



Faculty of Information and Communication Technology

**CANCER PATIENTS' RECORD QUERIES USING ANSWER ON-
DEMAND DATA SHARING ARCHITECTURE**

Aiza Syahida binti Zakaria

Master of Computer Science in Database Technology

2014

**CANCER PATIENTS' RECORD QUERIES USING ANSWER ON- DEMAND DATA
SHARING ARCHITECTURE**

AIZA SYAHIDA BINTI ZAKARIA

**A thesis submitted
in fulfillment of the requirements for the degree of
Master of Computer Science in Database Technology**

Faculty of Information and Communication Technology

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2014

BORANG PENGESAHAN STATUS TESIS*

JUDUL : CANCER PATIENTS' RECORD QUERIES USING ANSWER ON – DEMAND DATA SHARING ARCHITECTURE

SESI PENGAJIAN : 2013 - 2014

Saya : AIZA SYAHIDA BINTI ZAKARIA

Mengaku membenarkan tesis Sarjana ini disimpan di Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dengan syarat - syarat kegunaan seperti berikut:

1. Tesis dan projek adalah hakmilik Universiti Teknikal Malaysia Melaka.
2. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. ** Sila tandakan (/)

_____ SULIT (Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)

_____ TERHAD (Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/ badan di mana penyelidikan dijalankan)

___/___ TIDAK TERHAD

(TANDATANGAN PENULIS)

Alamat tetap: No 33a, Jalan Pokok Mangga 1, Taman Pokok Mangga, 75250 Melaka.

Tarikh : _____

(TANDATANGAN PENYELIA)

DR.MOHD SANUSI BIN AZMI

Tarikh : _____

CATATAN: * Tesis dimaksudkan sebagai Laporan Akhir Projek Sarjana (PS).

** Jika tesis ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa.

DECLARATION

I declare that this thesis entitled “Cancer patients’ record queries using Answer On-Demand Data Sharing Architecture” is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature :

Name : Aiza Syahida binti Zakaria

Date : 18 August 2014

APPROVAL

I hereby declare that I have read this thesis and in my opinion this thesis is sufficient in term of scope and quality for the award of Master of Science Computer in Database Technology.

Signature :

Supervisor Name : Dr. Mohd Sanusi bin Azmi

Date : 18 August 2014

DEDECATION

I lovely dedicate this thesis to my beloved parents, “Zakaria bin Amin”, “Amni binti Awi”,

who supported me each step of the way.

To my siblings, “Izmir Azraf bin Zakaria”, “Aidil Syaheed bin Zakaria”, “Nuraini binti

Mohd”,

who inspired me with their love of learning, supporting and teaching.

To my nephew, “Iznur Zaquan bin Izmir Azraf”,

who give me passion and strength with his smile.

ABSTRACT

In medical tourism environment, one prominent concern is completeness of patients' medical records. In this research, the concern is on the completeness of medical tourism patients' medical records, which falls under the tuple – based completeness problem. The focus is on cancer patients whose complete medical records are highly crucial. One way to acquire complete medical records for all cancer patients is by gathering the medical records from every individual hospital. The method used is full data integration approach that provides unified view and access of data from heterogeneous data sources. In details, the proposed framework applies on - demand data sharing techniques to the architecture in medical tourism healthcare system. This results in reducing the search time and data anomalies improve the efficiency of DBMS operations, build the architecture with data integration friendly, number of data sources is counted and standardize the database product. As a support to the results gained, there are strategies on the practicality of on – demand data sharing in the process to improve completeness and access of patients' medical records.

ABSTRAK

Dalam persekitaran pelancongan perubatan, satu kebimbangan yang utama adalah kesempurnaan rekod perubatan pesakit. Dalam kajian ini, kebimbangan itu menjurus kepada kesempurnaan rekod perubatan pesakit yang termasuk dalam permasalahan “tuple”. Tumpuan adalah kepada pesakit kanser yang rekod perubatan lengkapnya adalah sangat penting. Salah satu cara untuk memperolehi rekod perubatan pesakit kanser yang lengkap adalah dengan mengumpul rekod perubatan daripada setiap hospital yang mana pesakit telah menerima rawatan. Kaedah yang digunakan adalah integrasi data penuh yang memberi satu pandangan yang bersatu dan akses data daripada sumber data yang banyak. Secara terperinci, rangka kerja yang dicadangkan menggunakan teknik perkongsian data di atas permintaan bagi seni bina dalam sistem penjagaan kesihatan pelancongan perubatan. Usaha ini telah berjaya mengurangkan masa carian dan anomali data, meningkatkan kecekapan operasi DBMS, membina seni bina dengan integrasi data mesra, beberapa sumber data dikira dan menyeragamkan produk pangkalan data. Sebagai sokongan kepada keputusan yang diperolehi, terdapat strategi asas kepada amalan perkongsian data di atas permintaan untuk meningkatkan lagi kesempurnaan dan akses rekod perubatan pesakit.

ACKNOWLEDGEMENTS

First and foremost, I would like to take this opportunity to express my sincere acknowledgements to my supervisor Dr. Mohd Sanusi bin Azmi from the Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM) for his essential supervision, support and encouragement towards the completion of this thesis.

I would like to sincerely appreciate the extraordinary effort given by the AJK of Project 1 and Project 2. They had arranged a briefing to explain the Master project and giving more information to carry out the project.

My appreciation also goes to my friends especially Siti Hajar Saam and classmate for their valuable opinion and solution that contributed in the development of this project.

Finally, I owe special thanks to my family Zakaria bin Amin, Amni binti Awi, Izmir Azraf bin Zakaria, Aidil Syaheed bin Zakaria, Nuraini binti Mohamed and Iznur Zaquan bin Izmir Azraf who always support me during performing this project.

Last but not least, I would like to thank those who were directly and indirectly involved in helping me out.

TABLE OF CONTENTS

	PAGE
DECLARATION	
DEDICATION	
ABSTRACT	i
ABSTRAK	ii
ACKNOWLEDGEMENTS	iii
TABLE OF CONTENTS	iv
LIST OF TABLES	vii
LIST OF FIGURES	viii
LIST OF APPENDICES	ix
LIST OF ABBREVIATIONS	x
CHAPTER	
1.0 INTRODUCTION	1
1.2 Background of Study	3
1.3 Problem Statement	5
1.4 Research Question	5
1.5 Research Objective	5
1.6 Research Scope and Limitation	6
1.7 Significant and Research Limitation	6
1.8 Organization of the Thesis	7
2.0 LITERATURE REVIEW	9
2.1 Introduction	9
2.2 Data Integration	9
2.3 Related Work	10
2.3.1 Thailand : SocialNTT	11
2.3.1.1 Functional Design and Implementation	13
2.3.2 Taiwan : Situation – Aware Medical Tourism Service Search System (SAMTS)	16

2.3.3 China : Medical Care Management System (MCMS)	18
2.3.4 Summarization on Medical Healthcare System Service in Asian countries	19
2.4 Data Quality	23
2.5 Data Completeness	25
2.6 Conclusion	27
3.0 MATERIALS AND METHODS/METHODOLOGY	28
3.1 Introduction	28
3.2 Research Framework	28
3.2.1 Investigation Phase	29
3.2.2 Implementation Phase	30
3.3 Proposed Framework	32
3.4 Research Requirement	34
3.5 Conclusion	34
4.0 IMPLEMENTATION AND TESTING	35
4.1 Introduction	35
4.2 Implementation of Proposed Framework	35
4.2.1 Data Preparation and Representation	36
4.2.2 Conceptual and Logical Database Design	37
4.2.2.1 Conceptual Database Design	37
4.2.2.2 Logical Database Design	38
4.2.2.3 Physical Database Design	41
4.2.3 Database Management System (DBMS) Selection	42
4.2.4 User Interface Design for Proposed System	42
4.2.5 Data Integration	44
4.2.5.1 Database link between Oracle 10g Express Edition to MySQL Server 5.6.19	44
4.2.5.2 Connection between Oracle to Java application Eclipse Kepler	45
4.2.6 Test Result and Analysis	45
4.3 Conclusion	45

5.0 RESULTS AND DISCUSSION	47
5.1 Introduction	47
5.2 Results of PHIS Development Using Answer On-Demand Data Sharing Technique	47
5.3 The Challenges and Strategies to Data Integration	48
5.3.1 Heterogeneous Data	49
5.3.2 Bad Data	50
5.3.3 Lack of Storage Capacity	50
5.3.4 Unanticipated Costs	51
5.3.5 Lack of Management Expertise	53
5.3.6 Perception of Data Integration as an Overwhelming Effort	54
5.4 Conclusion	54
6.0 CONCLUSION AND FUTURE WORK	56
6.1 Introduction	56
6.2 Research Summary	56
6.3 Limitations of Research	58
6.4 Future Research	58
REFERENCES	60
APPENDIX A – Data Definition Language	65
APPENDIX B – Database link between Oracle 10g Express Edition to MySQL Server 5.6.19	67
APPENDIX C – Connection between Oracle 10g Express Edition to Java Application	72
APPENDIX D – Sample Output of Patient Health Information System (PHIS)	74

LIST OF TABLES

TABLE	TITLE	PAGE
2.1	The details of SocialNTT functional modules	14
2.2	The details of SAMTS five-tier system	17
2.3	Summarization of three related work of Medical Tourism in Asian countries	20
2.4	The comparison of the application architecture in Medical Tourism	22
3.1	Software and hardware requirement	34
4.1	The details on 12 factors with description and example	36
4.2	Data dictionary for table Doctor	39
4.3	Data dictionary for table Patient	40
4.4	Data dictionary for table Treatment	41
4.5	The details on the main component in user interface design for PHIS	43

LIST OF FIGURES

FIGURE	TITLE	PAGE
1.1	A typical data integration architecture in Healthcare	4
2.1	Overall SocialNTT application	11
2.2	SocialNTT application architecture	12
2.3	SocialNTT functional design diagram	13
2.4	The proposed of SocialNTT client application	15
2.5	The architecture of SAMTS	16
2.6	The proposed architecture applying Fuzzy Technique to select suitable scenic spots for Medical Care	18
2.7	The data integration process	26
3.1	Research framework	29
3.2	Proposed framework	32
4.1	ERD for proposed system	38
4.2	User interface design for PHIS	43

LIST OF APPENDIXS

APPENDIX	TITLE	PAGE
A	Data Definition Language	65
B	Database link Oracle Database 10g Express Edition to MySQL Server 5.6.19	67
C	Connection Oracle Database 10g Express Edition to Java application Eclipse Kepler	72
D	Sample output of Patient Health Information System (PHIS)	74

LIST OF ABBREVIATIONS

JCI - Joint Commission International

MOH - Ministry of Health's Malaysia

SocialNTT - Social Network for Thailand Tourism application

TAT - Tourism Authority of Thailand

MVVM - Model View View Model

POI - Point of Interest

GPS - Global Positioning System

SAMTS - Situation-Aware Medical Tourism Service Search System

MUs - Mobile Users

UDDIRs - UDDI Registries

MTSPs - Medical Tourism Service Providers

MTSS - Medical Tourism Services Server

DS - Database Server

MCMS - Medical Care Management System MCMS

DBMS - Database Management System

DDL - Data Defination Language

CHAPTER 1

INTRODUCTION

Medical tourism in Malaysia has grown rapidly and forming a profitable industry to Malaysian government. Recent research on Malaysia Asia (David Jr, 2013) has shown, since 2008, some of the states in Malaysia have been active in medical healthcare which are Selangor, Kuala Lumpur, Penang and Melaka. The main factor behind this is that the country has great infrastructure, medical expertise, low cost, language compatibility, multiple choices of food, shopping and of course the many beautiful places to visit around Malaysia (David Jr, 2013).

In order to give assurance to the patient regarding the quality of service provided, there are international accreditations, JCI (Joint Commission International) which likely become the benchmarks for all the hospitals that targeting to participate in medical tourism (Rhenu Bhuller, 2014). In Malaysia, there has currently 8 JCI accredited hospitals listed; i) IJN – National Heart Institute, ii) International Specialist Eye Centre, iii) Gleneagles Medical Centre Penang, iv) Penang Adventist Hospital, v) Gleneagles Intan Medical Centre, vi) Prince Court Medical Centre, vii) Sime Darby Medical Centre Subang Jaya and viii) Pantai Hospital Kuala Lumpur (Anonymous, n.d). To achieve the clinical and operational quality standards, the stringent operational and accreditation requirement are in line with the MOH (Ministry of Health's Malaysia) strict medical tourism ruling.

In 2005, the MOH had established MHTC (The Malaysia Healthcare Travel Council) that act as the ultimate purpose to promote and position Malaysia as a unique destination for world-class healthcare services (Anonymous, n.d). To step up the promotional campaign, the MOH then took the initiative to brand Malaysian medical tourism by launching the Malaysia

Healthcare logo and tagline “Quality Care for Your Peace of Mind” in June 2009. This indirectly facilitates private and government sector collaboration. All the issues affecting this industry can be effectively addressed to ensure that foreign patients’ have a seamless experience with Malaysia healthcare service.

In medical tourism environment, one of the prominent concerns is the completeness of patients’ medical records. Any doctors who treat a foreign patient are able to obtain the patients’ complete medical records regardless of the hospitals where the treatment was given. Medical record consist of a variety of types of ‘notes’ entered over time by doctors, recording observations and administration of drugs and therapies, orders for the administration of drugs and therapies, test results, x-rays, reports, allergies and basic information. All the information is crucial for the doctors to diagnose certain disease before suitable type of treatments and prescriptions can be recommended. There are situation where the medical records are incomplete. This problem will cause the patients to experience similar tedious, time-consuming and painful medical procedures every time they are admitted for treatment.

Many healthcare systems nowadays, operate independently and autonomously due to the reasons such as lack of networking capabilities and legacy issues. One way to acquire complete medical records for all patients is by gathering medical records from each hospital where the medical history of patients’ is stored. This method maps the full data integration approach that provides unified view and access of data from different data sources. Nevertheless, one requirement of full data integration is to resolve differences called heterogeneity among the participating systems. This requirement makes the task of gathering complete records expensive and impractical.

This study investigates how completeness the medical records can be improved through on-demand data sharing where the participating systems are heterogeneous and limited. A case study will be conducted as means to understand and verify the problem of data completeness and data access faced by local healthcare system practitioners. The architecture of the model will be developed to portray the interaction among main elements of the healthcare systems that act as client and server, ubiquitous data access devices and web infrastructure. Finally, this study will conclude with recommendations on the practicality of

the on-demand data sharing to improve the completeness and access of medical records for medical tourism and healthcare system in general.

1.1 Background of Study

This study concern is on the data completeness of medical tourism patients' medical records, which falls under the tuple-based completeness problem. In details, the study focuses on cancer patients' whose complete medical record is highly crucial. As the consequence of incomplete medical records, cancer patients are required to experience similar tedious, time-consuming and usually painful medical procedures every time they admitted for treatment in the hospitals that they attended. This is common scenario that has been experienced by many cancer patients' due to the fact that many healthcare systems worldwide unfortunately operate independently and autonomously. Mphatswe et al., (2012) found the i) technical reason, ii) management reason and iii) legacy reason are the factors that become the obstacles for seamless systems and data integration.

One way to acquire complete medical records for all cancer patients is by gathering medical records from every individual hospital where the patients' medical records are stored. This method maps the typical full integration approach which is already adopted in Healthcare as shown in Figure 1.1 below.

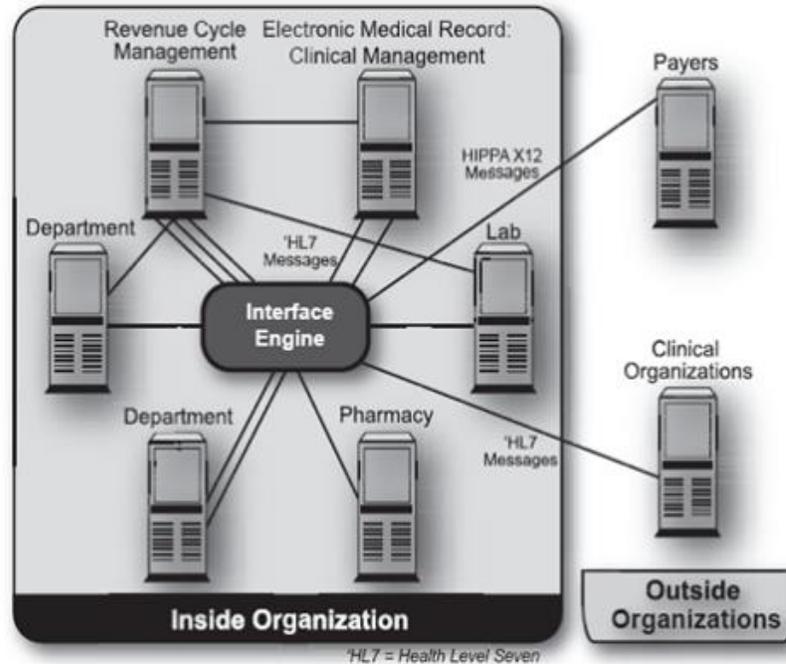


Figure 1.1: A typical data integration architecture in Healthcare (Motro 1989).

It combines data residing at different sources and provides unified access of data (Lenzerini et al. n.d.). Nevertheless, (Laboratories 1994) claim that one requirement of full data integration is to resolve differences called heterogeneity among the participating systems for example in terms of i) hardware and operating systems, ii) data management software, iii) middleware, iv) user interfaces and v) business rules.

A method which is a departure from the traditional data integration can improve completeness of medical records and data access by relaxing unrealistic integration requirement to be fulfilled by the participating data sources (system) is needed. In this research, the aimed is to test the hypothesis that completeness and accessibility of medical records can be improved through “on-demand data sharing” where participating system are heterogonous. By depth investigation of this form of on-demand data sharing from a fundamental point of view is crucial to improve completeness and access of medical records while promoting ‘virtual’ cooperation among medical tourism healthcare systems that are physically independent and heterogonous.

1.2 Problem Statement

Nowadays, in medical tourism, many healthcare systems operate independently. This poses difficulties for obtaining patient medical records. The doctors and patients are those who will be affected by this issue. It is because, based on the observation, both of them is relying on the medical records to give and get treatment. This study will integrate the data sources into unified view and allows access from different data sources. There is a system that been through the process of improvement. The system can act as a platform that can be easily used to retrieve the medical records from different data sources and response directly to patients' record queries.

1.3 Research Question

Based on the problem statements as stated in an earlier section, the research questions of this study are as follows:

- a. How to integrate the medical records from heterogeneous data sources?
- b. How to develop a platform that allows extracting patients' medical records queries?
- c. How to ensure that the requested patients' medical record is available every time when it is being queries?

1.4 Research Objective

The objectives of this study are listed below:

- a. To enhance the integrated architecture of on – demand data sharing for medical tourism healthcare system.
- b. To apply a technique of data integration in order to answer on – demand cancer patients' records queries.

1.5 Research Scope and Limitation

This research takes into consideration the ethical and privacy issues. Official approval from Public Health Physician, Ministry of Health Malaysia, Putrajaya was obtained and help in order to get the sample data.

The scopes of this study are;

- a. Using the full data integration approach to have unified view and access of data from heterogeneous data sources.
- b. Using the proposed framework to show the improvement of data completeness.
- c. Using the prototype system to prove the ideas of on – demand data sharing.

1.6 Significant and Research Contribution

This study offers a significant contribution that is not limited to the medical tourism where the improvement of records completeness is the goal. The proposed architecture could be benefit for the domains where tight and expensive data integration is the barrier for data completeness, such as bioinformatics. In addition, the model is also applicable in many electronic healthcare systems that require complete and ubiquitous access of patients' records that supports virtual cooperation.

1.7 Organization of Thesis

This thesis consists of six chapters and structured as follows:

a. Chapter 1: Introduction

Chapter 1 is the introductory of this study. This chapter provides information about the origins of the study, problem statement, research question, objectives of the study, scope and limitation and the significant and research contribution.

b. Chapter 2: Literature Review

Chapter 2 mainly focuses on the literature review. It presents brief introduction of data integration, data completeness and on - demand data sharing approach. There are also related works in medical tourism services that already implemented in several Asian countries. Each of the technology, techniques and methods used is identified and discussed.

c. Chapter 3: Research Methodology

Chapter 3 is discusses on the research methodologies that used in this study in order to achieve the research objectives. It comprises an introduction, research framework, proposed framework, research requirement and also milestone.

d. Chapter 4: Implementation and Testing

Chapter 4 can be classified as a main task towards the successful of this study. It involves the design, coding, configuration, implementation and testing for the proposed framework illustrated in the Chapter 3.

e. Chapter 5: Results and Discussion

In Chapter 5, the results obtain from the implementation and testing process is compared with the previous related work. There are also problems with strategies to data integration is discussed.

f. Chapter 6: Conclusion and Future Work

Chapter 6 summarizes the overall of this study. The summarization includes the result obtained, answers of the research questions and the achievement of research objectives. This chapter also includes some suggestions for future work.