

[Learn About Computer Science Careers]

[Get a Career Recommendation]

In the conversation above, during their interaction with CALICO Student A triggered the universal default response because CALICO was unable to match one of its AIML categories with the student's input. Because this interaction was logged and recorded, the CALICO development team was able to add an AIML category for the keyword "site" and "sites", which will then trigger CALICO to proceed down the recommendation system pathway for "Web Developer" and will provide the student with information about the web development career path as well as information as to how to set up an appointment with the career advisor specializing in this space.

### C. OPTIMIZATION

The optimization of CALICO occurred in two separate phases. The first phase of optimization occurred during the beta deployment as feedback was received from the test users. Based on initial interactions of our test users with CALICO during the beta, the first optimization taken in the development of CALICO was the decision to focus on guiding users through their interaction with CALICO using a questionnaire as opposed to allowing a more free-form conversation. This optimization had two positive impacts on



development of CALICO. First, it allowed the development team to maintain their development timeline and avoid over allocation of resources. By transitioning from trying to write categories for all possible inputs to instead writing a question-answer interface, we were able to deliver a usable product to the advisory team on schedule. In addition, the question-answer interface guides the user through their interaction with the chatbot which streamlines their conversation and allows them to receive a recommendation from CALICO more quickly.

The second phase of optimization is continuous and ongoing. Like the team from Penn Medicine mentioned in section A.2, our development team periodically reviews logs of interactions with CALICO with a focus on interactions that result in the user receiving the universal default response. These interactions are targeted for implementation into CALICO's knowledgebase. Once implemented, the development team monitors similar interactions to determine if further modification of CALICO is necessary to address the issue.

#### D. INSTALLATION

To install CALICO, please follow each of the steps below:

1. Navigate to <https://home.pandorabots.com/home.html>
2. If you already have a Pandorabots account, click the "Sign In" button in the upper-right hand corner of the page. If you do not already have a Pandorabots account, click the "Sign Up Free" button in the center of the page and follow the on-screen instructions to create an account.
3. Once signed in, select "Go To Dash" from the upper-right hand corner of the page.
4. In the left sidebar of your dashboard, click the "+" icon next to "My Bots" to create a new chatbot.
5. Specify a name for your chatbot, select "English" as the language and "Blank Bot" as the content type.
6. In the left sidebar of your dashboard, click the name of your new chatbot underneath "My Bots" → "Code Editor".
7. From the toolbar above the code editor, select "File" → "Upload".
8. Navigate to the directory containing the companion files associated with this document. Select all of the files underneath the directory entitled "calico" and click "Open".
9. Once the files have been successfully uploaded, click the "Done" button.
10. At this point, you have successfully installed CALICO into your own Pandorabots environment. You can now choose to interact with CALICO through the development environment or deploy CALICO to a production server.

#### E. BOT EFFECTIVENESS

CALICO has become quite an effective tool for the Student Advisory Department at WGU University. Because CALICO provides students with a career advisor that has expertise in the recommended career path, the Student Advisory Department has seen a reduction in the number of appointments that must be rescheduled with another advisor. In addition, career advisors have reported that students are coming to their appointments much better prepared to discuss careers of interest, which has resulted in shorter meetings with students and increased bandwidth career advisors to schedule meetings. Finally, a poll of computer science students has shown that students are generally satisfied with their



interactions with CALICO. However, the main frustration point for students is that they frequently receive the default response from CALICO when they type a response as opposed to selecting a button. As detailed in Section C, the second phase of our development team's optimization approach will be to continue to review CALICO's logs to determine which inputs commonly result in the universal default response from CALICO. As our development team uses this process to optimize CALICO's knowledge base, we expect that student satisfaction will continue to increase.

## F. CHALLENGES

The primary challenge faced during development of CALICO was the initial implementation of the chatbot, which was to try and predict all possible common inputs to the system, and have CALICO respond appropriately. This initial approach was problematic for several reasons. The development team quickly fell behind schedule as they tried to come up with AIML categories for all possible situations that they could think of. As more and more AIML categories were added to the system, the flow of a user through the recommendation system became more and more convoluted. This resulted in erroneous recommendations as well as frequent receipt of the universal default response. To resolve this issue, the decision was made to transition to a questionnaire-based approach for the recommendation system and focus on optimization of CALICO's response to textual input after deployment. This allowed the development team to get back on schedule, better understand the flow of a user through the system, and reduce unnecessary work.

## G. DEVELOPMENT ENVIRONMENT STRENGTHS AND WEAKNESSES

The Pandorabots development environment provided several useful features that supported the development of CALICO. The file generation feature allowed for easy generation of AIML files, map files, set files, and substitutions and kept these files organized into their own directories. This organization allowed our developers to clearly conceptualize and associate files with their given purpose. In addition, the embedded development environment chatbot was crucial for obtaining rapid feedback on the implementation and functionality of new AIML categories. By opening up the development environment chatbot, our developers could quickly ascertain if their new features were working properly which sped up iterations of CALICO. Finally, the ease with which CALICO can be both deployed to staging and production environments, as well as pushed to a code repository allowed our engineering team to forego development of a continuous integration and continuous deployment process for CALICO and instead focus on the development of new features. The only negative aspect of the Pandorabots development environment was that it would sometimes randomly log out our developers, causing them to lose valuable progress. Overall, our development team was satisfied with the Pandorabots development environment.

## H. SOURCES

1. Bhavika, R., Raghuwasnshi, N., Singh S. (2017). Chatbot for university related FAQs. 2017 *International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 1525-1530. doi: 10.1109/ICACCI.2017.8126057
2. Reshmi, S., Balakrishnan, K. (2016). Implementation of an inquisitive chatbot for database supported knowledge bases. *Sādhana*, 41(10), 1173-1178. doi:10.1007/s12046-016-0544-1





3. Nahdatul, A., Mohamad, H., Azaliza, Z., Zirawani, B. (2019). UNISEL Bot: Designing Simple Chatbot System for University FAQs. *International Journal of Innovative Technology and Exploring Engineering Regular Issue*, 9(2), 4689-4693. doi:10.35940/ijitee.b9067.129219
4. Cuayahuitl, H., Lee, D., Ryu, S., Choi, S., Hwang, I., Kim, J. (2019). Deep Reinforcement Learning for Chatbots Using Clustered Actions and Human-Likeness Rewards. *2019 International Joint Conference on Neural Networks (IJCNN)*. doi:10.1109/ijcnn.2019.8852376
5. Herriman, M., Meer, E., Rosin, R., Lee, V., Washington, V., Volpp, K. G. (2020, June 18). Maguire Herriman. Retrieved September 16, 2020, from <https://catalyst.nejm.org/doi/full/10.1056/cat.20.0230>

