

**TRANSLATED AL-QURAN RETRIEVAL SYSTEM EMPLOYING BIGRAM
USING APACHE**

NORSHARIZAN BINTI MUHAMAD

**THESIS SUBMITTED IN PARTIAL FULFILMENT FOR THE DEGREE OF
BACHELOR OF SCIENCE**

**FACULTY OF INFORMATION TECHNOLOGY
AND QUANTITATIVE SCIENCES
MARA UNIVERSITY OF TECHNOLOGY
SHAH ALAM**

2004

ACKNOWLEDGEMENTS

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ
الْحَمْدُ لِلَّهِ رَبِّ الْعَالَمِينَ
الصَّلَاةُ وَالسَّلَامُ عَلَى أَشْرَفِ الْأَنْبِيَاءِ وَالْمُرْسَلِينَ

Praise to Allah the Almighty for providing me a superior knowledge throughout this period of study and for all that has bestowed on me. It is with His power the study is completed.

I would like to express a very special gratitude and thankfulness to my supervisor Assoc. Prof. Dr. Zainab Abu Bakar who is giving me reliance and shrewd guidance that have profited me very much. She who is also gave lots of support in this knowledge quest.

My special appreciation is due to the Dean of the Faculty of Information Technology and Quantitative Sciences and the Programme Head of Bachelor (Hons.) Information Technology, Puan Salwani who is giving good management services throughout my degree program. My credits are also due to the Puan Noorazzah binti Abd Rahman and computer laboratory staff of the faculty for their most generous assistance.

To my friends, Kamariah Kamal, Hasmiza Yob, Nisrin Ahmad Ramly especially Mohd Khairul Nizam Abd. Karim and others in the faculty, who are simply too numerous to indicate, thanks for your time, understanding and buttress. May Allah will reciprocate you for all those things.

Last but not least, to my parents, Haji Muhammad bin Rakidin, and my siblings, I would like to forward my obliged to them for their continuous support during my study period, their patience and benevolence. Thank you.

وَاللَّهُ عَالِمُ بِالصَّوَابِ
أَمِينَ

ABSTRACT

Information retrieval is one of the fields in information technology. The main function of information retrieval is to provide the tools to perform searching effectively and efficiency. This project is to broad the study on information retrieval based on Malay query words and databases. This project is the enhancement of current project which had used thesauri. It is build to help users who are interested to know about certain subject matters from the Al-Quran, where the document retrieval system will automatically retrieve all relevant documents in response to the user queries. This project will employ bigram algorithm that will be used to retrieve translated Al-Quran documents using Malay keywords. Then evaluate the efficiency and effectiveness of the bigram method by comparing it with other conflation method that had been carried out by previous student in retrieving translated Al-Quran documents. This project will run on Apache and will be implement on-line so that it can be installed by users on personal computer. Apache provides many features that make it a very powerful local Web server.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	ii
TABLE OF CONTENTS	iii
LIST OF TABLE	vii
LIST OF FIGURES	ix
ABSTRACT	x
CHAPTER	
1.0 PROBLEM DESCRIPTION	1
1.1 Background Of The Problem	1
1.2 Problem Description	2
1.3 Problem Scope	3
1.4 Problem Significance	3
1.5 Summary	4
2.0 LITERATURE REVIEW	5
2.1 Introduction	5
2.2 An Information Retrieval System	6
2.3 Conflation Method	8
2.4 Stemming	9
2.5 Thesaurus	10
2.6 N-Gram	13
2.7 Images	19
2.8 On-line Retrieval System	20
2.9 Al-Quran Online	21
2.9.1 The Humanities Text Initiatives, Uni Of Michigan, USA	22
2.9.2 Brown University's Scholarly Technology Group, UK	23
2.9.3 Bandung Institute Of technology, Indonesia	24
2.9.4 Muslim Students Association, Oregon State Uni, USA	24
2.9.5 Prime Minister Department Malaysia	25
2.9.6 Al-Quran on CD-ROM	25

CHAPTER 1

PROBLEM DESCRIPTION

1.1. Background of the Problem

Information retrieval is one of the fields in information technology. The main function of information retrieval is to provide the tools to perform searching effectively and efficiency. An information retrieval system may retrieve written texts, spoken utterances, graph, or images.

Currently, online search engine for Malay database that are available are limited. Due to this problem, it is important to broad the study on information retrieval based on Malay query words and databases.

There are several Al-Quran retrieval systems that employ exact match available on the Internet. However, these retrieval methods retrieve word that did not satisfy the users. The retrieve documents are not related to users need. Thus, this project will develop bigram algorithm. Then evaluate the efficiency and effectiveness of the bigram method by comparing it with other conflation method that had been carried out by previous student.