

# dotEDU: Streamlining the College Experience via Chatbot Dialogue

Ben Fertig<sup>1</sup>, Priyam Mazumdar<sup>2</sup>, Daisy Mohiuddin<sup>3</sup>, Sarah Saltzman<sup>4</sup>

1. Computer Science | Psychology & Linguistics, 2. Quantitative Theory and Methods, 3. Psychology & Linguistics,

4. Quantitative Theory and Methods

Emory University, Atlanta, GA 30322

{ben.fertig, pmazumd, tmohiud, sarah.saltzman}@emory.edu

## Abstract

dotEDU is a socialbot that streamlines the college experience for prospective and current undergraduate students of U.S. colleges and universities through meaningful and empathetic dialogue. Navigating college from start to finish while balancing external commitments can induce significant stress in students at all stages of their undergraduate careers. This stress can stem from a variety of factors, such as applying to undergraduate programs, declaring a major, landing an internship, and securing post-graduate plans. Thus, we have designed a dialogue state machine using the *Emora STDM* Python package to facilitate informed decision-making and conflict resolution at various checkpoints throughout college. Specifically, the system caters to 6 audience groups: prospective undergraduate students, freshmen, sophomores, juniors, seniors, and college graduates. However, as previously mentioned, this bot is primarily designed for current and prospective students of undergraduate programs and not high school graduates who are not planning to attend college. Nevertheless, we have designed a brief conversation path for this latter demographic to acknowledge their current state in life. To best serve prospective students, we devised a recommender system that takes into account numerous factors, such as school size, location, and cost, to assist the user in deciding which colleges to apply to or attend. We incorporated an external dataset from the Integrated Postsecondary Data System by the National Center for Education Statistics into this portion of the bot to capture a variety of schools. This bot reads in school characteristics from the dataset, equipping the bot to recommend schools based on user preferences and provide further information should users ask for any college or university descriptions. To best serve current students, we evaluated the defining moments of each year in college based on personal experience and observations and hardcoded a variety of potential situations and challenges that users may want to discuss. Topics include but are not limited to adjusting to college life, student involvement, degree planning, and career preparation.

It is important to note that the bot was programmed by college students for college students. As a result, it not only establishes heightened credibility but also forges strong connections with users. This chatbot is novel in its offering of a convenient, efficient, and interactive supplement or alternative to in-person college, academic, and career counseling. Our hope is for users to feel confident in their abilities to tackle challenges faced during their higher education.

## 1. Introduction

dotEDU is a chatbot that simplifies the college search process, eases the transition into college, and aids with degree and career planning, among other functions. Depending on the user's stage in their higher education journey, dotEDU can recommend prospective students colleges and universities that best embody their preferences or guide a user who is a current undergraduate student through obstacles they have already encountered at university. If the user is a current undergraduate student who is not

experiencing any challenges, they can simply discuss with the bot their plans as they would with a friend or family member and the bot will provide suggestions or validations. The value of the bot largely stems from its ability to help the user determine their academic and professional priorities by brainstorming fitting environments and discussing other pertinent abstract and opinion-based topics.

Furthermore, there is a growing need for services that help reduce the stress of the college selection process. Many high school students have to conduct extensive research on dozens of universities while also attending to more immediate obligations. These large transitions should be handled with care. All students, even those who are already *in* college, face dilemmas, such as declaring their majors, finding internships, deciding which campus organizations to get involved in, and making post-graduate plans. Such stressors can quickly spiral out of control if left unattended. Thus, dotEDU helps students organize their thoughts and then recommends a course of action to help them reach their goals.

Our chatbot also offers a level of accessibility and simplicity that is not offered by more traditional solutions. While campus advisors can provide similar services to those of the bot, these advisors are often operating at capacity and demand that students schedule appointments well in advance of when they actually intend to meet. Unfortunately, the fast-paced nature of college life can often lead to unforeseen challenges that require more immediate attention. To help combat stress, pressure, and even mental illness among college students, dotEDU helps users plan for major milestones in an effort to make responsibilities feel less paralyzing. With the aforementioned challenges in mind, dotEDU seeks to achieve three goals: discuss with users the various academic, social, and professional opportunities that may be available to them; boost self-esteem; and navigate through difficult decision-making. One of our chatbot's core strengths is picking up on users' uneasiness and responding in a way that helps users logically evaluate their options. Because each conversation is completely customized, users can leave the conversation feeling confident about their upward trajectory.

dotEDU's novelty stems from its ability to show compassion, lend advice, and validate a user's decisions and choices. Current chatbots are largely information-based, which, although useful in time-constraining circumstances, remain emotionally disconnected from the user. To our knowledge, our largest competitor is an artificial intelligence-powered bot, Ivy, that partners with universities to answer college-related questions. Contrary to dotEDU, Ivy's main goal is to provide timely responses to a user's questions and direct them to relevant campus resources if applicable. As with most bots on the market, Ivy's predominant function is to share information with the user. dotEDU, however, is novel in that it infuses sympathy in its responses to avoid the robotic tone that Ivy and many other bots on the market currently elicit. dotEDU's resourcefulness, usefulness, and understanding maximize user engagement and positively impact the lives of current and prospective college students. Success may be quantified via high retention rates. For example, our bot would be considered a success if a user who chatted with it as a prospective student were to again chat with it as a first year student.

## **2. Related Work**

As previously mentioned, a current leader in the field of college life chatbots is *Ivy.ai*. As written by Ivy's creators, "Ivy's Natural Language Processing (NLP) is powered by a human-supervised deep learning algorithm called a Convolutional Neural Network (CNN). Using a massive set of highly curated training data, Ivy is able to accurately predict what your users are asking and get them to the answers they need." Ivy has some extremely powerful abilities, including the ability to understand context and return

immediate answers to multiple questions asked by the user in a single statement. During later stages of conversation, *Ivy* can also refer back to details mentioned earlier to answer future questions.

*Ivy* targets students that are already in college by aiding them on campus, whether that be by helping them get around or by offering important information. To implement this assistance, *Ivy*'s creators have laid out a 4 stage plan detailing how they go from nothing all the way to a fully functioning chatbot. First, they crawl universities' web pages and extract as much information as possible. Next, they connect *Ivy*'s extracted information to prebuilt vendor integration that allows for personalized answers to questions. The team then transcodes these features into conversation format. Lastly, the team verifies the correctness of the bot by making sure it functions as needed.

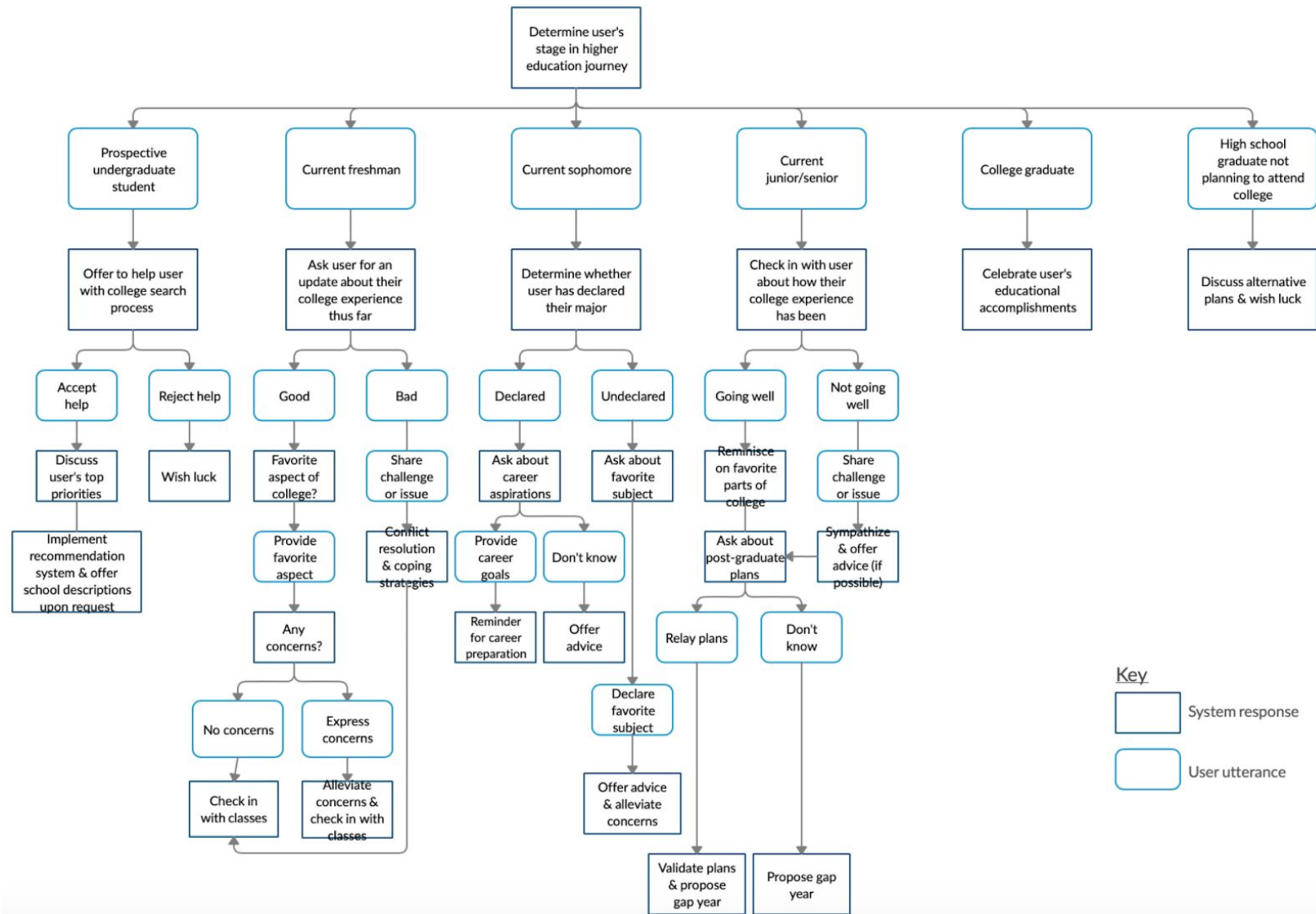
To keep the bot up-to-date with the latest information, *Ivy* re-crawls university websites daily. Another interesting feature about the bot is that it acknowledges that it is still a machine and that it may get stumped. We implemented a similar feature in our chatbot, error successors, which keep the conversation moving even if the user provides a response that the bot does not recognize. Sometimes, however, we cannot skip over a question if, for example, our bot's understanding of a user's response is necessary to progress the conversation. *Ivy*'s creators say that, if the bot gets stuck or if it is asked a question which cannot be sourced from the school's website, it will route the question back to the school so that it can be answered in future conversations, provided that other users ask the same or similar questions.

Lastly, *Ivy*'s creators have greatly simplified *Ivy*'s implementation to where the university does not have to deal with any backend of the program. After personalizing the chatbot for their school (which takes an estimated two to three weeks), universities can place a single line of *HTML* code within their website which will embed the chatbot on that page. Universities also have the ability to customize the appearance and brand of the chatbot to make it look like a part of their website.

Overall, the undertaking of *Ivy*'s creators is much more complex and complete than our own, but it demonstrates the potential that exists in this field. Universities are increasingly taking advantage of chatbots to optimize their workflow and to enable faster retrieval of information for their clients, bypassing the need for meetings or phone calls. *Ivy*'s team is also trying to implement its technology for more high-risk tasks, such as communication with healthcare providers, which demonstrates that many essential organizations are already taking this chatbot's technology into serious consideration.

### 3. Conversation Design

**Below is a state-machine diagram that depicts the simplified conversation flow of the bot.**



**Below is a sample conversation with a first-year student.**

*System:* I want to know more about your thoughts on higher education. Have you ever been to college?

*User:* im in my first year

*System:* First year is exciting! But, how has your transition been so far?

*User:* terrible

*System:* I'm sorry to hear that. Why so?

*User:* i dont have any friends

*System:* Don't worry. Some people take longer to make friends than others. There are lots of ways to do it. For example, have you tried making friends with your roommate or signing up for greek life or clubs?

*User:* it doesnt work

*System:* Well, I hope you seek out help if you need it. How are your classes though?

*User:* classes are going well

*System:* Wow, you must be a pro if you're enjoying your college classes already! I think you're going to do just fine. It was nice talking with you. Hope you have a nice day.

**Below is a sample conversation with a second-year student.**

*System:* I want to know more about your thoughts on higher education. Have you ever been to college?

*User:* im actually currently in college

*System:* Cool! What year are you in?

*User:* my sophomore year

*System:* Nice, you are no longer the youngest at the university. Sophomore year is an exciting time academically and socially. Have you declared your major yet?

*User:* nope, im not sure what to major in

*System:* Totally understandable. It can take students a while to find the right major for them. Perhaps I can assist you. What has been your favorite subject so far?

*User:* hmm id have to say theatre. i love the performing arts

*System:* That is great that you love theatre. Majoring in the arts will enable you to expand your creativity, which is definitely a valuable asset to have when applying for jobs.

*User:* thats what i try to tell myself, but im worried i wont make enough money when i graduate

*System:* It may seem that way now, but students with a college degree typically have a high earning potential. Right now I would just focus on what you are interested in.

*User:* alright, but im worried that i wont have any job opportunities if i pursue this major

*System:* I suggest speaking with your career center. Career counselors are very knowledgeable about employment opportunities in your areas of interest.

*User:* ok, i suppose i will try that. thanks for the advice

*System:* Anytime! Good luck with your classes. I look forward to hearing more about your experiences!

*User:* talk to you later

*System:* It was nice talking with you. Hope you have a nice day.

**Below is a sample conversation with a third or fourth-year student.**

*System:* I want to know more about your thoughts on higher education. Have you ever been to college?

*User:* yeah im a junior right now.

*System:* How has the college experience been for you so far?

*User:* great. i love college

*System:* That's great! What are your favorite things you have been involved in until now?

*User:* too many to choose from but ill go with my sorority

*System:* Greek life is an amazing way to make lots of good friends with similar interests to you! I hope you had enjoyed that truly unique experience! Now that your college experience is about to come to an end, what kinds of things do you plan on pursuing?

*User:* i think ill go directly into the workforce

*System:* That's really cool! Is the job you want to pursue related to what you did during university or will you be pursuing something else to expand your experience?

*User:* i want to try something new

*System:* It is totally ok to keep learning more things outside of your field, even after college. Have you ever considered taking a gap year to pursue a passion or a hobby before you go back to work or school?

*User:* nah. gotta make some money

*System:* After so many years of rigorous schooling, it is important to take breaks so you don't feel burnt out. If you decide not to take any gap year, it's important to find other things that give you breaks and de-stress.

*User:* totally agree, thanks

*System:* It was nice talking with you. Hope you have a nice day.

#### 4. Approach

Through our creative use of *Emora STDM*, a dialogue management framework that simulates a smooth conversation flow through the use of regular expressions and state machines, we were able to elicit natural, emotional, and feedback-filled system responses.

In terms of our novel recommendation system, our bot assesses user preferences through a series of questions and informs them of schools that would likely satisfy their needs. This conversation is punctuated with seamless transitions to make it less like a questionnaire. Through the use of *Emora STDM*'s *NatexNLU* and *NatexNLG* packages, our chatbot is able to search user responses for predetermined clusters of expected words. If there is a transition function that contains a Natex that overlaps with the user's input, it will be activated and the conversation will shift to the respective successor state. If the user's response satisfies multiple Natexes across multiple transition functions, an optional *score* argument can be provided for transition functions at the same state. Ties are then broken by the highest score across all flagged transitions. Should the user enter a response which flags none of the provided transitions' Natexes, error successors retain the fluidity of the conversation by automatically transitioning to a default successor state.

Research suggests that ontology-based chatbots are effective at correctly interpreting user input and responding appropriately (Vegesna et al. 2018). Thus, we have incorporated ontologies to ensure that the bot can understand and respond to a breadth of user utterances. For example, if we were to include an ontology containing a list of positive emotions, we could ensure that our bot provides the same response to users who exclaim that they are "ecstatic" as it does to those who exclaim that they are "elated." Such functions are needed to account for different styles of speech.

We have incorporated dictionaries and frozen sets into our ontologies to allow the recommender system to more effectively navigate user preferences. Furthermore, this design allows for smooth information retrieval by our Macros, which enable the system to recall user input from previous states even as the conversation proceeds. The hallmark Macro of this version of the chatbot is the

*CollegeRecommenderMacro* which retrieves previously stored variables and categorizes them into meaningful attributes, such as size, location, and setting. Matched variables will be set to specific college characteristics, such as large size, in-state location, and suburban setting. These attributes are paired with subsets of the master college dataset. Each attribute corresponds to a set of applicable colleges, and the final list of recommended colleges is a sampling from the intersection of those sets. One current limitation of the system is its inability to understand certain responses. For example, our system is unable to understand any search criteria regarding a school's reputation that do not pertain to it being either private, public, for-profit, historically black, tribal, Ivy League, or prestigious (i.e. ranked among the top 50 national universities by U.S. News & World Report). This shortcoming is mainly due to our dataset not being completely comprehensive and therefore unequipped to handle uncommon user queries. It is possible that the constellation of needs that a user provides may not match a specific list of universities from our database. To address this possibility, we have added a feature to the recommender system that

allows it to pull back specific metrics, loosening the user's constraints while still returning colleges and/or universities that satisfy as many of the users' needs as possible.

Additional Macros are implemented to control system transitions. For example, the aforementioned emotion ontology is paired with a Macro named *EmotionMacro*, allowing the chatbot to respond enthusiastically to positive emotions and reassuringly to negative emotions within a single dialogue state system transition. This allows the system to streamline the conversation so that even when the conversation diverges, all paths will eventually reconvene to the same question.

Research also suggests that part-of-speech tagging (POS) can facilitate high quality and precise syntactic category recognition and recall without compromising efficiency (Brill 1995). If the user is asked about the type of activity they are most interested in, there are simply too many possibilities for us to feasibly create a complete list of potential matches. Instead, we can allow for more generalizable input matching by using POS tagging from the *spaCy* package and then formulating an appropriate response by searching for specific grammatical categories in the user's input. Overall, these tools allow for a more realistic conversation and a broader level of understanding. Thus, we implemented POS tagging where possible to capture important information and respond directly to the user. However, as previously mentioned, most of our design is rooted in hard-coding specific responses to expected inputs. For instance, we have hard-coded several responses to accommodate users who may not have had any college experience in addition to including an optional conversation on financial options, which attends to students who express interest in receiving financial advice. Incorporating natural, humanistic expressions, such as "yeah" and "so", into our bot's speech recognition capabilities enhances the user's experience by creating more personal, empathetic, and human-like interactions. Periodic references to a user's answers, combined with various opinion-based suggestions, make the user feel that their emotions have been heard and validated.

To more effectively advise students about specific universities that match their preferred characteristics, we make use of public datasets that contain school locations, populations, and reputations, among other statistics. By distilling through the list of colleges, we can more successfully recommend a university tailored to the individuals' needs. The initial dataset used was the "Integrated Postsecondary Education Data System" by the National Center for Education Statistics, which defined some of the more concrete attributes about colleges and universities but lacked other important features (e.g. tuition). Thus we joined this dataset with another dataset entitled "University Rankings 2017" (Sidberry and Greenleigh, 2016), which provides text snippets from U.S. News that describes some of the more subjective attributes of colleges and universities.

A major challenge we faced in creating dotEDU was our dependence on hard coding and our inability to hit all the potential cases. The issues that enrolled undergraduates experience are essentially endless. We tried to target a handful of issues that we personally had undergone, as well as common issues that we know many undergraduates have to face, but to be a proper college chatbot, the breadth of our bot must be drastically improved so that it can serve a wider audience. Most of this challenge lies in the limitations of hardcoding, the limited number of cases we can manually build, and the ability of our Natexes to capture predicted responses. Addressing this challenge may require a package other than *Emora STDM* or different techniques, which could simplify the creation and capture of diverse cases.

Another challenge we faced during the creation of this bot was the disruption of conversation flow resulting from the occasional throwing of error messages by *spaCy*'s POS and NER tags. For example, many universities were not captured by NER tagging, while seemingly random words were. As a result, we have restricted NER tagging to a single transition in our conversation, specifically the transition in which the user responds with the college that they will be attending, provided they have already indicated

that they know where they will be going to college. However, this transition runs the risk of being flagged by any string containing the name of an *institution*, rather than only being flagged by a string containing the name of an institution of *higher education*. We would be interested in expanding our usage of NER tags should we find a way to address this error.

One challenge we faced in designing the college recommender system was keeping system responses concise, as we had to acknowledge previous user responses and gather additional information about the user. Furthermore, our bot must return a lot of information, but lengthy system responses run the risk of making the conversation less engaging, less intuitive, and more robotic. Thus, dotEDU briefly acknowledges the user's response to the previous question at the beginning of each subsequent question and maintains the cadence of the conversation while reaching all essential questions in a timely manner.

Moreover, the inflexibility of *Emora STDM*'s dialogue state machine presented several issues. For instance, it would have been useful to block off a future state based on a previous response. Instead, we had to duplicate branches for conversations which differed only slightly, while deleting the states that were inapplicable to each branch. This process was tedious and inefficient. We could have streamlined the quality of the conversation flow if we had the ability to assign conditions to certain states. Another drawback of *Emora STDM* is its inability to have a single user transition point to multiple system transitions. To enable the recognition of several classes of responses, we sometimes concatenated multiple Natexes within a single user transition, but this forced us to make the responses of the system transitions more generic. We were able to circumvent some of these issues using Macros, but this led to increased code complexity, which could have been avoided if *Emora STDM* were to have offered a more efficient solution. *Emora STDM* had also undergone multiple updates throughout the progression of our project, requiring us to relearn certain functions and denying us the time to take advantage of all the functionalities that *Emora STDM* had to offer.

## 5. Analysis

The success of this bot stems from its ability to capture a breadth of user responses across many conversation paths. As previously explained, this bot is primarily geared towards current and prospective undergraduate students of U.S. colleges and universities. This bot skillfully converses with the user at whatever stage they are in to empower them to shape their futures and pursue their passions. Our expertise on the bot's subject matter is reflected in its ability to handle a breadth of topics and guide the user down a fulfilling academic or career-oriented path. As current undergraduates, we are able to impart our wisdom to users on the coping mechanisms and conflict resolution strategies for various college-related challenges.

An example of where this knowledge shines through in our bot is in its ability to recommend colleges based on a user's preferences. These preferences include the cost, location, and prestige of the user's ideal college. Our bot not only spearheads the user's college search process but it also offers full descriptions of top colleges, giving the user a better understanding of each college's characteristics. Our understanding of the need for such a bot stems from our firsthand experience with the stress of the college search process. These experiences helped us realize that prospective students should have access to more resources that can help them balance the college search and application processes with difficult high school coursework and extracurricular commitments.



First-year undergraduate students often experience a range of emotions adjusting to college life. Many students must leave their hometowns, families, and friends and live in a foreign territory with strangers. Students in these situations, especially those who find comfort in routines and familiarity, are often prevented from experiencing the excitement and freedom that college life has to offer. To lessen the severity of this transition, our bot effectively supports first year students through challenges such as roommate drama, campus involvement, healthy eating, challenging coursework, homesickness, and mental health complications. Our bot sympathizes with the user and validates their emotions.

Many second-year undergraduates are tasked with declaring a major. While some students are certain about what they want to study, many are undecided or are experiencing a change of heart after two or three semesters worth of courses. Whichever of these stages a second-year student may be in, our bot helps them either decide their major or plan their professional trajectory. If the user has already declared a major, our bot initiates a subsequent conversation regarding the user's career goals. If the user is unsure of what to major in, the bot asks the user about their favorite subject in school and encourages the user to follow their passion. This bot is equipped to alleviate some common concerns, such as a lack of employment opportunities, low average salaries in the students' areas of interest, and familial pressure.

Lastly, juniors and seniors have likely already begun internships, full-time job searches and/or graduate school applications. Our bot discusses with the user the necessary steps they must take to secure future opportunities and helps them determine what those next steps should be. The bot asks users who are unsure of what path to pursue if they have considered taking a gap year to pursue something that interests them. Our bot also discusses necessary testing requirements for medical and law schools, should the user indicate that they are interested in pursuing these opportunities.

## **6. Conclusion and Future Work**

With the intensifying stresses of college and career planning, university students need more robust psychological and academic resources. Consequently, many universities are struggling to meet these high demands. If students were to have access to an interactive, compassionate, and dynamic bot like dotEDU, universities would be able to allocate their resources more effectively and efficiently. Students could use this bot as a preliminary step to fulfill their most immediate needs, and those who need additional help can either revisit the bot later or schedule an in-person appointment with a counselor at their university. Rather than using limited university resources, students could first consult this bot, which would allow university officials to deal with time sensitive and severe cases more promptly. Likewise, students who have interacted with dotEDU would have attempted to deal with the issue on their own before seeking outside help. The success of this bot can be measured by the number of returning users (i.e. number of people who are consistently visiting dotEDU as they face new issues).

Despite the comprehensiveness of dotEDU, its features can be improved upon. For example, the college recommendation system currently relies on user utterances to provide recommendations. If user input does not match university characteristics, the system proposes schools that meet the *majority* of user preferences. However, it is important to note that in this situation, the user is unable to select which criteria to eliminate. Future improvements can focus on determining the relative importance of various metrics to the user so the system can remove those metrics first. Another important aspect of the college application process is the categorization of universities into safety, match and reach schools based on the individual's performance in high school. It is crucial to be able to recommend a variety of schools with different entrance standards to give the user access to a wider array of options. Such a task would require

revamping the dataset to include information on admissions selectivity, average test scores, and average high school grade point average of admitted students. Currently, dotEDU covers a wide breadth of college experiences; however, it somewhat lacks specificity. For example, a current undergraduate student would benefit from in-depth coaching on graduate school preparation, including the testing timeline and application process. Future work would involve integrating multiple datasets with examination dates, campus events, and application deadlines. Another shortcoming of this bot is its relative inability to comprehend user's questions. Recognizing question words using POS tags would enable the system to adequately respond to any questions the user poses.

dotEDU has a bright future, but it also has numerous features that can already be improved on. We can use the feedback provided by its users to further develop its capabilities. Our hope for this bot is that it will help students with planning and organization so they can spend more time on their studies, extracurricular involvement, and personal development.

## 7. References

- Brill, Eric. "Transformation-based error-driven learning and natural language processing: A case study in part-of-speech tagging." *Computational linguistics* 21.4 (1995): 543-565.
- "Chatbot Features: State of the Art Deep Learning Algorithms." *IVY*, ivy.ai/features/chatbot.
- Clark, Marcia Roe. "Negotiating the freshman year: Challenges and strategies among first-year college students." *Journal of College Student Development* 46.3 (2005): 296-316.
- Kitzrow, Martha Anne. "The mental health needs of today's college students: Challenges and recommendations." *NASPA journal* 41.1 (2003): 167-181.
- Leong, Frederick TL, Margaret H. Bonz, and Peter Zachar. "Coping styles as predictors of college adjustment among freshmen." *Counselling Psychology Quarterly* 10.2 (1997): 211-220.
- Mattanah, Jonathan F., Gregory R. Hancock, and Bethany L. Brand. "Parental attachment, separation-individuation, and college student adjustment: A structural equation analysis of mediational effects." *Journal of Counseling Psychology* 51.2 (2004): 213.
- National Center for Education Statistics. "Integrated Postsecondary Education Data System." Feb. 2020.
- Paul, Elizabeth L., and Sigal Brier. "Friendsickness in the transition to college: Precollege predictors and college adjustment correlates." *Journal of Counseling & Development* 79.1 (2001): 77-89.
- Sidberry, Dan, and Ian Greenleigh. "US News University Rankings 2017 Edition." 13 Sept. 2016.
- Vegesna, Anusha, et al. "Ontology Based Chatbot (For E-Commerce Website)." *International Journal of Computer Applications*, vol. 179, no. 14, 2018, pp. 51-55.