

**PENGINTEGRASIAN DATA EXTENSIBLE MARKUP LANGUAGE (XML)  
KE DALAM PANGKALAN DATA HUBUNGAN MENGGUNAKAN COMMON  
WAREHOUSE METAMODEL (CWM)**

WAN MOHD HAFFIZ BIN MOHD NASIR

Tesis ini dikemukakan  
sebagai memenuhi syarat penganugerahan  
ijazah Sarjana Sains (Sains Komputer)

Fakulti Sains Komputer dan Sistem Maklumat  
Universiti Teknologi Malaysia

SEPTEMBER 2007

Ingatan Tulus Ikhlas Kepada Mereka yang Tersayang

## PENGHARGAAN

“Dengan Nama Allah Yang Maha Pemurah Lagi Maha Penyayang”  
Syukur yang seagungnya buat Dia yang Maha Memiliki segalanya, yang memberi sebelum diri meminta, yang memberi lebih daripada yang diminta, lalu dengan kemurahanNya serta dengan izinNya jua saya berjaya menyiapkan projek sarjana ini. Pertama-tamanya, setinggi penghargaan kepada Prof. Madya Dr. Shamsul bin Sahibuddin selaku penyelia utama dan Prof. Madya Dr. Md Nor bin Md. Sap selaku penyelia kedua, terima kasih di atas segala bimbingan dan tunjuk ajar yang diberikan semasa sesi perbincangan sepanjang proses menyiapkan tesis ini.

Salam penghargaan buat insan-insan yang tersayang yang sentiasa menyuntik semangat serta memberi sokongan dan dorongan tanpa jemu. Kepada sahabat-sahabat seperjuangan, samada yang terlibat secara langsung atau tidak, jasa baik dan tunjuk ajar daripada kalian amat dihargai.

Akhir kalam, salam penghargaan dan kasih sayang buat semua yang terlibat dalam membantu sehingga terhasilnya tesis ini. Semoga ada kebaikan dengan izinNya jua.

## ABSTRAK

Permintaan kepada pengintegrasian data secara pantas menjadi semakin tinggi dengan semakin banyak sumber-sumber maklumat yang terdapat di dalam perusahaan moden. *Extensible Mark-up Language* (XML) telah menjadi satu piawaian baru bagi perwakilan dan pertukaran data dalam *World Wide Web* (WWW), contohnya di dalam aplikasi *Business to Business* (B2B) pada e-dagang. Ini memerlukan alatan analisis data untuk mengendalikan data XML di samping format data tradisional. Tujuan penyelidikan ini adalah bagi meningkatkan kaedah pengintegrasian data XML ke dalam data hubungan berikutan berlakunya beberapa masalah daripada hasil proses pengintegrasian seperti kehilangan data. Kaedah yang dicadangkan daripada penyelidikan ini adalah melalui pengaplikasian *Common Warehouse Metamodel* (CWM) sebagai piawaian bagi pertukaran dan pengurusan metadata yang menggabungkan satu perkongsian metamodel bagi sintak dan semantik metadata. Hasil daripada penyelidikan ini adalah satu cadangan peningkatan senibina dan pendekatan pengintegrasian melalui pengaplikasian CWM serta satu perisian antaramuka yang telah dibangunkan bagi memudahkan proses pengintegrasian sebagai penyelesaian kepada masalah kehilangan data.

## **ABSTRACT**

The demand for rapid data integration is getting higher as more and more information sources appear in modern enterprises. Extensible Mark-up Language (XML) is fast becoming the new standard for data representation and exchange on the World Wide Web, e.g., in B2B e-commerce, making it necessary for data analysis tools to handle XML data as well as traditional data formats. The purpose of this research is to enhance the technique for XML data integration into relational data to solve integration problems such as missing data. The method that had been proposed is to apply Common Warehouse Metamodel (CWM) for metadata interchange and metadata management that incorporates a common shared metamodel for metadata syntax and semantics. The results of this research are the enhancement of the integration architecture and approach by applying CWM as well as the development of an interface software to simplify the integration process as a solution for the missing data problem.

## KANDUNGAN

<b>BAB</b>	<b>PERKARA</b>	<b>MUKA SURAT</b>
	<b>PENGESAHAN STATUS TESIS</b>	
	<b>PENGESAHAN PENYELIA</b>	
	<b>JUDUL</b>	<b>i</b>
	<b>PENGAKUAN PENULIS</b>	<b>ii</b>
	<b>DEDIKASI</b>	<b>iii</b>
	<b>PENGHARGAAN</b>	<b>iv</b>
	<b>ABSTRAK</b>	<b>v</b>
	<b>ABSTRACT</b>	<b>vi</b>
	<b>KANDUNGAN</b>	<b>vii</b>
	<b>SENARAI JADUAL</b>	<b>viii</b>
	<b>SENARAI RAJAH</b>	<b>ix</b>
	<b>SENARAI SINGKATAN</b>	<b>xiii</b>
	<b>SENARAI ISTILAH</b>	<b>xxi</b>
	<b>SENARAI SIMBOL</b>	<b>xxiii</b>
	<b>SENARAI LAMPIRAN</b>	<b>xxiv</b>
<b>1</b>	<b>PENGENALAN</b>	<b>1</b>
	1.1 Pengenalan	1
	1.2 Latarbelakang Masalah	2
	1.3 Penyataan Masalah	4

1.4	Matlamat Utama	5
1.5	Objektif	5
1.6	Skop Penyelidikan	6
1.7	Kepentingan Penyelidikan dan Sumbangan Ilmiah	6
1.8	Struktur Thesis	8
<b>2</b>	<b>KAJIAN LITERASI</b>	<b>10</b>
2.1	Pengenalan	10
2.2	Penyelidikan Pengintegrasian Data	11
2.3	XML	12
2.4	Metadata	12
2.5	Perbincangan	14
2.6	Pengintegrasian Data	16
	2.6.1 Sejarah dan Contoh Pengintegrasian Data	16
	2.6.2 Contoh Pengintegrasian Data	18
2.7	Common Warehouse Metamodel (CWM)	19
	2.7.1 Struktur CWM	20
2.8	Meta Integration	23
2.9	Pendekatan dan Senibina Pengintegrasian Sedia Ada	24
2.10	Kesimpulan	27
<b>3</b>	<b>METODOLOGI PENYELIDIKAN</b>	<b>28</b>
3.1	Pengenalan	28
3.2	Rangka Kerja Penyelidikan	30
	3.2.1 Formulasi Masalah (Fasa 1)	30
	3.2.1.1 Kajian Literasi	32
	3.2.1.2 Analisis Sistem Semasa	32
	3.2.1.3 Proposal Penyelidikan	33
	3.2.2 Pembangunan Sistem (Fasa 2)	33
	3.2.3 Implementasi dan Integrasi (Fasa 3)	34
	3.2.4 Penulisan Laporan (Fasa 4)	35

3.3	Sumber Data dan Peralatan	35
3.4	Proses Pergerakan Data	36
3.5	Perolehan Maklumat	37
3.6	Langkah-Langkah Proses Pergerakan Data	38
	3.6.1 Penakrifan Stor Sumber Data	38
	3.6.1.1 Pengurangan Sumber Data	39
	3.6.1.2 Penganalisaan Kandungan dan Struktur Data	40
	3.6.2 Penakrifan Stor Destinasi	41
	3.6.3 Aplikasi Transformasi Data	42
	3.6.3.1 Kemungkinan-Kemungkinan Aplikasi Transformasi Data	42
	3.6.3.2 Penentuan Transformasi-Transformasi Data	43
	3.6.4 Penentuan Antaramuka	44
	3.6.5 Penakrifan Keteguhan	44
	3.6.6 Senibina Aplikasi	45
	3.6.7 Penghasilan Senario Pengujian	45
	3.6.8 Migrasi Akhir dan Kriteria Kejayaan	46
3.7	Aplikasi Kepada Maklumat Data Elektronik	46
3.8	Cadangan Senibina Pengintegrasian	53
3.9	Rumusan	55
<b>4</b>	<b>PERLAKSANAAN</b>	<b>57</b>
	4.1 Pengenalan	57
	4.2 Perbandingan Perlaksanaan Pengintegrasian	58
	4.3 Antaramuka Pengintegrasian	59
	4.4 Penyediaan Data dan Metadata	64
	4.5 Pergerakan Data XML	69
	4.6 Implementasi Pengintegrasian Data eBusiness	70
	4.6.1 Langkah-Langkah Awal	71
	4.6.2 Elemen-Elemen yang Digunakan	72



4.6.3	Proses Pemetaan	74
4.6.4	Penjanaan Migrasi Data	82
4.7	Kesimpulan	87
<b>5</b>	<b>PENGUJIAN DAN ANALISIS</b>	<b>88</b>
5.1	Pengenalan	88
5.2	Persekitaran Pengujian	89
5.3	Kriteria Perbandingan	90
5.4	Contoh Penyataan Pertanyaan	91
5.5	Matrik Pengujian Perbandingan	93
5.6	Pengujian Perbandingan Sebelum Peningkatan	94
5.6.1	Kajian Kes 1: e-Business	94
5.6.2	Kajian Kes 2: Sumber Manusia	100
5.7	Perbandingan Sebelum dan Selepas Peningkatan Pendekatan	102
5.8	Perbincangan	109
<b>6</b>	<b>PERBINCANGAN DAN KESIMPULAN</b>	<b>110</b>
6.1	Pengenalan	110
6.2	Hasil Penyelidikan	111
6.3	Pencapaian Objektif Kajian	112
6.4	Kebaikan dan Kelemahan Pendekatan Pengintegrasian	113
6.5	Cadangan Pembaikan	115
6.6	Penyelidikan Masa Hadapan	115
6.7	Kesimpulan	116
	<b>RUJUKAN</b>	<b>117</b>
	<b>Lampiran A - C</b>	<b>121 - 134</b>

**SENARAI JADUAL**

<b>NO JADUAL</b>	<b>TAJUK</b>	<b>MUKA SURAT</b>
3.1	Skema Data dan Contoh Data bagi Pengujian	46
4.1	Perbandingan Pengintegrasian Sebelum dan Selepas Peningkatan	58
5.1	Contoh Penyataan Pertanyaan	91
5.2	Matrik Perbandingan Pengujian	93
5.3	Bilangan Kehilangan Data bagi P1	95
5.4	Bilangan Kehilangan Data bagi P2	96
5.5	Bilangan Kehilangan Data bagi P3	97
5.6	Bilangan Kehilangan Data bagi P4	98
5.7	Bilangan Kehilangan Data Kajian Kes 1	99
5.8	Bilangan Kehilangan Data Kajian Kes 2	101

5.9	Bilangan Kehilangan Data P1, P2, P3, dan P4 selepas Peningkatan	103
5.10	Peraturan Perbandingan Kehilangan Data bagi P1, P2, P3, dan P4	104
5.11	Bilangan Kehilangan Data Kajian Kes 2 selepas Peningkatan	106
5.12	Peraturan Perbandingan Kehilangan Data bagi P5,P6, dan P7	107
5.13	Peraturan Pengurangan Kehilangan Data Keseluruhan	108

**SENARAI RAJAH**

<b>NO RAJAH</b>	<b>TAJUK</b>	<b>MUKA SURAT</b>
2.1	Skema bagi Gudang Data	17
2.2	Common Warehouse Metamodel	21
2.3	Senibina Pengintegrasian Data Mikael R. Jensen	24
3.1	Rangka Kerja Operasi	31
3.2	Sumber Data dan Instrumentasi	36
3.3	DTD bagi Dokumen Jualan	48
3.4	Contoh Dokumen XML yang menuruti DTD	49
3.5	DTD bagi Dokumen Pemetaan	50
3.6	Dokumen XML bagi Pemetaan	50
3.7	DTD bagi Dokumen Komponen	51
3.8	Dokumen Komponen XML	53

3.9	Senibina Pengintegrasian Data and Metadata XML	54
4.1	Skrin ‘Splash’ bagi Data Integration Interface	59
4.2	Menu Utama bagi Data Integration Interface	60
4.3	Antaramuka Data Integration Interface	61
4.4	Menu “About” bagi Data Integration Interface	62
4.5	Menu “Setting Manager” bagi Data Integration Interface	63
4.6	Antaramuka bagi Meta Integration Model Bridge (MIMB)	64
4.7	Dokumen XML bagi Tempahan Belian	66
4.8	DTD bagi Tempahan Belian	67
4.9	Model CWM bagi Pangkalan Data Jualan	68
4.10	Penggunaan MIMB dalam Penukaran Model Data	69
4.11	Pengintegrasian Data eBusiness	72
4.12	Operasi pada Pemetaan Atribut “ShipAddress”	76
4.13	Klaus ‘Where’ pada Pemetaan Kelas “Orders”	78
4.14	Pemetaan Kelas “ <i>Orders_Details</i> ”	81

4.15	Persediaan bagi Penjanaan Kod	82
4.16	Penentuan Destinasi	83
4.17	Set Paramater Masa Larian	84
4.18	Pemindahan Perpustakaan yang diperlukan bagi Penjanaan Kod	85
4.19	Hierarki Hasil Pemindahan Perpustakaan	85
4.20	Penjanaan Kod daripada <i>Command Line</i>	86
5.1	Graf Perbandingan Bagi Pertanyaan P1	96
5.2	Graf Perbandingan Bagi Pertanyaan P2	97
5.3	Graf Perbandingan Bagi Pertanyaan P2	98
5.4	Graf Perbandingan Bagi Pertanyaan P4	99
5.5	Graf Perbandingan Bagi Pertanyaan P1,P2,P3, dan P4	100
5.6	Graf Perbandingan Bagi Pertanyaan P5, P6, dan P7	102
5.7	Graf Perbandingan Sebelum dan Selepas Peningkatan bagi Kajian Kes 1	105
5.8	Graf Perbandingan Sebelum dan Selepas Peningkatan bagi Kajian Kes 2	107

5.9	Graf Perbandingan Purata Kehilangan Data	108
-----	--	-----

## SENARAI SINGKATAN

CWM	-	Common Warehouse Metamodel
XML	-	Extensible Mark-up Language
DTD	-	Document Type Definition
MIW	-	Meta Integration Works
MIMB	-	Meta Integration Model Bridge
B2B	-	Business To Business
OLAP	-	Online Analytical Processing
ETL	-	Extract, Transform, & Load
BI	-	Business Intelligence
OMG	-	Object Management Group
UML	-	Unified Modeling Language
URL	-	Uniform Resource Locator
OIM	-	International Organization for Migration
WWW	-	World Wide Web



**SENARAI ISTILAH**

Sistem Pengurusan Pangkalan Data	-	<i>Database Management System</i>
Capaian semula maklumat	-	<i>Information retrieval</i>
Pertanyaan	-	<i>Query</i>
File Teks	-	<i>Text file</i>
Hubungan	-	<i>Relation</i>
Jadual	-	<i>Table</i>
Bahasa Pertanyaan Berstruktur	-	<i>Structured Query Language</i>

**SENARAI LAMPIRAN**

<b>NO LAMPIRAN</b>	<b>TAJUK</b>	<b>MUKA SURAT</b>
A	Aturcara kod bagi “Data Integration Interface”	121
B	Kod Aturcara Menu Utama Data Integration Interface	126
C	Contoh Sebahagian Kandungan Model Data CWM	129

## **BAB 1**

### **PENGENALAN**

#### **1.1 Pengenalan**

XML secara pantas telah menjadi satu piawaian bagi perwakilan dan pertukaran data. Ia menyediakan satu format yang diterima bagi mempersembahkan kedua-dua struktur dan kandungan data. Oleh itu ia boleh digunakan bagi membantu pengintegrasian data berstruktur, separa berstruktur dan tidak berstruktur. Penyelidikan ini pula bertitik tolak kepada hakikat bahawa XML itu sendiri secara tunggal tidak dapat menyediakan penyelesaian komprehensif kepada kesinambungan masalah daripada pengintegrasian data. Terdapat beberapa cabaran yang perlu ditangani.

Penyelidikan ini melihat kepada isu pengintegrasian data XML dan data hubungan bagi menyiasat dan menangani beberapa cabaran yang timbul daripada pelaksanaan pengintegrasian data XML dan data hubungan dengan memberikan fokus kepada metadata-metadata yang terlibat dalam pelaksanaan pengintegrasian tersebut.

## 1.2 Latar Belakang Masalah

Pengintegrasian Sumber Data XML memerlukan lebih dari satu alatan bagi menyusun data ke dalam sintak yang sama. Pengintegrasian data adalah aktiviti kompleks yang melibatkan peringkat yang berbeza iaitu dari segi model data, skema data dan misalan data. Penyesuaian model data memerlukan model data yang biasa bagi memetakan maklumat daripada sumber data yang berbeza. Setelah dapat diterima satu model data biasa yang harus digunakan, masalah yang berbangkit pula adalah penyesuaian perwakilan berbeza bagi entiti yang sama. Sebagai contoh dua sumber menggunakan nama yang berbeza bagi konsep yang sama (“harga” dan “kos”) atau nama yang sama bagi mewakili konsep yang berbeza. (Bertino dan Ferrari, 2001)

Menurut Bertino dan Ferrari (2001) lagi, kunci utama pengintegrasian data adalah satu bahasa bagi menentukan semantik yang berkait dengan kandungan data. Matlamat utama adalah bagi menentukan senibina, model-model, dan piawaian bagi menyediakan keterangan semantik yang boleh difahami. Beberapa isu memerlukan penyiasatan lanjut. Antara isu yang relevan bagi pengintegrasian data adalah pembangunan satu asas yang formal bagi metadata yang piawai. Seterusnya adalah pembangunan teknik dan alatan bagi penghasilan, pengekstrakan, dan penyimpanan metadata. Seterusnya adalah menyiasat persekitaran rangka kerja *interoperability* bagi semantik dan pembangunan alatan berasaskan semantik bagi tujuan penjelajahan pengetahuan. Pembangunan alatan yang sesuai bagi pengintegrasian berasaskan XML juga adalah penting. Alatan-alatan tersebut mesti menyokong pengintegrasian dengan lebih banyak automasi.

Satu contoh aplikasi pengintegrasian yang boleh dilihat di sini adalah sebuah pangkalan data yang mengandungi maklumat tempahan pelanggan perlu dikemaskini dengan maklumat tempahan belian. Masalahnya disini ialah maklumat tempahan belian tersebut ditulis dalam format XML. Kes ini memperlihatkan kepada kita kepada perlunya pengintegrasian data dilakukan supaya segala maklumat yang diperlukan boleh dicapai dan dianalisa khususnya oleh aplikasi pengguna akhir seperti sistem OLAP.

Menurut Do dan Rahm (2000), pengurusan metadata dan pengintegrasian adalah masalah nombor satu bagi pengintegrasian di dalam gudang data dan *business intelligence* kerana kebiasaannya ia melibatkan penggunaan pelbagai alatan dan produk dengan setiap satunya memiliki definisi dan format metadatanya yang tersendiri. Oleh yang demikian penghasilan, perkongsian, dan pengurusan metadata bagi alatan dan produk ini mengambil masa dan boleh mencetuskan masalah.

Menurut Mimno (2002), terdapat beberapa pendekatan bagi pengintegrasian data yang digunakan sekarang bagi menyelesaikan masalah tersebut tetapi semua pendekatan tersebut menggunakan metadata-metadata yang tersendiri. Ketiadaan pengurusan metadata yang piawai mengakibatkan masalah lain yang seringkali berlaku seperti kehilangan data yang mewujudkan persoalan kepada konsistensi data.

### **1.3 Penyataan Masalah**

Penyataan masalah yang utama dalam penyelidikan ini ialah:

“Apakah pendekatan pengintegrasian data XML dan metadata XML yang sesuai dan efektif bagi memastikan hasil pengintegrasian data yang konsisten dan tepat?

Berikut adalah beberapa penyataan masalah lain yang membawa kepada perlunya penyelidikan ini dilakukan:

- i. Bagaimanakah untuk menghasilkan satu senibina pengintegrasian data yang dapat menyediakan satu piawaian bagi pengintegrasian dan pemetaan metadata?
- ii. Apakah piawaian yang dapat digunakan di dalam pengurusan metadata bagi senibina yang dihasilkan?

### **1.4 Matlamat Penyelidikan**

Matlamat utama penyelidikan ini adalah meningkatkan dan mengimplementasi satu senibina perisian bagi pengintegrasian data XML dan metadata XML melalui penggunaan piawaian bagi pengurusan, pertukaran, dan perkongsian metadata.

## 1.5 Objektif

Objektif penyelidikan ini adalah seperti berikut:

- i. Mencadangkan satu senibina pengintegrasian data yang merupakan peningkatan daripada senibina pengintegrasian sedia ada iaitu daripada segi piawaian kepada pertukaran dan *interoperability* metadata.
- ii. Mengaplikasikan piawaian pengurusan metadata daripada industri dalam senibina yang dicadangkan bagi mendapatkan pertukaran dan perkongsian metadata yang meluas.
- iii. Mengimplementasikan senibina pengintegrasian data dan menggunakan metadata dalam persekitaran gudang data supaya dapat digunakan oleh aplikasi pengguna akhir seperti OLAP.

## 1.6 Skop Penyelidikan

Skop penyelidikan adalah bertumpu kepada:

- i. Pengintegrasian data XML dan data hubungan melalui pemetaan terhadap kedua-dua model data. Data XML dipilih sebagai sumber data pengintegrasian kerana penggunaannya yang meluas sebagai medium pertukaran data.

- ii. Model data yang digunakan bagi XML adalah model data DTD berbanding penggunaan XML Schema kerana penggunaan DTD adalah lebih meluas berbanding XML Schema.
- iii. Cadangan kepada satu senibina perisian bagi pengintegrasian data XML dengan data hubungan dengan CWM sebagai piawaian bagi pengurusan metadata bagi meningkatkan pengintegrasian data dan metadata.
- iv. Pengimplementasian adalah sehingga kepada hasil pemetaan kedua-dua model data dengan model data CWM sebagai model bagi pangkalan data serta kerana analisis yang akan dijalankan adalah terhadap hasil pengintegrasian data sahaja.
- v. Pengimplementasian juga adalah sehingga kepada migrasi atau perpindahan data daripada sumber data ke destinasiya iaitu daripada sumber data XML kepada pangkalan data hubungan kerana analisis yang akan dijalankan adalah terhadap hasil pengintegrasian data sahaja
- vi. Pengujian dilakukan bagi menilai dan membuat pengukuran kepada pendekatan dan senibina pengintegrasian data yang dicadangkan sebagai perbandingan kepada pendekatan dan senibina pengintegrasian data yang lain.

## **1.7 Kepentingan Penyelidikan dan Sumbangan Ilmiah**

Penyelidikan yang dijalankan adalah berdasarkan masalah yang timbul daripada isu ketiadaan penggunaan sistem pengurusan metadata yang seragam. Penyelidikan ini memfokuskan kepada pengintegrasian data serta metadata XML dan data hubungan di dalam sistem gudang data. Penyelidikan ini penting bagi mengatasi masalah di dalam pengintegrasian data iaitu dari segi keupayaan untuk mengekstrak data daripada sumber data yang berlainan dengan menyelesaikan perbezaan dalam definisi serta konsistensi



data. Oleh yang demikian satu pendekatan dan senibina proses pengintegrasian data yang baik diperlukan untuk bertindak sebagai enjin mengekstrak, mengubah, dan memuat yang berpandukan metadata bagi memastikan proses ini dilaksana secara optima dan terurus.

Penyelidikan turut melihat kepada aspek aplikasi pengguna akhir yang akan menggunakan data yang telah diproses. Aplikasi-aplikasi seperti Sistem Bantuan Keputusan, Kumpulan Sistem Bantuan Keputusan, dan Sistem Maklumat Eksekutif serta aplikasi lain yang akan menggunakan data OLAP memerlukan data yang lengkap bagi membantu pengguna akhir membuat analisis dan keputusan berdasarkan bantuan daripada aplikasi pengguna akhir tersebut. Ini adalah kerana data-data tersebut adalah bersifat kritikal dan sebarang kehilangan data boleh mempengaruhi analisis yang akan dijana oleh aplikasi pengguna akhir. Oleh itu, adalah amat penting untuk menghasilkan satu sistem pembangunan yang sempurna bermula daripada pemrosesan data itu sendiri sehingga kepada bahagian akhir pembangunan aplikasi.

Sememangnya telah ada pendekatan dan senibina sedia ada sebagai penyelesaian kepada pengintegrasian data. Melalui kajian, pemerhatian, dan analisis kepada senibina dan pendekatan tersebut mendapati wujudnya masalah lain seperti kehilangan data akibat perbezaan format dan definisi. Penyelidikan ini penting kerana dengan peningkatan kepada pendekatan dan senibina tersebut akan dapat menyelesaikan isu-isu yang telah dikenal-pasti. Kajian ini akan dapat memberi panduan kepada pembangun-pembangun pengintegrasian data bagi membina satu sistem gudang data yang baik bagi menghasilkan aplikasi pengguna akhir yang efisien dan lengkap.

Selain itu, penyelidikan ini turut memberikan sumbangan ilmiah yang lain seperti penggunaan piawaian Common Warehouse Metamodel (CWM) dan pemetaan model

data daripada dua sumber data yang berbeza. Penyelidikan ini menghuraikan proses lengkap pengintegrasian bermula daripada pemuatan model data sehingga kepada migrasi akhir. Kolaborasi pengintegrasian data bersama CWM akan menjadi satu pendekatan dan senibina baru sebagai rujukan dan panduan kepada pembangun-pembangun pengintegrasian data di dalam persekitaran gudang data.

## **1.8 Struktur Thesis**

Laporan ini membincangkan mengenai penyelidikan yang telah dijalankan dengan secara keseluruhannya mengandungi tujuh bab.

Bab 1 adalah merupakan bab pengenalan kepada penyelidikan yang dilaksanakan. Bab ini menjelaskan mengenai latar belakang masalah penyelidikan, pernyataan masalah yang terhasil daripada masalah penyelidikan, objektif dan matlamat penyelidikan. Seterusnya dalam bab ini, dinyatakan mengenai skop penyelidikan dan kepentingan penyelidikan.

Bab 2 adalah perbincangan mengenai kajian literasi yang telah dijalankan. Ia melihat kepada pendekatan dan senibina pengintegrasian data yang sedia ada. Piawaiian bagi pengurusan metadata juga turut dibincangkan.

Bab 3 pula menerangkan mengenai metodologi penyelidikan. Dalam bab ini, penulis membincangkan kaedah-kaedah yang terlibat dalam penyelidikan ini.

Bab 4 akan membincangkan mengenai pendekatan dan senibina pengintegrasian data yang ditingkatkan. Segala proses-proses yang terlibat akan dihuraikan satu persatu di dalam bab ini.

Bab 5 membincangkan proses implementasi terhadap senibina yang dicadangkan. Proses implementasi antara lain yang terkandung dalam bab ini adalah seperti proses pemuatan sumber data, penetapan elemen data, pemetaan model data sehinggalah kepada migrasi data daripada sumber data ke destinasi.

Bab 6 membincangkan mengenai pengujian yang dilaksanakan dalam penyelidikan ini. Proses seperti query dilaksanakan bagi menentukan bahawa proses pengintegrasian dan migrasi data berjaya dilaksanakan di samping membuat pengukuran dari segi ketepatan data. Selain itu, jadual dan graf bagi membuat perbandingan bagi memperlihatkan kelebihan senibina dan pendekatan ini turut dimuatkan.

Bab 7 adalah bab perbincangan yang cuba menghuraikan secara keseluruhan aspek-aspek penyelidikan seperti hasil penyelidikan, kerja-kerja peningkatan bagi masa hadapan, dan kesimpulan bagi keseluruhan laporan thesis ini.