Software Design Document

Pineapple Chatbot

Prepared by

Kuong Thong, Enrique Castillo, Julio Aguilar, Jose Lopez, Yilin Ruan

CSULA ITS Department

Version History

Version	Date	Summary of Changes
apple	12/4/18	Added Introduction sections (1.1–1.2).
banana	12/5/18	Added early DFD diagrams and transferred initial design content.
cantaloupe	12/6/18	Added database design and sections 8–9.1.
dates	12/7/18	Completed document draft and added validation sections.

Contents

1 Introduction

1.1 Purpose

Pineapple Chatbot is an informational chatbot for a standalone web application that helps users find information related to the CSULA website. The chatbot processes user queries, responds with useful answers, and can receive user feedback.

1.2 Document Conventions

This document follows a standard Software Design Document (SDD) structure with numbered sections and subsections.

1.3 Intended Audience

This document is for developers, testers, designers, project managers, and documentation staff.

1.4 System Overview

Pineapple Chatbot receives natural-language queries from users, analyzes the text to determine intent and relevant keywords, retrieves matching information from a backend data store, and returns an appropriate response.

2 System Architecture

2.1 Level 0 DFD

Figure 1: DFD – Level 0 for Pineapple Chatbot

At a high level, the system receives user input, processes it through internal modules, and returns a response.

2.2 Level 1 DFD

Figure 2: DFD – Level 1 for Pineapple Chatbot

The system consists of five major modules:

1. **Input Module** – Collects user inputs.
2. **Preprocessor Module** – Cleans and normalizes text.
3. **Logic Module** – Identifies user intent, keywords, and selects a response.
4. **Storage Module** – Stores and retrieves crawled data and responses.

5. **Output Module** – Formats and returns the final answer to the user.

Data Flow Summary

User query → Preprocessing → Intent recognition → Keyword lookup → Data retrieval → Response → User.

3 Database Design

A MySQL database stores the crawled website data and related information required by the chatbot.

Key Table: `termUrlKeywords` (approximately 300,000 records)

- **Columns:**
  - `term`
  - `URL`
  - `keywords`

Additional tables store faculty information and labels for the Bayesian classifier.

The database enables:

- Fast keyword matching
- Efficient URL lookup
- Long-term storage of crawled data

4 User Interface

The Pineapple Chatbot user interface is designed as a simple, web-based chat interface.

Overview

- Simple chat-style interface built with React.
- Input box for user queries.
- Response window for answers and links.

Future Enhancements

- Richer UI components (e.g., quick-reply buttons, suggested queries).
- Improved accessibility and compliance with web accessibility standards.

5 Glossary

- **Statement**: Object representing a processed user input.
- **Intent**: The interpreted purpose of a user query.

- **Similarity Score**: Metric used to choose the closest matching response.

6 References

Original Pineapple Chatbot design documents, tutorials for Python NLP libraries, and MySQL documentation.