A HIERARCHICAL TREE STRUCTURE FOR
NATIONAL ARCHIVES OF MALAYSIA WEB SITE

MOHAMAD TAHARA IJAB

Faculty
Of
Computer Science and Information Technology

1999
A HIERARCHICAL TREE STRUCTURE FOR NATIONAL ARCHIVES OF MALAYSIA WEB SITE

MOHAMAD TAHA IJAB

A dissertation submitted in partial fulfillment of the requirements for the degree of Master of Computer Science

Faculty of Computer Science and Information Technology

UNIVERSITI MALAYA

June 1999
Declaration

I certify that this thesis submitted for the degree of Masters is the result of my own research, except where otherwise acknowledged, and that this thesis (or any part of the same) has not been submitted for higher degree to any other university or institution.

Signed:........................................
Mohamad Taha Ijab

Date:..............................
12th August 1999
ACKNOWLEDGEMENTS

First of all, I would like to praise God for His Guidance, Alhamdulillah. Next in line, to my parents, Father and Mother, thank you for your love and that never ending supports. Brothers and sisters, thank you so much. My great and funny little nephews and nieces, for inspiring me to achieve the best. To close friends Norizam and Nazeri, thanks for the friendship. To Kak Inson and Kak Faiezah, for giving ideas and encouragement.

Not forgetting my wonderful project supervisor, Dr. Siti Salwah Salim for her friendly help and invaluable knowledge in the field of Human-Computer Interaction and also all faculty lecturers especially Dr. Yew Kok Meng, Dr. Lee Sai Peck, Dr. Sapiyan Baba, Dr. Selvanathan and En. Abdullah Ghani. Also to fellow friends in the faculty: especially Mohd. Azul, Awang and other UNITAR IT Fast Trackers. Thank you very much.

Last but not least, Finally, to UNITAR IT Fast Track programme coordinator and Prof. Dr. Khairuddin Hashim, thank you for sponsoring my study here.

Heartfelt Thanks,

Mohamad Taha Ijab
Faculty of Computer Science and Information Technology
Universiti Malaya
50609 Kuala Lumpur
MALAYSIA
ABSTRACT

This thesis studies and implements the content visualization for National Archives of Malaysia web site. This web site can be categorized into a content-rich type as it stores large number of documents on its site. The two major problems currently faced by Internet users are disorientation and cognitive overload. Prior to the implementation phase, the research is centered on information visualization, user interface design principles, and hierarchical tree structure. Several examples of visualization interfaces such as TileBars, WebTOC, and LifeLines are reviewed. Related concepts and tools used in this thesis are studied. They include Internet, World Wide Web (WWW), web browsers, Java, applets, and Java Virtual Machine (JVM). The development process of the new interface called CRIVE (Content-Rich Website Visualizer) is done using two Java integrated software environments: Symantec Visual Café and MS-Visual J++. Two applets are developed which are the TreeApp Navigator and the Control applet. The TreeApp Navigator applet is used for navigating the site’s content and it is implemented as a hierarchical tree structure. Another applet called Control is written for controlling the way users view the site’s contents. A new web site called CRIVE web site is created and uploaded onto an independent, free web hosting server available from Tripod.com. Another copy of the CRIVE web site is uploaded onto a MS-Internet Information Server (IIS) running on Windows NT machine. The usability testing is conducted with fourteen Internet users. They were asked to compare the CRIVE web site with the existing National Archives web site in terms of its usability and subjective satisfaction. From the findings, it is found that all subjects were satisfied with the new interface and responded affirmatively. This indicates that interface such as CRIVE is in demand by users, and is motivated for further research in the area.
# TABLE OF CONTENTS

Declaration  
Acknowledgements  
Abstract  
Table of Contents  
List of Figures  
List of Tables

| Chapter 1 – Introduction |  
|--------------------------|---|---|---|---|
| 1.1 Problem Statement | 3  
| 1.1.1 Disorientation | 3  
| 1.1.2 Cognitive Overhead | 4  
| 1.2 Scope of Research | 5  
| 1.3 Thesis Goals | 6  
| 1.4 Significance of the Research | 8  
| 1.5 Methodology | 8  
| 1.6 Thesis Organization | 9  

| Chapter 2 – Literature Review | 11  
|-------------------|---|---|---|
| 2.1 Web Site: A Definition and Structure | 11  
| 2.2 Types of Web Sites | 13  
| 2.3 Content-Rich Web Sites Examples | 15  


2.3.1 Library of Congress 15
2.3.2 National Archives of Australia 18
2.3.3 American Museum of Natural History 20
2.3.4 National Museum of American Art 23
2.3.5 National Archives of Malaysia 26

2.4 Summary 28

Chapter 3 - Information Visualization 29

3.1 HCI and Information Visualization 29
3.2 The Importance of Information Visualization 31
3.3 Information Visualization and Information Retrieval 33
3.4 Information Visualization Framework 36
3.5 Examples of Information Visualization Interfaces 48
3.6 Summary 62

Chapter 4 - Design 63

4.1 User Interface Design Principles 63
4.2 User Interface Structure 66
4.3 Summary 69
Chapter 5 – Implementation

5.1 Internet 70
5.2 Java 73
5.3 Internet Server 81
5.4 Algorithm and CRIVE Interface 84

Chapter 6 – Testing and Results 107

6.1 Purpose, Hypothesis and Variables 107
   6.1.1 Purpose 107
   6.1.2 Hypothesis 107
   6.1.3 Variables 108

6.2 Testing 108
   6.2.1 Participants 108
   6.2.2 Materials 111
   6.2.3 Procedures and Materials 112

6.3 Results 114
   6.3.1 Task Completion Time 114
   6.3.2 Subjective Satisfaction 119

6.4 Discussion 121
   6.4.1 Discussion of Task Completion Times 121
   6.4.2 Discussion of Subjective Satisfaction 123
Chapter 7 - Conclusion

7.1 Information Visualization Challenges 125
7.2 Lessons Learnt 126
7.3 Achievements 126
7.4 Thesis Constraints 128
7.5 Future Enhancements 128

REFERENCES

APPENDICES

APPENDIX A Experiment Consent Form
APPENDIX B Preliminary Questionnaire
APPENDIX C Experiment Questionnaire Form
APPENDIX D User Satisfaction Form
LIST OF FIGURES

Figure 1.1    Internet Model
Figure 1.2    Overall Thesis Process
Figure 2.1    A Typical Web Site Structure
Figure 2.2    The Library of Congress Web Site
Figure 2.3    The Library of Congress Web Site – Site Map
Figure 2.4    The National Archives of Malaysia Web Site
Figure 2.5    The National Archives of Malaysia Web Site – Site Map
Figure 2.6    The American Museum of Natural History Web Site
Figure 2.7    The American Museum of Natural History Web Site – Site Map
Figure 2.8    The National Museum of American Art Web Site
Figure 2.9    The National Museum of American Art Web Site – Site Map
Figure 2.10   The National Archives of Malaysia Web Site
Figure 2.11   The National Archives of Malaysia Web Site – Site Map
Figure 3.1    Information Retrieval Model
Figure 3.2    Information Visualization Model
Figure 3.3    TileBars embedded in Scatter/Gather interface
Figure 3.4    Pad++
Figure 3.5    WebBook
Figure 3.6    SPIRE Themescape
Figure 3.7    SeeSoft showing overview of a software project
Figure 3.8  LifeLines in medical history records
Figure 3.9  Hyperbolic Tree
Figure 3.10  WebTOC
Figure 3.11  Searching in WebTOC
Figure 3.12  Navigational View Builder
Figure 4.1   A Hierarchical Structure In A Web Site
Figure 5.1   Symantec Visual Café Pro V2.5
Figure 5.2   MS-Visual J++ 1.1
Figure 5.3   Free Web Hosting Server : Tripod.com
Figure 5.4   User flyhigh_3 logged in Tripod.com
Figure 5.5   TreeApp Navigator applet
Figure 5.6   Control applet
Figure 5.7   Control applet: Search tab
Figure 5.8   Control applet: Control Panel tab
Figure 5.9   Control applet: Legend tab
Figure 5.10  Control applet: About tab
Figure 5.11  The Main Page of CRIVE Web Site
Figure 5.12  The Welcome Page of CRIVE Web Site
Figure 5.13  The interface shows a hyperlink titled “View General Site Map of the Site”
Figure 5.14  The General Site Map shown as image file
Figure 5.15  The CRIVE Web Site
Figure 5.16  The TreeApp Navigator applet in Expand mode
Figure 5.17  The TreeApp Navigator applet in Contract mode

Figure 5.18  The new web site shows document displayed in relations to its file type (text file)

Figure 5.19  The new web site shows document displayed in relations to its file type (image file)

Figure 6.1  The Distribution of Subjects Per Session

Figure 6.2  Time Taken To Complete Tasks for National Archives Web Site

Figure 6.3  Time Taken To Complete Tasks for CRIVE Web Site
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>DTTT Framework</td>
</tr>
<tr>
<td>3.2</td>
<td>Information Visualization Interface Examples</td>
</tr>
<tr>
<td>4.1</td>
<td>GUI Characteristics</td>
</tr>
<tr>
<td>6.1</td>
<td>Testing Distribution By Number of Subjects Per Session</td>
</tr>
<tr>
<td>6.2</td>
<td>Time Taken To Complete Tasks for National Archives Web Site</td>
</tr>
<tr>
<td>6.3</td>
<td>Time Taken To Complete Tasks for CRIVE Web Site</td>
</tr>
<tr>
<td>6.4</td>
<td>Subjective Satisfaction Results</td>
</tr>
</tbody>
</table>