Requirements Document

JabberJack

Team Member:

Sara Harris, Tyler Zimmerman, Jiasheng Yang, Gabriel Proctor

Team Mentor: Felicity H. Escarzaga

Client: Dr. Andy Wang

Northern Arizona University

Accepted a baseline for the project:

Client Signature: _(signature)_ 12/7/2021

Team Signatures: _JabberJack_

3 December, 2021

Contents

1.0 Introduction

JabberJack

Chatbots and other helpful bots are becoming increasingly more common throughout both the world of business and within academia [1]. Websites use chatbots to answer simple questions that visitors may have about products or services. There are even chatbots that are present on some Northern Arizona University (NAU) websites such as the one present on the Campus Health Services page [2]. Chatbots are not as complex as many people believe them to be; chatbots are not able to hold a conversation or answer extremely complex questions[1]. Even though chatbots cannot interact with humans in a human-like manner, there are a great deal of research teams working to improve the interactivity between chatbots and humans[1]. Chatbots can either be simple such as rule based chatbots which operate based on a number of rules, or they can be complex with artificial intelligence (AI) such as chatbots similar to Amazon's Alexa and Apple's Siri which learn from interactions to improve future interactions. The world runs on data, and the more data that a chatbot has available to it, the more helpful they become [4].

Chatbots are popular for customer consumption too, meaning that a great deal of chatbots can be found within regular people's homes. Amazon's Alexa is by far the most popular and the company holds about 62% of the market share within this industry [5]. There are also other options such as Apple's HomePod and Google's Google Home hold the remaining market shares of the industry. All of these chatbots are built utilizing complex computing called artificial intelligence. While artificial intelligence (AI) is the most popular type of chatbot, the more common chatbots are the ones that people can find located on websites. These chatbots are usually basic and only hold a small amount of information with little to no learning ability. They can answer basic questions about the website or about the company's service or product, so the company does not need to employ a real person for that role.

Our client Dr. Andy Wang is the dean of NAU's college of engineering, informatics, and applied sciences (CEIAS). Dr. Wang maintains and runs the programs contained within CEIAS and facilitates growth for both the college and programs in CEIAS. Dr. Wang is in the business of computing, engineering, and most importantly people. Dr. Wang deals with lots of people on a daily basis and is interested in helping them find answers to their questions efficiently.

2.0 Problem Statement

JabberJack

Dr. Andy Wang manages the entire college of engineering, informatics, and applied sciences; he wants the building to operate as efficiently as possible. The engineering building at Northern Arizona University has faculty that do all sorts of tasks from teaching classes to assisting students in the creation of their schedules. There are also frequent tours that come through the building that are given to prospective students of CEIAS programs. The operations of the building have to function every day in order to keep the college running smoothly.

Visitors and students of the college alike always have questions; sometimes these questions cannot even be answered by an internet search. Unfortunately, the answers are not found as easily. Northern Arizona University has many websites that have similar information, but sometimes the answer is unclear and difficult to find. Parents and students have to scour the internet for the answers to the questions that they have; the only other option to get this information is to either ask a faculty member or email the department. While asking a faculty member a question is not inherently problematic it takes time away from their normal day to day tasks to answer these questions. Even finding someone to answer the question is difficult, it takes too long and can leave parents, students, and prospective students waiting for hours and even days.

The severe scattering of information and the sheer amount of time it could take for questions to be answered causes the day to day flow of the engineering building/department to be interrupted. This lowers the efficiency of the department as a whole and hurts the flow of the entire operation and can even affect the retention of prospective students. Some specific issues are:

- Questioning faculty takes time away from the job that they were hired for and can cause them to trail behind in their work
- Information is scattered and can be contradicting at times; A basic search, unless extremely specific will not usually give a straight answer

In the next section, this document will describe a solution that will solve all of these issues that are hindering the flow of information. This solution could allow the college to run more smoothly and leave visitors and students with their questions answered.

3.0 Solution Vision

JabberJack

To enable CEIAS faculty and staff to complete their daily tasks without interruption Team JabberJack will create a chatbot called ChatterJack to answer questions that students, visitors, and prospective students have. The ChatterJack will be a rule-based model chatbot and will answer text inputted questions of users based on a preconceived set of question/answer pairs within a database. The ChatterJack will be localized on a system but will have the ability to connect to the internet (server) in order to access an updatable database that will contain the question/answer pairs. This will solve both the issue of unanswered questions as well as Northern Arizona University's problem of scattered information, improving both the lives of visitors as well as staff. In order to solve the problem comprehensively, the ChatterJack will meet these core features:

- A centralized updatable database containing at minimum 500 question/answer pairs
- Reasonable, quick, and intelligent responses to users' questions via text and audio
- User Authentication and information storage

_____Each core feature will integrate with one another to create a fully functional chatbot system. Figure 3.1 The ChatterJack Workflow shows how the ChatterJack is capable of giving an answer when a user asks a question.
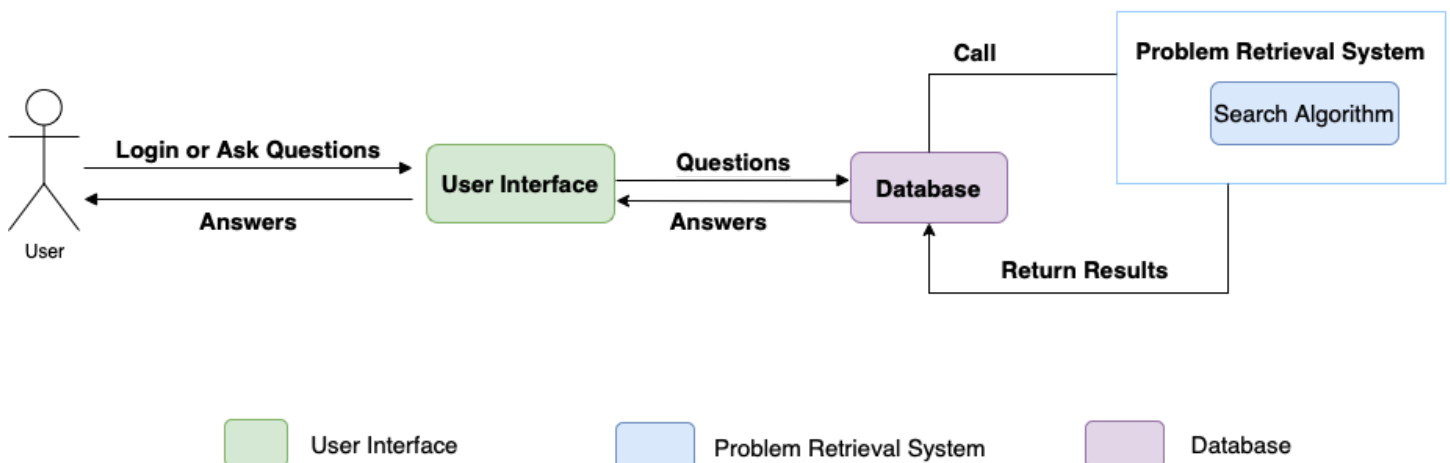

Figure 3.1 The ChatterJack Workflow

The system will use a database that has manually entered question and answer pairs. All questions will only pertain to Northern Arizona University, so the ChatterJack will not know animal facts or answers to non-scope questions. There is a simple flow of data from the ChatterJack to the user; when the user prompts the ChatterJack by asking a question, it will search the saved database to find a corresponding answer. The answer will flow from the database to the user interface and then be displayed to the user. The ChatterJack will allow staff to focus on their work because visitors will not ask them questions; instead, they will ask the ChatterJack. The solution will also centralize much of the scattered information about Northern Arizona University so that there will be a one-stop shop for all NAU questions. It will also allow visitors/students to get the questions they have answered in a timely manner without tracking someone down or emailing a department and waiting; visitors/students will receive answers to the questions that sometimes go unanswered. This project, while it may not revolutionize the academic chatbot industry, will set the foundation for a chatbot system at Northern Arizona University.

4.0 Project Requirements

JabberJack

In order to create an intelligent rule-based model chatbot, JabberJack must focus on a set of project requirements that should target higher-level key features of the solution from the client perspective. The followings are the core features:

- A centralized updatable database containing at minimum 500 question/answer pairs
- Reasonable, quick, and intelligent responses to users' questions via text and audio
- User Authentication and information storage

JabberJack will meet all above functions by the project requirements by Spring 2022. These requirements will be divided into three different categories: functional requirements, non-functional requirements, and environmental requirements.

4.1 Functional Requirements

The ChatterJack should be capable of meeting the following functional requirements that define its system capabilities.

4.1.1 User Level Requirements

These requirements are the things that users need to be able to do with the software. The users of this chatbot will be students, faculty, university visitors, and database administrators. Users will be able to:

- Type their message using a text based interface
- Speak their message using an audio based interface
- View and hear the chatbot's response
- Submit feedback as to how helpful the answer was
  - Yes, it was helpful
  - No, it was not helpful
  - Unsure, don't know if the answer was helpful or not
- Edit Q/A pairs within the database (As database administrators)

4.1.2 User Interface/Frontend

The user interface will be the portion of the chatbot where users can interact with the software itself. There are two parts that need to be addressed regarding the user interface:

- User to chatbot question/answer interaction interface
- Administration/Management to database interaction interface

These will look different as they are not performing the same functions.

The user side must have a space for an input of questions. This is the most important function of the user interface as this interface is the face of the chatbot. Users will be able to input their question textually into a chatbot. There is also a secondary option for user input as the chatbot will be able to accept voice input. There will also be a way to alert the chatbot that the user's input is complete and that the chatbot should start searching for the answer. There will be a prompt indicating that the chatbot is searching for the answer. Once an answer has been found the chatbot will display the answer to the user in an appropriate manner.

- Text box for user textual input
- Prompt for spoken questions
  - Able to view the entire question at once
  - If speech input, users can edit text
- Question Submit button
- A "thinking" indication for the search
- "Pop up" for the answer
  - Displays tables when appropriate
  - Textual answers when appropriate
  - Will speak answers

As for the administration side there will be more elements as administrators are able to do more. For starters there will be a login screen that will contain two input fields; one for the username and one for the password. Also on this login screen there will be an enter or okay button to alert the system that the administrator is finished inputting information. Once the admin is logged in, there will be another screen that displays all of the actions administrators can take. These actions will be:

- Create new user accounts
  - Page to input information
- View the system's database of questions and answers (Q/A database)
  - Retrieve tables from the database
  - Display tables retrieved from database
- Add new question/answer pairs to the Q/A database
  - Page to input new question/answer pairs
- Remove question/answer pairs from the Q/A database
  - View the tables from the database
  - Delete Q/A pairs
- Edit existing question/answer pairs within the database
  - View the tables from the database
  - Edit information contained within the tables

4.1.3 Authentication and Security

Authentication and security play an important role in any secure system; the ChatterJack chatbot is no exception. The authentication will only occur on the server side of the application as there will be no login required for "normal" users of the chatbot. Administrators will need to be authenticated before they are able to access the chatbot's data. The data will be the **most** secure part of the chatbot as people, aside from admins, should not be able to edit the database.
- Ability for administrators to access the database
  - Includes viewing and editing
- Only authorized users should be allowed to access the server side

4.1.4 Database Management System

The database management system will need to be able to facilitate the introduction of new information as well as remove information pertaining to both question and answer pairs and administrative users and their passwords. These requirements can be broken down as:
- The ability to add a question/answer pair with administrator validation
- The ability to remove a question/answer pair with administrator validation
- The ability to add a user to the users database with administrator approval
- The ability to remove a user from the user database with administrator approval

4.1.5 Speech Conversion

The chatbot will be able to perform basic speech conversion practices. The purpose of these functionalities is to one, increase compatibility with the Mechanical Engineering project of the same name, and two, reduce the burden placed on the user. These functions will include:
- Text to speech conversion
- Speech to text conversion
- Natural language processing

4.2 Non-Functional Requirements

To create an intelligent chatbot the JabberJack must consider three attributes as the specific performance requirements to evaluate the functional requirements. These attributes include: *speed*, *security*, and *usability*. Here are the analysis of these requirements:

4.2.1 Speed

The ChatterJack chabot will match user questions to questions stored within the database which will take time. This will take a maximum of 10 seconds; if an answer is not found within that time period the chatbot will respond with an unknown answer statement. This is done in order to ensure that users are satisfied and do not become impatient.

4.2.2 Security

The ChatterJack chatbot will not allow unauthorized users to access any of the data contained within the database. An unauthorized user will just continue to get the same error and be unable to access the data. This will be done in order to protect both user data as well as the Q/A database.

4.2.3 Usability

The simplicity is mainly pertinent to the User Interface (UI). A simple UI will allow non tech savvy users to utilize it . The user interface should be so simple that anyone who can read and write will be able to use it to its full capacity. The chatbot should respond to questions in a perceivably polite tone; although the chatbot itself cannot be "polite" the language implemented is able to be written in such a tone.

4.3 Environmental Requirements

There are no explicitly stated environmental requirements requested by the client. This project should set a foundation for future chatbot projects at Northern Arizona University.

4.3.1 Operating System

Since there are no project restrictions put into place by the client there is only a single self-imposed requirement. The project will have the ability to run smoothly on every operating system. It will work on Windows, IOS (Apple), and Linux. There is a bit of ambiguity regarding the linux environment since there are so many different versions or distributions; the software should run smoothly on Ubuntu, an extremely popular distribution of linux.

4.3.2 Question & Answer Database

_____The question and answer database will contain all of our preconceived question and answer pairs. These question/answer pairs will only pertain to questions concerning Northern Arizona University and specific questions about the College of Engineering, Informatics, and Applied Sciences. This database will be easily updatable as administrators can manually update it through the online portal. The chatbot will be able to answer basic frequently asked questions about Northern Arizona University. This database will have at minimum 500 pairs so at it's conception it will be able to answer 500 **different** questions.

5.0 Potential Risks

JabberJack

A chatbot does not have too many inherently harmful risks to users. There are still concerns that need to be addressed to ensure that implementation is flawless.

- Unauthorized user access to the Q/A database
- Incorrect answer return to user
- Ensuring Q/A database contains correct and up-to-date information

While none of these risks may cause user death they are all still detrimental to the operations of the chatbot. These risks all have a low, moderate, or a high probability of occurring on the system.

Unauthorized user access to the system has a **moderate** chance of occurring. This is a concern since unauthorized users could access the Q/A database and can change question/answer pairs causing the data to be incorrect. These malicious actors would be able to change the data to whatever they want; this could be inappropriate and even mentally harmful to users of the chatbot. This risk is important to address since this chatbot is supposed to be helpful and should not cause any form of distress. The security measures taken will be as strict as possible in order to mitigate this risk. Administrators are the only users that should be able to access the information stored on the database and they will receive an email alerting them to any changes made within the database.

Returning incorrect answers to the user has a **high** chance of occurring as chatbots can misinterpret questions. While it does have a high chance of occurring the damage is likely to be minimal. If the question is "What time is the basketball game?" and the answer is incorrect then the student would be missing out on a school event. This is a problem since the student could potentially blame the chatbot for the incorrect answer that caused them to miss the game, this will cause bad publicity for the chatbot and could potentially lead to low usage. While this issue is detrimental to the chatbot it is easily avoided; administrators need to keep

up with updating information. This should not happen as long as administrators update the online database often and ensure that current information is accurate.

Administrators need to ensure that information is up-to-date and correct; this is another risk to the chatbot. Since all of the information is inputted manually it will become an issue of manpower and commitment. The probability of this occurring is **high**, it is valued at this as ensuring that the information is correct is the most crucial part of the chatbot. The chatbot would be useless if the information contained within its database is outdated. This team is unable to make such a commitment it becomes up to the client to ensure that updates happen. It is important for the same reasons as the risk above and also for the fact that we cannot ensure the continuation of maintenance on the database so it becomes likely that it will never be followed up on. To address this risk the chatbot will be able to do a few things like web scrape for answers it does not know, allow faculty access to their own information within the database, and will also allow users to give feedback regarding the accuracy of the answers.

The potential risks that the project faces could be extremely harmful to the chatbot's operations. However, with proper mitigation, these risks will not hinder the success of the chatbot.

6.0 Project Plan

JabberJack

The following figures, figures 6.1 and 6.2, offer a brief overview of the team's plan to design and implement the software. Figure 6.1 describes the schedule and process that occurred in the Fall semester of 2021; this is the primary documentation and preliminary design stage--starting from basic team standards and moving on to a more complex and detailed requirements document. The Fall 2021 semester will end with a complete demonstration of a prototype.

After the Fall 2021 semester, there is a short university break, winter break, where there is perceived to be no progress being made on the project itself. Then, as depicted in figure 6.2, the Spring semester picks back up in January 2022 where the team will focus on implementation and testing the software that was designed in Fall 2021. The team will follow this plan in order to completely design and implement the ChatterJack chatbot project.
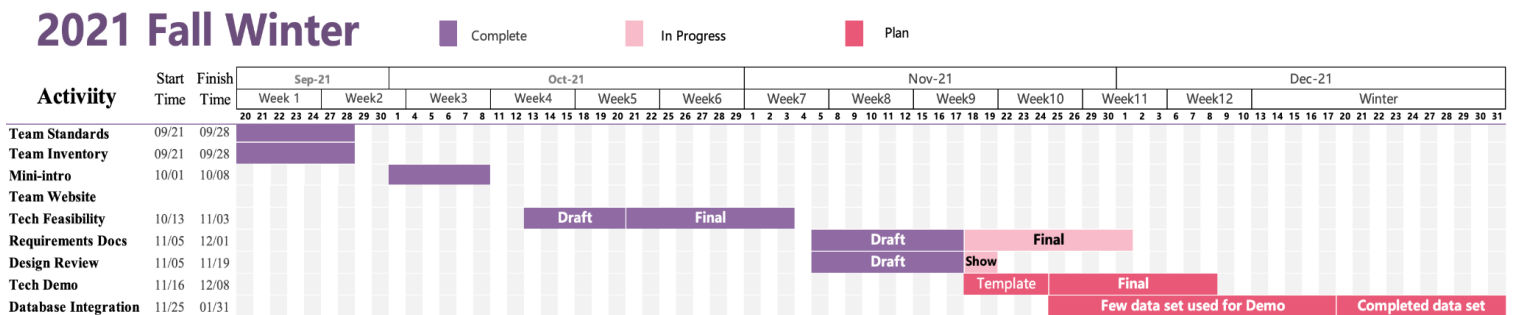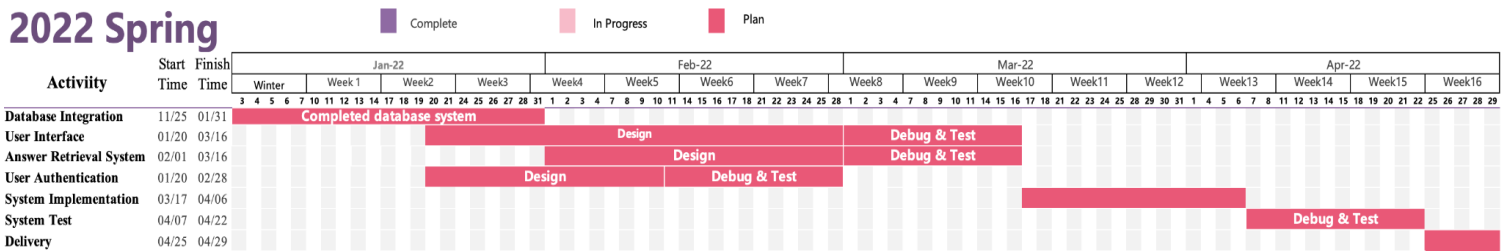
2021 Fall and Winter Semester Schedule:

Figure 6.1: 2021 Fall and Winter Semester Schedule

2022 Spring Semester Schedule:

2022 Spring

| Complete | In Progress | Plan |

Activity	Start Time	Finish Time	Jan-22				Feb-22				Mar-22				Apr-22			
			Winter	Week1	Week2	Week3	Week4	Week5	Week6	Week7	Week8	Week9	Week10	Week11 Week12	Week13	Week14	Week15	Week16
Database Integration	11/25	01/31	Completed database system															
User Interface	01/20	03/16			Design						Debug & Test							
Answer Retrieval System	02/01	03/16				Design					Debug & Test							
User Authentication	01/20	02/28			Design		Debug & Test											
System Implementation	03/17	04/06																
System Test	04/07	04/22														Debug & Test		
Delivery	04/25	04/29																

Figure 6.2: 2022 Spring Semester Schedule

Here we continue with a brief description of each of the activities that we plan to complete:

- **Database Integration (11/25/2021 - 01/31/2022)**
  As the database is the first step to implementing the key requirements, JabberJack will begin to collect and integrate the data set of Question and Answer (QA) from the end of 2021 Fall. This will be finished by January 1st.

- **User Interface (01/20 - 03/16)**
  JabberJack will start designing the user interface (UI) from January 20th. The main design should focus on the login window and chat window. This process will integrate the user authentication.

- **Answer Retrieval System(02/01 - 03/16)**
  JabberJack will design the optimal string-searching algorithm after collecting and integrating the database. The development and testing of the question retrieval system revolves around the saved database. JabberJack will allow about two weeks for debugging and testing.

- **User Authentication (01/20 - 02/28)**
  The implementation of user authentication will not take a very long time. And it will start at the same time as the design of the UI, but it should be completed before finishing the UI design.

- **System Implementation (03/17 - 04/06)**
  The system implementation will integrate *Database*, *Question Retrieval System*, *User Interface*, *User Authentication* together.

- **System Test (04/07 - 04/22)**
  JabberJack will test the system by motivating some conversations with the chatbot. Then the ChatterJack will be deployed on the NAU website by the dialog box iron button.

7.0 Conclusion

JabberJack

A university functions like a well-oiled machine, but even well-oiled machines get stuck sometimes and slow down. Information is the oil, and without it, the university would not run as efficiently. Currently, at Northern Arizona University, there is an information problem; there are too many websites with information, and answers to questions are buried under complex queries and portals. The ChatterJack chatbot could solve this problem; it will provide both a consolidation of information as well as quick answers to visitor/student questions. This document has outlined the problem and solution as well as some of the functional, non-functional, and environmental restrictions that the project will follow during the implementation phase. The project will have an intuitive user interface as well as a secure user authentication service for administrators. While the project will not reach Alexa or Siri levels of Q and A, the chatbot will have the ability to answer user questions about the CEIAS department quickly and accurately. This project will set a sound foundation for future capstone projects to build upon to truly create a well-rounded chatbot capable of helping students and staff at NAU.

8.0 References

JabberJack

---

[1] Miklosik, Andrej & Evans, Nina & Qureshi, Athar. (2021). The Use of Chatbots in Digital Business Transformation: A Systematic Literature Review. IEEE Access. 9. 106530-106539. 10.1109/ACCESS.2021.3100885.

[2] NAU. 2021.(2021). Retrieved November 3, 2021 from https://in.nau.edu/campus-health-services/

[3] NAU. 2021. Facts and stats. (2021). Retrieved November 3, 2021 from https://nau.edu/about/facts-and-stats/

[4] Rooein, Donya. (2019). Data-Driven Edu Chatbots. 46-49. 10.1145/3308560.3314191.

[5] Smilijanic Stasha. 2021. Amazon Alexa Statistics, facts, and Trends. (February 2021). Retrieved November 16, 2021 from https://policyadvice.net/insurance/insights/amazon-alexa-statistics/