UNIVERSITI TEKNOLOGI MARA

DESIGN AND IMPLEMENTATION OF MALAYSIAN MICROBIAL STRAIN DATABASE

HASDIANTY ABDULLAH

Thesis submitted in fulfillment of the requirements
for the degree of
Master of Science (Information Technology)

Faculty of Information Technology & Quantitative Sciences

November 2008

ABSTRACT

Malaysian Microbial Strain Database (MMSD) is a database of microorganisms collection from many places in Malaysia. It is developed specially to organize microbial data generated from two projects granted by MOSTI. Relational database modeling has been applied during the design phase in order to organize the microbial data. MMSD is also assisted with graphical user interface (GUI) query and retrieval system to facilitate its users in organizing and regaining the data stored. The system is implemented on three-tier architecture with MySQL as the DBMS at the bottom layer, PHP as server side script in the middle layer and the GUI as the first layer. The relational database system enables three main functions, one is to provide the researchers with web interface to perform data management task for data generated (sampling record, culture information, identified microbes information and 16S rRNA sequence). Secondly, the database system allows the users to issue query to obtain information from the database and third, the system allows users to view results in a user friendly interface.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.0   Introduction

One of the hallmarks of modern genomics research is the generation of enormous amounts of raw sequence data (Xiong, 2006). Biologists today are swimming in a rapidly rising sea of data. Efficient experimental techniques, primarily DNA sequencing and microarrays, have led to an influx of basic biological data. The exploitation of these biological data creates the need for databases that are easy to maintain and can be easily managed by users who are most likely to be biologist. The current research in biology depends heavily on the effective exploitation of huge amounts of data (Bornberg & Paton, 2002). However, biologists have been dealing with problems of information management since 17th century (Gibas & Jambeck, 2001). Thus, the very first challenge in the genomics era is to store and handle the staggering volume of information through the establishment and use of computer databases. Biological databases have been considered as such a tool to assist scientists in data management (Chen, 2006). The development of databases to handle the vast amount of molecular biological data is thus a fundamental task of bioinformatics.

Bioinformatics is a new discipline of science that incorporate biology, computer science and information technology (Barnum, 2005). Bioinformatics represents a new field at the interface of the twentieth-century revolutions in molecular biology and computers (Pevsner, 2003). In its broadest sense, the term bioinformatics can be considered to mean information technology applied to the management and analysis of biological data (Attwood & Parry-Smith, 1999).