Comets Engine: Cyberminer Search Engine

Final Project Plan

Team Members:

Shanjida Khatun (sxk200130@utdallas.edu)
Miao Miao (mxm190091@utdallas.edu)
Mohit Anand (mohit.anand@utdallas.edu)
Zilong Wang (zxw200004@utdallas.edu)
Jeya Visshwak Jeyakumar (jxj190055@utdallas.edu)
Mitha Alshammary (mi.alshammary@utdallas.edu)

Team Website:
https://personal.utdallas.edu/~zxw200004/SA/SE6362_21F.html

For SE6362.001
Professor: Dr. Lawrence Chung

1. Introduction

1.1 Project Overview

1.1.1 Phase 1:

The project described within this document for phase 1 is a system designed with the specific goal of creating architecture for a simple KWIC software system and implementing it. An Object-Oriented architectural style will be used for this project, and, in the first phase, the architecture will be implemented as a Java applet (or an equivalent). The resulting implementation will be made accessible through the team's web page.

As its major functional requirements, the KWIC index system shall accept an ordered set of lines, where each line is an ordered set of words, and each word is an ordered set of characters. Any line shall be "circularly shifted" by repeatedly removing the first word and appending it at the end of the line. The KWIC index system shall output a listing of all circular shifts of all lines in ascending alphabetical order.

It is also required that the KWIC system be easily understandable, portable, enhanceable and reusable with good performance. The KWIC system shall also be user-friendly, responsive, and adaptable. These form the non-functional requirements for the projects. The team is responsible for analyzing these requirements to detect any problems that may exist with the requirements.

This document gives a preliminary plan for how the company aims to achieve the above stated aims. The first section gives an overview, describes project deliverables, and itemizes the evolution of this document. Lastly the first section gives the meaning of acronyms that may be encountered in the rest of the document and lists references from which guidelines have been drawn. In the second section, the organizational structure of the executing team is given, and the third section shows how the team as well as the project will be managed from inception to completion. Technical Processes used are described in Section 4 while the fifth and final section details the work elements, schedule, and budget for the project.

1.1.2: Phase 2:

The project described within this document for phase 2 is a web search engine, Cyberminer, designed with the specific goal of creating architecture for it and implementing it. An Object-Oriented architectural style will be used for this project, in the second phase, the architecture will be implemented as a website. The resulting implementation will be made accessible through the team's website.

Cyberminer, as a web search engine, shall accept a list of keywords and return a list of URLs whose descriptions contain any of the given keywords. Cyberminer shall use another software system, the KWIC system, which we created in phase 1, as a component, to efficiently maintain a database of URLs and

the corresponding descriptions. The KWIC system shall accept an ordered set of lines, where each line consists of two parts: the URL part and the descriptor part. The descriptor part of any line shall be "circularly shifted" by repeatedly removing the first word and appending it at the end of the line. The KWIC index system shall output a listing of all circular shifts of the descriptor parts of all lines in ascending alphabetical order, together with their corresponding URLs. No line in the output list shall start with any noise work such as "a", "the", and "of". The KWIC system shall allow for two modes of operation: i) for building initial KWIC indices; and ii) for growing the indices with later additions.

Cyberminer also shall allow for some other features such as: Case sensitive, the system shall store the input as given and retrieve the input also as such; Hyperlink enforcement: When the user clicks on the URL, which has been retrieved as the result of a query, the system shall take the user to the corresponding web site; Deletion of out-of-date URL and corresponding description from the database.

It is also required that Cyberminer shall be easily understandable, portable, enhanceable and reusable with good performance. Cyberminer shall also be user-friendly, responsive, and adaptable. Also, Cyberminer shall run with recent versions of popular web browsers. These form the non-functional requirements for the projects. The team is responsible for analyzing these requirements to detect any problems (including, but not limited to inconsistencies, ambiguities, and incompleteness) that may exist with the requirements.

This document gives a preliminary plan for how the company aims to achieve the above stated aims. The first section gives an overview, describes project deliverables, and itemizes the evolution of this document. Lastly the first section gives the meaning of acronyms that may be encountered in the rest of the document and lists references from which guidelines have been drawn. In the second section, the organizational structure of the executing team is given, and the third section shows how the team as well as the project will be managed from inception to completion. Technical Processes used are described in Section 4 while the fifth and final section details the work elements, schedule, and budget for the project.

1.2 Project Deliverables

Deliverable Name	Description	Date of Completion
Preliminary Project Plan	High-Level project document which includes scope of the project, timelines, and team members etc.	09/02/21

Project I Presentation	PPT of preliminary project definition	09/30/21
Final Preliminary Definition	Requirements are stated in this document clearly in a systematic way.	10/12/21
Detailed Design	This document contains a more detailed description of the processing logic and data structures such as pseudocode	10/19/21
Source Code	Source Code of the application	10/26/21
Test Plan	This document describes the scope and approach of testing activities	11/12/21
User Manual	This document contains how to access and use the system	11/20/21
Final Project Plan	Project Plan which includes end to end details right from scope of the project to summary	11/25/21
Project II Presentation	PPT of Project II	11/29/21

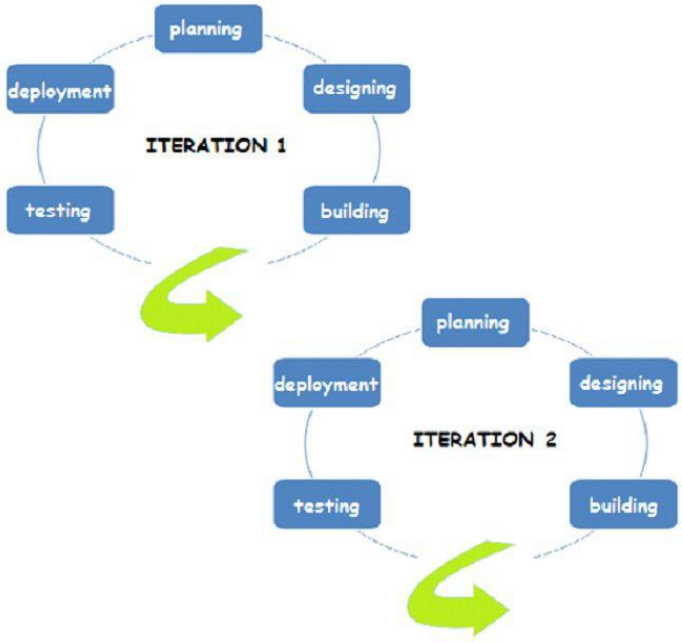
1.3 Evaluation of the Project Plan

Date	Description	Member(s) Involved
09/01/21	Preliminary project description done.	All
10/11/21	Updated project description for Project 1.	All
10/11/21	Updated PowerPoint slides for Interim Project	All

2. Project Organization

Process Model

We will be using Agile methodology where the product will be built in two phases with a combination of iterative and incremental process models. We as a team will be working simultaneously in each iteration that involves planning, design, develop, test, and deploy. With each phase the product deliverables will be improved and updated for missing areas.



2.2 Organizational Structure

Instead of having a single team lead, we will rotate the team leadership role each week. The team leader for a given week will be responsible for reserving the meeting room, coming up with the week's agenda, keeping notes of the meeting, and leading the meeting.

2.3 Organizational Boundaries and Interfaces

Team leaders during each phase will be responsible for coordinating team meetings, updates, communications, and team deliverables.

2.4 Project Responsibilities

All the team members will be involved in all phases of the project life cycle. Each team member will be contributing to each deliverable mentioned in section 1.2, according to the roles below:

- Project Plan - Project Manager
- Requirement Specification - Project Team
- Design - Team
- Develop - Team
- Test - Team

3. Managerial Process

3.1 Management objectives and priorities

3.1.1. Objectives

3.1.1.1. Keep records of all documents and work done (realized through version control systems).
3.1.1.2. Make sure every team member is on the same page about goals and progress of the project.
3.1.1.3. Make sure every team member is aware of all aspects of the project.
3.1.1.4. Ensure tasks and goals are clear.

3.1.2. Priorities

3.1.2.1. Finish all work in advance of deadlines.
3.1.2.2. Working efficiently, through smart use of organization.
3.1.2.3. Taking advantage of each member's individual skills and experience.
3.1.2.4. Starting all work early and spacing it out over a long period of time.

3.2 Assumptions, Dependencies, and Constraints

3.2.1 All members have a basic understanding of software engineering processes, and experience writing functional code.
3.2.2 All members are familiar with word processors, version control systems, and standard IDEs.
3.2.3 All members must be available for meetings.
3.2.4 All members will contribute in equal measure to the completion of the project.

3.3 Risk management

3.3.1 **Risk:** Accidental or malicious deletion of data.
Mitigation: Version control.

3.3.2 **Risk:** Member contracts illness or is incapacitated.
Mitigation: All resources are online, so remote work is possible.

3.3.3 **Risk:** Member's computer breaks.
Mitigation: All members bring laptops to meetings, and all resources are online, so any computer is usable.

3.3.4 **Risk**: Miscommunication between team members.
Mitigation: All communication is performed in group forums.

3.4 Monitoring and Controlling Mechanisms

Work logs on files, so that who works on what can be determined.

4. Technical Process

4.1 Methods, Tools, and Techniques

i. Rational Rose will be used to describe the constituents of the architecture including components, connections, and constraints.
ii. Java 2 Enterprise Edition (J2EE) will be used to develop the KWIC Search Engine. Microsoft Word and Presentation will be used for documentation and presentation slides, respectively.
iii. We will use Microsoft Word for mockup.
iv. We will make a Java Applet.
v. We will use Microsoft Word for documentation and Google Slides for slide preparation.

4.2 Software Documentation

Documentation will be written before any code is written, to facilitate better understanding of the requirements and constraints before any code is written. Documentation will be written in Microsoft Word. Test Driven Development will be used in writing as much of the codebase as possible to ensure proper functionality is well documented and defined for all code units.

4.3 Project Support Functions

Data and functional validation will be performed together in a group meeting. Project communication will be handled via Google Groups.

5. Work elements, Schedule, and Budget

5.1 Work Elements

Phase	Phase Deliverable	Start Date	End Date	Resources
Preliminaries	Preliminary Software Project Management Plan	8/26/21	9/1/21	Microsoft Word
Phase 1.1	Interim KWIC Requirements & Architectural Specification Document & User Manual, Implementation Report	9/9/21	9/29/21	Microsoft Word
Phase 1.2	Revised Software Project Management Plan	10/7/21	10/11/21	Microsoft Word
	KWIC System Prototype			
Phase 2.1	Interim Revised Requirements & Architectural Specification Status Report	10/21/21	11/15/21	Microsoft Word
Phase 2.2	Final Cyberminer Requirements & Architectural Specification Document & User Manual	11/18/21	11/25/21	Microsoft Word
	Cyberminer Search Engine Prototype	11/25/21	11/26/21	

5.2 Schedule

Deliverable	Due Date
Preliminary Project Plan	09/02/21
Interim Project I	09/30/21
Final Project I	10/12/21
Interim project II	11/16/21
Final Project II	11/30/21

5.3 Budget

This project does not have a budget. All the team members are students working for free. Also, the licenses for the development tools and computational resources are provided by UTD.