CS 25-314 Chatbot for Undergraduates
Project Proposal

Prepared for

Caroline Budwell

By

Kennedy Martin

Israel Agoe-Sowah

Antony Fuentes

Eric Simoni

Date

9/26/2024

Executive Summary

Virginia Commonwealth University is committed to enhancing the student experience and ensuring that every student has access to the resources necessary to be successful. In line with this commitment, the Computer Science Department at VCU has identified a significant issue; students in the Computer Science program, both incoming and current, are facing difficulties in finding reliable information regarding their degree track. While the VCU website holds all the necessary information, its navigation is complex and often leads students to seek assistance from academic advisors and faculty for basic guidance. This reliance burdens departmental staff and leads to inefficiencies and increased workloads.

The proposed solution led by the director of the department, Caroline Budwell, is a chatbot that will be implemented on the VCU website. This chatbot will assist students in accessing degree-related information reducing their reliance on faculty for routine questions.

The main requirements are as follows: develop a chatbot that answers undergraduate Computer Science frequently asked questions and integrate the chatbot into the VCU website.

Thus far, this team has done research into the tools and techniques that could be utilized in the completion of the project and has chosen the most agreeable route. Furthermore, planning and construction of the database using SQLite has begun as well as the initial construction of the front and back end programs.

Completion of this project is expected by May 2025. Major deliverables before then include the completion of a prototype by December 2024. This involves a database as well as front/back end programs that communicate properly with each other and with a model for natural language processing.

Table of Contents

Section A. Problem Statement

At Virginia Commonwealth University (VCU), we place a strong emphasis on enhancing the student experience and ensuring that every student has access to the resources needed for academic success. However, it has recently come to our attention that students in the Computer Science program, both current and incoming, are facing significant challenges in efficiently obtaining accurate information about their degree and subsequent requirements.

Although the VCU website provides the necessary information, students often find it difficult to navigate. As a result, many students frequently turn to academic advisors or faculty members for assistance, creating additional workloads for staff and leading to inefficiencies. This issue is particularly concerning given the recent introduction of the Bachelor of Arts (BA) in Computer Science, which is expected to increase the number of enrolled students and further exacerbate these difficulties.

If left unaddressed, this problem will likely lead to increased faculty time dedicated to answering routine questions, incurring indirect costs for the department. Additionally, students may experience confusion regarding their academic paths, which could result in unnecessary course enrollments, delays in graduation, or even changes in major due to misinformation or misunderstanding.

This project aims to resolve these issues by streamlining the process for students to access degree-related information and reducing the reliance on faculty and advisors for basic inquiries. Leading the project is the Undergraduate Director of the Computer Science department, Caroline Budwell.

The history of chatbots began in 1966 with ELIZA, one of the first programs to simulate conversation and project an understanding of human language. ELIZA laid the foundation for how chatbots operate today. Over the years, many chatbots were developed, each utilized in different ways but all simulating human interaction.

The next big milestone occurred in the early 2000s with the uptick of machine learning. This advancement allowed chatbots to evolve from simple, template-based systems to more sophisticated models using neural networks. By the late 2010s, chatbots became more widespread as the tools and technologies needed to build them became more accessible.

Today, chatbots have reached new levels of capability thanks to transformer architectures, which allow for human-like interaction like never seen before. Chatbots are now seen in various settings and are implemented across the web.

This project focuses on implementing a chatbot on the VCU website to assist with answering general questions about curriculum and degree tracks. In the future, this project could be scaled to encompass more ambitious goals, such as dynamically creating web pages based on user queries or even acting as a teaching assistant to guide students through course material with step-by-step explanations.

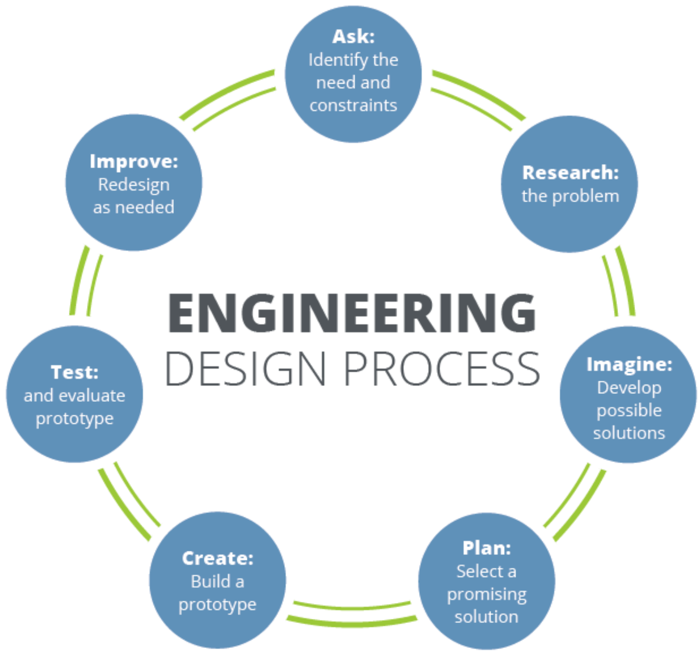

Figure 1. The iterative nature of the engineering design process [2].

Section B. Engineering Design Requirements

B.1 Project Goals (i.e. Client Needs)

Our project aims to fulfill the needs of undergraduate computer science students by providing an efficient and user-friendly chatbot that answers FAQs. Students will face challenges in accessing academic information, either because of the information being dispersed across different resources or because they must contact an advisor/administrative staff. This often leads to delays in decision-making, especially during peak times like course registration. This chatbot will address this issue by consolidating the most common inquiries into a single, accessible platform where students can reduce the time they spend searching for answers and easing the burden of advisors having to answer similar questions multiple times a semester. The chatbots 24/7 availability ensures that students have access to this information outside of staff business hours.

- To enhance accessibility by providing 24/7 access to information and reducing reliance on faculty and department office hours.
- To increase efficiency in answering common queries about course registration, prerequisites, office hours, academic requirements, etc. by automating FAQs.
- To allow students to improve their decision-making by giving access to information for course planning, which can help with avoiding registration delays and missing deadlines.
- To improve the student's engagement with the university's website, by providing it with a responsive and helpful chatbot feature.

B.2 Design Objectives

- Provide Real-Time Information:
    - The chatbot will be able to provide answers to 90% of FAQs within a reasonable time window. Which will be achieved through internal knowledge base covering of topics such as course registration, deadlines, prerequisites, etc.
    - The time between a student's question and the chatbot's response should remain under a reasonable time threshold.
    - This feature should be done and fully operational by the end of this semester.
- Handle High Traffic Efficiency:
    - The chatbot will be designed to handle a large number of concurrent users without performance degradation. This will allow the system to function well during high-traffic periods like registration week.

    - We will have to conduct tests during peak usage to ensure that the system maintains an acceptable performance level with us wanting to keep response times for all users to be under a reasonable amount of time.
- Integrating with University Systems:
    - The chatbot will integrate with the university's database to provide dynamic answers.
    - Integration success will be measured by the chatbot's ability to pull real-time data from the university's database
    - APIs will be used to fetch this data, and integration will be tested with student queries to ensure accuracy.
- Adaptability:
    - The chatbot will be designed to allow for easy updates and the ability to expand beyond the Computer Science department.

B.3 Design Specifications and Constraints

- Response Time Constraint:
    - The chatbot must respond to the user's question within a reasonable amount of time. Given that the system will have to handle peak load scenarios of a large number of concurrent users.
    - We will have to simulate traffic conditions to verify that performance under load, ensuring response time stays under a certain amount of time.
- Data Privacy Constraint:
    - The chatbot must adhere to the university's data privacy standards.
    - The chatbot will not collect or store any personal information from users, its functionality is to answer FAQS and provide information that can be found on the VCU website.
- Budget constraint:
    - Total development cost for the chatbot should not exceed $1000. This includes expenses related to cloud hosting, APIs, and other external services that may be required for development.
- Maintenance and Scalability Constraint:
    - Chatbot has to be designed for easy maintenance, allowing updates to the FAQs database or the chatbot's knowledge base without requiring a complete system overhaul.
    - Admin users must be able to update FAQ database or add new topics through a simple content management interface
    - Scalability: The chatbot should be able to allow future expansion into other academic departments.

- Platform Compatibility Constraint:
  - The chatbot has to be compatible with VCUs existing web infrastructure and has to be seamlessly integrated into the website. It also has to be able to function across different devices, such as desktops, phones, and tablets.
  - Compatibility tests will have to be conducted across multiple browsers and device types to ensure smooth functionality.
  - The chatbots interface will have to adapt to various screen sizes, ensuring that every user will be able to access this feature equally without any constraints.
- Error Handling and Reliability Constraint:
  - The chatbot has to have error-handling features, in which the chatbot will be able to direct the user to the appropriate department or offer an email or site in which they can reference, if the user's question couldn't be answered.
  - If the chatbot fails to retrieve a response within the given time window, the system will suggest alternate actions.
  - Reliability will be tested by intentionally creating failure scenarios to verify the system can handle these questions smoothly.

Section C. Scope of Work

C.1 Deliverables

Preliminary Design Report: Details of the chatbot's design framework, including its architecture, functional requirements, and user interface plans.

Functional Chatbot Prototype: A working chatbot that answers basic queries regarding the degree program, deployed on a website.

Final Design Report: A comprehensive report covering all aspects of the chatbot's development, functionality, and technical details.

User Manual: Create a user manual explaining how the chatbot works for both students and admins responsible for maintaining it.

Fall Poster and Presentation: Presentation summarizing the project's progress and key milestones to date.

Capstone Poster and Presentation: Presentation showcasing the completed product, and its utility.

C.2 Milestones

Milestone	Description	Est. Time
Preliminary Design Report	Document initial designs for the chatbot.	1-2 week
Development of Computational Model	Set up the back-end logic, including NLP or rule-based systems.	3 weeks
Frontend UI Design	Build a chatbot user interface (UI) for interaction.	2 weeks
Prototype Completion	Finalize a working prototype for the chatbot.	3 weeks
Data Acquisition and Analysis	Gather Real user data from beta testing.	2-3 weeks
Create a user manual	Document how users can interact with the chat bot, and how to maintain it for admins.	1 week
Capstone EXPO Poster & Presentation	Final poster and presentation for project demonstration.	1-2 week

C.3 Resources

Software

1. Python to code the backend in combination with Flask.
2. Flask to add support for web-based interactions.
3. HTML to create the front end web pages for users to interact with.
4. Visual Studio Code as an IDE to edit any code with.
5. Javascript to code how the web page works.

Database

1. SQLite to create and populate database
2. Visual Studio utilizes with SQLite

Version Control

1. Github to keep a repository of our code and the changes we make to it.

Section D. Concept Generation

A number of methods can be used to help generate design concepts from simple reflection and brainstorming, to working the problem backwards, using reverse thinking techniques, and looking to nature for inspiration (i.e. biomimicry). Existing solutions, or components of existing solutions, can be substituted, combined, adapted, modified, put to other uses, eliminated, or rearranged to meet new design objectives and specifications. A minimum of 3 overall design concepts is required for this section although more are welcome. Provide a brief description of how each design concept addresses the design problem. Discuss the potential pros and cons, including and potential risks of failure, of each of these concepts.

It is likely that each design concept may consist of several components. In this case, one or more of these components may offer a sub-problem that can be further explored, modified, or otherwise improved upon. These sub-problems may lead to the addition of several additional design concepts and may require the inclusion of a design concept chart or matrix to organize all ideas and potential solutions.

Provide any initial design sketches, drawings, 3D renderings, or conceptual models such as dataflow diagrams, process flows, etc. developed during the concept ideation phase. All hand drawings should be drawn to scale using basic engineering drafting tools (i.e. ruler, protractor, and compass). Geometric stencils can also be used to help produce quality hand drawings. Drawings should be presented in a professional manner, preferably done on engineering graph paper and using a high-quality scan. All sketches should be labeled to identify major components and different drawing views or projections if applicable. Basic dimensions should be provided to give a general sense of scale. Label each sketch or drawing with the name of the team member responsible for the sketch, the date it was drawn, and the drawing scale.

Section E. Concept Evaluation and Selection

Using a systematic decision-making process, evaluate each of the design concepts and choose the one that is most likely to succeed in meeting the design objectives and constraints. A Decision Matrix, or Pugh Matrix, helps to analyze alternatives, eliminate biases, and make rational decisions through thought and structure. First, work to develop a set of selection criteria for which to evaluate the previously generated design concepts. Selection criteria often include concepts of performance, cost, safety, reliability, risk, etc. Note that the selection criteria developed here will likely be more general than the project design objectives. As with the design objectives, conversations with the client help define appropriate selection criteria.

In many cases, the client may value the selection criteria differently, preferring that more emphasis be placed on some than others. In this case, weighting factors may be used to place more or less importance on the various criteria in the decision making process. Again, conversations with the client can be used to define criteria weighting factors. Often times, these conversations must be analyzed and interpreted by the team to determine which criteria are more important to the client and by how much. Feel free to discuss the assigned weighting factors with the client to see if they seem accurate.

Next, define an associated metric to represent each criteria. Metrics should be specific and quantifiable, providing numerical values that quantify the often vague concepts of the selection criteria. Metrics can be obtained, generated, or estimated through a number of methods including simple background research, preliminary design calculations, or basic analyses. Note that these metrics do not need to specifically align with the design specifications although there may be some commonality between the two. Provide a brief discussion of the rationale for selecting each of the assigned metrics.

Using the defined metrics, evaluated each design concept against all selection criteria by filling out a Decision Matrix. Design concepts can be compared by using simple rank scoring, raw scoring, or weighted scoring techniques and design concept with which to move forward can be selected. This type of process provides a meaningful, unbiased means for choosing a preliminary design concept prior to moving forward with more comprehensive, detailed analyses as provided in the design methodology section below. The results of this process should be discussed with the project client prior to moving forward with the selected design. Table 1 provides an example of a simple decision matrix.

Table 1. Example of a Decision Matrix.

	Design Concept A	Design Concept B	Design Concept C	Design Concept D
Criteria 1				
Criteria 2				
Criteria 3				
Criteria 4				
Criteria 5				
Total Score				

Note: Weights can be assigned to each criterion if desired.

Section F. Design Methodology

Provide a detailed explanation of the methods that will be used to help evaluate, improve, and evolve the design through the iterative engineering design process. Consider that ultimately, the final design must be verified and validated to ensure that it meets all of the previously developed and listed design objectives and specifications. Verification ensures that the design meets all specification, while validation confirms that the design functions as intended such to meet the client's needs. While it is common for initial design concepts to first be evaluated using simplified design criteria and metrics, the chosen design should be advanced, and later verified, using engineering calculations, computational models, experimental data, and/or testing procedures.

Use this section to describe any underlying physical principles and mathematical equations that govern the design. Provide details of any computer-aided modeling techniques used to evaluate the design including the software used, prescribed boundary conditions, and assumptions. Include a detailed description of any experimental testing methods including required testing equipment, test set-up layout, data acquisition and instrumentation, and testing procedures. If one or more prototypes is to be produced and tested, provide a detailed description of how each will be evaluated.

Note: The contents of this section are expected to vary from project to project. Subsections may be appropriate for providing details of analytical, computational, experimental, and/or testing methods. Some potential subsections that may be included in this section are provided. While critical design equations may be provided here, lengthy mathematical derivations may be included in an appendix. Validation procedures are critical and all projects should address such topic.

F.1 Computational Methods (e.g. FEA or CFD Modeling, example sub-section)

F.2 Experimental Methods (example subsection)

F.3 Architecture/High-level Design (example subsection)

F.5 Validation Procedure

Describe how the design team will validate that the final design meets the client's needs. This section should include a plan to meet with the client towards the end of the project to discuss final design details and demonstrate a prototype, experimental test, and/or simulation results. Provide a relative time frame for this validation to occur (e.g. "mid-March" or "early-April"). Include a brief discussion on how client feedback will be captured, such as a formal survey, interview, or observation notes of the client using the prototype. It may also include plans to solicit feedback from other stakeholders and/or potential users.

Section G. Results and Design Details

Use this section to highlight the major results of the design methodology described above including important analytical, computational, experimental, modeling, assembly, and testing results. This section should be one of the most substantial sections of the report showcasing all of the hard work and effort that went into the completion of the final design and delivery of the project deliverables. Show how the identified problem was solved.

Highlight the prominent features of the final design through analysis results, modeling, drawings, renderings, circuit schematics, instrumentation diagrams, flow and piping diagrams, etc. to show that the design functions as intended and meets all design objectives and constraints. Overview designs such as dataflow diagrams, process flow, swim lane diagrams, as well as presentation-layer designs (e.g. storyboards for front-ends) should be included here. Detailed designs such as database designs, software designs, procedure flowcharts, or pseudocode should be included here. Support computational and experimental results with key plots and figures. All supporting figures should be clearly labeled and annotated to highlight the most important points of the figure (i.e. explicitly point out what the reader should focus on or understand about the image).

Note that while all results should be used to help inform design decisions, not all results may be necessary to include in the main body of the report. Extraneous supporting results (e.g. graphs, data, design renderings, drawings, etc.) that are not necessary for presenting the fundamental findings can be placed in one or more appendices. Detailed documentation of each program module can be provided as appendix.

G.1 Modeling Results (example subsection)

G.2 Experimental Results (example subsection)

G.3 Prototyping and Testing Results (example subsection)

G.4. Final Design Details/Specifications (example subsection)

Note that while the design constraints and specifications may have provided minimum or maximum values, or ranges or values, that the design needed to meet, the final design specifications should be listed here showing that the required design values were met. A list of final design details can also be included demonstrate fulfillment of the design objectives.

Note: Preliminary results should be included in the Preliminary Design Report to show the progress made of the selected design concept to-date. This section should be updated for the Final Design Report to include documentation of all of the work that was completed on the project throughout the entirety of the academic year.

Section H. Societal Impacts of Design

In addition to technical design considerations, contemporary engineers must consider the broader impacts that their design choices have on the world around them. These impacts include the consideration of public health, safety, and welfare as well as the potential societal, political/regulatory, economic, environmental, global, and ethical impacts of the design. As appropriate for the project design, discuss how each of these considerations influenced design choices in separate subsections. How will the design change the way people interact with each other? What are the political implications of the design? Does the technology have the potential to impact or shift markets? Does the design have any positive or negative effects on the environment? Don't forget to consider unintended consequences such as process or manufacturing byproducts. What impacts might the design have on global markets and trade? Are there any ethical questions related to the design?

While it is hard to forecast the various impacts of a technology, it is important to consider these potential impacts throughout the engineering design process. When considered during the early stages of the design phase, consideration of these impacts can help determine design objectives, constraints, and specifications and help drive design choices that may mitigate any potential negative impacts or unintended consequences.

Note: A minimum of 4 of these design considerations, including the consideration of public health, safety, and welfare, are required for the Preliminary Design Report while a section for all considerations must be included in the final design report.

H.1 Public Health, Safety, and Welfare

Provide a list of all design safety features and provide a brief description of each. Discuss the potential effects the design may have on public health, safety, and welfare. References to the codes and standards previous provided and the organizations that produced them may be summarized or referenced here.

H.2 Societal Impacts

H.3 Political/Regulatory Impacts

H.4. Economic Impacts

H.5 Environmental Impacts

H.6 Global Impacts

H.7. Ethical Considerations

Section I. Cost Analysis

Provide a simple cost analysis of the project that includes a list of all expenditures related to the project. If an experimental test set-up or prototype was developed, provide a Bill of Materials that includes part numbers, vendor names, unit costs, quantity, total costs, delivery times, dates received, etc. Do not forget to include all manufacturing costs incurred throughout the completion of the project. If the design is expected to become a commercial product, provide a production cost estimate including fixed capital, raw materials, manufacturing (including tooling and/or casting), and labor costs to produce and package the device. Note that this type of detailed cost analysis may be listed as a project deliverable.

Note: The Preliminary Design Report should include all costs incurred to date. It is expected that this section will be expanded and updated between the preliminary and final design reports.

Section J. Conclusions and Recommendations

Use this section to summarize the story of how the design team arrived at the final design. Focus on the evolution of the design through the use of the engineering design process including lessons learned, obstacles overcome, and triumphs of the final design. Revisit the primary project goals and objectives. Provide a brief summary of the final design details and features paramount to the function of the design in meeting these goals and objectives.

A discussion may be included to discuss how the design could be further advanced or improved in the future. If applicable, summarize any questions or curiosities that the final results/design of this effort bring to mind or leave unanswered. If this project might continue on as a future (continuation) senior design project, detail the major milestones that have been completed to date and include any suggested testing plans, relevant machine drawings, electrical schematics, developed computer code, etc. All relevant information should be included in this section such that future researchers could pick up the project and advance the work in as seamless a manner as possible. Documents such as drawings, schematics, and codes could be referenced here and included in one or more appendix. If digital files are critical for future work, they should be saved on a thumb drive, external hard drive, cloud, etc. and left in the hands of the project advisor and/or client.

Appendix 1: Project Timeline

Provide a Gantt chart of similarly composed visual timeline showing the start and end dates of all completed tasks and how they are grouped together, overlapped, and linked together. Include all senior design requirements including design reports and Expo materials (i.e. Abstract, Poster, and Presentation). All major milestones should be included in the timeline.

Appendix 2: Team Contract (i.e. Team Organization)

Copy and paste the content from the completed Team Contract here starting with Step 1 of the Team Contract and including all content following the ‘Contents’ list.

Appendix 3: [Insert Appendix Title]

Note that additional appendices may be added as needed. Appendices are used for supplementary material considered or used in the design process but not necessary for understanding the fundamental design or results. Lengthy mathematical derivations, ancillary results (e.g. data sets, plots), and detailed mechanical drawings are examples of items that might be placed in an appendix. Multiple appendices may be used to delineate topics and can be labeled using letters or numbers. Each appendix should start on a new page. Reference each appendix and the information it contains in the main text of the report where appropriate.

Note: Delete this page if no additional appendices are included.

References

Provide a numbered list of all references in order of appearance using APA citation format. The reference page should begin on a new page as shown here.

[1] VCU Writing Center. (2021, September 8). *APA Citation: A guide to formatting in APA style.* Retrieved September 2, 2024. https://writing.vcu.edu/student-resources/apa-citations/

[2] Teach Engineering. *Engineering Design Process*. TeachEngineering.org. Retreived September 2, 2024. https://www.teachengineering.org/populartopics/designprocess

[3] *The Evolution and History of AI Chatbots - Just Think AI*. (n.d.). Www.justthink.ai. https://www.justthink.ai/blog/the-evolution-and-history-of-ai-chatbots