ENTER XML: THE FUTURE MARKUP LANGUAGE

by

HAMDAN BIN ABDUL MAAD

An independent study submitted to the

Universiti Teknologi Mara

in partial fulfillment of the requirement for the

Master of Science in Information Technology

October 2000

ACKNOWLEDGEMENT

First, I would like to take this opportunity to express my gratitude to Course Coordinator Associate Professor Dr. Mohd Isa Mohd Samat, for his guidance and patience throughout my study. A special gratitude and appreciation to my supervisor Associate Professor Dr. Saadiah Yahya for her valuable time, suggestion and guidance in the completion of this independent study.

I also owe a large debt of gratitude to many of my FTMSK colleagues especially to Associate Professor Azizi Ngah Tasir (Dean), Dr. Adnan Ahmad, Dr. Nor Laila Md Noor, Puan Rosslina Mohamed Nawi, Puan Aishah Ahmad, and Puan Marina Ismail, and my course mates who have helped me in my never-ending quest of knowledge. All cooperation and attention given are highly appreciated, and hence I wish to thank them whole-heartedly.

Last but not least to my lovingly wife and kids, for their patience and encouragement throughout my study. May Allah bless you all!

CONTENTS

ABSTRACT

XML and HTML are both subsets of SGML. XML is a markup language for describing structured data where the content is separated from presentation. To process an XML document, a software program is needed: XML parser (an XML processor) – currently IE5 support XML and Nestcape through Morzilla.org. XML provides a file format for representing data and a schema for data to include a description of its own structure. XML is a standard, extensible, universal format for Web-based data. With its powerful expressiveness and flexibility, XML promises to add structure to data on the Internet, bringing the Web one step closer to realizing the potential for universal communication with anyone, anywhere. Using XML with XSL or CSS, Web site's content and style can be managed, and changed in one place (the style sheet) instead of editing piles of HTML files. Using SAX or DOM, the Web document can be treated as object structures and process them in general and clean way. Developers who learn XML now will find it a powerful tool for data representation, storage, modeling, and interoperation. XML will not be one of those heralded technological advances that falls after huge investments have been made. It is here, and here to stay. So, enter XML, the future markup language.

1. Introduction

Most documents on the Webs are stored and transmitted in HyperText Markup Language (HTML). HTML is a simple language well suited for hypertext, multimedia, and the display of small and reasonably simple documents. Bill Gates Microsoft's well-known founder and President said in an interview in February 1996, "HTML is our data". The pronouncement was an emblem for the impact that the Internet and its friendly interface, the World Wide Web (WWW), had exerted on government, institution, corporations, and people. The success of the Web and its original means for presenting information, have been amply demonstrated in all aspects of live using HTML.

Web browsers were viewed as potential application platforms, but Java technology needs (frankly) more to chew than HTML offered in order to fulfill that vision. The web-based applications will relied too much on Common Gateway Interface (CGI) script at the server to process the data in web pages if HTML is used as the data standard. This will contributes mightily to Internet traffics and makes the Web slow for many users.

One of the most significant developments on the WWW and in electronic publishing and electronic commerce (e-commerce) can be said was XML. XML stand for eXtensible Markup Language, is a more recent extension of Markup Language. It has been well accepted as the next important Internet technology, and the natural worthy companion to Java™ programming language itself. Most enterprises (especially in USA) have enormously used XML.