|  |  |  |
| --- | --- | --- |
| **D:\PROJECT\rabiproject\DOCUMENT\images.jpg** |  | **docu0005** |

**EXPLORATORY SEARCH FOR RETRIEVING UNAWARE FIELDS FOR USERS USING ONTOLOGY CLUSTERING**

**A PROJECT REPORT**

***Submitted by***

**M.CHELLAMMAI (211413205025)**

**S.PREETHI (211413205074)**

**M.SAVITHA (211413205092)**

***in partial fulfillment for the award of the degree***

***of***

**BACHELOR OF TECHNOLOGY**

***in***

**INFORMATION TECHNOLOGY**

**PANIMALAR ENGINEERING COLLEGE, POONAMALLEE**

**ANNA UNIVERSITY : CHENNAI 600 025**

**APRIL 2017**

**ANNA UNIVERSITY: CHENNAI 600 025**

**BONAFIDE CERTIFICATE**

Certified that this project report **“EXPLORATORY SEARCH FOR RETRIEVING UNAWARE FIELDS FOR USERS USING ONTOLOGY CLUSTERING”** is the bonafide work of**“M.CHELLAMMAI(211413205025) S.PREETHI(211413205074)** and **M.SAVITHA(211413205092)”** who carried out the project under my supervision.

**SIGNATURE SIGNATURE**

**Dr. M. HELDA MERCY M.E., Ph.D., Mrs. R.MANJULA M.E., (Ph.D.,)**

**HEAD OF THE DEPARTMENT SUPERVISOR**

Associate Professor

Department of Information Technology Department of Information

Technology

Panimalar Engineering College Panimalar Engineering College

Poonamallee, Chennai - 600 123 Poonamallee, Chennai - 600 123

Submitted for the project and viva-voce examination held on \_\_\_\_\_\_\_\_\_\_\_\_\_\_

**SIGNATURE SIGNATURE**

**INTERNAL EXAMINER EXTERNAL EXAMINER**

**DECLARATION**

I hereby declare that the project report entitled “**EXPLORATORY SEARCH FOR RETRIEVING UNAWARE FIELDS FOR USERS USING ONTOLOGY CLUSTERING** ” which is being submitted in partial fulfilment of the requirement of the course leading to the award of the ‘Bachelor of Technology in Information Technology ’ in **Panimalar Engineering College, Affiliated to Anna University- Chennai** is the result of the project carried out by me under the guidance and supervision of **Mrs. R.MANJULA M.E., (Ph.D.,) Associate Professor in the Department of Information Technology**. I further declared that I or any other person has not previously submitted this project report to any other institution/university for any other degree/ diploma or any other person.

Date**:**

Place**:** Chennai Signature of batch members

(M.CHELLAMMAI)

(S.PREETHI)

(M.SAVITHA)

It is certified that this project has been prepared and submitted under my guidance.

Date: **Mrs. R.MANJULA M.E., (Ph.D.,)**

Place: Chennai (Associate Professor / IT )

**ACKNOWLEDGEMENT**

A project of this magnitude and nature requires kind co-operation and support from many, for successful completion . We wish to express our sincere thanks to all those who were involved in the completion of this project.

We would like to express our deep gratitude to **Our Honorable Secretary and Correspondent, Dr. P. CHINNADURAI, M.A., Ph.D.,** for his kind words and enthusiastic motivation which inspired us a lot in completing this project.

We also express our sincere thanks to **Our Respected Directors Mrs. C. VIJAYA RAJESHWARI and Mr. C. SAKTHI KUMAR, M.E.,** for providing us with the necessary facilities for completion of this project.

We also express our appreciation and gratefulness to **Our Principal Dr. K. MANI, M.E., Ph.D.,**  who helped us in the completion of the project. We wish to convey our thanks and gratitude to our head of the department, **Dr. M. HELDA MERCY, M.E., Ph.D.,** Department of Information Technology, for her support and by providing us ample time to complete our project.

We express our indebtedness and gratitude to our staff in charge, **Mrs. R.MANJULA M.E.,(Ph.D.,)** Associate Professor, Department of Information Technology for her guidance throughout the course of our project.

We thank our parents and friends for providing their extensive moral

Support and encouragement during the course of the project.

**ABSTRACT**

As deep web grows at a very fast pace, there has been increased interest in techniques that help efficiently locate deep-web interfaces. An exploratory search may be driven by a user’s curiosity or desire for specific information. When users investigate unfamiliar fields, they may want to learn more about a particular subject area to increase their knowledge rather than solve a specific problem. a matching query style has significant limitations. Search results are satisfactory only when users give the right search words. To achieve more accurate results for a Exploratory search, Smart Crawler ranks websites to prioritize highly relevant ones for a given topic. This will be achieved by using query which is submitted to the application will be preprocessed, after pre-processing only root words will be taken and it will find Synonym, Hypernym and Hyponym and it will listed to the user. So this is the reason that all possible links can be found related to search. If any words in that displayed list are selected then all the website links, images and news feeds will be given as final output to the user. Then the book mark concept is included that is the book marked link will be added to the application directly not to the browser so the bookmarked content will visible globally.

**LIST OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| **CHAPTER NO.** | **TITLE** | **PAGE NO** |
|  | **ABSTRACT** |  |
|  | **LIST OF TABLES** |  |
|  | **LIST OF FIGURES** |  |
|  | **LIST OF ABBREVIATIONS** |  |
| **1** | **INTRODUCTION** | **1** |
|  | 1.1. OVERVIEW OF THE PROJECT | 1 |
|  | 1.2 SCOPE OF THE PROJECT | 1 |
|  | 1.3 OBJECTIVE OF THE PROJECT | 2 |
|  | 1.4 EXISTING SYSTEM  1.5 NEED FOR THE SYSTEM  1.6 PROPOSED SYSTEM | 2  3  3 |
| **2** | **LITERATURE SURVEY** | **4** |
|  | 2.1 SURVEYS  2.2 FEASIBILITY STUDY | 4  7 |
| **3** | **SYSTEM DESIGN SPECIFICATION** | **8** |
|  | 3.1 SYSTEM ARCHITECTURE DESIGN | 8 |
|  | 3.2 UML DIAGRAMS | 18 |
|  | 3.2.1. Use case diagram | 18 |
|  | 3.2.2 Sequence diagram | 19 |
|  | 3.2.3 Class diagram | 20 |
|  | 3.2.4. State chart diagram  3.2.5 Collaboration diagram  3.3 OVERVIEW OF SYSTEM DIAGRAM | 21  22  23 |
| **4** | **MODULE DESCRIPTIONS** | 25 |
|  | 4.1 MODULES  4.2 MODULE DESCRIPTION  4.2.1 User interface  4.2.2 Data pre processing  4.2.3 Ontology clustering  4.2.4 Multi term search  4.2.5 Clustering the most relevant content  4.2.6 Bookmarks | 25  26  26  27  28  29  29  29 |
| **5** | **REQUIREMENT SPECIFICATION** | **30** |
|  | 5.1 HARDWARE REQUIREMENTS | 30 |
|  | 5.2 SOFTWARE REQUIREMENTS  5.3 LANGUAGE SPECIFICATIONS | 30  31 |
| **6** | **IMPLEMENTATION** | **35** |
|  | 6.1 SAMPLE CODE | 35 |
|  | 6.2 SAMPLE SCREEN SHOTS | 40 |
| **7** | **TESTING AND MAINTENANCE** | **54** |
|  | 7.1 TESTING OBJECTIVE | 54 |
|  | 7.2 TEST CASE DESIGN | 55 |
|  | 7.3 TESTING STRATEGIES | 56 |
|  | 7.3.1 Integration Testing | 57 |
|  | 7.3.2 White Box Testing  7.3.3 Black Box testing | 57  57 |
|  | 7.3.4 Interface Testing  7.3.5 Module Testing  7.3.6 Smoke Testing  7.4 MAINTENANCE  7.5 SCREENSHOTS FOR TESTING  7.6 TEST CASE TEMPLATE | 57  57  58  58  59  63 |
| **8** | **CONCLUSION AND FUTURE ENHANCEMENTS** | **67** |
|  | 8.1 CONCLUSION | 67 |
|  | 8.2 FUTURE ENHANCEMENT | 67 |
| **9** | **REFERENCES** | **68** |
| **10** | **APPENDIX** | **70** |

**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| **Figure No.** | **Name Of the Figure** | **Page No.** |
| 3.1 | Proposed System Architecture | 8 |
| 3.2.1 | UML Use Case Diagram | 18 |
| 3.2.2 | UML Sequence Diagram | 19 |
| 3.2.3 | UML Class Diagram | 20 |
| 3.2.4 | UML State chart diagram | 21 |
| 3.2.5 | UML Collaboration diagram | 22 |
| 3.3 | Overall system diagram | 23 |
| 6.2.1 | Screen Shot For homepage | 40 |
| 6.2.2 | Screen Shot For registration | 41 |
| 6.2.3 | Screen Shot For login | 42 |
| 6.2.4 | Screen Shot For search engine | 43 |
| 6.2.5 | Screen Shot for exploratory search | 44 |
| 6.2.6 | Screen Shot For websites | 45 |
| 6.2.7 | Screen Shot for bookmarking | 46 |
| 6.2.8 | Screen Shot which displays bookmark manager | 47 |
| 6.2.9 | Screen Shot About us | 48 |
| 6.2.10 | Screen Shot For enquiry | 49 |
| 6.3.1 | Screen Shot for Ontology database | 50 |
| 6.3.2 | Screen Shot for Bookmark database | 51 |
| 6.3.4 | Screen Shot for user database | 52 |
| 6.3.5 | Screen Shot for keywords database | 53 |
| 7.5.1 | Screen Shot for recording | 60 |
| 7.5.2 | Screen Shot for test runner | 61 |
| 7.5.3 | Screen Shot for successful test cases | 62 |
| 7.6.1 | Test case table | 63 |

**LIST OF ABBREVATIONS**

TF Term Frequency

TF-IDF Term Frequency and Inverse Document Frequency

IR Information Retrieval

VSM Vector Space Model

SSA Successor Stemming Algorithm

STD Suffix Tree Document

CBD Clustering-By Directions