

## **Chapter 1: Introduction**

### **1.1 Preface**

A spreadsheet is defined as a computer program into which one enters sets of figures in columns or rows which can be moved around and used for calculating (refer to the Oxford Advanced Learning Dictionary 5<sup>th</sup> Edition). Nowadays, the most common spreadsheet program used worldwide would be the Microsoft Excel application.

The Microsoft Excel application is being used in storing records of accounts, contract tenders, daily business transactions and some other forms of data which are best stored in tabular format. Moreover, with the emergence of the Information Communication Technology (ICT) and computer hardware performance, more and more industries are heading towards transforming their business data into digital format especially the spreadsheet format.

This project is regarding the manipulation of data contained in a Microsoft Excel file by a Java-written computer program. Two types of manipulation will be focused in the proposed system. The two types of manipulation are segregating a Microsoft Excel document and combining several Microsoft Excel documents.

In the rest of this chapter, an introduction of the Object-Oriented Data Automation (OODA) System for Microsoft Excel Files will be presented. Besides that, the purpose,

objectives, scopes, contributions of the proposed system will also be explained in this chapter.

## **1.2 Introduction to the Object-Oriented Data Automation System for Microsoft Excel Files**

The computerization of corporate data for operational and analytic business applications has experienced severe setbacks over the past few decades. As groundbreaking technologies are introduced and new levels of performance are achieved, the need of getting relevant and accurate piece of information on the right time is a critical issue now. In response to this critical issue, many forms of data storage were developed such as data warehouses and data marts.

A common file type that is widely used nowadays would be the spreadsheet file, where the data is stored in several worksheets. This project is regarding the automation involved in dealing with any Microsoft Excel file (\*.xls) by using Object-Orientation (O-O) approach. An Excel file can be manipulated by using an O-O based language namely Java. Given an Excel file, the Java program can open it and read all the contents in it. Furthermore, this program can segregate the Excel worksheets into several sub-files for the convenience of the user to distribute it to different groups of readers.

Besides that, the OODA System for Microsoft Excel Files can be used for automatically combining several Microsoft Excel files. There are two types of combination: vertically combined and horizontally combined. The purpose of this function is to create a tool for

user to combine any Microsoft Excel documents into one major file so that a comprehensive view of those Excel documents can be achieved.

An Object-Oriented programming language namely Java is used for developing this project. Each element found in a Microsoft Excel document could be represented by different object. All objects are packaged for the ease of system deployment in the system development stage.

Since the OODA System for Microsoft Excel Files is developed using Java language, therefore it is a platform-independent system. However, since Microsoft Excel application is a Windows System 32 program, the dynamic link library (DLL) for Microsoft Excel is needed and should be placed in the system directory for running in other operating systems such as LINUX and UNIX.

### **1.3 Problem Statement**

Automation is a way of reducing the time used for executing any task. But, before the invention of computer, all tasks were being executed manually and it took a longer period of time for completion. However, with today's technology, spreadsheet system has been invented and used as a way of recording business data. The spreadsheet system such as the Microsoft Excel application has been used widely and it speeds up the completion of well-sorted business documents.

Nevertheless, people nowadays are heading towards a highly speed of working environment. For example, if a manual process could be automated 50%, it would contribute to a faster process life cycle.

This project has proposed a system whereby the users can use it to segregate an Excel document or combine several Excel documents. Traditionally, the “cut, copy & paste” feature provided by Microsoft will be used to perform data duplication and relocation. However, it is a tedious thing to do if a person needs to browse through thousands of Excel documents and combine it into one Excel document.

The same tedious situation will also occur when a person needs to segregate a large/long Microsoft Excel file into a number of small files. Again, the “cut, copy & paste” feature is the only command that could be used to execute this task.

Hence, the idea of developing the OODA System for Microsoft Excel Files was originated and implemented. The proposed system could help to reduce the amount of work involved for segregating a large Excel document and combining a number of Excel documents as well.

## **1.4 Project Mission**

There are two missions for this project.

- i. Develop Java classes which can represent the Microsoft Excel elements.
- ii. Produce a system that can access and manipulate each data cell in Microsoft Excel document to automate the process of segregating an Excel document and combining several Excel documents.

## **1.5 Objectives**

The objectives of developing an Object-Oriented Data Automation System for Microsoft Excel files are listed as below:

- i. To produce a system written in Java programming language that can access elements provided in Microsoft Excel application.
- ii. To build a user interface that enables users to manipulate Microsoft Excel documents without invoking Microsoft Excel application. The user interface utilises the use of existing browser applications such as Microsoft Internet Explorer and Netscape Communicator.
- iii. To create Java classes that use the Component Object Model (COM) bridge for representing each element provided in Microsoft Excel application.
- iv. To provide a way for users to segregate an Excel document without manually “cut/copy & paste” the Excel data.
- v. To provide a way for users to combine several Excel documents (reside in several Excel files) into one file.

- vi. To provide users a link to the output files for each action of segregation as well as combination of the Excel documents.

## **1.6 Project Scopes**

This project will focus on the following scope:

- i. Support only Microsoft Excel version 2000 which could be found in the Microsoft Office Suite version 2000.
- ii. Four Microsoft Excel built-in elements are studied. These four elements are the Application object, Worksheet object, Workbook object and the Range object.
- iii. The OODA System for Microsoft Excel Files could only be executed under Microsoft Windows platforms (Microsoft 9x, Microsoft ME, Microsoft XP)
- iv. Tomcat Web Server version 3.3 is used as the Java Servlet Container. The Java Servlet Container is responsible for the invocation of the Java Servlet pages which act as the Interface layer for this project.
- v. Provide a way for users to segregate an Excel document into sub files. The segregation is based on the exact words entered by users. The exact words must be located in the original Excel document and they must be in the same font case (in order for system to search the Excel cell that contains the words). User may choose to provide at least a pair of words or the name of the Excel cell (e.g A1, B1...) to indicate the starting point and the ending point of the sub file which will be the output of the segregation process.

- vi. Provide a way for users to combine up to seven Excel files and save them in one file name. Users can choose whether to combine the files horizontally or vertically. The combined Excel file will be saved in a new Excel file.
- vii. A hyperlink is provided in the Java Servlet page for user to link to the output Excel files generated from the Excel files segregating process or the combining process.

## **1.7 Project Contributions**

As the workload for office administration increases tremendously nowadays, more and more organizations will be relying on the usage of spreadsheet program. Microsoft Excel application is a dominant program in the market today. Hence, with the introduction of this project, shorter time of Excel data manipulation can be achieved.

With the OODA System for Microsoft Excel Files, the “cut/copy & paste” command will be less used. In segregating an Excel document into several sub files, user will just need to provide the exact words/phrases found in the document and click the submit button. The words/phrases will become the starting point and the ending point that indicate to the system from where it should duplicate the data and save it into a new file.

Besides that, this system also provides an easier way for user to combine several Microsoft Excel documents. User can choose to combine all the data in the Excel documents in vertical format or horizontal format.

## **1.8 Project Assumptions**

The assumptions made in developing the OODA System for Microsoft Excel files are listed below:

- i. Users will need to have the Microsoft Excel installed in their personal workstations.
- ii. Microsoft Excel bundled in Microsoft Office Suite version 2000 is supported by this system. Earlier versions of Microsoft Excel have yet to be tested.
- iii. The original Excel document that is needed for segregation process must have unique words/phrases in all used data cells. This is to ensure that the starting point and the ending point for the segregation can be located.
- iv. Users must have a default web browser application installed in their workstation. The system will use the web browser as the medium of user interface.

## **1.9 Report Organization**

The report consists of eight chapters. First, an introduction to this project is prepared in Chapter 1. A literature review is prepared in Chapter 2. Then, followed by the methodology used in developing this project (Chapter 3).

Chapter 4 will explain the Object-Oriented Analysis in realizing this project. The Object-Oriented Design is included in later chapter, Chapter 5. After this, a description regarding system development is presented in Chapter 6.



Next, documentation regarding System Deployment and System Testing are presented in Chapter 7. This chapter will focus on all steps involved in deploying the OODA System for Microsoft Excel Files. A system testing procedure is also included in this chapter. System testing is important for achieving the correctness and validity of the developed system.

Last, Chapter 8 contains an overall conclusion of this project. There will be a discussion on the expansion of this project and also the limitation faced while deploying the system into the real-world situation.