

**VISVESVARAYA TECHNOLOGICAL UNIVERSITY**

JNANA SANGAMA CAMPUS, BELGAVI-590018



**PROJECT REPORT**

**On**

**CREATING THE KNOWLEDGE BASE USING WIKIPEDIA**

**Submitted by**

**DEEPASHREE V**

**4AL15IS008**

**GANESH PRASAD E**

**4AL15IS009**

**KAVANA M G**

**4AL15IS012**

**In partial fulfillment of the requirements for the degree of**

**BACHELOR OF ENGINEERING**

**In**

**INFORMATION SCIENCE AND ENGINEERING**

**Under the Guidance of**

**Mr. Manjunath H R**

**Associate Professor**



**DEPARTMENT OF INFORMATION SCIENCE AND ENGINEERING**  
**ALVAS INSTITUTE OF ENGINEERING AND TECHNOLOGY**

**Moodbidri-574225, Karnataka**

**2018– 2019**



**ALVAS INSTITUTE OF ENGINEERING AND  
TECHNOLOGY MIJAR, MOODBIDRI D.K. -574225  
KARNATAKA**



**DEPARTMENT OF INFORMATION SCIENCE AND ENGINEERING  
CERTIFICATE**

*Certified that the project work entitled "CREATING THE KNOWLEDGE BASE USING WIKIPEDIA" is a bonafide work carried out by*

**DEEPASHREE V**

**4AL15IS008**

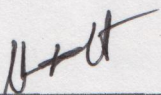
**GANESH PRASAD E**

**4AL15IS009**

**KAVANA M G**

**4AL15IS012**

in partial fulfilment for the award of BACHELOR OF ENGINEERING in **INFORMATION SCIENCE AND ENGINEERING** of the **VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELGAUM** during the year 2018-2019. It is certified that all corrections/suggestions indicated for Internal Assessment have been incorporated in the report deposited in the departmental library. The project report has been approved as it satisfies the academic requirements in respect of project work prescribed for the Bachelor of Engineering Degree.

  
Mr. MANJUNATH H R

**Project Guide**

  
Mr. JAYANTKUMAR A. RATHOD

**H.O.D.**

**Dept. of Information Science & Engineering  
Alva's Institute of Engineering & Technology  
Mijar, MOODBIDRI**

  
Dr. PETER FERNANDES

**PRINCIPAL**

**Alva's Institute of Engg. & Technology,  
Mijar, MOODBIDRI - 574 225, D.K**

**Name of the Examiners**

**Signature with Date**

1.

2.



## TABLE OF CONTENTS

Wikipedia is the largest repositories in the Web. The term knowledge base was in connection with the expert systems as it is the part of Artificial Intelligence. A knowledge base can be created for any entity. The existing system like YAGO, MediaWiki tries to convert Wikipedia into a structured database to provide a vast knowledge base across the domains. It is very difficult to get the information which we want across the domains. So, the solution would be to get a systematic automated approach to build a knowledge base using Wikipedia on entity which we are interested in. The proposed system provides a knowledge base built upon the location as its entity. The system is feeded with seed data, by using these seed data it traverses through the Wikipedia graph and builds knowledge base using similarity measurement between seed data and traversed upcoming pages of wiki graph. Any expert AI systems uses gold standard knowledge base to take any decisions.

1.1 SCOPE	3
2. LITERATURE SURVEY	4
3. PROPOSED SYSTEM	9
4. REQUIREMENT SPECIFICATION	10
4.1 FUNCTIONAL REQUIREMENTS	10
4.2 NONFUNCTIONAL REQUIREMENTS	11
4.3 HARDWARE REQUIREMENTS	12
4.4 SOFTWARE REQUIREMENTS	12
5. SYSTEM DESIGN	15
5.1 HIGHLEVEL DESIGN	15
5.2 SYSTEM ARCHITECTURE	19
5.3 DATABASE DESIGN	20
5.4 DATAFLOW DIAGRAM	21
5.5 UML CASE DIAGRAM	22
5.6 UML SEQUENCE DIAGRAM	23
6. IMPLEMENTATION	24
7. SYSTEM TESTING	26
7.1 TESTING OBJECTIVES	26