UNIVERSITI TEKNOLOGI MARA

DEVELOPMENT OF SPIDER USING JAVA FOR BROKEN LINKS CHECKING

BY

MAHFUDZAH BT ISHAK

2007123857

BACHELOR OF SCIENCE (HONS) DATA COMMUNICATION AND NETWORKING

FACULTY OF COMPUTER AND MATHEMATICAL SCIENCES

NOVEMBER 2009

ACKNOWLEDGEMENT

Firstly I would like to thank Allah SWT because I have completed this project along with the report.

Next a big thanks to Puan Zolidah Kasiran as the supervisor for my project and the opportunity given to me doing my final year project with her so that my study in Universiti Technology MARA (UiTM) will be completed. Thanks you for the guidance you gave to me.

Moreover I'm glad to thank Encik Adzhar, Coordinator for Final Year Project and lecturer of ITT580 (Research Proposal) subject for his advices and guidance during class. Not forgetting to Universiti Technology MARA (UiTM) especially Faculty of Computer and Mathematic Sciences including all lecturers

Last but not least to my family. Big thanks for the support and courage you give to me and for my colleagues finally we've come to the end. Hopefully what we learnt and attempt will gave benefit to us and have blessing from Allah SWT.

Thank you.

ABSTRACT

This thesis describes the implementation of Spider program using JAVA programming language for broken links checking. The main objective of the research is to develop a Spider program to check broken links on the websites using JAVA. Java is a particularly good choice as a language to construct a spider. Java has built-in support for the HTTP protocol, which is used to transfer most Web information. Java also has an HTML parser built in. Both of these two features make Java an ideal choice for spiders. Besides, this Spider program provides a proper arrangement of result from the Spider program to make it easy to read by users.

TABLE OF CONTENTS