

**INTEGRATION OF STUDENT INFORMATION DATA
USING XML**

MOHD AMAR MAAROF BIN AHMAD

UNIVERSITI UTARA MALAYSIA

M. Sc. (ICT) 2008

Q4-11
11/11/08
11/11/08

**INTEGRATION OF STUDENT INFORMATION DATA
USING XML**

**A thesis submitted to the
Division of Applied Sciences, College of Arts and Sciences
in partial fulfillment of the requirements for the degree of
Master of Science (Information and Communication Technology),
Universiti Utara Malaysia**

By

Mohd Amar Maarof bin Ahmad



KOLEJ SASTERA DAN SAINS
(College of Arts and Sciences)
Universiti Utara Malaysia

PERAKUAN KERJA KERTAS PROJEK
(Certificate of Project Paper)

Saya, yang bertandatangan, memperakukan bahawa
(I, the undersigned, certify that)

MOHD. AMAR MAAROF AHMAD
[86862]

calon untuk Ijazah
(candidate for the degree of) **MSc. (Information Communication Technology)**

telah mengemukakan kertas projek yang bertajuk
(has presented his/her project paper of the following title)

INTEGRATION OF STUDENT INFORMATION DATA USING XML

seperti yang tercatat di muka surat tajuk dan kulit kertas projek
(as it appears on the title page and front cover of project paper)

bahawa kertas projek tersebut boleh diterima dari segi bentuk serta kandungan dan meliputi bidang ilmu dengan memuaskan.
(that the project paper acceptable in form and content, and that a satisfactory knowledge of the field is covered by the project paper).

Nama Penyelia Utama
(Name of Main Supervisor): **MR. NURNASRAN PUTEH**

Tandatangan
(Signature)

: 

Tarikh
(Date)

: 20 Nov 2008

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a postgraduate degree from Universiti Utara Malaysia, I agree that the university library may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purpose may be granted by my supervisor or, in her absence by the Assistant Vice Chancellor of the College of Arts and Sciences. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be given to me and to Universiti Utara Malaysia for any scholarly use which may be made of any material from my thesis.

Requests for permission to copy or to make other use of materials in this thesis, in whole or in part, should be addressed to

Assistant Vice Chancellor
College of Arts and Sciences
Universiti Utara Malaysia
06010 UUM Sintok
Kedah Darul Aman.

ABSTRAK (BAHASA MELAYU)

Sekolah-sekolah di Malaysia telah dibekalkan dengan beberapa sistem untuk meningkatkan lagi keupayaan pengurusan data pelajar di dalam persekitaran sekolah. Sistem-sistem berkenan telah banyak memudahkan kerja-kerja para guru yang sentiasa dibebani dengan kerja-kerja di sekolah. Kajian ini dibuat bagi menghasilkan model integrasi data bagi semua sistem maklumat yang berkaitan dengan pelajar dengan menggunakan XML. Tiga fasa utama dalam kajian ini ialah analisa dan pemodelan, pembangunan prototaip dan pengujian. Model integrasi data tersebut berjaya dihasilkan. Prototaip telah dibangunkan berasaskan model tersebut. Akhirnya prototaip tersebut telah menjalani pengujian dan keputusannya sangat memberangsangkan. Beberapa cadangan telah disyorkan untuk kajian yang lebih lanjut pada masa akan datang.

ABSTRACT

Schools in Malaysia have been provided with few systems to increase their capabilities to manage student data in the schools environment. With so many workloads that teachers have to face, those systems really give a relief to the teachers. This study was carried out to formulate the integration model of student information data using XML. There are three main phases involved in the study which are analysis and modeling, prototype development and testing. Upon completion, models of data integration were formulated and prototype was developed based on the model. Finally the prototypes were tested and the results are promising. There are some recommendations also for future work.

ACKNOWLEDGMENTS

In the name of ALLAH, The Most Beneficent, The Merciful

I am thankful to Allah for granting me the strength to complete this project. I received lot of encouragement and support from many individuals.

I would like to express my heartfelt thanks to my supervisor, Mr. Nurnasran bin Puteh for his guidance and full support. I thank him for the assistance, inspiration and motivation to overcome the difficulties in completing the project. Also my gratitude goes to evaluator Madam Nur Fadziana binti Faisal Mohamed and Associate Professor Abdul Bashah bin Mat Ali for the valuable and constructive comments. To all the people in Sekolah Menengah Kebangsaan Padang Siding, Arau, Perlis, thank you for your cooperation.

It would not been possible for me to complete the program and this project without the support from my family. First and foremost, my gratitude goes to my loving wife, Nurhashila binti Abdul Rahman. To my daughter, Auni Balqis, thank you for understanding your father's life long passion for learning. To my father and mother, thank you for your pray, support and love.

Finally, I would like to thank other lecturers, librarians, course mates and friends who have in some way or other given me encouragement, motivation and assistance throughout the journey of completing this writing.

TABLE OF CONTENTS

PERMISSION TO USE	i
ABSTRAK (BAHASA MELAYU)	ii
ABSTRACT	iii
ACKNOWLEDGMENTS	iv
TABLE OF CONTENTS	v
LIST OF FIGURES	vii
LIST OF TABLES	ix
LIST OF ABBREVIATIONS	x
LIST OF APPENDICES	xi

CHAPTER ONE: INTRODUCTION

1.1	Problem Statement	4
1.2	Research Objective	6
1.3	Scope and Constraint	6
1.4	Significance of the Study	7
1.5	Report Organization	7

CHAPTER TWO: LITERATURE REVIEW

2.1	Introduction	8
2.2	Data Integration	9
2.3	Significant of XML	10
2.3.1	Well-formed XML	11
2.3.2	Extended Stylesheet Language Transformation (XSLT)	12
2.3.3	XPath	13
2.4	Related Work	14
2.5	Summary	15

CHAPTER THREE: METHODOLOGY

3.1	Introduction	16
3.2	Phase I: Analysis and Modeling	16
3.2.1	Analysis of the Current Systems	16
3.2.1.1	Observation	17
3.2.1.2	Interviews	17
3.2.1.3	Analyzing Software and Related Problem	17
3.2.1.4	Analyzing Tables and Attributes	18
3.2.3	Data Modeling	18
3.3	Phase II: Prototype Development	19
3.4	Phase III: Testing	20
3.5	Summary	21

CHAPTER FOUR: FINDINGS AND DISCUSSION

4.1	Introduction	21
4.2	Phase I: Analysis and Modeling	21
4.2.1	Analysis of the Current Systems	21
4.2.1.1	Observation	22
4.2.1.2	Interviews	23
4.2.1.3	Analyzing Software and Related Problem	23
4.2.1.4	Analyzing Tables and Attributes	24
4.2.2	Data Modeling	26
4.3	Phase II: Prototype Development	30
4.3.1	Create an XML data from selected table (TMurid)	30
4.3.2	Import an XML document into SISTEK	35
4.3.3	Import an XML document into SSDM	41
4.4	Phase III: Testing	47
4.5	Summary	56

CHAPTER FIVE: CONCLUSION AND FUTURE WORK

5.1	Introduction	57
5.2	Research Contribution	58
5.3	Recommendation for Future Work	58

REFERENCES	60
-------------------	-----------

LIST OF FIGURES

Figure 1.1: The operation of current systems	5
Figure 2.1: The XSLT Process	12
Figure 3.1: Prototyping Process	19
Figure 4.1: The attributes in SMM, SISTEK and SSDM tables	25
Figure 4.2: Overall system model	26
Figure 4.3: Import XML process model	27
Figure 4.4: Detail process in SMM model	28
Figure 4.5: Detail Import Process in SISTEK Model	28
Figure 4.6: Detail Import Process in SSDM model	29
Figure 4.7: Code to create and export XML data	30
Figure 4.8: Added button 'Eksport XML' on top existing main menu in SMM	31
Figure 4.9: Click 'Eksport XML' button to run the export process	31
Figure 4.10: Browse the location to export the XML document	32
Figure 4.11: Choose the location to export XML	32
Figure 4.12: Click the 'Export XML' button to export the XML document	33
Figure 4.13: Process of exporting XML data was successful	33
Figure 4.14: The Well-formed XML	34
Figure 4.15: Code to import XML	36
Figure 4.16: From drop down menu, select appropriate list to open XML menu	37
Figure 4.17: Click the browse button to locate the XML data	38
Figure 4.18: Select the appropriate XML file to import	38
Figure 4.19: Click Load XML button to run the import process	39
Figure 4.20: Response said it was a successful process	39
Figure 4.21: The data imported appear in the TMurid subform table	40

Figure 4.22: Code to import XML	41
Figure 4.23: From main menu, click the 'Import data dari XML' button	43
Figure 4.24: Click browse button to locate the XML file	44
Figure 4.25: Select the appropriate XML file to import	44
Figure 4.26: Click the 'Load XML' button to execute the import process	45
Figure 4.27: Response said it was a successful process	45
Figure 4.28: The data imported appear in the TMurid subform table	46
Figure 4.29: Adding discipline record	48
Figure 4.30: Changed student name in SMM	49
Figure 4.31: The student name had been changed in SISTEK	50
Figure 4.32: The result before changing the data	51
Figure 4.33: The result after changing the data	51
Figure 4.34: Student IC before changing data	53
Figure 4.35: Adding discipline record to the student	54
Figure 4.36: Changing Student IC in SMM	54
Figure 4.37: Changing data was successful in SISTEK after import process	55
Figure 4.38: The existing record about student discipline had lost	55
Figure 4.39: Result after changing IC data in SMM	56

LIST OF TABLES

Table 4.1: Comparison between three systems

LIST OF ABBREVIATIONS

DTD	Document Type Definition
HTML	HyperText Markup Language
IC	Identity Card
IE	Internet Explorer
IT	Information Technology
LAN	Local Area Network
MOE	Ministry of Education
PC	Personal Computer
SSDM	Sistem Sahsiah Disiplin Pelajar
SISTEK	Sistem Pinjaman Buku Teks
SMM	Sistem Maklumat Murid
USM	Universiti Sains Malaysia
VBA	Visual Basic Application
W3C	World Wide Web Consortium
XML	Extended Markup Language
XSL	Extended Stylesheet Language
XSLT	Extended Stylesheet Language Transformation

LIST OF APPENDICES

Appendix A: List of Attributes in TMurid, TPelajar and peribadi Tables	66
Appendix B: List of Codes from SMM	69
Appendix C: XML Document with One Student Data	73
Appendix D: List of Codes from SISTEK	76
Appendix E: List of Codes from SSDM	80
Appendix F: XSL Transformation Document	87
Appendix G: New XML Document Created After XSL Transformation	89
Appendix H: XML Document in Testing Phase	91
Appendix I: New XML Document Created After XSL Transformation In Testing Phase	98

CHAPTER ONE

INTRODUCTION

Information Technology is really helping teachers nowadays in their routine works. It makes things easier, faster and systematic. So teachers nowadays have to increase their knowledge in Information Communication technology (ICT) literacy to improve their teaching in class and also in managing student information system.

In Malaysia, managing student information has changed significantly. Previously teachers have to record all students' data in 101 card and 102 card manually. Lots of writing processes were involved to fill in both the cards. Now teachers only need to key in the data in the 'Sistem Maklumat Murid' (SMM), print out and keep them in the students files. SMM is a student information system developed by Mr.Md Yusoff Alaudin and Mr. Amat Sazali Abu Hassan from the Penang Education Department. This system has been accepted by Ministry Of Education to use by all schools in Malaysia since 2003.

In the process of borrowing text books, previously teachers need to fill in G Form with students names, tick the books they had borrowed and sign the form. Now with Sistem Pinjaman Buku Teks (SISTEK), using bar code reader, all the books need to be scanned and the system will record the books borrowed by the students. So the process becomes faster and more effective then before. SISTEK was developed by the Textbook Division under MOE. This system has been tested for the first time in

The contents of
the thesis is for
internal user
only

REFERENCES

- Advantage, e. (2002). Understanding XQuery and its potential. Retrieved 26 October, 2008, from http://articles.techrepublic.com.com/5100-10878_11-1059465.html
- Ahmad, R. b. K. (2005). Managing the use of ICT in schools: Strategies for school leaders, *Jurnal Pengurusan Kepimpinan Pendidikan* (Vol. 15). Genting Highland: IAB.
- Amador Durán, A. R. C., Rafael Corchuelo, Miguel Toro. (2002). *Engineering Web applications with XML and XSLT*. Paper presented at the Proceedings of the IEEE Joint International Conference on Requirements Engineering.
- Ambur, O. (2008). Advantages of XML [Electronic Version]. Retrieved 26 October from <http://xml.gov/presentations/gsa/sld006.htm>.
- Andrea R. de Andrade, E. V. M., Maria da G. C. Pimentel. (2004). Engineering Web applications with XML and XSLT. *IEEE*.
- Bryan, M. (1997). Guidelines for using XML for Electronic Data Interchange. Retrieved 27 October, 2008, from <http://xml.coverpages.org/xml-ediGuide970912.html>
- Chee-Yong Chan, P. F., Minos Garofalakis, Rajeev Rastogi. (2002). *Efficient Filtering of XML Documents with XPath Expressions*. Paper presented at the Proceedings of the 18th International Conference on Data Engineering (ICDE'02).
- Contributor, G. (2008). Introduction to XSLT. Retrieved 26 October, 2008, from http://articles.techrepublic.com.com/5100-10878_11-1044797.html

David Hunter, J. R., Joe Fawcett, Eric van der Vlist, danny Ayers, Jon Duckett, Andrew Watt, Linda McKinnon. (2007). *Beginning XML, 4th Edition* (4 ed.). Indianapolis: Wiley Publishing, Inc.

Dennis Pedersen, T. B. P. (2004). *Synchronizing XPath Views*. Paper presented at the Proceedings of the International Database Engineering and Applications Symposium (IDEAS'04).

Elisa Bertino, E. F. (2001). XML and Data Integration [Electronic Version]. *November-Disember 2001*, 75-76. Retrieved 10 July 2008 from <http://www.sei.cmu.edu/isis/pdfs/xml-data.pdf>.

Ernesto Damiani, L. T. (2000). *Flexible Query Techniques for Well-formed XML Documents*. Paper presented at the Fourth International Conference on knowledge-Based Intelligent Enginem'ng Systems & Allied Technologies

Glossary of NLP Terms. (2007). Retrieved 25 October, 2008, from <http://www.inspiritive.com.au/glossary.htm>

Graham, I. (2002). <XML> and the Future of Internet-based Computing Retrieved 27 October, 2008, from <http://72.14.235.104/search?q=cache:owmzyveaKPEJ:www.iangraham.org/talks/11mar02/>

H.M.Deitel, P. J. D., A.B.Goldberg. (2004). *Internet & World Wide Web How To Program* (3rd Edition ed.): Pearson Education International.

Hiroko Kinutani, M. Y., Shunsuke Uemura. (2001). Identifying Result Subdocuments of XML Search Conditions. *IEEE*

Ho-pong Leung, F.-I. C., Stephen C.F. Chan, Robert Luk. (2005). *XML Document Clustering Using Common XPath*. Paper presented at the Proceedings of the 2005 International Workshop on Challenges in Web Information Retrieval and Integration

Inc, B. (2002). Use XPath to locate information in XML documents. Retrieved 26 October, 2008, from http://articles.techrepublic.com.com/5100-10878_11-1054416.html

Jeffrey L Whitten, L. D. B. (1998). *System Analysis and Design Method* (4th ed.). Boston: McGraw-Hill Education.

Jelena Jovanović, D. G. (2005). *Bridging knowledge bases' heterogeneity using XML/XSLT approach*. Paper presented at the e-Technology, e-Commerce and e-Service, 2005. IEEE '05. Proceedings. The 2005 IEEE International Conference

Kay, M. (2005). Saxon: Anatomy of an XSLT processor [Electronic Version]. Retrieved 27 October from <http://www.ibm.com/developerworks/library/x-xslt2/>.

Korol, J. (2005). *Access 2003 Programming by Example with VBA, XML, and ASP*. Texas: Wordware Publishing, Inc.

Lenzerini, M. (2002, 20 October 2008). Data Integration: A Theoretical Perspective. Retrieved 11 Nov, 2008, from http://en.wikipedia.org/wiki/Data_integration#cite_note-refone-0

M.Sneed, H. (2002). *Using XML to Integrate existing Software Systems*. Paper presented at the Proceedings of the 26th Annual International Computer Software and Applications Conference.

Mahmood, M. b. (2008). The Teaching of English. Retrieved 27 October, 2008, from

http://www.rasdian.edu.my/portal/index.php?option=com_content&task=view&id=89&Itemid=82

N.A. (2007). Transforming XML files with XSLT in Access. Retrieved 27 October 2008, from <http://office.microsoft.com/en-us/access/HA010345761033.aspx>

N.A. (2008a, 30 October 2008). How Data Integration Works. Retrieved 12 Nov, 2008, from <http://communication.howstuffworks.com/data-integration2.htm>

N.A. (2008b). Sistem Analisis Peperiksaan. Retrieved 27 October, 2008, from http://www.moe.gov.my/jpnjohor/index.php?option=com_content&task=view&id=64

O'Reilly. (2002). XML Schema. Retrieved 26 October, 2008, from <http://www.unix.com.ua/oreilly/xml/schema/gloss.htm>

Options Trading Glossary. (2002). Retrieved 25 October, 2008, from http://www.trading-futures-markets.com/Options/options_glossary.htm#m

Pawson, D. (2007). XSLT Terminology Clarification. Retrieved 25 October 2008, from www.dpawson.co.uk/xsl/xslvocab.html

Quin, L. (2008, 14 October 2008). Extensible Markup Language (XML). Retrieved 26 October, 2008, from <http://www.w3.org/XML/>

Robert. (2002). Defination. Retrieved 26 October, 2008, from http://www.mcs.k12.ny.us/pages/srobert/new_page_1.htm

Robert E. Kraut, S. T. D., Susan Koch. (1989). Computerization, productivity, and quality of work-life. *Communications of the ACM*, 32(2), 220-238.

Schmitter, R. (2003). An XML Portable Rich Client Application. Retrieved 27

October, 2008, from

http://www.idealliance.org/papers/dx_xml03/html/abstract/07-00-09.html

Shirota, Y. (2004). *Applying XML and XSLT Techniques to a Personalized Distance*

Learning System for Business Mathematical Education. Paper presented at

the Proceedings of the 18th International Conference on Advanced

Information Networking and Application.

Simon Bennett, S. M., Ray Farmer. (2002). *Object Oriented Systems Analysis and*

Design using UML (2nd ed.). London: McGraw-Hill Education.

Sommerville, I. (2004). *Software Engineering* (7th edition ed.): International

Computer Science Series.

Tennison, J. (2005). *Beginning XSLT 2.0: From Novice to Professional*. New York:

APress.

Timperley, H. (2000). Workload and the Professional Culture of Teachers Retrieved

27 October, 2008, from <http://ema.sagepub.com/cgi/content/abstract/28/1/47>

Wahab, A. R. (2007). Requirement Model for School Online Examination. *Universiti*

Utara Malaysia, 8.

Wikipedia. (2008a). Computer Software. Retrieved 26 October, 2008, from

http://en.wikipedia.org/wiki/Computer_software

Wikipedia. (2008b). Diagram. Retrieved 26 October, 2008, from

<http://en.wikipedia.org/wiki/Diagram>

Wikipedia. (2008c, 10 Nov 2008). iGoogle. Retrieved 12 Nov, 2008, from

<http://en.wikipedia.org/wiki/IGoogle>

Wikipedia. (2008d, 24 October 2008). Problem. Retrieved 26 October, 2008, from

<http://en.wikipedia.org/wiki/Problem>

Wikipedia. (2008e). Software prototyping. Retrieved 19 October, 2008, from
http://en.wikipedia.org/wiki/Software_prototyping

Wikipedia. (2008f, 24 Sept 2008). Table (database). Retrieved 26 October 2008,
2008, from [http://en.wikipedia.org/wiki/Table_\(database\)](http://en.wikipedia.org/wiki/Table_(database))

Wikipedia. (2008g). XML. Retrieved 27 October, 2008, from
<http://en.wikipedia.org/wiki/XML#Parsers>

WordNet a Lexical Database fo the English Language. (2006). Retrieved 25
October, 2008, from <http://wordnet.princeton.edu/>

Zurinah Zainol, R. A. S., Rosni Abdullah, Nur' Aini, Wahidah Husain. (2005, June).
*A System to Integrate and Manipulate Protein Databasae Using BioPerl and
XML*. Paper presented at the Proceedings of World Academy of Science,
Engineering and Technology.