

**A STUDY ON INFORMATION EXTRACTION FROM
HETEROGENEOUS, DISTRIBUTED AND AUTONOMOUS DNA
DATABASES USING ONTOLOGY-BASED AGENT**

**RUHAIDAH SAMSUDIN
MUHAMAD RAZIB OTHMAN
SAFAAI DERIS
ROSLI MD. ILLIAS**

**FAKULTI SAINS KOMPUTER DAN SISTEM MAKLUMAT
UNIVERSITI TEKNOLOGI MALAYSIA**

2005

ABSTRACT

WWW is an information resource with virtually unlimited potential. However, this potential is relatively untapped because it is difficult for machine to process and integrate this information meaningfully. This is where the development of semantic web comes in, because unlike the existing WWW, where data is primarily intended for human consumption, the Semantic Web provides data that is also machine processable. Simple HTML Ontology Extension (SHOE), with its clear and efficient capabilities, is best applied in support the development of semantic web by allowing web authors to annotate their pages with machine-readable knowledge that can then be used by agents and query engines. Ontology-design methodology (Noy and McGuinness 2001) is a promising approach to demonstrate SHOE, since; all of SHOE's implicit knowledge arise from its use of ontologies. It is expected that this project will bring the web to its full potential, as well as expanding the application of SHOE in order to support the development of semantic web.

ABSTRAK

WWW mengandungi koleksi halaman maklumat di dunia yang boleh dicapai oleh pengguna dengan cepat dan mudah serta merupakan sumber maklumat dengan potensi yang hampir tidak terhad. Namun, potensi ini secara relatifnya masih belum dapat diaplikasikan sepenuhnya kerana adalah sukar bagi sebuah mesin untuk memproses dan mengintegrasikan maklumat yang besar dan kompleks ini dengan lebih bermakna. Kelemahan yang menjadi alasan kepada pembentukan semantik web ini merupakan satu contoh masalah di *World Wide Web* (WWW) di mana semantik web berperanan menyediakan data yang berupaya diproses oleh mesin selain menyediakan data untuk penggunaan manusia semata-mata. Sambungan Mudah HTML Ontologi (SHOE) yang merupakan bahasa ontologi prototaip untuk web digunakan sebagai pendekatan utama di dalam menyelesaikan isu semantik web amnya dan khususnya di dalam isu pencarian di web ini di mana, pengarang web dapat menganotasi laman web mereka dengan pengetahuan yang difahami oleh mesin dan kemudian dapat digunakan oleh agen dan enjin carian. Bagi membenarkan SHOE diaplikasikan di dalam projek ini, metodologi bagi analisa dan rekabentuk ontologi (Noy dan McGuinness, 2001) telah digunakan, ini adalah kerana pembangunan ontologi merupakan teras kepada usaha ini. Projek ini diharap dapat meningkatkan potensi carian web, di samping memperluaskan penggunaan SHOE bagi menyokong usaha pembangunan semantik web.

KANDUNGAN

BAB	PERKARA	MUKA SURAT
	ABSTRAK	ii
	ABSTRACT	iii
	KANDUNGAN	iv
	SENARAI JADUAL	viii
	SENARAI RAJAH	ix
	SENARAI SINGKATAN	xi
	SENARAI ISTILAH	xii
	SENARAI LAMPIRAN	xiv
BAB I	Pengenalan	1
	1.1 Pendahuluan	1
	1.2 Latar Belakang Masalah	2
	1.3 Pernyataan Masalah	3
	1.4 Rangka Kerja Teoritikal	3
	1.5 Matlamat dan Objektif	4
	1.6 Skop	5
	1.7 Kepentingan Projek	6
	1.8 Definisi Istilah	6
	1.9 Susunan Tesis	7
BAB II	Kajian Literatur	8
	2.1 Pendahuluan	8
	2.2 Semantik Web	9
	2.2.1 Konsep Semantik Web	9
	2.2.2 Faedah Semantik Web	11
	2.2.3 Aplikasi Semantik Web	11

2.3	Ontologi	12
2.3.1	Konsep Ontologi	12
2.3.2	Faedah Ontologi	13
2.3.3	Aplikasi Ontologi	14
2.4	Sambungan Mudah HTML Ontologi	14
2.4.1	Konsep Sambungan Mudah HTML Ontologi	15
2.4.2	Faedah Sambungan Mudah HTML Ontologi	17
2.4.3	Aplikasi Sambungan Mudah HTML Ontologi	18
2.5	Carian Web	19
2.5.1	Isu dan Masalah Carian Web	21
2.5.2	Contoh-Contoh Enjin Carian	22
2.5.3	Signifikan Sambungan Mudah HTML Ontologi dalam Carian Web	22
2.6	Ringkasan	23
BAB III	METODOLOGI PROJEK	24
3.1	Pendahuluan	24
3.2	Reka Bentuk Projek	25
3.3	Rangka Kerja Operasi	26
3.3.1	Model Kitar Hayat Pembangunan Perisian Berasaskan Pengetahuan	26
3.3.2	Fasa Analisa Masalah Domain Ontologi	27
3.3.3	Fasa Reka Bentuk Ontologi	27
3.3.4	Fasa Implementasi dan Pengujian Ontologi	28
3.3.5	Fasa Integrasi dan Pengujian Aplikasi Ontologi	29
3.4	Instrumentasi dan Teknologi Kajian	30
3.5	Perancangan Projek	30
3.6	Ringkasan	31
BAB IV	ANALISA MASALAH DOMAIN ONTOLOGI	32
4.1	Pendahuluan	32
4.2	Rangka Kerja Fasa Analisa Masalah Domain Ontologi	33
4.2.1	Membina Gambarajah Kes Gunaan	34
4.2.2	Membina Gambarajah Aktiviti	34

4.2.3	Membina Gambarajah Jujukan	34
4.2.4	Membina Gambarajah Kelas	34
4.2.5	Mengenal Pasti Domain dan Skop Ontologi	35
4.2.6	Mengguna Semula Ontologi Sedia Ada	35
4.2.7	Menyenarai Istilah Penting	36
4.3	Implementasi Fasa Analisa Masalah Domain Ontologi Laman Web Jabatan Kejuruteraan Perisian	36
4.3.1	Gambarajah Kes Gunaan	37
4.3.2	Gambarajah Aktiviti	38
4.3.3	Gambarajah Jujukan	39
4.3.4	Gambarajah Kelas	39
4.3.5	Domain dan Skop Ontologi	40
4.3.6	Guna Semula Ontologi	42
4.3.7	Istilah Utama	42
4.4	Ringkasan	43
BAB V	REKA BENTUK ONTOLOGI	44
5.1	Pendahuluan	44
5.2	Rangka Kerja Fasa Reka Bentuk Ontologi	45
5.2.1	Mentakrif Kelas dan Hirarki Kelas	46
5.2.2	Mentakrif Slot	46
5.2.3	Mentakrif Faset	47
5.2.4	Mentakrif Tika	47
5.2.5	Membina Gambarajah Pakej Ontologi	48
5.2.6	Membina Gambarajah Komponen Ontologi	48
5.2.7	Membina Gambarajah Penempatan Ontologi	48
5.3	Implementasi Fasa Reka Bentuk Ontologi Laman Web Jabatan Kejuruteraan Perisian	48
5.3.1	Kelas dan Hirarki Kelas	49
5.3.2	Slot	49
5.3.3	Faset	51
5.3.4	Tika	52
5.3.5	Gambarajah Pakej Ontologi	53
5.3.6	Gambarajah Komponen Ontologi	54

5.3.7	Gambarajah Penempatan Ontologi	55
5.4	Ringkasan	56
BAB VI	IMPLEMENTASI DAN PENGUJIAN ONTOLOGI SERTA INTEGRASI DAN PENGUJIAN APLIKASI ONTOLOGI	57
6.1	Pendahuluan	57
6.2	Rangka Kerja Fasa Implementasi dan Pengujian Ontologi	58
6.3	Rangka Kerja Fasa Integrasi dan Pengujian Aplikasi Ontologi	58
6.4	Hasil Implementasi	59
6.5	Hasil Integrasi	68
6.6	Hasil Pengujian	72
6.6.1	Pengujian Pengejaan Perkataan	73
6.6.2	Pengujian Format HTML	74
6.6.3	Pengujian Pautan	74
6.6.4	Pengujian Kesesuaian Pengimbas	75
6.6.5	Pengujian Masa Muatan	76
6.7	Ringkasan	77
BAB VII	PERBINCANGAN DAN KESIMPULAN	78
7.1	Pendahuluan	78
7.2	Sumbangan Dan Pencapaian Projek	79
7.3	Masalah dan Kelemahan Projek	81
7.4	Cadangan Pembaikan dan Peningkatan Projek	82
7.5	Kesimpulan	83
	SENARAI DOKUMEN RUJUKAN DAN BIBLIOGRAFI	86
	LAMPIRAN A – D	89

SENARAI JADUAL

NO. JADUAL	TAJUK	MUKA SURAT
4.1	Senarai istilah bagi taksonomi jabatan	43

SENARAI RAJAH

NO. RAJAH	TAJUK	MUKA SURAT
1.1	Rangka kerja teoritikal	4
2.1	Asas seni bina Sambungan Mudah HTML Ontologi	16
3.1	Reka bentuk projek berdasarkan rangka kerja Tiga Peringkat	25
3.2	Model kitar hayat pembangunan perisian berasaskan pengetahuan	26
3.3	Fasa analisa masalah domain ontologi	27
3.4	Fasa reka bentuk ontologi	28
3.5	Fasa implementasi dan pengujian ontologi	29
3.6	Fasa integrasi dan pengujian aplikasi ontologi	29
4.1	Aktiviti dalam fasa analisa masalah domain ontologi	33
4.2	Gambarajah kesgunaan bagi gelintaran web menggunakan SHOE	38
4.3	Gambarajah kelas bagi kesgunaan Tapis Pengetahuan	40
4.4	Petikan paparan utama laman web Jabatan Kejuruteraan Perisian	41
4.5	Petikan paparan domain dan skop ontologi bagi laman web Jabatan Kejuruteraan Perisian	41
4.6	Petikan paparan ontologi universiti yang boleh diguna semula oleh ontologi jabatan	42
5.1	Aktiviti dalam fasa reka bentuk ontologi	45
5.2	Petikan paparan kelas dan hirarki kelas bagi laman web Jabatan Kejuruteraan Perisian	49
5.3	Petikan paparan slot bagi kelas pensyarah	50
5.4	Petikan paparan slot bagi kelas perundingan	50
5.5	Petikan paparan faset bagi slot nama	51
5.6	Petikan paparan faset bagi slot nilai perundingan	51
5.7	Petikan paparan tika bagi kelas pensyarah	52
5.8	Petikan paparan tika bagi kelas perundingan	53

5.9	Gambarajah pakej ontologi	54
5.10	Gambarajah komponen ontologi	54
5.11	Gambarajah penempatan ontologi	55
6.1	Aktiviti dalam fasa implementasi dan pengujian ontologi	58
6.2	Aktiviti dalam fasa integrasi dan pengujian aplikasi ontologi	59
6.3	Petikan paparan utama laman web Jabatan Kejuruteraan Perisian berbentuk maklumat	60
6.4	Petikan paparan senarai program pengajian bagi Jabatan Kejuruteraan Perisian berbentuk maklumat	60
6.5	Petikan paparan senarai subjek yang ditawarkan bagi Jabatan Kejuruteraan Perisian berbentuk maklumat	61
6.6	Petikan paparan senarai pensyarah bagi Jabatan Kejuruteraan Perisian berbentuk maklumat	61
6.7	Petikan paparan pentadbiran bagi Jabatan Kejuruteraan Perisian berbentuk maklumat	62
6.8	Petikan paparan senarai pelajar bagi Jabatan Kejuruteraan Perisian berbentuk maklumat	62
6.9	Petikan kod atur cara HTML bagi laman web Jabatan Kejuruteraan Perisian berbentuk maklumat	63
6.10	Petikan paparan laman web Jabatan Kejuruteraan Perisian berbentuk ontologi	64
6.11	Petikan kod atur cara SHOE bagi laman web Jabatan Kejuruteraan Perisian berbentuk ontologi	66
6.12	Petikan kod atur cara SHOE bagi kategori pensyarah	67
6.13	Petikan kod atur cara SHOE bagi kategori kumpulan penyelidikan	67
6.14	Petikan paparan Expose	69
6.15	Petikan paparan Parka	69
6.16	Petikan paparan Antara Muka Parka untuk Pertanyaan	70
6.17	Contoh ringkasan laporan pengujian NetMechanic	73
6.18	Contoh laporan pengujian pengejaan perkataan	73
6.19	Contoh laporan pengujian format HTML	74
6.20	Contoh laporan pengujian pautan	75
6.21	Contoh laporan pengujian kesesuaian pengimbas	76
6.22	Contoh laporan pengujian masa muatan	76

SENARAI SINGKATAN

DAML	DARPA Agent Markup Language
FSKSM	Fakulti Sains Komputer dan Sistem Maklumat
HTML	HyperText Markup Language
HTTP	HyperText Transfer Protocol
KP	Kejuruteraan Perisian
OIL	Ontology Interchange Language
PIQ	Parka Interface for Queries
PLUS	Parallel Understanding System
RDF	Resource Description Framework
SHOE	Sambungan Mudah HTML Ontologi
SUO	Standard Upper Ontology
UML	Unified Modeling Language
URL	Universal Resource Locator
URI	Universal Resource Identifier
UTM	Universiti Teknologi Malaysia
W3C	World Wide Web Consortium
WWW	World Wide Web
XML	Extensible Markup Language

SENARAI ISTILAH

Agen Pintar	Intelligent Agent
Agen	Agent
Akses	Access
Borang	Form
Carian	Searching
Catatan	Annotation
Dedaun	Leaves
Editor Teks	Text Editor
Enjin Carian	Search Engine
Faset	Facet
Heuristik	Heuristic
Hiper Pautan	hyperlink
Hiperteks	Hypertext
Hirarki	Hierarchy
Indeks	Index
Istilah	Term
Kata Kunci	Keyword
Katalog	Catalog
Kelas	Class
Masa Muatan	Load Time
Maya	Virtual
Melayari	Browse
Model Air Terjun	Waterfall Model
Ontologi	Ontology
Ontologikal	Ontological
Parent Category	Kategori Iubapa
Pautan	Link

Pelayar	Browser
Pencatat Pengetahuan	Knowledge Annotator
Penerapan	Assertion
Pengimbas	Browser
Penjejak Web	Web Crawler
Penjejakan	Crawler
Pertanyaan	Query
Rangka Kerja Tiga Peringkat	Three-Tier Framework
Repositori	Repository
Sauh	Anchor
Semantik	Semantic
Sintak	Syntax
Slot	Slot
Taksonomi	Taxonomy
Teks	Text
Tika	Instance

SENARAI LAMPIRAN

NO. LAMPIRAN	TAJUK	MUKA SURAT
A	Carta Gantt bagi Perancangan Projek I	89
B	Carta Gantt bagi Perancangan Projek II	90
C	Gambarajah Aktiviti	91
D	Gambarajah Jujukan	94

BAB I

PENGENALAN

1.1 Pendahuluan

Kemajuan teknologi berasaskan web telah banyak mempengaruhi perkembangan pembangunan pelbagai jenis laman web seperti e-pendidikan, e-saham, e-bank dan sebagainya. Fenomena ini membolehkan pengguna memperolehi pelbagai jenis maklumat pada bila-bila masa secara efisien dan cepat. Namun, masalah pengguna ialah mendapatkan maklumat yang bertepatan dengan keperluan mereka kerana alatan yang digunakan seperti katalog atau enjin carian tidak dapat memahami konteks bagi makna perkataan dan hubungan semantik di antara perkataan bagi kata kunci yang diinputkan oleh pengguna. Topik-topik yang dikupas dalam bab ini adalah seperti berikut:

- (a) Latar belakang dan motivasi projek.
- (b) Pernyataan masalah dan isu-isu projek.
- (c) Rangka kerja teoritikal dan struktur projek.
- (d) Matlamat projek dan bagaimana ianya dicapai.
- (e) Kepentingan dan impak projek.
- (f) Skop sempadan dan batasan projek.
- (g) Definisi perkataan utama yang digunakan dalam projek.
- (h) Susunan dan bentangan tesis.

1.2 Latar Belakang Masalah

World Wide Web (WWW) adalah satu sistem hipermedia yang mengandungi koleksi maklumat laman web yang boleh dicapai oleh pengguna dengan cepat dan mudah. Ia juga merupakan sumber maklumat dengan potensi yang hampir tidak terhad. Namun, potensi ini secara relatifnya masih belum dapat diaplikasikan sepenuhnya kerana sukar bagi sesebuah mesin untuk memproses dan mengintegrasikan maklumat yang banyak dan teragih ini secara lebih bermakna. Baru-baru ini, para penyelidik telah memulakan penyelidikan ke atas potensi dalam mengaitkan kandungan web dengan makna yang lebih jelas dan tidak taksa untuk menghasilkan semantik web. Carian web yang dilakukan kerap memaparkan hasil keputusan yang tidak bertepatan dengan kehendak pengguna, walaupun teknologi enjin carian pada masa kini mengalami peningkatan dalam servis masing-masing, ia boleh diperbaiki sekiranya bahasa penanda semantik digunakan untuk menerangkan kandungan halaman web tersebut.

Tidak dinafikan bahawa perwakilan pengetahuan mampu menyelesaikan pelbagai masalah berkaitan dengan web. Namun, penyelidikan yang sedia ada tidak boleh diaplikasikan secara terus kepada semantik web. Berbeza dengan kebanyakan pangkalan pengetahuan tradisional, web sentiasa berubah dengan cepat dan mengandungi jumlah maklumat yang banyak. Satu kajian telah dibuat bagi menyelidik bagaimana perwakilan pengetahuan perlu berubah secara dinamik untuk disesuaikan dengan ciri-ciri yang dimiliki oleh web (Jeff Heflin, 2001). Kajian ini menyampaikan kaedah baru untuk mengintegrasikan sumber data web dengan menggunakan ontologi. Selain daripada memperincikan semantik dari set istilah, ontologi juga boleh memperluaskan dan melakukan semakan di antara ontologi yang berlainan.

Semantik web yang mempunyai potensi penyelidikan dan pembangunan yang besar ini didemonstrasi menggunakan Sambungan Mudah HTML Ontologi (*Simple HTML Ontology Extension*, SHOE) yang berasaskan bahasa perwakilan pengetahuan untuk menjadikan carian di laman web lebih tepat. Ianya juga merupakan bahasa ontologi prototaip untuk web dan digunakan untuk membangunkan ontologi yang boleh diguna semula dan dikongsi.

1.3 Pernyataan Masalah

Seperti yang tersedia maklum, WWW adalah gudang maklumat terbesar yang pernah dihimpunkan oleh manusia. Individu di seluruh pelosok dunia boleh mengakses maklumat daripada kandungan laman web. Namun, perkembangan saiz web telah mendatangkan masalah di dalam membuat carian. Masalah primer bagi projek ini ialah:

“Bagaimana ontologi digunakan untuk membuat carian web yang bersifat dinamik.”

Bagi mengatasi masalah ini, perkara-perkara berikut perlu diambil perhatian dalam menjadikan capaian laman web yang diinginkan dapat diperolehi walaupun saiz web itu sendiri semakin berkembang dari sehari ke sehari:

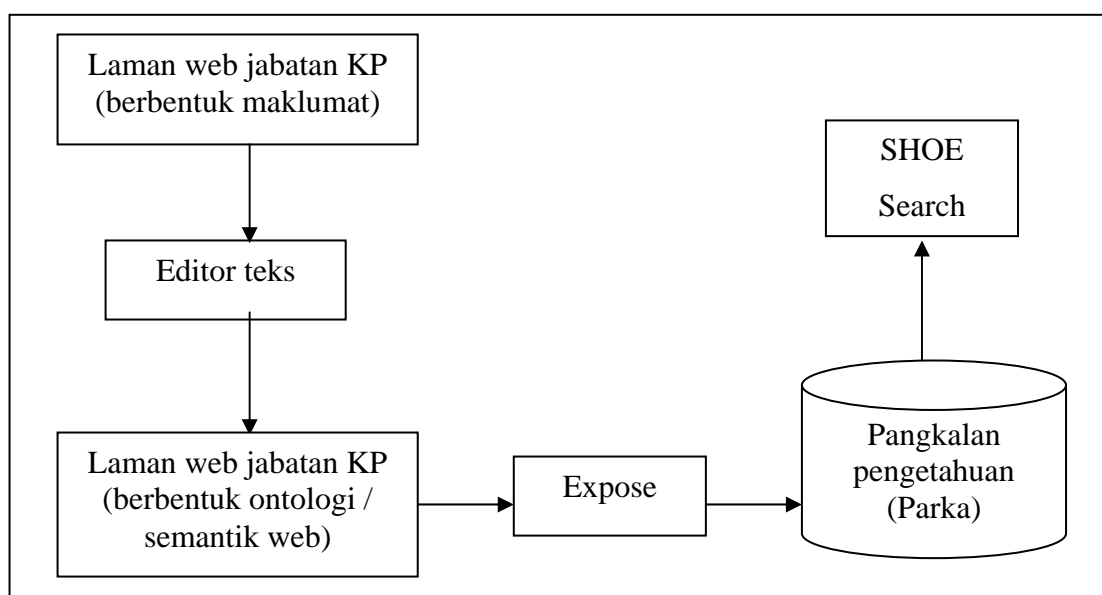
- (a) Bagaimanakah mekanisme untuk menukar suatu laman web yang berbentuk maklumat kepada laman web yang berbentuk pengetahuan? Di mana kandungan laman web yang berbentuk maklumat ini mempunyai penerangan yang panjang dan amat sukar bagi mesin untuk memahaminya.
- (b) Apakah mekanisme yang digunakan untuk mengendalikan masalah mengenai perwakilan pengetahuan di dalam web yang berciri dinamik dan mempunyai persekitaran teragih? Di mana ianya akan diaplikasi kepada semantik web.
- (c) Apakah mekanisme yang digunakan untuk membuat carian ke atas laman web yang berbentuk pengetahuan?

1.4 Rangka Kerja Teoritik

Laman web Jabatan Kejuruteraan Perisian (KP) yang merupakan salah satu jabatan di Fakulti Sains Komputer dan Sistem Maklumat (FSKSM), Universiti Teknologi Malaysia (UTM) digunakan sebagai kajian kes. Projek ini dimulakan dengan membangunkan laman web jabatan KP berbentuk maklumat dan kemudian, menterjemahkannya kepada laman web jabatan KP berbentuk pengetahuan dengan

menggunakan sintaks dan notasi SHOE yang bersesuaian merujuk kepada laman web berbentuk maklumat yang telah dibangunkan tadi. Proses ini dilakukan secara manual dengan menggunakan editor teks. Apabila penganotasian ke bentuk pengetahuan telah dilakukan, laman web yang berbentuk SHOE ini akan diindeks ke dalam pangkalan pengetahuan menggunakan Expose.

Pertanyaan terhadap pengetahuan ini kemudiannya boleh dilakukan dengan menggunakan SHOE Search. Melalui penggunaan SHOE Search dan Expose, antara muka yang dihasilkan memberi cara baru kepada pengguna untuk melayari web dengan membenarkan mereka menghantar pertanyaan yang kompleks dan membuka dokumen berkenaan dengan hanya mengklik kepada *Universal Resource Locator* (URL) yang dipaparkan hasil daripada pertanyaan yang telah pengguna lakukan. Gambarajah rangka kerja teoritikal bagi projek ini ditunjukkan pada Rajah 1.1.



Rajah 1.1 : Rangka kerja teoritikal

1.5 Matlamat dan Objektif

Matlamat projek ini adalah untuk membina laman web yang berinformasi dengan menggunakan SHOE agar carian yang lebih berkesan dan bermakna dapat

dilakukan. Bagi merealisasikan matlamat ini dan secara tidak langsung menjawab persoalan utama kajian dan isu-isu yang timbul, beberapa objektif perlu dicapai, iaitu:

- (a) Membangunkan laman web jabatan KP menggunakan *HyperText Markup Language* (HTML) sebagai laman web berbentuk maklumat.
- (b) Mengimplementasi ontologi yang digunakan untuk mengendalikan masalah mewakili pengetahuan di dalam web dan menukarkan web yang berbentuk maklumat kepada web berbentuk pengetahuan.
- (c) Mengimplementasi SHOE yang merupakan bahasa perwakilan pengetahuan berasaskan web yang digunakan untuk membangunkan semantik web.
- (d) Mengimplementasi Parka yang digunakan untuk menyimpan pengetahuan.
- (e) Mengimplementasi Expose dan SHOE Search yang digunakan untuk membuat carian kepada pengetahuan di dalam web.

1.6 Skop

Skop yang menjadi batasan dan sempadan dalam melaksanakan projek adalah seperti berikut:

- (a) Laman web jabatan KP yang berbentuk maklumat dibangunkan sebagai kajian kes untuk carian web dalam bidang pendidikan.
- (b) HTML dipilih sebagai bahasa untuk membangunkan laman web jabatan KP yang berbentuk maklumat.
- (c) Ontologi dipilih sebagai pendekatan untuk mewakili pengetahuan.
- (d) SHOE dipilih sebagai bahasa untuk membangunkan laman web yang berbentuk pengetahuan.
- (e) Parka dipilih sebagai alatan yang berfungsi sebagai pangkalan pengetahuan.
- (f) Expose dan SHOE Search dipilih sebagai alatan untuk membantu proses carian ke atas semantik web dan pangkalan pengetahuan.

1.7 Kepentingan Projek

Kajian di dalam carian web menggunakan SHOE merupakan bidang yang baru dan mempunyai potensi untuk terus berkembang di masa hadapan. Keyakinan ini turut diperkatakan oleh James Hendler (2001), di mana beliau menjangkakan dalam beberapa tahun yang akan datang setiap syarikat, universiti atau agensi kerajaan akan menukarkan sumber web mereka ke dalam bentuk pengetahuan dan disambungkan kepada pustaka ontologi atau ontolingua untuk diguna semula dan dikongsi. Ianya membolehkan maklumat bertukar di antara aplikasi dan membenarkan program komputer untuk mengumpul dan memproses kandungan web serta membuat penukaran secara bebas di antara satu sama lain tanpa sempadan. Selain itu, SHOE juga mampu berkembang di dalam persekitaran yang dinamik dan memberi banyak faedah kepada agen pintar di dalam web untuk membuat carian dan perlombongan pengetahuan.

1.8 Definisi Istilah

Maklumat	Merupakan fakta dan data yang disusun untuk menerangkan situasi atau keadaan sesuatu.
Pengetahuan	Merupakan fakta, kebenaran atau kepercayaan, perspektif atau konsep, pertimbangan atau jangkaan dan metodologi atau kaedah yang diaplikasi untuk menterjemah maklumat mengenai situasi dan menentukan bagaimana untuk mengatasinya.
Ontologi	Spesifikasi formal yang jelas mengenai istilah di dalam sesuatu domain dan perhubungan di antara istilah-istilah berkenaan.
Pangkalan pengetahuan	Repositori yang mengandungi ontologi berserta dengan set tika bagi kelas ontologi.

SHOE	HTML yang berasaskan bahasa perwakilan pengetahuan.
Expose	Penjejak Web (<i>Web-crawler</i>) yang membuat carian terhadap laman web berasaskan penanda SHOE.
SHOE Search	Antara muka yang membenarkan pengguna untuk membuat pertanyaan untuk mendapatkan jawapan bagi carian yang dilakukan daripada pangkalan pengetahuan.

1.9 Susunan Tesis

Terdapat tujuh bab yang membentuk kandungan tesis, keterangan bagi setiap bab adalah seperti berikut:

- (a) Bab I – menghuraikan latar belakang, pernyataan masalah, matlamat, rangka kerja teroritikal, kepentingan dan skop projek.
- (b) Bab II – menghuraikan latar belakang, isu dan keperluan membangunkan laman web yang berbentuk pengetahuan menggunakan ontologi.
- (c) Bab III – menghuraikan metodologi, reka bentuk, rangka kerja operasi dan teknologi yang digunakan dalam pelaksanaan projek.
- (d) Bab IV – membentangkan aktiviti dan artifak yang berkaitan dalam menganalisa masalah domain menggunakan Protégé dan Bahasa Pemodelan Tergabung (*Unified Modeling Language, UML*).
- (e) Bab V – membentangkan aktiviti dan artifak yang berkaitan dalam mereka bentuk laman web berbentuk maklumat dan semantik web menggunakan Protégé dan UML.
- (f) Bab VI – membentangkan hasil implementasi projek serta pengujian yang telah dilakukan.
- (g) Bab VII – membentangkan kesimpulan, sumbangan projek serta kerja masa hadapan.

BAB II

KAJIAN LITERATUR

2.1 Pendahuluan

Kajian literatur dilaksanakan bagi menerangkan kemajuan yang dicapai dalam penyelidikan ontologi dan sistem berasaskannya sama ada dari segi teori atau praktikal. Tumpuan utama diberikan kepada pendekatan-pendekatan yang digunakan untuk menangani dan menyelesaikan isu-isu serta masalah yang dihadapi dalam mewakili pengetahuan menggunakan ontologi dan membangunkan enjin carian berasaskan pengetahuan. Perkara signifikan yang perlu diselidiki ialah penjelasan mengenai keperluan dan kepentingan mengimplementasi ontologi. Selain itu, metodologi yang digunakan bermula daripada kajian konsep sehingga kepada implementasi juga perlu dijelaskan secara terperinci. Topik-topik yang dikupas dalam bab ini adalah seperti berikut:

- (a) Latar belakang, konsep, motivasi dan kepentingan semantik web serta contoh-contoh aplikasi berasaskan semantik web.
- (b) Latar belakang, konsep, motivasi dan kepentingan ontologi serta contoh-contoh aplikasi berasaskan ontologi.
- (c) Latar belakang, konsep, motivasi dan kepentingan SHOE.
- (d) Latar belakang dan isu carian web serta contoh-contoh enjin carian.
- (e) Signifikan SHOE dalam carian web.

2.2 Semantik Web

Semantik web telah diperkenalkan oleh Tim Bernes-Lee yang merupakan pencipta WWW, URIs, HTTP dan HTML. Ia merupakan jaringan maklumat berskala global yang saling berpaut antara satu sama lain supaya ianya mudah diproses dan difahami oleh mesin. Semantik web juga merupakan satu pendekatan yang efisien bagi mewakili data di dalam WWW sebagai satu pangkalan data yang saling berpaut. Menurut Bernes-Lee (1999), semantik web ditakrifkan sebagai:

“Semantic web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation.”

Bagaimanapun, masih ramai yang kabur mengenai semantik web. Seajar dengan kekeliruan ini, *Semantic Web Agreement Group* memberi penjelasan yang lebih mudah mengenai semantik web, iaitu:

“Semantic web is a web that includes documents or portions of documents, describing explicit relationships between things and containing semantic information intended for automated processing by our machines.”

Terdapat sekumpulan pengkaji yang berdedikasi di *World Wide Web Consortium* (W3C), diketuai oleh Tim Bernes-Lee berusaha untuk memperbaiki, memperluas dan mempiawaikan pelbagai jenis bahasa, penerbitan, alatan dan berbagai-bagai perkara lagi berkaitan dengan semantik web. Walaupun, teknologi semantik web masih berada di peringkat permulaan namun, masa hadapan bagi semantik web secara umumnya adalah cerah.

2.2.1 Konsep Semantik Web

Pada masa kini, WWW telah dibangunkan secara pantas dengan kewujudan hampir tujuh juta laman web dalam sehari sebagai medium dokumen kepada manusia berbanding ia bertindak sebagai maklumat yang boleh dimanipulasi secara automatik. Keadaan ini perlu diubah dengan mentransformasikan web kepada semantik web supaya komputer dapat mencari semantik sesuatu data dengan

mengikuti hiper pautan kepada definisi sesuatu istilah dan peraturan-peraturan tertentu bagi membolehkan proses penaakulan mengenainya dibuat secara logikal.

Terdapat beberapa bahasa semantik web yang piawai dengan semantik, sintak dan pendekatan yang berbeza, antaranya ialah:

- (a) Ontobroker (Decker, *et.al*, 1999) – merupakan bahasa berasaskan logik rangka yang merangkumi bahasa penakrifan ontologi, bahasa catatan laman web, perangkak web, enjin taabir dan satu set antara muka pertanyaan. Ciri utama bahasa ini adalah ianya padat dan mudah untuk mengekspresikan ontologi dan data secara bersepadu.
- (b) RDF (Lassila, 1998) – atau dikenali sebagai *Resource Description Framework* yang merupakan satu model yang mudah untuk mewakili sebarang bentuk dan jenis data. Ianya direkomendasi oleh W3C untuk mengatasi limit semantik XML.
- (c) OIL (Fensel, *et.al*, 2000) – atau dikenali sebagai *Ontology Interchange Language* yang merupakan bahasa bagi mendiskripsikan ontologi di dalam web. Semantik bahasa ini adalah berdasarkan logik diskripsi dan sintaksnya pula berlapiskan RDF.
- (d) DAML (Hendler dan McGuinness, 2000) – atau dikenali sebagai *DARPA Agent Markup Language* yang merupakan bahasa semantik web yang paling tinggi profilnya kerana ia menggabungkan ciri-ciri terbaik yang dimiliki oleh pelbagai bahasa semantik web seperti RDF, OIL dan SHOE.
- (e) SHOE (Heflin dan Hendler, 2000b) – menggabungkan ciri-ciri bahasa penanda hiperteks, perwakilan pengetahuan, log data dan ontologi untuk menangani masalah semantik di dalam web.

2.2.2 Faedah Semantik Web

Studer, *et.al* (1996), menyatakan bahawa di antara kepentingan dan signifikan semantik web adalah:

- (a) Menangani masalah carian, capaian, senggaraan dan janaan maklumat yang semakin berkembang dan besar amaunnya di dalam web.
- (b) Membantu pengguna mencapai sesuatu gol atau matlamat dengan mengakses dan menyediakan maklumat yang boleh difahami dan diproses oleh mesin bagi membolehkan mesin menyediakan perkhidmatan berasaskan taakulan.
- (c) Membolehkan pencapaian dan kejuruteraan pengetahuan serta perwakilan pengetahuan diaplikasikan untuk menangani masalah pembentukan pengetahuan untuk menjadikan aplikasi dan laman web lebih bermakna dan jelas.

2.2.3 Aplikasi Semantik Web

Menurut Heflin (2001), semantik web boleh diaplikasikan dan diadaptasikan dalam pelbagai domain, antaranya ialah:

- (a) Carian web – semantik web membantu pengguna mengorganisasi dan melayari web serta mendapatkan maklumat daripada web dengan lebih tepat dan bermakna.
- (b) Agen internet – semantik web membantu agen berkongsi maklumat secara global, memohon dan mengamati laman web serta melarikan perkhidmatan web bagi membolehkan agen berupaya untuk membuat perbandingan ketika membeli-belah, mengatur percutian dengan lebih sempurna dan mengambil bahagian di dalam jualan lelong.
- (c) Sistem tolakan – semantik web membantu pengguna memanipulasikan data yang terdapat di dalam web dengan lebih berkesan apabila melangani sesuatu perkhidmatan kerana hanya maklumat yang relevan sahaja diberikan kepada pengguna secara automatik.

2.3 Ontologi

Ontologi secara umumnya dianggap sebagai menyediakan makna bagi istilah yang digunakan untuk mewakilkan pengetahuan. Oleh itu, ontologi merupakan sebahagian daripada bahasa perwakilan. Membina ontologi merupakan usaha mewakilkan pengetahuan yang dimiliki oleh manusia ke dalam bentuk yang lebih berstruktur. Di dalam bidang kepintaran buatan definisi mengenai ontologi dirujuk kepada pernyataan yang dibuat oleh Gruber (1993), iaitu:

“An ontology is an explicit specification of a conceptualization.”

Manakala, menurut Guarino (1998) pula, beliau mentakrifkan ontologi sebagai:

“An ontology is a logical theory accounting for the intended meaning of a formal vocabulary, i.e., its ontological commitment to a particular conceptualization of the world. The intended models of a logical language using such a vocabulary are constrained by its ontological commitment. An ontology indirectly reflects this commitment (and the underlying conceptualization) by approximating these intended models.”

2.3.1 Konsep Ontologi

Istilah ontologi dipinjam dari bidang falsafah, di mana ontologi merupakan penerangan yang sistematik mengenai kewujudan. Di dalam bidang kepintaran buatan pula, sesuatu yang wujud itu boleh diwakili. Apabila pengetahuan mengenai sesuatu domain diwakili dalam bentuk yang formal, set objek yang boleh diwakilkan itu dikenali sebagai hujahan sejagat (seperti kelas, perhubungan, fungsi atau atribut). Pembangunan ontologi juga melibatkan pentakrifan set data dan strukturnya.

Struktur ontologi terdiri daripada kelas atau dikenali juga sebagai konsep. Sifat yang menerangkan pelbagai ciri dan atribut bagi konsep dipanggil slot atau dikenali juga sebagai peranan. Manakala, had ke atas slot dipanggil faset atau dikenali juga sebagai had peranan. Ontologi mengandungi satu set ahli (tika) bagi

sesuatu kelas untuk membentuk pangkalan pengetahuan. Mereka bentuk ontologi adalah sukar dan memakan belanja yang banyak di dalam usaha pembinaannya. Oleh itu, untuk mendapatkan hasil yang baik, maka pembinaannya perlulah bersaiz besar dan menyeluruh. Semakin banyak ontologi yang boleh dikongsi oleh pelbagai aplikasi, maka semakin berfaedah penggunaannya.

2.3.2 Faedah Ontologi

Peranan ontologi telah diselidiki dalam pelbagai bidang yang berbeza seperti berikut:

- (a) Perwakilan pengetahuan.
- (b) Kejuruteraan pengetahuan.
- (c) Reka bentuk pangkalan data.
- (d) Pemprosesan bahasa tabii.
- (e) Sistem pelbagai agen.
- (f) Semantik web.
- (g) Kejuruteraan perisian.
- (h) Sistem pengurusan maklumat.

Faktor yang menyebabkan kajian telah dijalankan ke atas ontologi adalah disebabkan oleh ontologi merupakan langkah yang mempunyai keupayaan yang tinggi di dalam menjelaskan secara formal sifat bagi domain berbanding dengan cara yang dibuat secara bebas dan tidak berstruktur. Disamping itu, ontologi juga digunakan untuk menggalakkan penggunaan semula dan pengintegrasian pengetahuan.

Di dalam bahagian aplikasi, keutamaan diberikan kepada penggunaan ontologi untuk perdagangan elektronik, integrasi perusahaan, pengurusan pengetahuan dan pengaksesan maklumat di dalam web. Bagi aplikasi perdagangan elektronik sebagai contoh, ianya merupakan bidang yang memerlukan data yang banyak kerana entiti seperti peruncit, pemborong dan pengeluar memerlukan aliran data di antara satu sama lain bagi melaksanakan transaksi sesuatu produk. Ontologi

diimplementasi ke dalam bidang ini bagi membolehkan perkongsian maklumat dan pengetahuan di antara aplikasi yang berbeza terutamanya ke atas istilah-istilah sepunya yang menerangkan domain aplikasi secara generik. Malahan, fleksibiliti ontologi boleh ditambah dengan menakrifkan hubungan di antara istilah tersebut.

2.3.3 Aplikasi Ontologi

Terdapat pelbagai ontologi formal yang berlainan dan ditulis dalam bahasa selain daripada SHOE, iaitu:

- (a) *Ontolingua* – servis ini disediakan oleh Knowledge System Laboratory di Universiti Stanford, di mana sejumlah ontologi disumbangkan oleh beberapa pengguna untuk dikongsi dan diguna semula.
- (b) *Ontosaurus* – pelayar ontologi dari USC's Information Sciences Institute yang menyediakan pelbagai ontologi yang ditulis menggunakan bahasa perwakilan pengetahuan Loom.
- (c) *WebOnto* – pelayar ontologi berasaskan Java dari Knowledge Media Institute yang mengandungi pelbagai ontologi.
- (d) *Cyc Upper Ontology* – merupakan projek yang dibangunkan dengan matlamat untuk membentuk ontologi berdasarkan kepada cara penaakulan yang optimum.
- (e) *IEEE Standard Upper Ontology (SUO)* – percubaan untuk membina dan menghasilkan piawaian bagi ontologi aras atas.

2.4 Sambungan Mudah HTML Ontologi

SHOE dibangunkan oleh kumpulan Parallel Understanding Systems (PLUS) iaitu ahli kumpulan penyelidikan yang diketuai oleh Professor Hendler daripada Universiti Maryland semenjak daripada tahun 1995. Sejak daripada awal pembangunannya sehinggalah sekarang, kumpulan PLUS ini telah memperbaiki

bahasa berkenaan dan mengeksperimen penggunaannya dari semasa ke semasa. SHOE merupakan bahasa HTML yang berasaskan bahasa perwakilan pengetahuan. SHOE menggabungkan ciri-ciri bahasa penanda, perwakilan pengetahuan, datalog dan ontologi di dalam usaha untuk menyelesaikan masalah yang unik mengenai semantik di dalam web. SHOE membantu enjin carian mendapatkan pengetahuan dengan menambah tag yang memberikan maksud semantik kepada sesuatu web. Struktur asas SHOE adalah sebagaimana struktur ontologi yang telah dibincangkan sebelum ini. Tag bagi SHOE dibahagikan kepada dua kategori, iaitu:

- (a) Tag untuk membina ontologi, di mana ontologi bagi SHOE adalah satu set peraturan yang menerangkan apakah jenis pernyataan dokumen SHOE yang boleh dibina dan apakah yang dimaksudkan oleh pernyataan tersebut.
- (b) Tag untuk catatan atau komen di dalam dokumen web.

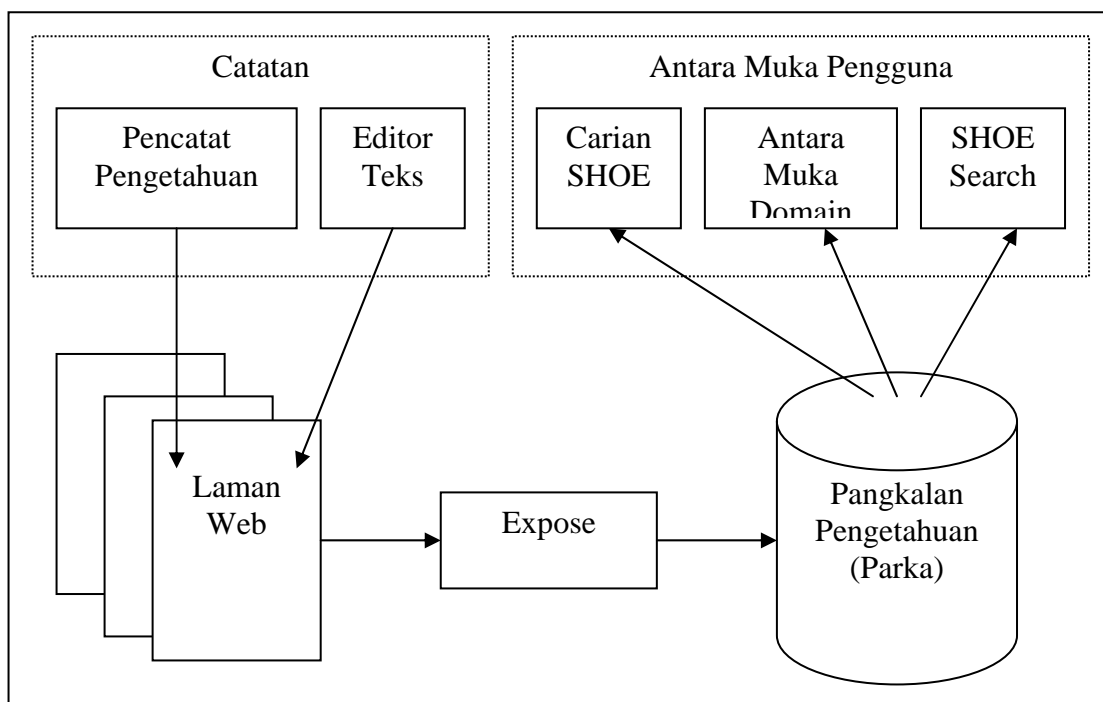
2.4.1 Konsep Sambungan Mudah HTML Ontologi

Sistem yang menggunakan bahasa SHOE perlu mengambil kira beberapa isu berkaitan dengan reka bentuk atau seni bina keseluruhan SHOE sebagaimana yang ditunjukkan oleh Rajah 2.1. Perkara yang perlu dipertimbangkan adalah bagaimana ontologi dibina dan kemungkinan penyediaan alat bantu untuk memprosesnya. Alatan tersebut perlu menyediakan kemudahan yang dapat membantu pengguna menambah pernyataan yang dikenali sebagai catatan kepada laman web pengguna. Disamping berupaya menentukan bagaimana pernyataan tersebut dicapai dan kemudian diproses. Reka bentuk yang umum mengenai SHOE adalah seperti berikut:

- (a) Reka bentuk ontologi – sebelum SHOE boleh digunakan, ontologi yang bersesuaian perlulah disediakan. Ontologi boleh sama ada direka bentuk oleh pengguna sendiri atau pengguna boleh mencari ontologi yang telah sedia untuk diguna pakai.
- (b) Catatan – tika SHOE ditambah kepada dokumen pengguna dengan menggunakan kaedah pencatat pengetahuan atau editor teks.
- (c) Pengaksesan maklumat – laman web yang telah direka oleh pencatat pengetahuan ini boleh ditanya dan diindeks dengan menggunakan

Expose, di mana ontologi disimpan dalam pangkalan pengetahuan yang dipanggil Parka.

- (d) Pemrosesan maklumat – apa sahaja alatan pertanyaan yang boleh mengakses pengetahuan SHOE yang tersimpan di dalam Parka.



Rajah 2.1 : Asas seni bina Sambungan Mudah HTML Ontologi

Beberapa alatan yang umum telah dibangunkan bagi menyokong pelaksanaan SHOE. Kebanyakan alatan dikodkan menggunakan bahasa Java. Antara alatan yang dimaksudkan itu adalah:

- (a) Pencatat pengetahuan – alatan ini menjadikan penambahan pengetahuan SHOE ke dalam laman web menjadi mudah melalui antara muka borang yang disediakan. Alatan ini juga mempunyai antara muka yang memaparkan tika, ontologi dan penerapan.
- (b) Expose – merupakan *web crawler* yang membuat carian terhadap laman web dengan penanda SHOE. Expose menyimpan pengetahuan yang telah dikumpulkan ke dalam pangkalan pengetahuan. Apabila Expose telah mengumpul pengetahuan daripada web, ia boleh menjawab pertanyaan yang sofistikated mengenai data yang telah dikumpulkan itu.

- (c) Parka – enjin pangkalan pengetahuan bagi Expose yang merupakan sistem perwakilan pengetahuan yang berprestasi tinggi. Parka membenarkan Expose menanyakan soalan yang sofistikated mengenai sejumlah besar data yang mempunyai perkaitan yang tinggi antara satu sama lain.
- (d) SHOE Search – membenarkan pengguna untuk membuat pertanyaan dan mengisukannya kepada pangkalan pengetahuan Parka. Apabila digunakan dengan Expose, pengguna dibenarkan untuk membuat pertanyaan dan kemudian membuka dokumen hasil daripada pertanyaan yang dilakukan dengan mengklik URL yang terdapat pada hasil paparan dari pertanyaan yang telah dilakukan.

2.4.2 Faedah Sambungan Mudah HTML Ontologi

SHOE memainkan peranan yang penting di dalam membantu pengguna membuat carian web dengan tepat dan lebih bermakna. Ia juga membantu di dalam memberi jawapan yang berkaitan berdasarkan kepada pengintegrasian automatik daripada pelbagai jenis sumber hasil daripada pertanyaan pengguna.

Bagaimanapun, penyelidikan ke atas SHOE masih diteruskan bagi mencari penyelesaian berkenaan dengan masalah di dalam semantik web. Skop bagi semantik web adalah sama luasnya jika diibaratkan dengan saiz web semasa. Oleh itu, kebolehskalaan adalah kritikal kepada sistem berasaskan semantik web. Selain daripada itu, walaupun SHOE adalah bahasa web berasaskan ontologi yang pertama dihasilkan tetapi terdapat pelbagai hala tuju untuk mencapai kemajuan ke atas bahasa web yang berasaskan ontologi ini. Para pengkaji perlu untuk membangunkan satu set kriteria bagi membuat penilaian terhadap bahasa semantik web tersebut.

2.4.3 Aplikasi Sambungan Mudah HTML Ontologi

SHOE telah dipilih sebagai pendekatan kepada usaha pembangunan semantik web di dalam projek ini merujuk kepada penggunaannya di dalam beberapa domain seperti berikut :

- (a) Domain sains komputer – kumpulan PLUS telah menganotasi sebanyak lima belas halaman web jabatan sains komputer yang menduduki senarai tertinggi di Amerika Syarikat. Sebagai contoh, berikut merupakan senarai beberapa halaman web universiti-universiti yang terlibat berserta URLnya. Universiti Maryland boleh dilayari di www.cs.umd.edu/projects/plus/SHOE/domains/cs/umd.html, MIT di www.cs.umd.edu/projects/plus/SHOE/domains/cs/mit.html, Stanford di www.cs.umd.edu/projects/plus/SHOE/domains/cs/stanford.html, Princeton di www.cs.umd.edu/projects/plus/SHOE/domains/cs/princeton.html, Universiti Michigan di www.cs.umd.edu/projects/plus/SHOE/domains/cs/umich.html, Berkeley di www.cs.umd.edu/projects/plus/SHOE/domains/cs/berkeley.html dan Cornell di www.cs.umd.edu/projects/plus/SHOE/domains/cs/cornell.html. Objektifnya adalah untuk menilai sejauh mana keberkesanan di dalam menambah SHOE ke dalam laman web yang mengandungi jumlah data yang banyak.
- (b) Maklumat penerbitan dari ResearchIndex.com - ResearchIndex.com merupakan enjin carian bagi penerbitan sains komputer yang telah diterbitkan secara atas talian. Kumpulan PLUS telah menghasilkan wrapper bagi halaman web ini di mana ia boleh menterjemah halaman penerbitan yang terdapat di ResearchIndex.com kepada pernyataan SHOE secara terperinci. Kumpulan ini kemudiannya menggunakan wrapper ini untuk menghasilkan pernyataan SHOE bagi kesemua penerbitan merujuk kepada setiap lima belas jabatan sains komputer berkaitan yang telah dianotasi sebelum ini. Halaman ini mengandungi pernyataan SHOE seperti tajuk, pengarang dan kategori bagi penerbitan yang telah dipilih.

- (c) Domain Transmissible Spongiform Encephalopathy – SHOE telah diaplikasi ke dalam domain ini yang merupakan hasil kerjasama di antara kumpulan PLUS dan ahli Joint Institute for Food Safety and Nutrition (JIFSAN). JIFSAN pula merupakan usahasama di antara Food and Drug Administration (FDA) dan Universiti Maryland, yang berusaha untuk memperluaskan pengetahuan dan sumber sedia ada bagi menyokong analisis risiko di dalam bidang keselamatan pemakanan. Laman web mengenainya boleh dirujuk di <http://www.cs.umd.edu/projects/fsrac/>. Matlamat utama adalah membangunkan laman web yang akan bertindak sebagai pusat maklumat risiko keselamatan pemakanan. Laman web ini juga dikhaskan untuk kegunaan pelbagai lapisan pengguna termasuk, penyelidik, pembuat polisi, pentaksir risiko dan masyarakat yang lain umumnya. Maka, ia perlulah berupaya untuk memberi tindak balas terhadap pertanyaan yang dilakukan di mana, terminologi, kompleksiti dan ketepatan pertanyaan itu sendiri boleh berubah-ubah. Keadaan ini agak sukar sekiranya dikendalikan oleh pendekatan berasaskan kata kunci tetapi ia agak mudah dikendalikan sekiranya menggunakan SHOE.

2.5 Carian Web

WWW telah dibangunkan oleh Timothy Berners-Lee iaitu seorang ahli fizik dan saintis komputer yang berbangsa British sebagai projek bersama European Organization untuk penyelidikan nuklear di Geneva, Switzerland. Berners-Lee memulakan penghasilan hiperteks di awal tahun 1980-an. Ianya begitu cepat tersebar ke universiti-universiti di seluruh dunia. Kaedah penyimpanan maklumat di dalam web secara perhubungan, perolehan dokumen melalui sambungan hiperteks dan menamakan halaman web dengan URL menjadikan sambungan yang lancar bagi keseluruhan internet. Ini membenarkan pengaksesan maklumat yang mudah di antara bahagian yang berbeza di dalam internet.

Pengguna mempunyai dua alatan utama untuk membantu mereka mengenal pasti sumber yang berkaitan di dalam web, iaitu katalog dan enjin carian. Katalog dibentuk daripada kepakaran manusia. Oleh itu, ianya cenderung kepada ketepatan yang tinggi. Namun, untuk menyelenggaranya adalah sukar disebabkan oleh web sentiasa berkembang. Enjin carian pula telah dibina bagi mengikuti perkembangan yang terus berlaku di dalam web. Enjin carian merupakan perisian komputer yang menyusun satu senarai dokumen yang biasanya diletakkan di dalam WWW berserta dengan kandungan bagi dokumen-dokumen tersebut. Enjin carian bertindak balas terhadap kemasukan atau pertanyaan dari pengguna dengan membuat carian terhadap indeksnya dan memaparkan satu senarai dokumen yang berpadanan dengan pertanyaan carian yang telah dibuat. Sesetengah enjin carian memasukkan teks bahagian permulaan bagi laman web di dalam senarai mereka. Tetapi, sesetengah enjin carian yang lain pula hanya memasukkan tajuk atau alamat laman web sahaja.

Enjin carian yang berasaskan-penjejakan (*crawler*) mempunyai tiga elemen yang utama. Elemen pertama adalah *spider* atau dikenali juga sebagai *crawler*. Elemen ini akan melawati laman web, membacanya dan kemudian mengikuti pautan yang terdapat di laman web tersebut satu persatu. *Spider* akan kembali ke laman web tersebut mengikut jangka masa yang tetap seperti setiap satu atau dua bulan untuk mengesan sebarang perubahan yang berlaku terhadap laman web tersebut. Kesemua maklumat yang diperolehi oleh *spider* akan diletakkan ke elemen kedua enjin carian iaitu indeks. Indeks juga dikenali sebagai katalog, ia adalah seperti sebuah buku yang besar dan mengandungi salinan bagi setiap laman web yang telah dilawati oleh *spider*. Perisian enjin carian merupakan elemen ketiga bagi enjin carian. Ia merupakan program yang menyemak kesemua jutaan laman yang telah direkodkan ke dalam indeks untuk mencari padanan kepada carian dan menyenaraikan laman-laman web tersebut bersesuaian dengan apa yang ia rasakan paling relevan.

Selain daripada pelbagai prinsip heuristik bagi enjin carian elektronik (seperti carian mengikut tajuk, indeks, tajuk berita atau kata kunci), agen perisian adalah alternatif dalam menangani proses carian yang sukar. Carian dilakukan terhadap laman web yang sama ada dihubungkan kepada agen atau yang didaftarkan dengan agen. Baru-baru ini enjin carian mengaplikasikan penggunaan ontologi bagi menjadikan carian lebih bermakna.

2.5.1 Isu dan Masalah Carian Web

Enjin carian mampu untuk mengindeks sebahagian besar daripada web, namun pengguna seringkali menghadapi masalah di dalam membuat carian terhadap sumber yang diinginkan. Terdapat dua masalah yang dihadapi oleh pengguna apabila membuat carian. Masalah pertama yang sering dihadapi adalah hasil yang dipaparkan oleh enjin carian setelah melakukan carian adalah terlalu banyak sehingga pengguna terpaksa memeriksa setiap satu laman web bagi mendapatkan sumber yang benar-benar memenuhi kehendak pengguna tersebut. Sekiranya terdapat berpuluh atau beratus laman web yang dipaparkan, sudah tentu ini akan mendatangkan masalah kepada pengguna. Masalah kedua pula, pengguna tidak memperolehi apa yang dikehendakinya apabila tiada hasil yang dipaparkan setelah pengguna membuat pertanyaan kepada enjin carian. Masalah pertama di atas disebabkan oleh penggunaan perkataan yang sama bagi makna yang berlainan di dalam konteks dan kata kunci yang berbeza menyebabkan pengindeksan tidak mempunyai pengertian bagi perhubungan di antara perkataan. Manakala, bagi masalah kedua yang jarang dihadapi oleh pengguna, ianya disebabkan oleh penggunaan istilah atau set istilah yang tidak diletakkan di laman web.

Isu lain di dalam carian laman web adalah kebolehan enjin carian untuk melaksanakan sesuatu tugas dan bukan terhad kepada melaksanakan carian sahaja. Tugas yang dimaksudkan adalah apabila pengguna membuat pertanyaan untuk membuat penempahan satu percutian ke Jamaica sebagai contohnya, maka timbul permasalahan untuk menempah tiket kapal terbang dan mengatur untuk penyewaan kenderaan serta penempahan bilik hotel, di mana kesemuanya berdasarkan harga penawaran yang paling murah. Kehendak pengguna seperti ini adalah sukar untuk dilaksanakan oleh sesuatu enjin carian kerana tiada mekanisme pintar yang diterapkan ke dalamnya.

2.5.2 Contoh-Contoh Enjin Carian

Enjin carian yang terletak di dalam WWW telah menjadi begitu popular sama seperti web itu sendiri yang telah menjadi pilihan di dalam persekitaran internet. Enjin carian di dalam web mempunyai kelebihan dalam menawarkan akses kepada pilihan sumber maklumat yang banyak yang berlokasi di internet. Enjin carian di dalam web selalunya dibangunkan oleh syarikat persendirian. Namun, kebanyakan enjin carian yang terdapat di dalam web adalah percuma bagi penggunaannya. Contoh-contoh enjin carian yang terdapat di dalam web adalah seperti berikut:

- (a) Google.
- (b) Direct Hit.
- (c) AltaVista.
- (d) Lycos.
- (e) HotBot.
- (f) Applied Semantic.

2.5.3 Signifikan Sambungan Mudah HTML Ontologi dalam Carian Web

SHOE yang merupakan bahasa HTML yang berasaskan bahasa perwakilan pengetahuan membenarkan pengarang kepada laman web untuk membuat catatan kepada laman web mereka dengan meletakkan pengetahuan semantik. Catatan ini dinyatakan dalam bentuk pengetahuan ontologikal, di mana ianya dijana dengan membuat penyambungan kepada ontologi piawai yang telah sedia ada di dalam web. Ini menjadikannya mudah untuk mengarahkan agen web melaksanakan pencarian ke atas pertanyaan yang dibuat oleh pengguna. Ianya juga bertindak sebagai penggantian kepada pencarian dengan menggunakan kata kunci yang mudah seperti yang dilakukan oleh kebanyakan enjin carian pada masa kini.

Hasil daripada carian yang telah dilakukan pula menjamin kepada ketepatan dan terdapat kaitan yang munasabah melalui penggunaan ontologi. Malahan, klasifikasi bagi laman web dan perincian mengenai perhubungan serta atribut bagi laman web dapat difahami oleh mesin. SHOE juga bukan sekadar membuat

pertanyaan ke atas pangkalan data tetapi apabila ianya digunakan secara langsung oleh perisian pelayar web, ia dapat membantu pengguna melayari laman web dengan lebih pantas dan memahami betapa bernilainya suatu laman web itu apabila ianya dilayari. Selain itu, SHOE boleh digunakan untuk menyimpan maklumat atau pengetahuan bagi agen. Iaitu maklumat yang diperolehi oleh agen ketika melayari web sebagaimana yang diarahkan oleh pengguna atau sistem. Agen tersebut juga boleh menggunakan maklumat itu untuk keperluan di masa yang lain atau menggunakan maklumat tersebut untuk membantu di dalam melayari web dengan lebih baik.

2.6 Ringkasan

Kajian yang dilakukan di dalam bab ini menunjukkan bahawa penggunaan ontologi di dalam sistem SHOE merupakan suatu usaha yang baik dalam menangani masalah yang wujud di dalam domain carian web. Walaupun SHOE tidak menjanjikan penyelesaian yang menyeluruh terhadap isu semantik web. Namun, dengan keberkesanan penggunaannya menjadikan SHOE wajar untuk diberi perhatian.

BAB III

METODOLOGI PROJEK

3.1 Pendahuluan

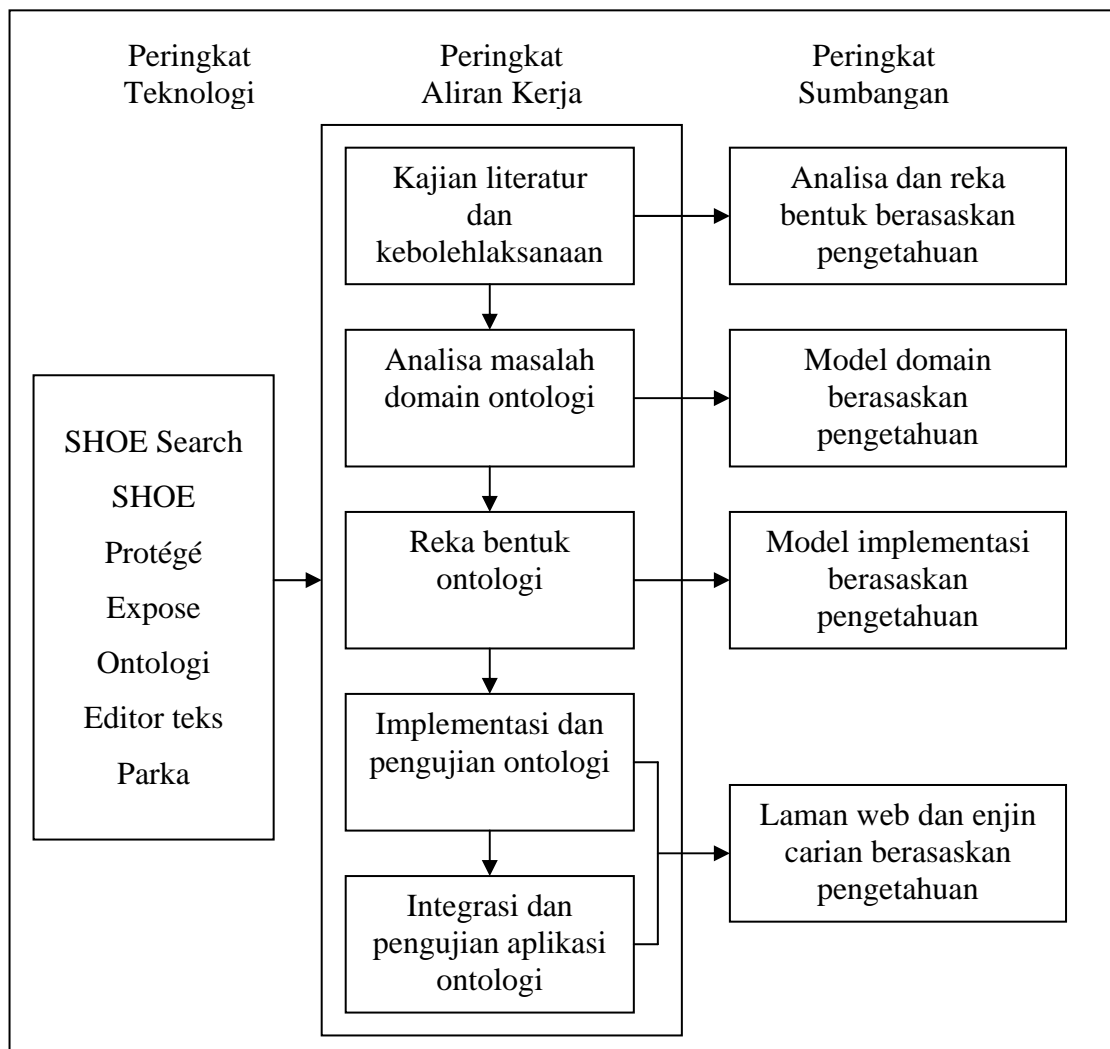
Terdapat pelbagai pendekatan telah dihasilkan oleh para penyelidik sebagai mekanisme bagi membangunkan ontologi dan aplikasi berasaskannya. Tujuannya ialah untuk memastikan atribut kualiti sesuatu ontologi seperti mudah senggara, ketidakbergantungan, boleh guna semula dan kecekapan dapat dicapai.

Bagaimanapun, kejuruteraan perisian berasaskan ontologi merupakan satu bidang baru yang memerlukan metodologi yang dapat menyokong teknologi ontologi berkenaan dan pengurusan pembangunan aplikasi berasaskannya. Topik-topik yang dikupas dalam bab ini adalah seperti berikut:

- (a) Reka bentuk kajian yang menggunakan konsep rangka kerja Tiga Peringkat (*Three Tier framework*).
- (b) Rangka kerja operasi yang menggunakan metodologi yang diperkenalkan oleh Noy dan McGuinness (2001).
- (c) Instrumentasi dan teknologi yang digunakan dalam projek.
- (d) Perancangan dan jadual pelaksanaan projek.

3.2 Reka Bentuk Projek

Projek ini berkonsepkan sains guna di mana terdapat pendekatan iaitu SHOE yang kemudiannya diaplikasikan ke dalam domain carian web bagi menyelesaikan isu yang dihadapi oleh domain tersebut. Bidang bagi projek ini pula adalah kejuruteraan perisian iaitu pembangunan enjin carian berasaskan pengetahuan dan selain itu ia juga merangkumi bidang kepintaran buatan iaitu penggunaan ontologi yang merupakan sebahagian daripada perwakilan pengetahuan. Teknologi yang digunakan, aliran kerja yang dilaksanakan dan sumbangan yang dihasilkan adalah sebagaimana yang dijelaskan dalam Rajah 3.1.



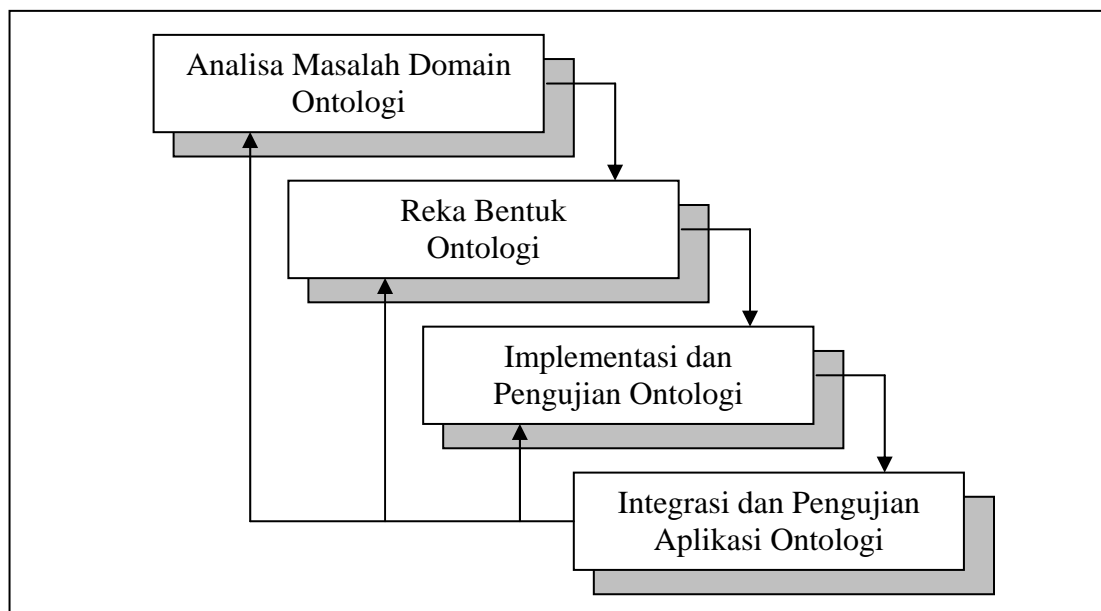
Rajah 3.1 : Reka bentuk projek berdasarkan rangka kerja Tiga Peringkat

3.3 Rangka Kerja Operasi

Rangka kerja operasi meliputi tiga perkara utama iaitu model bagi panduan pembangunan projek, proses-proses yang dilaksanakan dan artifak yang dihasilkan.

3.3.1 Model Kitar Hayat Pembangunan Perisian Berasaskan Pengetahuan

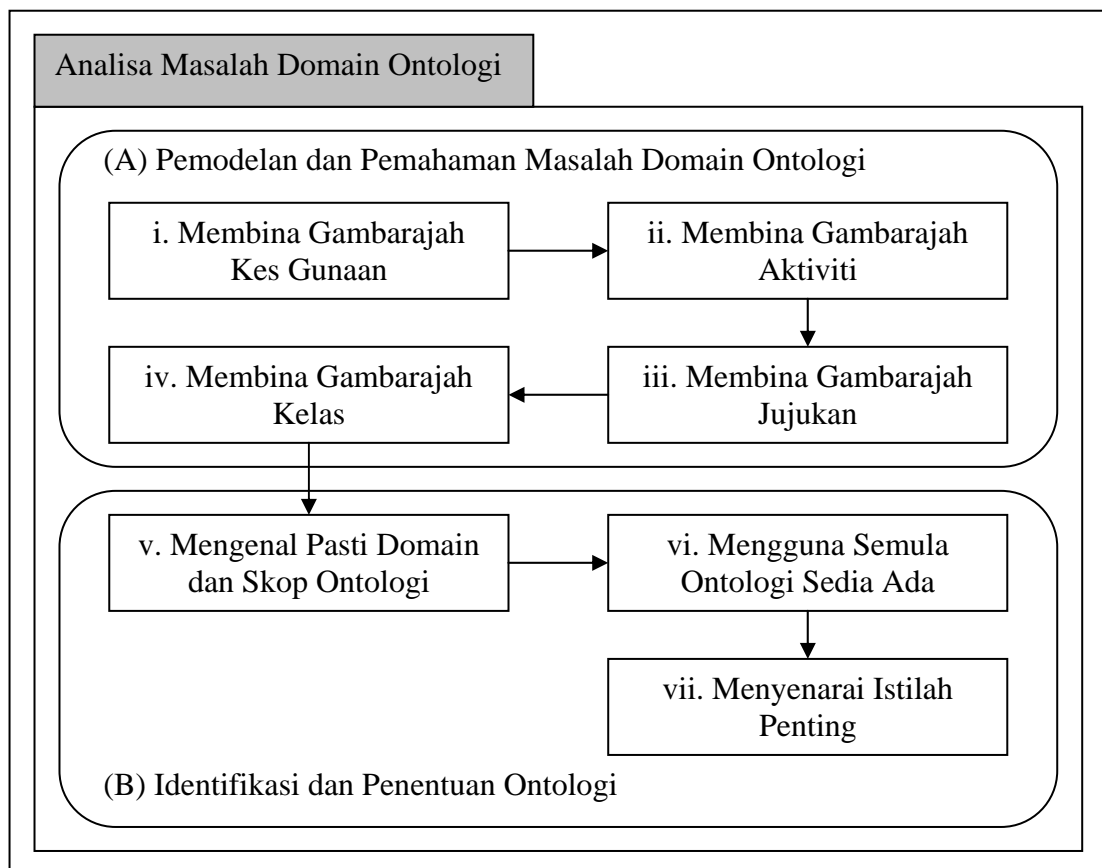
Metodologi bagi menganalisa dan mereka bentuk ontologi diadaptasi daripada metodologi yang diperkenalkan oleh Noy dan McGuinness (2001), di mana Protégé digunakan sebagai alatan untuk menganalisa dan mereka bentuk ontologi. Metodologi ini kemudiannya diubah suai dan ditambah agar terdapat kesinambungan untuk proses implementasi dan pengujian. Ianya juga diserasikan dengan model Air Terjun bagi membolehkan kombinasi unik ini berupaya menjadi model kitar hayat pembangunan perisian berasaskan pengetahuan sebagaimana yang diilustrasikan oleh Rajah 3.2.



Rajah 3.2 : Model kitar hayat pembangunan perisian berasaskan pengetahuan

3.3.2 Fasa Analisa Masalah Domain Ontologi

Pada fasa ini, domain ontologi dimodelkan dan difahami melalui pelaksanaan aktiviti yang terdapat di dalam Rajah 3.3, iaitu membina gambarajah kesgunaan, membina gambarajah aktiviti, membina gambarajah jujukan, membina gambarajah kelas, mengenal pasti domain dan skop ontologi, mengguna semula ontologi yang sedia ada dan menyenarai istilah penting yang terlibat di dalam domain.

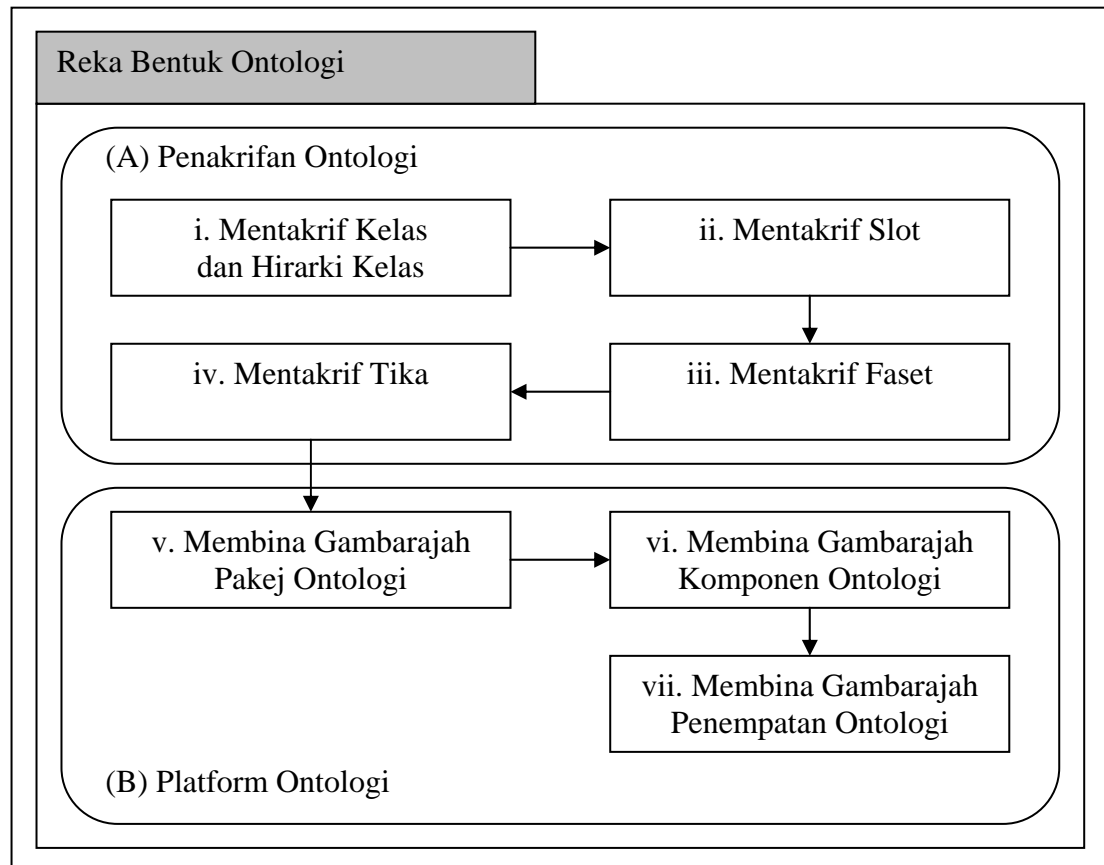


Rajah 3.3 : Fasa analisa masalah domain ontologi

3.3.3 Fasa Reka Bentuk Ontologi

Pada fasa ini, setiap taksonomi ontologi direka bentuk sehingga ianya boleh dijemakan kepada kod atur cara yang boleh difahami oleh mesin melalui

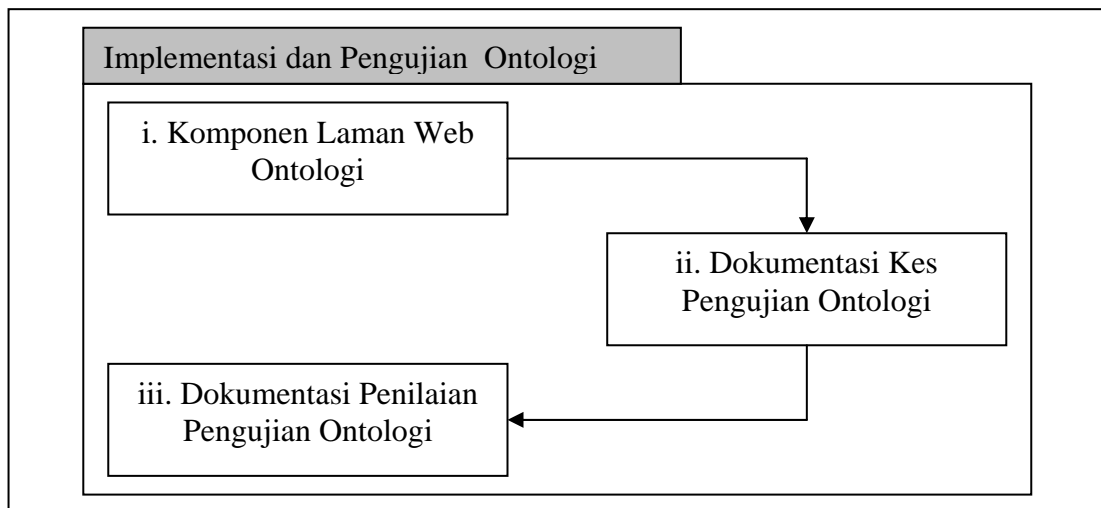
perlaksanaan aktiviti yang terdapat di dalam Rajah 3.4, iaitu aktiviti mentakrif kelas dan hirarki kelas, mentakrif slot, mentakrif faset, mentakrif tika bagi sesuatu ontologi, membina gambarajah pakej ontologi, membina gambarajah komponen ontologi dan membina gambarajah penempatan ontologi.



Rajah 3.4 : Fasa reka bentuk ontologi

3.3.4 Fasa Implementasi dan Pengujian Ontologi

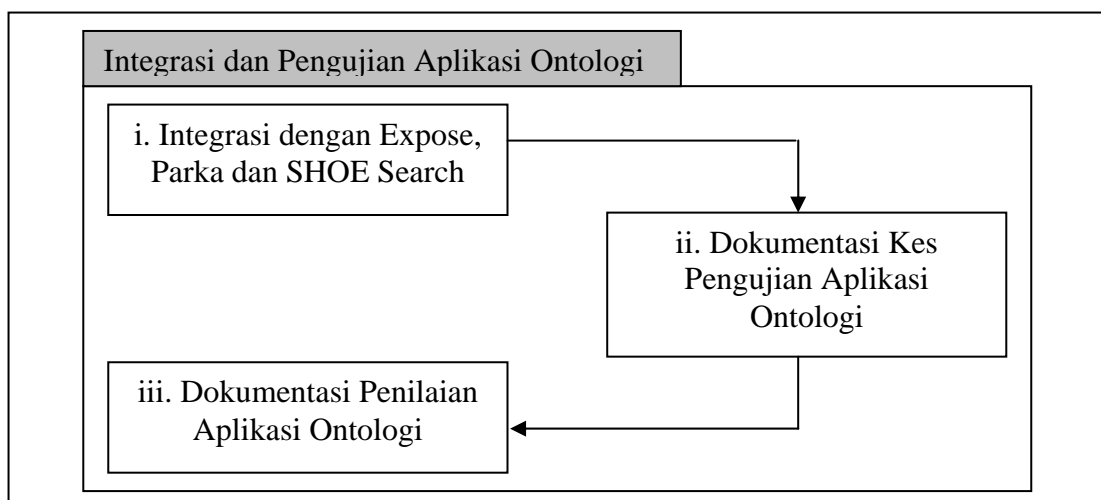
Pada fasa ini, reka bentuk ontologi diterjemah kepada satu set atur cara atau unit atur cara dan setiap satunya diuji secara berasingan melalui perlaksanaan aktiviti yang terdapat di dalam Rajah 3.5, iaitu aktiviti pembinaan komponen laman web berbentuk ontologi dan pengujian ontologi melalui penghasilan dokumentasi kes pengujian ontologi dan dokumentasi penilaian pengujian ontologi.



Rajah 3.5 : Fasa implementasi dan pengujian ontologi

3.3.5 Fasa Integrasi dan Pengujian Aplikasi Ontologi

Pada fasa ini, unit atur cara ontologi diintegrasikan dengan *web crawler*, pangkalan pengetahuan dan antara muka enjin carian serta diuji sebagai satu aplikasi ontologi yang lengkap melalui pelaksanaan aktiviti yang terdapat di dalam Rajah 3.6, iaitu aktiviti pengintegrasian dengan Expose, Parka dan SHOE Search serta pengujian ontologi melalui penghasilan dokumentasi kes pengujian aplikasi ontologi dan dokumentasi penilaian pengujian aplikasi ontologi.



Rajah 3.6 : Fasa integrasi dan pengujian aplikasi ontologi

3.4 Instrumentasi dan Teknologi Projek

Berikut adalah senarai instrumentasi dan teknologi bagi projek yang digunakan, iaitu:

- (a) Ontologi – pendekatan yang digunakan untuk mewakili pengetahuan.
- (b) Protégé – alatan yang digunakan untuk menganalisis dan mereka bentuk pengetahuan.
- (c) SHOE – bahasa berasaskan ontologi yang digunakan untuk melakukan anotasi dan transformasi terhadap laman web yang berbentuk maklumat kepada bentuk pengetahuan.
- (d) SHOE Search – alatan berbentuk antara muka pengguna yang digunakan untuk membenarkan pengguna membuat pertanyaan ke atas pangkalan pengetahuan Parka.
- (e) Expose – alatan yang digunakan sebagai *web crawler* untuk membuat carian terhadap laman web dengan penanda SHOE.
- (f) Parka – pangkalan pengetahuan yang digunakan untuk menyimpan pengetahuan yang telah dikumpulkan oleh Expose.
- (g) Editor teks – alatan yang digunakan untuk menambah tag SHOE yang bersesuaian.
- (h) UML – bahasa pemodelan yang digunakan untuk menganalisa dan mereka bentuk laman web berbentuk maklumat dan pengetahuan.
- (i) HTML – bahasa yang digunakan untuk membangunkan laman web berbentuk maklumat.

3.5 Perancangan Projek

Perancangan bagi pelaksanaan projek ini dibahagikan kepada dua fasa utama seperti berikut:

- (a) Projek 1.

Dilaksanakan pada semester II 2001/2002 dan meliputi aktiviti mengenal pasti projek, melaksanakan kajian literatur, menentukan metodologi projek, menganalisa masalah domain ontologi, mereka bentuk ontologi dan menghasilkan draf tesis. (Sila rujuk Lampiran A bagi carta gantt perancangan pelaksanaan projek 1).

(b) Projek 2.

Dilaksanakan pada semester II 2002/2003 dan meliputi aktiviti mengimplementasi dan menguji ontologi, mengintegrasikan dan menguji aplikasi ontologi serta menghasilkan tesis. (Sila rujuk Lampiran B bagi carta gantt perancangan pelaksanaan projek 2).

3.6 Ringkasan

Satu laman web yang berbentuk pengetahuan dibangunkan dengan menuruti dan mematuhi fasa-fasa yang telah digariskan oleh model pembangunan perisian berasaskan pengetahuan yang dihasilkan berdasarkan metodologi yang diperkenalkan oleh Noy dan McGuinness (2001). Hasil yang diperolehi daripada model ini turut mengambil kira aspek-aspek berikut:

- (a) Mengintegrasikan *web crawler* iaitu Expose ke dalam aplikasi.
- (b) Mengintegrasikan antara muka pengguna enjin carian yang ramah pengguna iaitu SHOE Search ke dalam aplikasi.
- (c) Mengintegrasikan pangkalan pengetahuan iaitu Parka ke dalam aplikasi.

BAB IV

ANALISA MASALAH DOMAIN ONTOLOGI

4.1 Pendahuluan

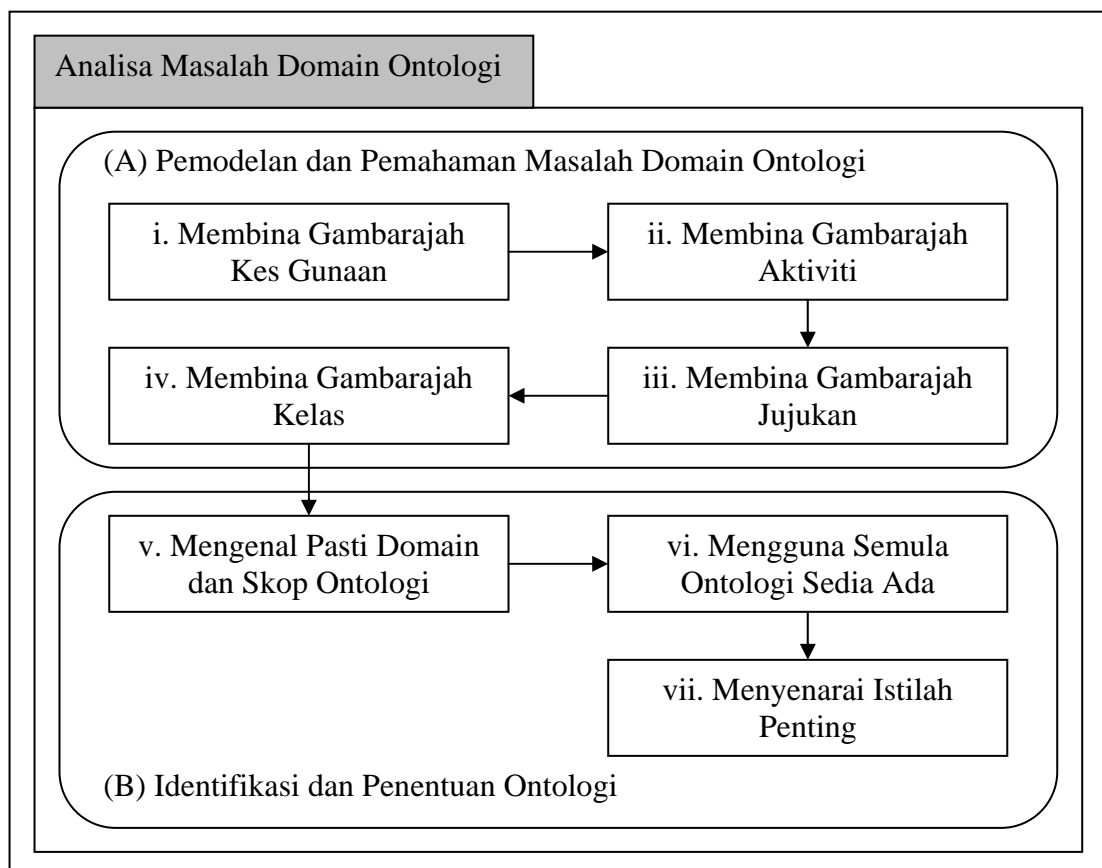
Kajian yang dilakukan oleh Noy dan McGuiness (2001), menunjukkan bahawa sebahagian daripada sebab mengapa ontologi dibangunkan ialah untuk membolehkan domain pengetahuan dianalisa dan diguna semula. Bagaimanapun, memandangkan teknologi internet yang semakin berkembang dan domain pengetahuan yang turut sama semakin kompleks serta besar, maka satu mekanisme yang menyokong pelaksanaan analisa masalah domain ontologi diperlukan. Salah satu penyelesaiannya ialah dengan menggunakan Protégé bagi mengidentifikasi dan menentukan ontologi serta istilah yang digunakan. Selain itu, domain bagi aplikasi ontologi juga perlu difahami dan dianalisa berdasarkan artifak UML dengan menggunakan Rational Rose. Topik-topik yang dikupas dalam bab ini adalah seperti berikut:

- (a) Konsep dan rangka kerja yang digunakan dalam fasa analisa masalah domain ontologi.
- (b) Implementasi fasa tersebut dalam menganalisa ontologi bagi laman web jabatan KP.

4.2 Rangka Kerja Fasa Analisa Masalah Domain Ontologi

Rajah 4.1 menggambarkan aktiviti di dalam menganalisa masalah domain ontologi yang melibatkan dua aktiviti utama iaitu pemodelan dan pemahaman masalah domain ontologi serta identifikasi dan penentuan ontologi:

- (a) Pemodelan dan pemahaman masalah domain ontologi melibatkan kerja-kerja membina gambarajah kesgunaan, gambarajah aktiviti, gambarajah jujukan dan gambarajah kelas.
- (b) Identifikasi dan penentuan ontologi melibatkan kerja-kerja mengenal pasti domain dan skop ontologi, mengguna semula ontologi sedia ada dan menyenarai istilah penting.



Rajah 4.1 : Aktiviti dalam fasa analisa masalah domain ontologi

4.2.1 Membina Gambarajah Kes Gunaan

Gambarajah kes gunaan dibangunkan kerana ia dapat menggambarkan satu set kes gunaan, pelaku dan hubungan serta interaksi di antara keduanya. Ianya juga membolehkan kelakuan sistem dari aspek yang statik dimodelkan supaya keperluan sistem dan bagaimana sistem seharusnya berfungsi dapat dikenal pasti.

4.2.2 Membina Gambarajah Aktiviti

Gambarajah aktiviti dibangunkan kerana ia dapat menggambarkan model aliran kerja bagi proses perniagaan dari satu aktiviti kepada satu aktiviti yang lain. Ianya membolehkan fungsi sistem dari aspek yang dinamik dimodelkan supaya pelaksanaan kerja atau tindakan dapat dikenal pasti.

4.2.3 Membina Gambarajah Jujukan

Gambarajah jujukan pula dihasilkan dengan tujuan untuk menggambarkan secara grafik alir senario hubungan serta interaksi di antara objek dengan penekanan terhadap urutan masa iaitu apakah yang mula-mula berlaku dan apakah yang berlaku seterusnya melalui penghantaran mesej. Ianya membolehkan kelakuan sistem dari aspek yang dinamik dimodelkan supaya peranan kelas dan antara muka dapat dikenal pasti.

4.2.4 Membina Gambarajah Kelas

Gambarajah kelas dibangunkan dengan tujuan untuk menggambarkan satu set kelas, antara muka, kolaborasi dan hubungan di antaranya. Ianya mengilustrasikan reka bentuk sistem dari aspek yang statik dengan memberi keterangan generik

mengenai sistem yang dibangunkan dan menjelaskan struktur penghantaran mesej serta struktur data sistem.

4.2.5 Mengenal Pasti Domain dan Skop Ontologi

Pembangunan ontologi dimulakan dengan mengenal pasti domain dan skopnya terlebih dahulu, iaitu dengan menjawab beberapa persoalan asas seperti:

- (a) Apakah domain yang akan diliputi oleh ontologi?
- (b) Apakah tujuan dan kegunaan ontologi berkenaan?
- (c) Apakah jenis maklumat yang perlu ada di dalam ontologi?
- (d) Siapakah yang akan menggunakan dan menyelenggara ontologi?

Jawapan kepada soalan-soalan di atas boleh berubah sepanjang proses menganalisa ontologi. Tetapi pada sesuatu masa tertentu, soalan-soalan tersebut mestilah membantu menghadkan skop dan sempadan ontologi bagi domain yang hendak dimodelkan. Sekiranya pihak yang akan menyelenggara ontologi berkenaan menerangkan domain di dalam bahasa atau istilah yang berlainan daripada apa yang digunakan oleh pengguna ontologi, maka pemetaan di antara istilah tersebut adalah diperlukan untuk membentuk persamaan. Manakala, cara untuk menentukan skop bagi ontologi adalah dengan mendrafkan senarai soalan yang berupaya dijawab oleh pangkalan pengetahuan yang berasaskan ontologi, iaitu dikenali sebagai soalan-soalan berkemampuan (Gruninger and Fox, 1995). Ianya bagi memastikan pangkalan pengetahuan tersebut memiliki maklumat yang mencukupi dan bersesuaian.

4.2.6 Mengguna Semula Ontologi Sedia Ada

Adalah wajar untuk mempertimbangkan apakah yang telah dihasilkan oleh pihak lain dengan menyemak jika ianya boleh diperbaiki dan ditambah untuk disesuaikan dengan tugas dan domain yang sedang dibangunkan. Penggunaan semula ontologi yang sedia ada mungkin menjadi satu keperluan sekiranya aplikasi yang

hendak dibangunkan perlu berinteraksi dengan aplikasi lain yang telah terikat kepada sesuatu piawaian ontologi tertentu atau perbendaharaan kata yang telah dianggap piawai (standard).

Terdapat beberapa pustaka ontologi yang telah sedia ada di dalam bentuk elektronik dan boleh diimport untuk diguna semula. Sebagai contohnya ialah *Ontolingua* (<http://www.ksl.stanford.edu/software/ontolingua/>), *DAML* (<http://www.daml.org/ontologies/>), *UNSPSC* (<http://www.unspsc.org>), *RosettaNet* (<http://www.rosettanel.org>) dan *DMOZ* (<http://www.dmoz.org>).

4.2.7 Menyengarai Istilah Penting

Penting bagi pembangun ontologi untuk menyenaraikan semua istilah yang hendak digunakan bagi membolehkan ianya difahami dan disahkan oleh pengguna serta membolehkan reka bentuk ontologi dapat dilaksanakan dengan tepat. Senarai istilah tersebut haruslah lengkap tanpa mengambil kira pertindihan di antara konsep yang diwakili, perhubungan antara istilah-istilah tersebut atau ciri dan sifat yang dimiliki oleh sesuatu konsep.

4.3 Implementasi Fasa Analisa Masalah Domain Ontologi Laman Web Jabatan Kejuruteraan Perisian

Laman web jabatan KP boleh dilayari pada alamat <http://kp.fsksm.utm.my>. Laman web ini menyediakan kemudahan kepada para pensyarah menyediakan kemudahan pembelajaran secara maya dan mempublikasi hasil kerja mereka. Selain itu, laman web ini juga membolehkan para pelajar dan pelayar luaran mendapatkan maklumat berkaitan dengan jabatan (seperti aktiviti, program pengajian, kakitangan dan sebagainya) dan pembelajaran (seperti jadual kuliah, nota, tugas dan sebagainya).

4.3.1 Gambarajah Kes Gunaan

Rajah 4.2 menunjukkan gambarajah kes gunaan yang menjelaskan masalah domain bagi gelintaran web menggunakan SHOE, di mana aplikasi berkomunikasi dengan pelayar bagi mendapatkan keperluan pengguna, kemudian mengarahkan SHOE Search melaksanakan gelintaran. Proses gelintaran melibatkan anotasi laman web berbentuk ontologi dengan merujuk kepada laman web berbentuk maklumat menggunakan Editor Teks dan penapisan pengetahuan di dalam laman web SHOE ke dalam pangkalan pengetahuan Parka yang dilaksanakan oleh Expose.

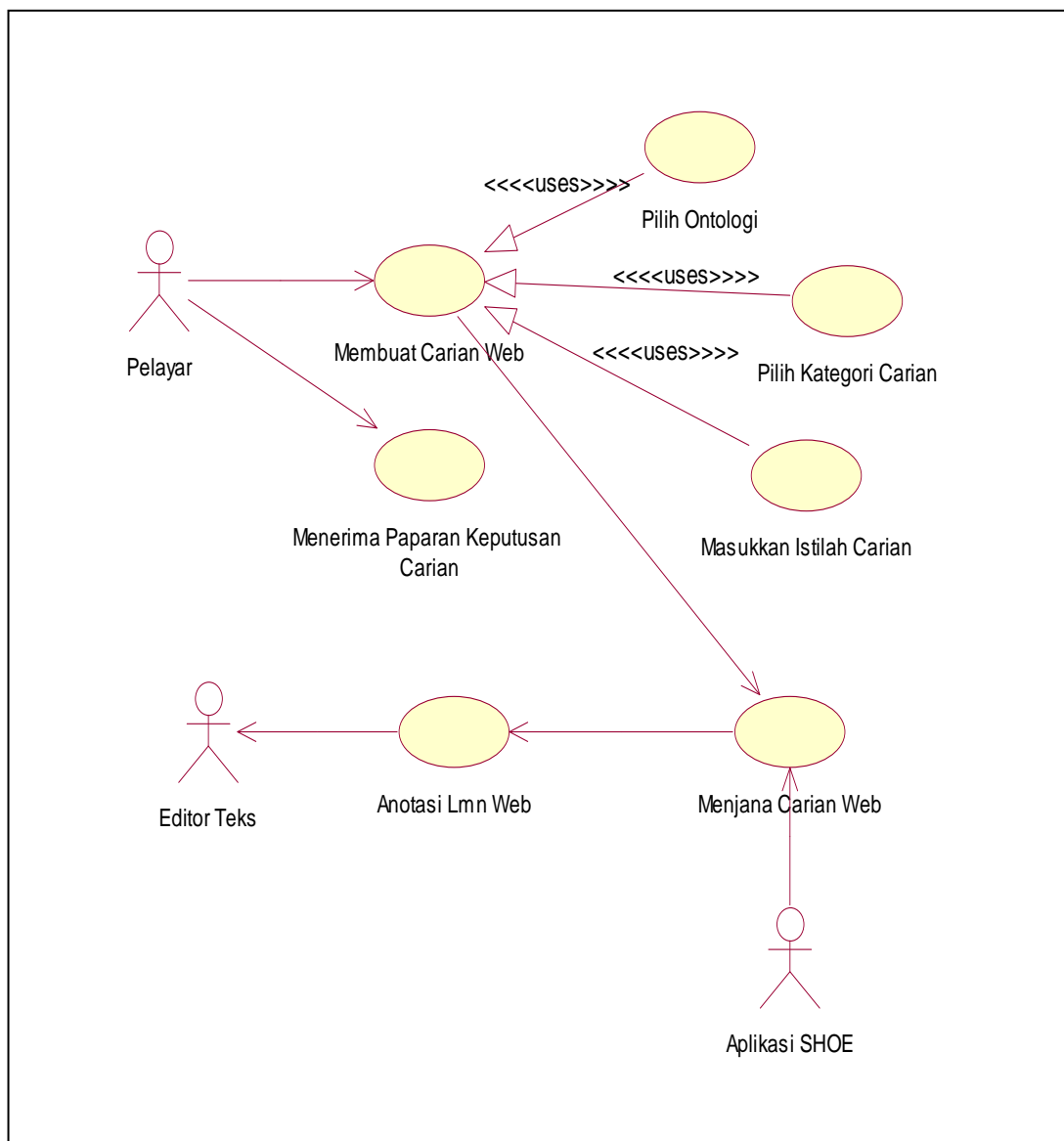
Terdapat tiga pelaku yang dikenal pasti, iaitu:

- (a) Pelayar yang bertanggungjawab memberikan dan membekalkan keperluan kepada enjin carian.
- (b) Aplikasi SHOE yang terdiri daripada Expose, SHOE Search dan Parka bertanggungjawab menapis pengetahuan yang terdapat di dalam laman web SHOE ke dalam pangkalan pengetahuan, melaksanakan carian di dalam pangkalan pengetahuan dan memaparkan keputusan kepada pelayar.
- (c) Editor Teks yang bertanggungjawab membentuk sintaks dan semantik laman web SHOE dengan merujuk kepada laman web berbentuk maklumat yang telah dibangunkan.

Manakala, kes gunaan yang terlibat adalah:

- (a) Kes gunaan Buat Carian Web yang berfungsi menerima dan menguruskan komunikasi bagi mendapatkan keperluan pelayar setelah pelayar memasukkan maklumat yang dikehendaki bagi melaksanakan carian. Maklumat yang diperlukan adalah seperti memilih ontologi, memilih kategori carian dan memasukkan istilah carian.
- (b) Kes gunaan Menjana Carian Web yang berfungsi mencari dan mendapatkan keputusan carian daripada pangkalan pengetahuan.
- (c) Kes gunaan Anotasi Laman Web yang berfungsi menukar laman web berbentuk maklumat kepada laman web berbentuk ontologi.

- (d) Kes gunaan Menerima Paparan Keputusan Carian yang berfungsi menyediakan laporan mengenai keputusan proses gelintaran web.



Rajah 4.2 : Gambarajah kes gunaan bagi gelintaran web menggunakan SHO E

4.3.2 Gambarajah Aktiviti

Gambarajah aktiviti pada Lampiran C menunjukkan aliran kerja yang dilaksanakan oleh tiga kes gunaan yang utama bagi Gelintaran Web Menggunakan SHO E dalam proses membuat carian web ke atas pangkalan pengetahuan untuk

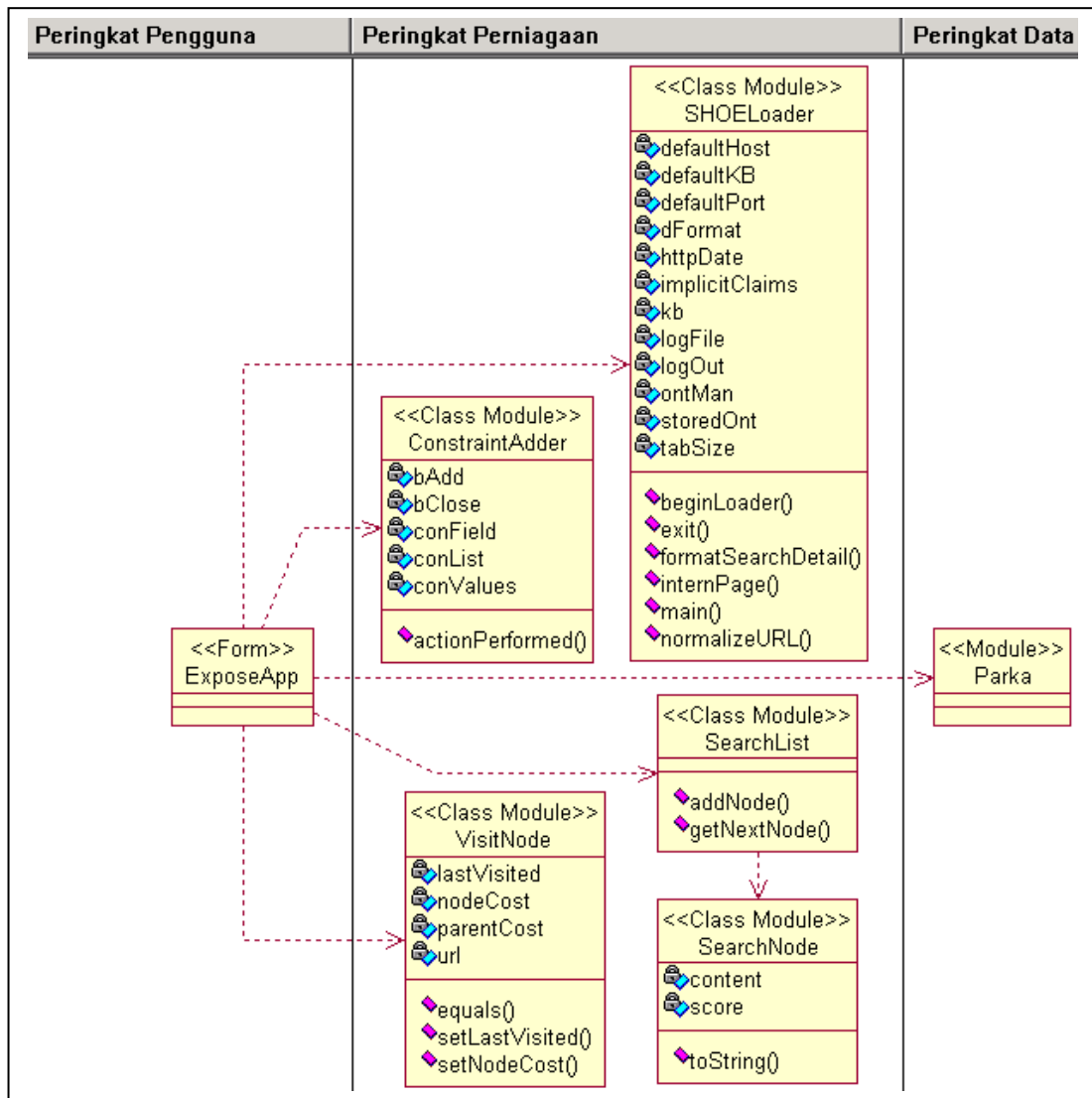
mendapatkan keputusan yang lebih bermakna bagi membolehkan pengguna membuat keputusan, memanipulasikan maklumat yang diterima dan melaksanakan proses pelayaran internet.

4.3.3 Gambarajah Jujukan

Gambarajah jujukan mengilustrasi turutan logik berdasarkan gambarajah aktiviti yang telah ditunjukkan sebelum ini. Sebanyak tiga senario yang dikenal pasti iaitu membuat carian web, menerima paparan keputusan carian dan transformasi maklumat. Senario membuat carian web menggabungkan dua gambarajah aktiviti iaitu gambarajah membuat carian web dan gambarajah menjana maklumat carian. Setiap senario ini adalah seperti yang ditunjukkan pada gambarajah jujukan yang terdapat di dalam Lampiran D.

4.3.4 Gambarajah Kelas

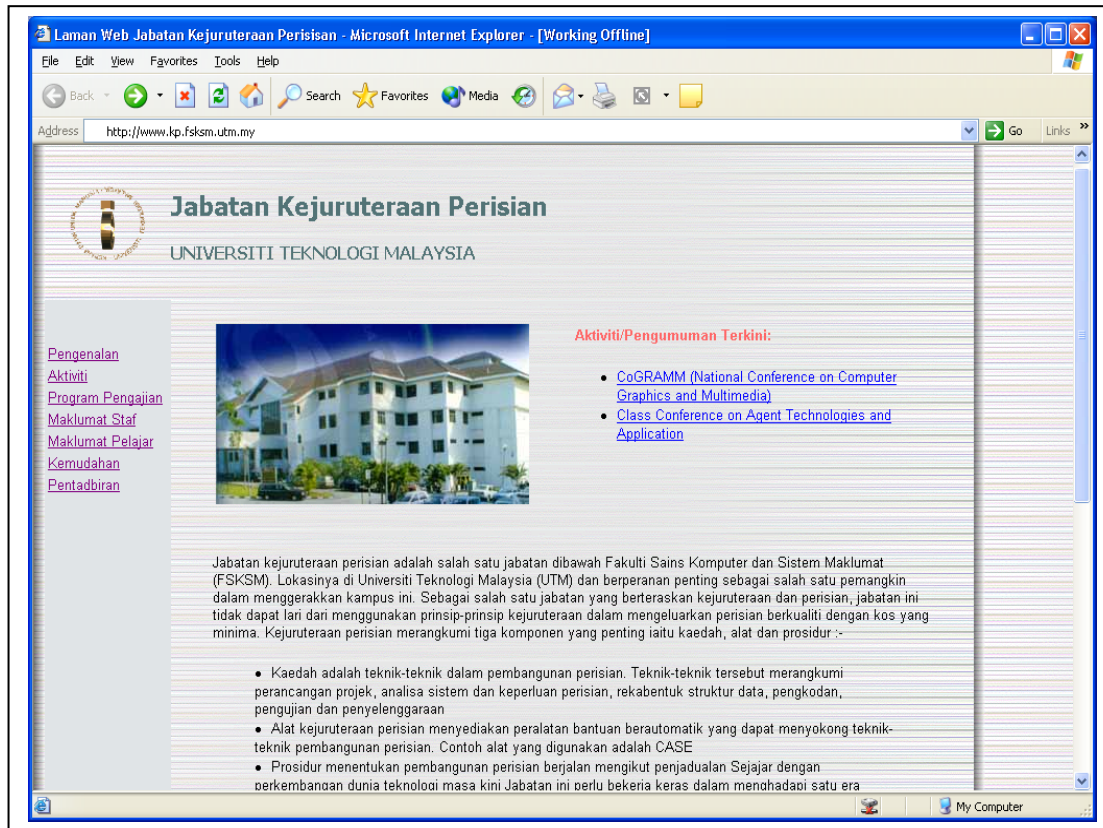
Gambarajah kelas dijana dengan mengekstrak objek yang dikenal pasti dari gambarajah kes gunaan dan gambarajah jujukan. Aspek yang statik dan dinamik bagi sesuatu objek diwakili oleh atribut dan operasi bagi objek tersebut. Rajah 4.3 menunjukkan contoh gambarajah kelas bagi kes gunaan Tapis Pengetahuan dengan menggunakan pendekatan rangka kerja Tiga Peringkat, di mana kelas-kelas dikategorikan kepada peringkat pengguna (antara muka pengguna), peringkat perniagaan (model domain) dan peringkat data (pangkalan data).



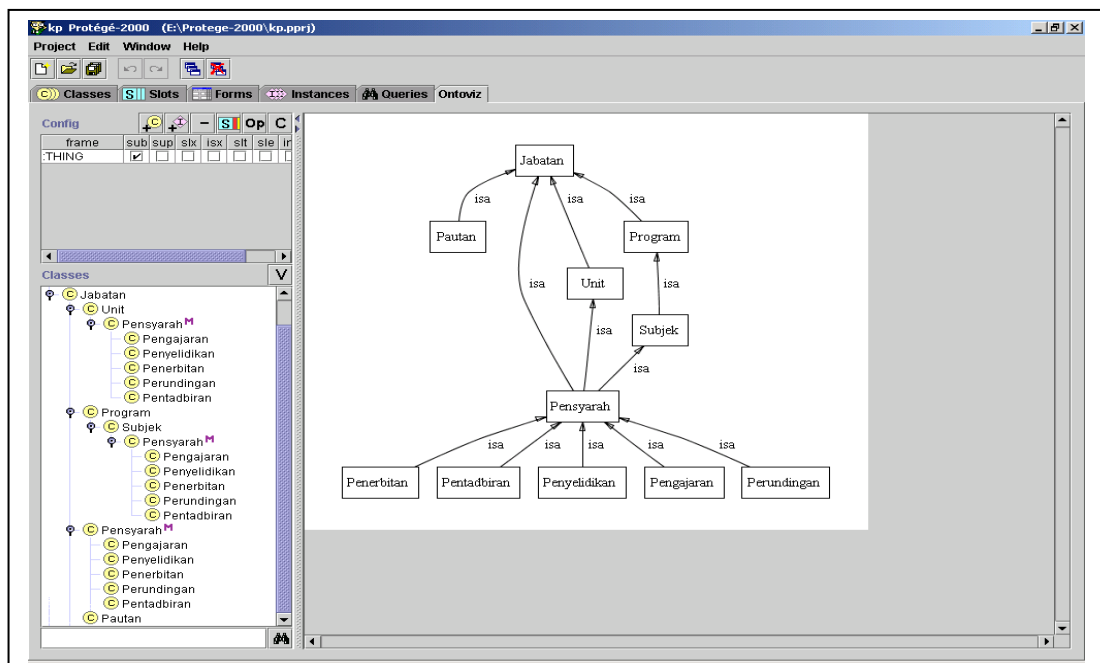
Rajah 4.3 : Gambarajah kelas bagi kesgunaan Tapis Pengetahuan

4.3.5 Domain dan Skop Ontologi

Rajah 4.4 menunjukkan petikan paparan laman utama jabatan KP yang menjadi kajian kes bagi projek yang dilaksanakan ini. Manakala, Rajah 4.5 memaparkan domain dan skop ontologi bagi laman web tersebut yang merangkumi ontologi seperti unit, program, subjek, pensyarah dan pautan serta hubungan di antara ontologi-ontologi berkenaan. Ontologi pensyarah misalnya boleh dikembangkan kepada ontologi penerbitan, pentadbiran, penyelidikan, pengajaran dan perundingan.



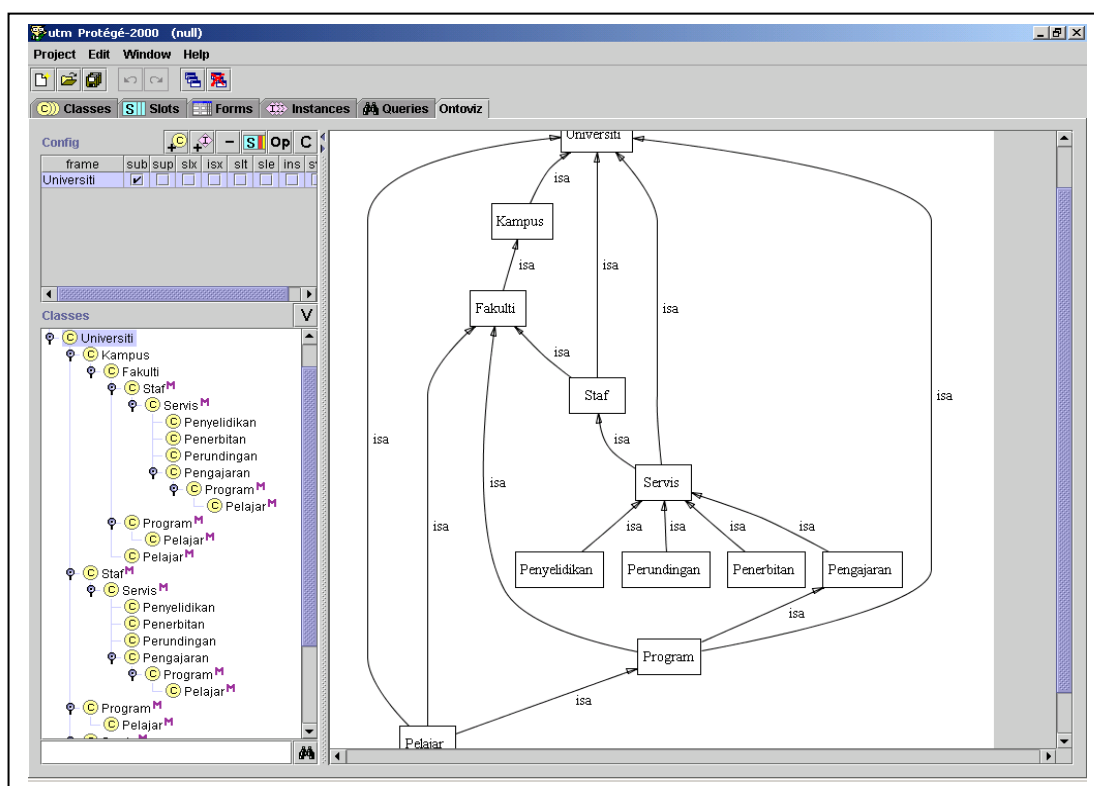
Rajah 4.4 : Petikan paparan utama laman web Jabatan Kejuruteraan Perisian



Rajah 4.5 : Petikan paparan domain dan skop ontologi bagi laman web Jabatan Kejuruteraan Perisian

4.3.6 Guna Semula Ontologi

Memandangkan adaptasi ontologi ke atas bidang pendidikan dan enjin carian merupakan perkara yang baru di arena tempatan. Maka, projek yang dilaksanakan ini berupaya menjadi perintis bagi membentuk satu laman web pustaka ontologi atau ontolingua. Ianya membolehkan ontologi yang sama khususnya tika digunakan semula untuk membangunkan laman web ontologi bagi jabatan-jabatan yang lain di UTM. Rajah 4.6 menunjukkan contoh ontologi universiti yang boleh diguna semula seperti kampus, fakulti, staf, servis, program dan pelajar.



Rajah 4.6 : Petikan paparan ontologi universiti yang boleh diguna semula oleh ontologi jabatan

4.3.7 Istilah Utama

Jadual 4.1 menunjukkan senarai istilah utama yang digunakan bagi menerangkan taksonomi jabatan.

Jadual 4.1 : Senarai istilah bagi taksonomi jabatan

Istilah	Diskripsi
Unit	Domain atau pengkhususan jabatan seperti bidang kepintaran buatan dan kejuruteraan perisian.
Program	Kursus-kursus yang ditawarkan oleh jabatan seperti kursus peringkat sarjana muda dan pasca ijazah.
Subjek	Mata pelajaran yang ditawarkan oleh jabatan.
Pensyarah	Staf akademik yang dipertanggungjawabkan oleh jabatan.
Pengajaran	Subjek yang diajar dan pelajar yang diselia oleh staf jabatan.
Penerbitan	Projek publikasi dan pres yang dilaksanakan oleh staf jabatan.
Penyelidikan	Projek penyelidikan dan pembangunan yang dilaksanakan oleh staf jabatan.
Perundingan	Projek khidmat rundingan teknikal yang dilaksanakan oleh staf jabatan.
Pentadbiran	Kerja-kerja penyeliaan dan pengurusan jabatan atau selainnya yang dilaksanakan oleh staf jabatan.
Pautan	Capaian kepada laman web luaran seperti halaman web perpustakaan dan pusat penyelidikan.

4.4 Ringkasan

Fasa analisa masalah domain ontologi ke atas laman web jabatan KP telah dihasilkan menggunakan Rational Rose melalui pelaksanaan aktiviti memodelkan dan memahami masalah domain ontologi dan menggunakan Protégé melalui pelaksanaan aktiviti mengidentifikasi dan menentukan ontologi. Hasil yang diperolehi dapat menjelaskan domain ontologi dan aplikasi ontologi yang hendak dibangunkan melalui penghasilan beberapa gambarajah kelas ontologi yang divisualisasikan menggunakan tab Ontovitz dan gambarajah kes gunaan, gambarajah aktiviti, gambarajah jujukan serta gambarajah kelas.

BAB V

REKA BENTUK ONTOLOGI

5.1 Pendahuluan

Noy dan McGuiness (2001), menyatakan bahawa pembangunan ontologi secara umumnya merangkumi kerja-kerja mentakrifkan kelas yang wujud di dalam skop ontologi yang hendak dibangunkan, menyusun kelas-kelas berkenaan sebagai satu hirarki kelas yang mempunyai hubungan kelas super dengan subkelas, mentakrifkan slot yang merupakan sifat bagi kelas dan mengumpukkan nilai kepada slot. Aktiviti pembangunan ini dapat dilaksanakan dengan sistematik dan efisien menggunakan Protégé kerana ia merupakan alatan untuk pembangunan aplikasi berasaskan pengetahuan dan pangkalan pengetahuan. Selain itu, platform ontologi juga perlu dibangunkan berdasarkan artifak UML menggunakan Rational Rose.

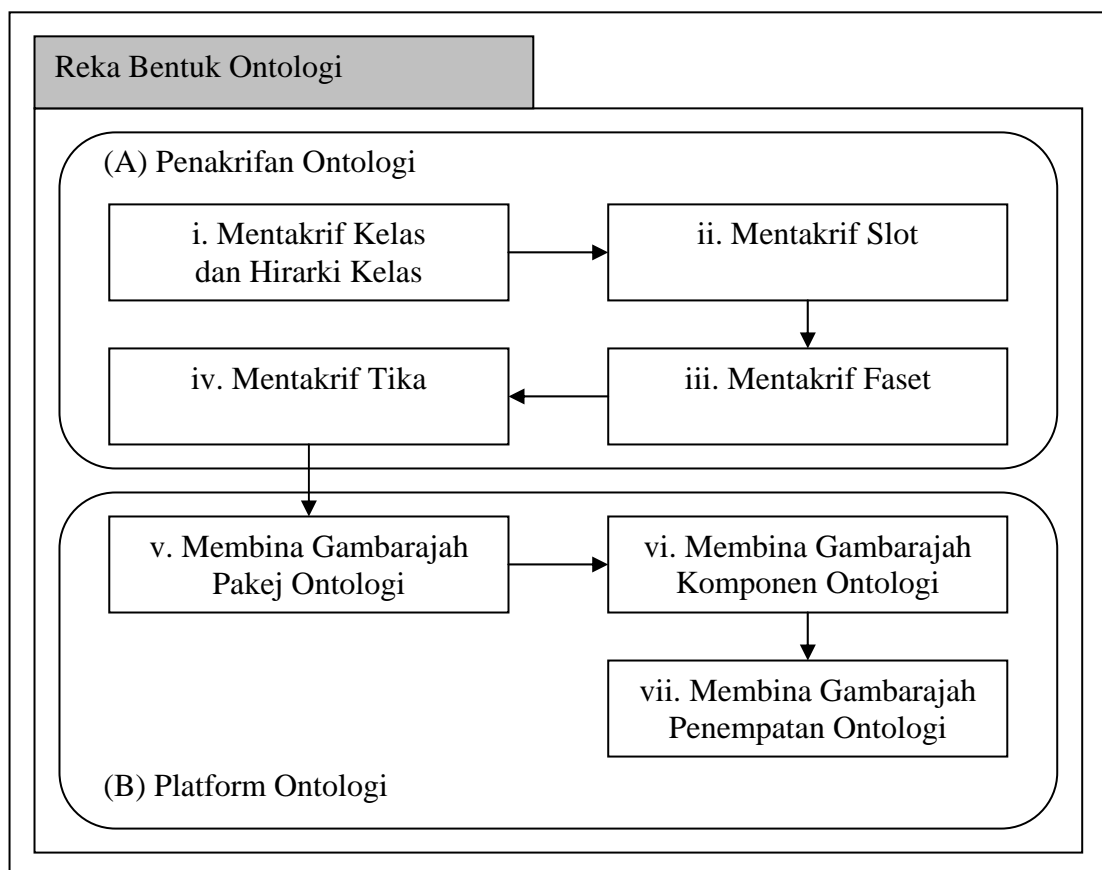
Topik-topik yang dikupas dalam bab ini adalah seperti berikut:

- (a) Konsep dan rangka kerja yang digunakan dalam fasa reka bentuk ontologi.
- (b) Implementasi fasa tersebut dalam reka bentuk ontologi bagi laman web jabatan KP.

5.2 Rangka Kerja Fasa Reka Bentuk Ontologi

Rajah 5.1 menunjukkan aktiviti di dalam mereka bentuk ontologi yang melibatkan dua aktiviti utama iaitu penakrifan ontologi dan platform ontologi:

- (a) Penakrifan ontologi melibatkan kerja-kerja mentakrif kelas dan hirarki kelas, slot, faset dan tika bagi sesuatu ontologi.
- (b) Platform ontologi melibatkan kerja-kerja membangunkan gambarajah pakej ontologi, gambarajah komponen ontologi dan gambarajah penempatan ontologi.



Rajah 5.1 : Aktiviti dalam fasa reka bentuk ontologi

5.2.1 Mentakrif Kelas dan Hirarki Kelas

Menurut Uschold and Gruninger (1996), terdapat beberapa pendekatan yang boleh digunakan untuk mengenal pasti dan membangunkan hirarki kelas, iaitu:

- (a) Proses pembangunan atas-bawah dengan mendefinisikan konsep yang umum di dalam domain dan diikuti dengan pengkhususan bagi setiap konsep.
- (b) Proses pembangunan bawah-atas dengan mendefinisikan kelas yang paling khusus atau dedaun bagi hirarki. Kemudian, mengumpulkan dan mengkategorikan kelas-kelas berkenaan kepada konsep yang lebih umum.
- (c) Proses pembangunan kombinasi yang merupakan integrasi pendekatan atas-bawah dan bawah-atas iaitu dengan mendefinisikan konsep yang lebih penting terlebih dahulu. Kemudian, membuat kesimpulan yang umum dan mengkhususkan konsep berkenaan dengan cara yang bersesuaian.

Tiada satu pun daripada ketiga-tiga kaedah di atas yang secara semulajadinya adalah lebih baik jika dibandingkan antara satu sama lain. Pendekatan yang hendak dipilih untuk dilaksanakan adalah bergantung kepada kompleksiti sesuatu domain dan persepsi pembangun ontologi berkenaan. Apa juga pendekatan yang dipilih, pembangunan ontologi selalunya bermula dengan memberi takrifan ke atas kelas melalui istilah yang menerangkan sesuatu objek yang wujud di dalam domain yang dikaji. Istilah ini akan dibentuk menjadi kelas di dalam ontologi atau sebagai sauh di dalam hirarki kelas.

5.2.2 Mentakrif Slot

Kelas sahaja tidak menyediakan maklumat yang mencukupi untuk menentukan skop bagi sesuatu ontologi. Apabila sesuatu kelas telah ditakrifkan, struktur dalaman bagi kelas tersebut juga perlu diterangkan. Slot merupakan sifat

sesuatu kelas yang menerangkan pelbagai ciri dan atribut kelas berkenaan. Slot juga dipanggil peranan dan mempunyai beberapa jenis seperti berikut:

- (a) Bersifat intrinsik seperti rasa dan bau bagi kelas minuman.
- (b) Bersifat ekstrinsik seperti nama dan kategori bagi kelas minuman.
- (c) Bersifat bahagian seperti kandungan kalsium dan natrium bagi kelas minuman.

5.2.3 Mentakrif Faset

Faset menerangkan nilai yang boleh dimiliki oleh suatu slot sama ada dari segi jenis, julat nilai yang dibenarkan, kardinaliti atau nilai lalainya. Parameter faset yang lazim digunakan adalah seperti berikut:

- (a) Kardinaliti slot – menentukan bilangan minimum dan maksimum nilai yang dimiliki oleh sesuatu slot.
- (b) Jenis nilai slot – jenis nilai yang boleh diisi ke dalam slot sama ada rentetan, nombor, boolean, bilangan atau tika.
- (c) Domain dan julat slot – kelas diumpukkan kepada domain slot apabila jenis nilai slot adalah tika. Ianya adalah menyerupai hubungan mandatori dalam pangkalan data tradisional. Tujuannya adalah untuk memastikan nilai slot yang hendak diinputkan kepada tika mestilah telah wujud di dalam kelas yang bertindak sebagai domain slot.

5.2.4 Mentakrif Tika

Langkah akhir bagi mereka bentuk ontologi adalah mencipta tika bagi kelas-kelas yang terdapat di dalam hirarki. Penakrifan tika bagi sesuatu kelas memerlukan kepada perkara-perkara berikut:

- (a) Membuat pemilihan kelas.
- (b) Mencipta tika secara individu bagi kelas berkenaan.
- (c) Mengumpukkan nilai slot bagi setiap tika.

5.2.5 Membina Gambarajah Pakej Ontologi

Gambarajah pakej ontologi dibangunkan dengan tujuan untuk mengumpul dan mengorganisasi elemen-elemen ontologi dan sistem ontologi ke dalam pakej-pakej tertentu serta menjelaskan hubungan dan kebergantungan antara pakej.

5.2.6 Membina Gambarajah Komponen Ontologi

Gambarajah komponen ontologi pula dibangunkan untuk menjelaskan komponen ontologi dan sistem ontologi serta kebergantungan antaranya sama ada ianya komponen kod sumber, kod perduaan atau kod boleh laku.

5.2.7 Membina Gambarajah Penempatan Ontologi

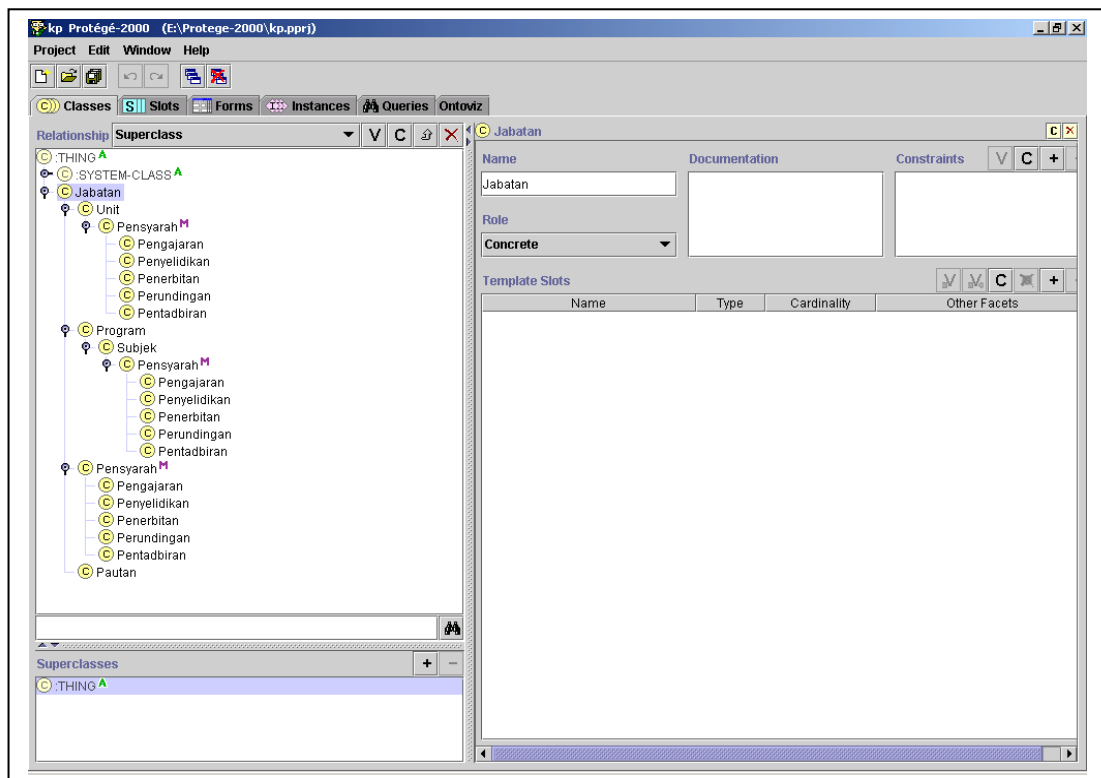
Gambarajah penempatan ontologi dihasilkan dengan tujuan untuk menerangkan konfigurasi bagi elemen-elemen pemrosesan masa larian serta proses ontologi dan sistem ontologi yang terdapat di dalamnya serta visualisasi pengagihan komponen di sepanjang perusahaan.

5.3 Implementasi Fasa Reka Bentuk Ontologi Laman Web Jabatan Kejuruteraan Perisian

Aktiviti reka bentuk ontologi dilaksanakan ke atas laman web jabatan KP bertujuan untuk mengidentifikasi dan mengklasifikasi kelas, slot, faset dan tika serta hubungan antaranya. Selain itu, platform ontologi juga dibangunkan bagi menjelaskan elemen sebenar reka bentuk ontologi dan sistem ontologi supaya ianya boleh dijelmakan kepada atur cara yang boleh dilarikan dan ditempatkan ke dalam persekitaran dunia sebenar.

5.3.1 Kelas dan Hirarki Kelas

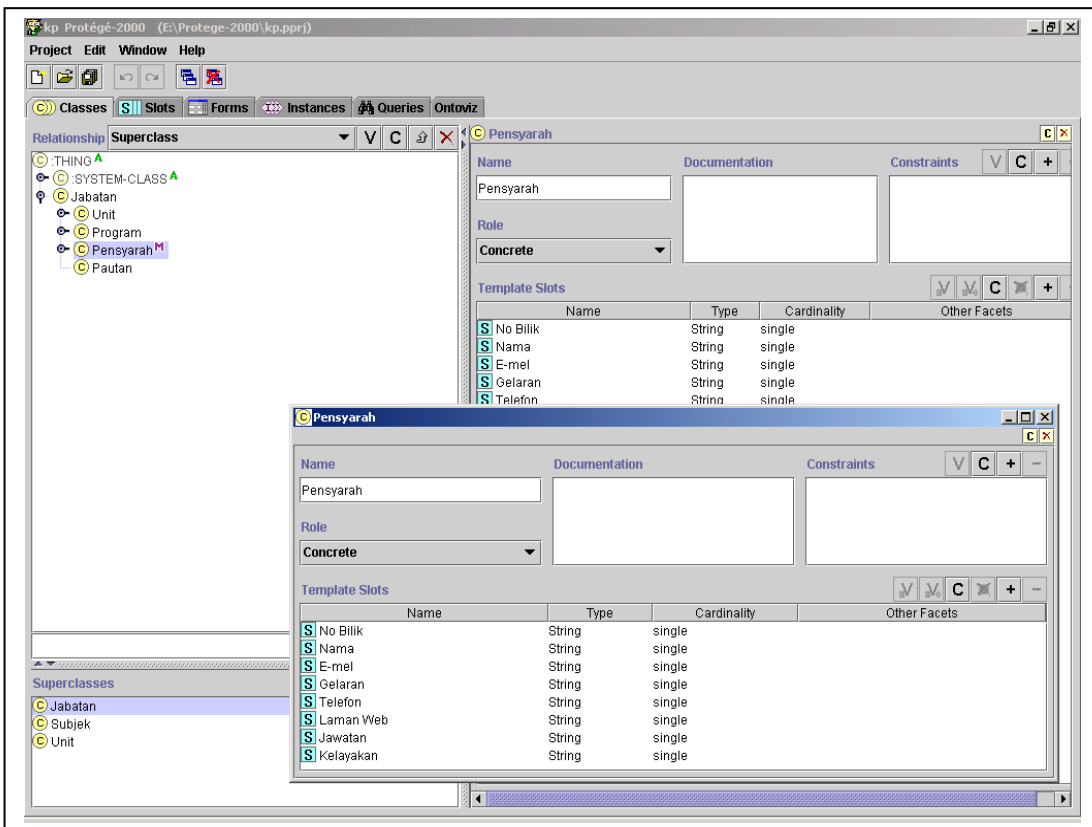
Rajah 5.2 menunjukkan pelbagai aras bagi taksonomi jabatan, di mana jabatan merupakan konsep yang umum. Manakala unit, program, pensyarah, subjek, pautan, pengajaran, penyelidikan, penerbitan, perundingan dan pentadbiran merupakan konsep aras tinggi, pertengahan dan bawah.



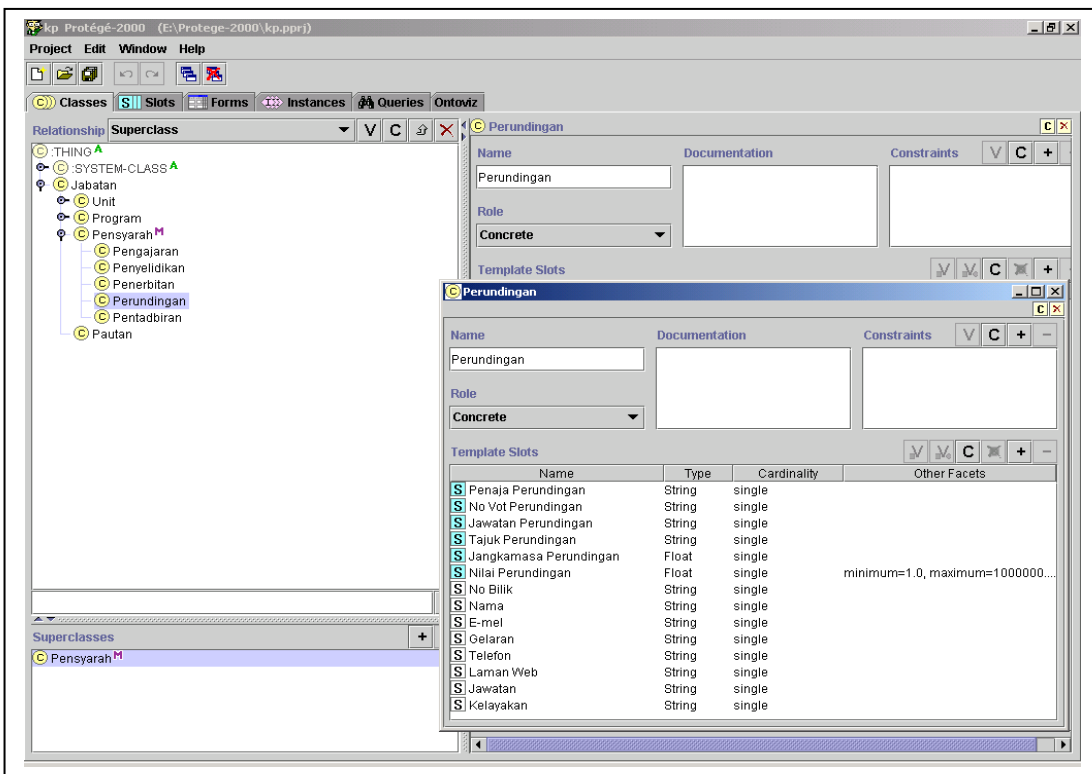
Rajah 5.2 : Petikan paparan kelas dan hirarki kelas bagi laman web Jabatan Kejuruteraan Perisian

5.3.2 Slot

Contoh slot adalah sebagaimana pada Rajah 5.3 dan 5.4. Slot bagi pensyarah adalah nama, jawatan, e-mel, nombor bilik, laman web, gelaran dan kelayakan. Manakala, slot bagi perundingan adalah nombor vot, penaja, tajuk, nilai, jangka masa dan jawatan. Kelas perundingan juga mewarisi semua slot yang dimiliki oleh kelas supernya iaitu kelas pensyarah.



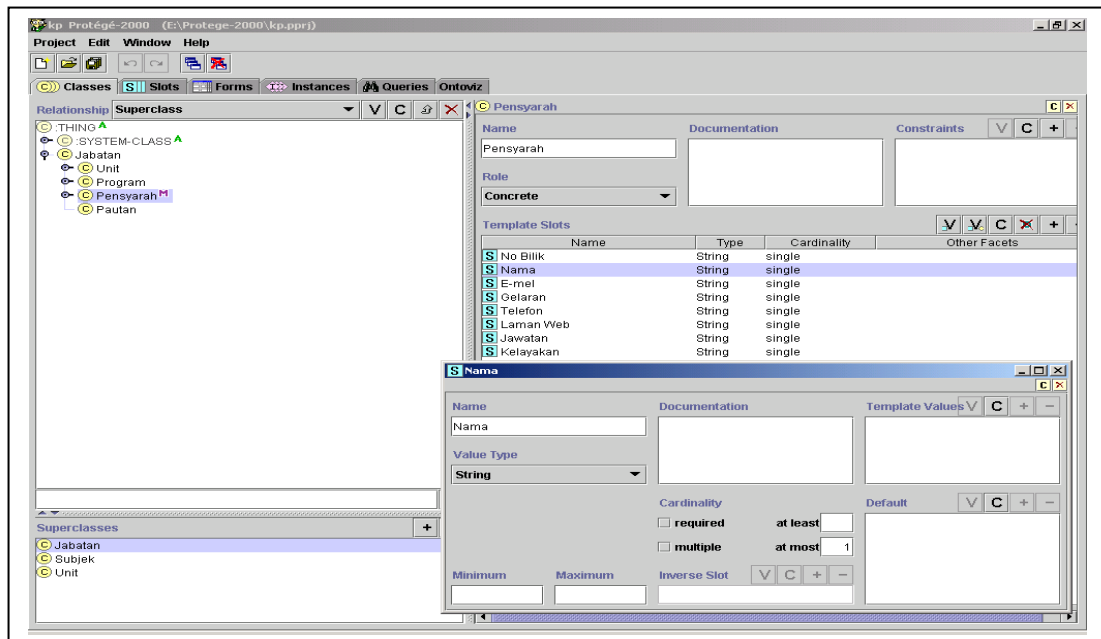
Rajah 5.3 : Petikan paparan slot bagi kelas pensyarah



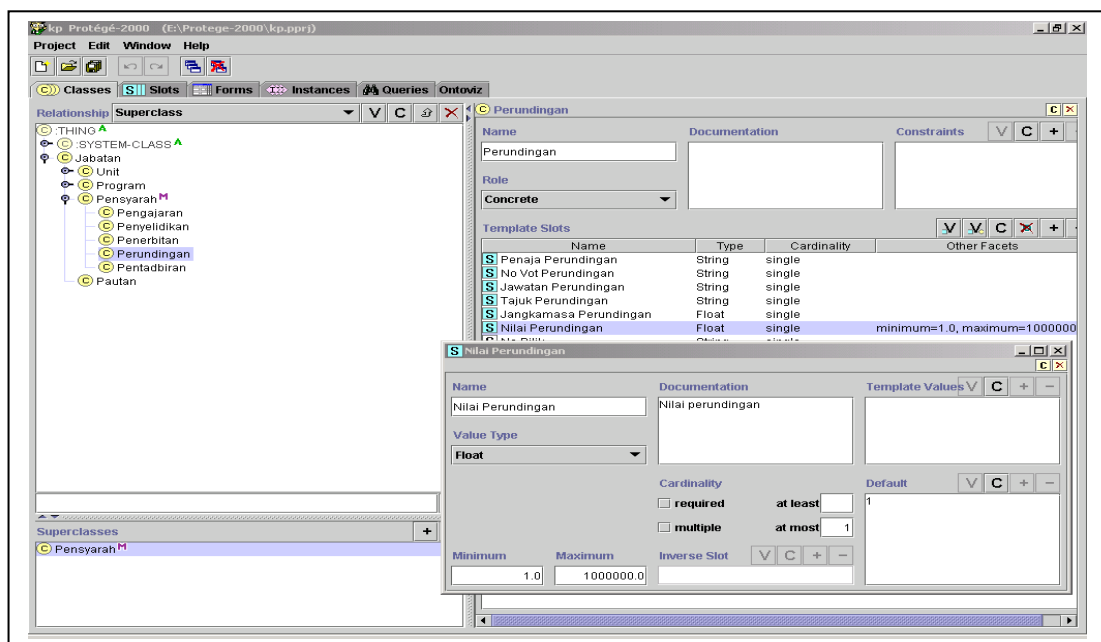
Rajah 5.4 : Petikan paparan slot bagi kelas perundingan

5.3.3 Faset

Rajah 5.5 dan 5.6 memaparkan contoh faset bagi slot nama dan nilai perundingan. Nama merupakan slot yang mempunyai jenis nilai rentetan. Manakala, faset bagi nilai perundingan adalah berjenis apung dengan julat dari satu sehingga sejuta. Nilai lalai bagi slot nilai perundingan adalah satu.



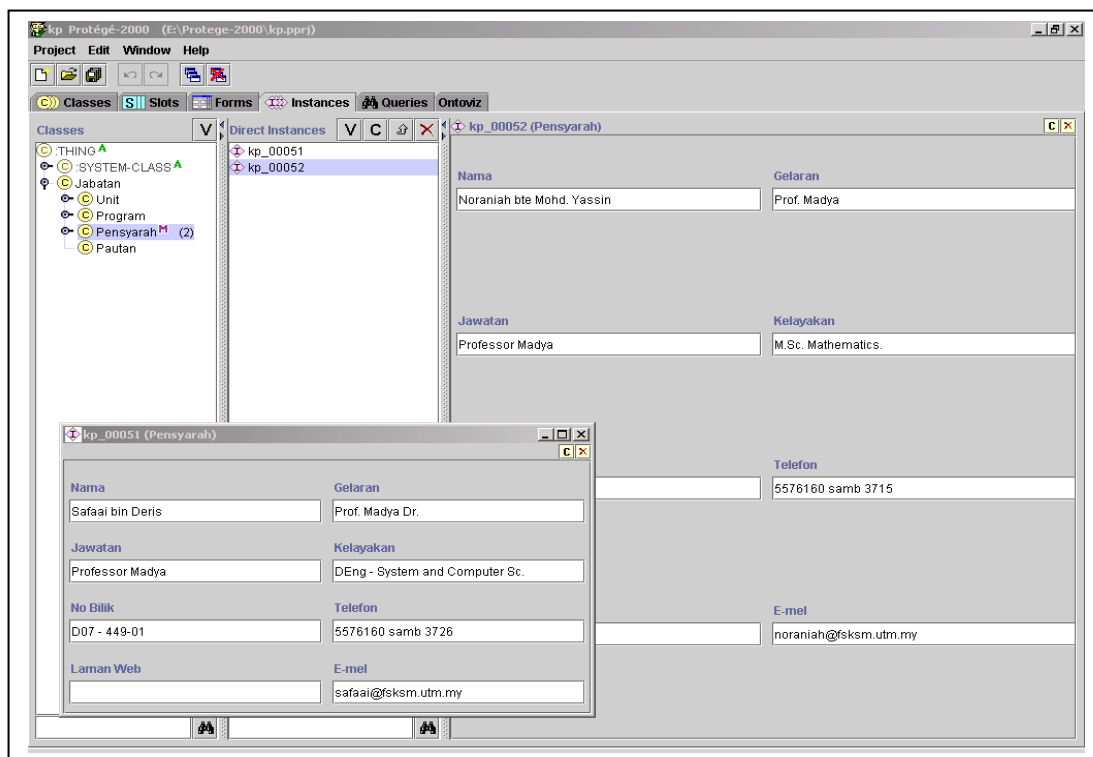
Rajah 5.5 : Petikan paparan faset bagi slot nama



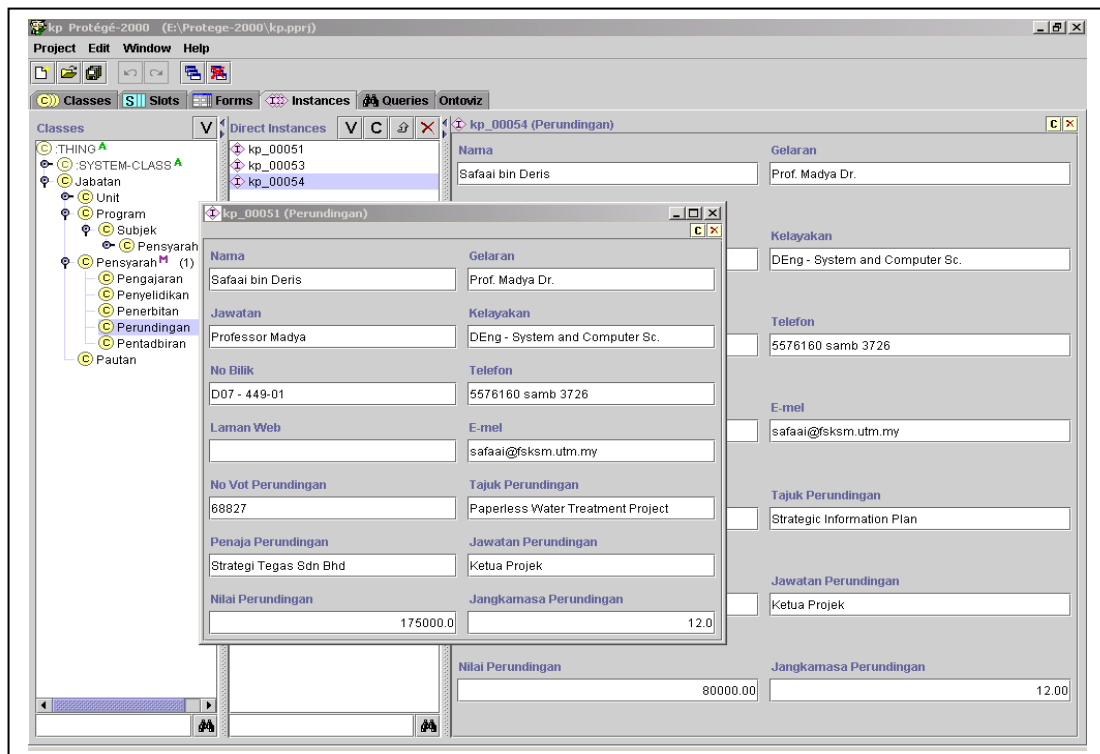
Rajah 5.6 : Petikan paparan faset bagi slot nilai perundingan

5.3.4 Tika

Tika bagi kelas dimasukkan melalui borang yang dijana secara automatik berdasarkan kepada slot dan faset sesuatu kelas. Rajah 5.7 dan 5.8 menunjukkan contoh tika bagi kelas pensyarah dan perundingan, di mana seorang pensyarah boleh mempunyai lebih daripada satu projek perundingan. Sebagaimana tika Safaai bin Deris yang mempunyai lebih daripada satu projek perundingan. Contohnya tika yang dilaksanakan oleh beliau ialah *Paperless Water Treatment* untuk Strategi Tegas Sdn Bhd dan *Strategic Information Plan* untuk Lembaga Kemajuan Perusahaan Pertanian Negeri Pahang.



Rajah 5.7 : Petikan paparan tika bagi kelas pensyarah

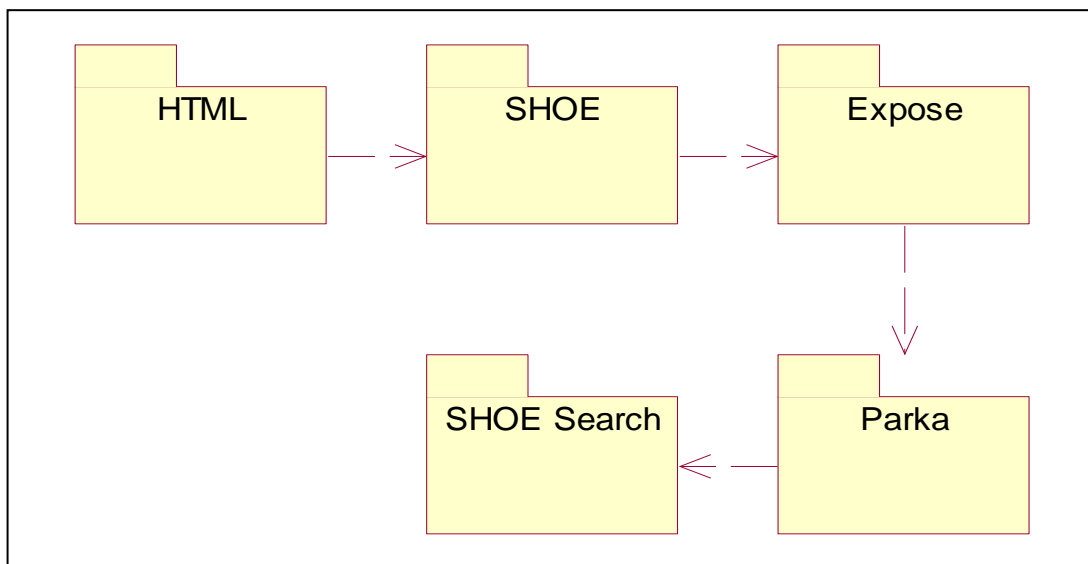


Rajah 5.8 : Petikan paparan tika bagi kelas perundingan

5.3.5 Gambarajah Pakej Ontologi

Gambarajah pakej ontologi pada Rajah 5.9 menunjukkan lima pakej yang dikenal pasti dalam membentuk ontologi dan sistem ontologi, iaitu:

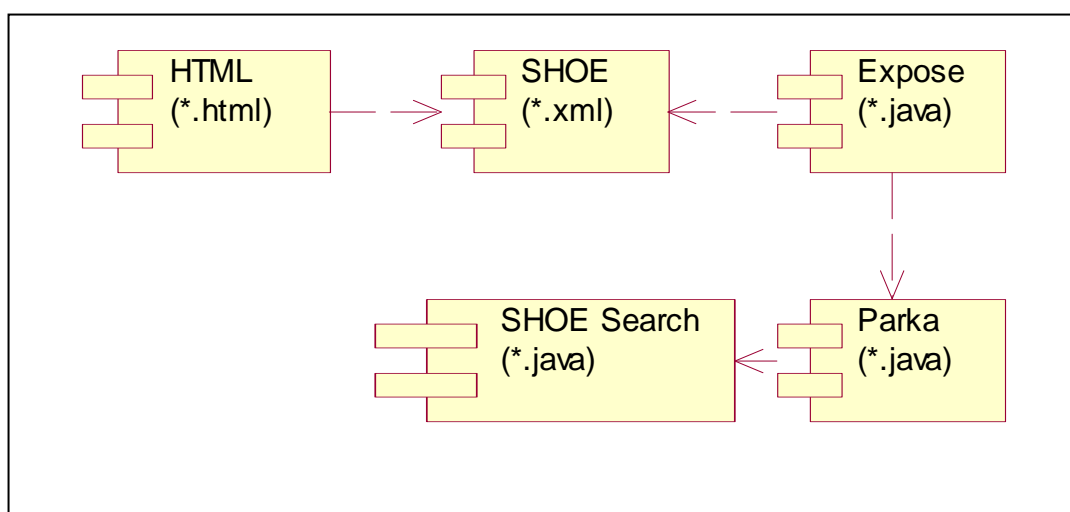
- HTML – pakej yang mengandungi laman web berbentuk maklumat.
- SHOE – pakej yang mengandungi laman web berbentuk ontologi.
- Expose – pakej yang mengandungi *web crawler*.
- Parka – pakej yang mengandungi pangkalan pengetahuan.
- SHOE Search – pakej yang mengandungi antara muka gelintaran web.



Rajah 5.9 : Gambarajah pakej ontologi

5.3.6 Gambarajah Komponen Ontologi

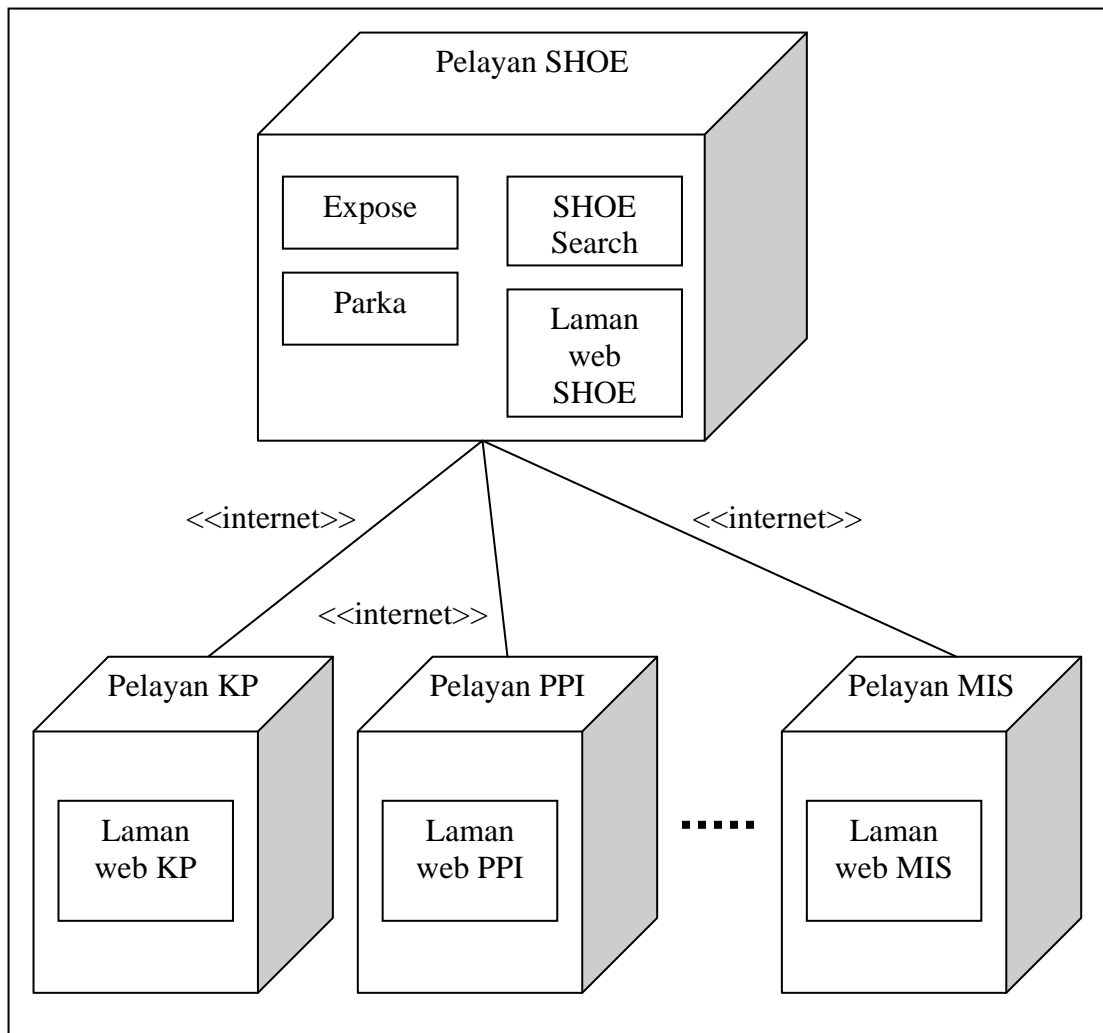
Keseluruhan komponen gelintaran web iaitu Expose, Parka dan SHOE Search dibangunkan menggunakan bahasa JAVA seperti yang ditunjukkan di dalam Rajah 5.10. Manakala, komponen laman web berbentuk maklumat dibangunkan menggunakan HTML dan komponen laman web berbentuk SHOE pula dibangunkan menggunakan HTML yang telah ditambah dengan tag-tag SHOE yang bersesuaian.



Rajah 5.10 : Gambarajah komponen ontologi

5.3.7 Gambarajah Penempatan Ontologi

Rajah 5.11 menunjukkan contoh gambarajah penempatan ontologi bagi enjin gelintaran web berasaskan SHOE, di mana terdapat beberapa laman web jabatan yang berbentuk maklumat di dalam mesin pelayan yang tersendiri. Laman web tersebut akan ditukarkan kepada bentuk ontologi dan diletakkan dalam laman web SHOE di dalam mesin pelayan SHOE. Bagi membolehkan gelintaran web dilakukan ke atas laman web SHOE ini, beberapa alatan seperti Expose, Parka dan SHOE Search digunakan. Pelayan SHOE akan berinteraksi dengan pelayan setiap jabatan melalui internet.



Rajah 5.11 : Gambarajah penempatan ontologi

5.4 Ringkasan

Fasa reka bentuk ontologi ke atas laman web jabatan KP telah dilaksanakan menggunakan Protégé melalui pelaksanaan aktiviti penakrifan ontologi dan menggunakan Rational Rose melalui pelaksanaan aktiviti platform ontologi. Hasil yang diperolehi dapat menerangkan taksonomi dan atribut ontologi serta struktur dan komponen enjin gelintaran web yang hendak dibangunkan dengan jelas dan tidak taksa.

BAB VI

IMPLEMENTASI DAN PENGUJIAN ONTOLOGI SERTA INTEGRASI DAN PENGUJIAN APLIKASI ONTOLOGI

6.1 Pendahuluan

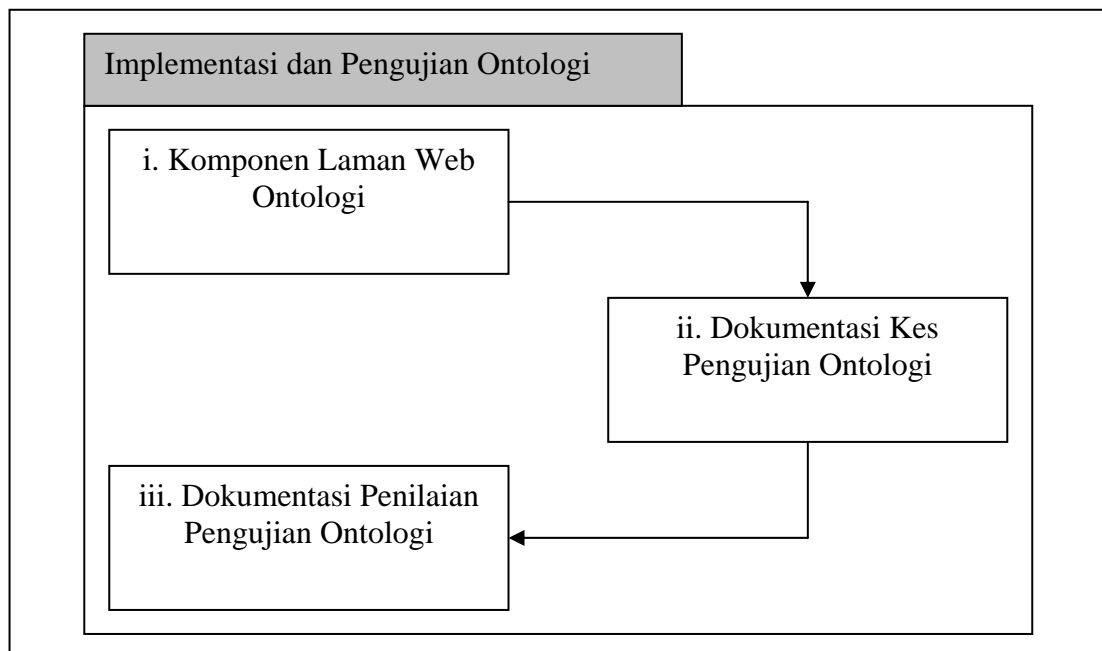
Perkara utama yang perlu diambil kira dalam fasa implementasi dan pengujian ontologi ialah menterjemahkan reka bentuk ontologi kepada satu set arahan atur cara. Di dalam projek ini, SHOE digunakan sebagai bahasa untuk menulis ontologi, di mana ianya disokong oleh sintaks dan notasi HTML. Manakala, reka bentuk laman web berbentuk maklumat dibangunkan menggunakan sintaks dan notasi HTML sepenuhnya. Fasa integrasi dan pengujian aplikasi ontologi pula menitikberatkan penghasilan produk akhir yang boleh dilarikan dalam persekitaran dunia sebenar. Di dalam projek ini, satu aplikasi carian web berasaskan semantik web atau ontologi dibangunkan dengan mengintegrasikan alatan-alatan yang menyokong SHOE seperti Expose, Parka dan SHOE Search dengan laman web berbentuk pengetahuan dan maklumat. Topik-topik yang dikupas dalam bab ini adalah seperti berikut:

- (a) Konsep dan rangka kerja yang digunakan dalam fasa implementasi dan pengujian ontologi.
- (b) Hasil daripada pelaksanaan fasa implementasi dan pengujian ontologi.
- (c) Konsep dan rangka kerja yang digunakan dalam fasa integrasi dan pengujian aplikasi ontologi.
- (d) Hasil daripada pelaksanaan fasa integrasi dan pengujian aplikasi ontologi.

6.2 Rangka Kerja Fasa Implementasi dan Pengujian Ontologi

Rajah 6.1 menunjukkan aktiviti di dalam mengimplementasi dan menguji ontologi. Tujuan fasa ini adalah untuk menghasilkan komponen semantik web dan komponen web. Terdapat tiga aktiviti utama di dalam fasa ini, iaitu:

- (a) Membangunkan komponen laman web berbentuk ontologi atau semantik web menggunakan SHOE dan laman web berbentuk maklumat atau web menggunakan HTML.
- (b) Menentukan kes pengujian bagi setiap komponen.
- (c) Menjalankan aktiviti pengujian dan menilai hasil pengujian bagi setiap komponen.

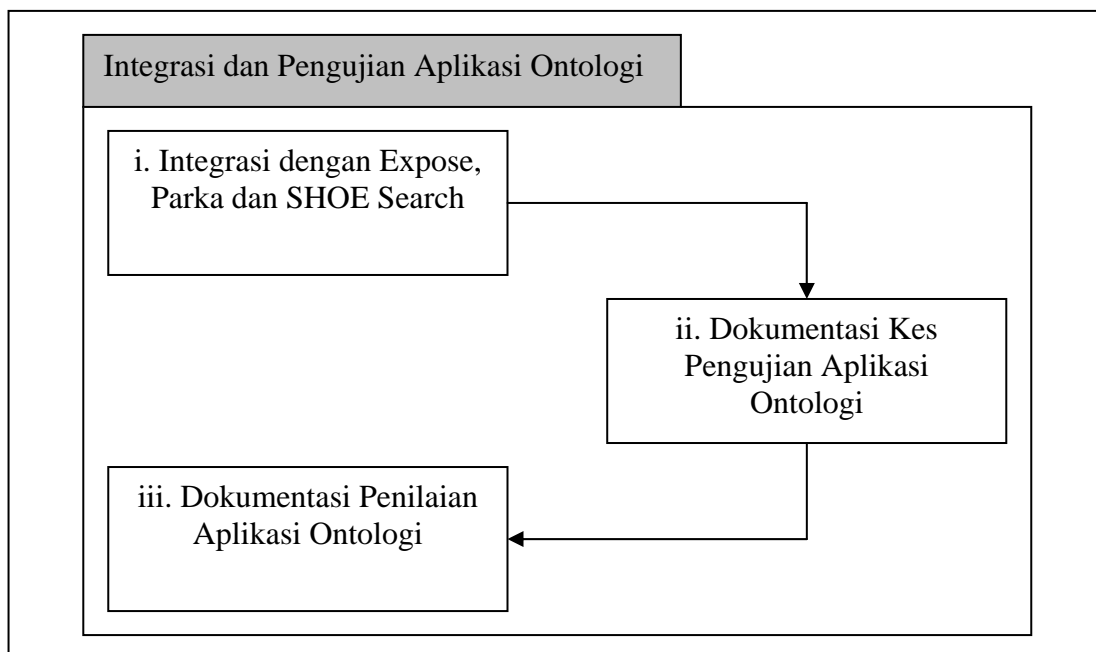


Rajah 6.1 : Aktiviti dalam fasa implementasi dan pengujian ontologi

6.3 Rangka Kerja Fasa Integrasi dan Pengujian Aplikasi Ontologi

Rajah 6.2 pula menggambarkan aktiviti di dalam mengintegrasikan dan menguji aplikasi ontologi. Tujuan fasa ini adalah untuk menghasilkan satu aplikasi berasaskan ontologi yang lengkap. Terdapat tiga aktiviti utama di dalam fasa ini, iaitu:

- (a) Mengintegrasikan komponen laman web berbentuk ontologi dan laman web berbentuk maklumat dengan alatan SHOE seperti Expose, Parka dan SHOE Search.
- (b) Menentukan kes pengujian bagi pengintegrasian setiap komponen.
- (c) Menjalankan aktiviti pengujian dan menilai hasil pengujian bagi pengintegrasian setiap komponen.



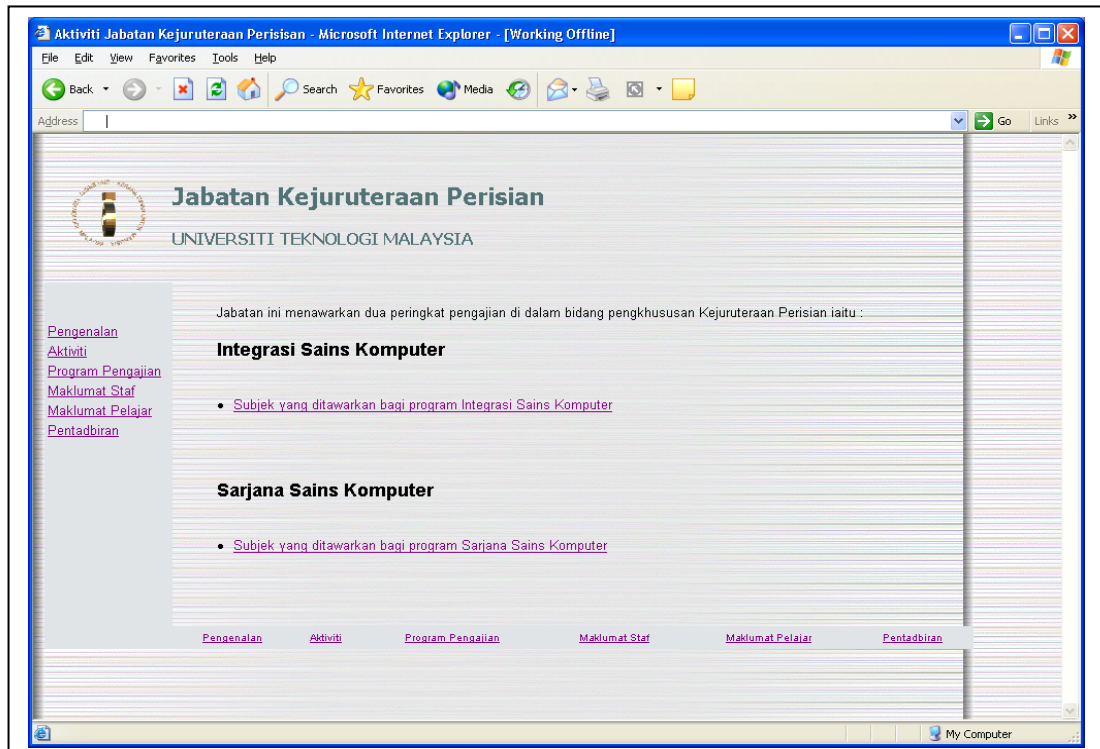
Rajah 6.2 : Aktiviti dalam fasa integrasi dan pengujian aplikasi ontologi

6.4 Hasil Implementasi

Hasil pertama implementasi bagi projek ini adalah bermula dengan pembangunan laman web jabatan KP yang berbentuk maklumat atau web menggunakan HTML. Rajah 6.3 sehingga Rajah 6.8 menunjukkan sebahagian petikan paparan bagi laman web jabatan KP yang berbentuk maklumat seperti laman utama, senarai program pengajian, senarai subjek yang ditawarkan, senarai pensyarah, pentadbiran jabatan dan senarai pelajar.



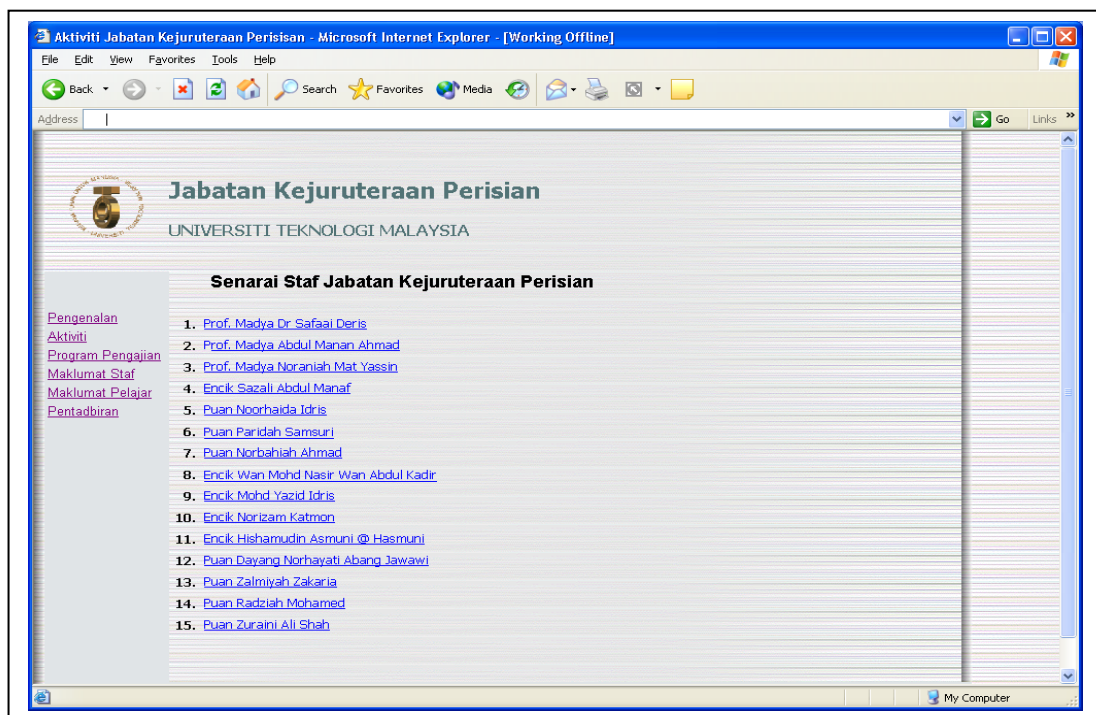
Rajah 6.3 : Petikan paparan utama laman web Jabatan Kejuruteraan Perisian berbentuk maklumat



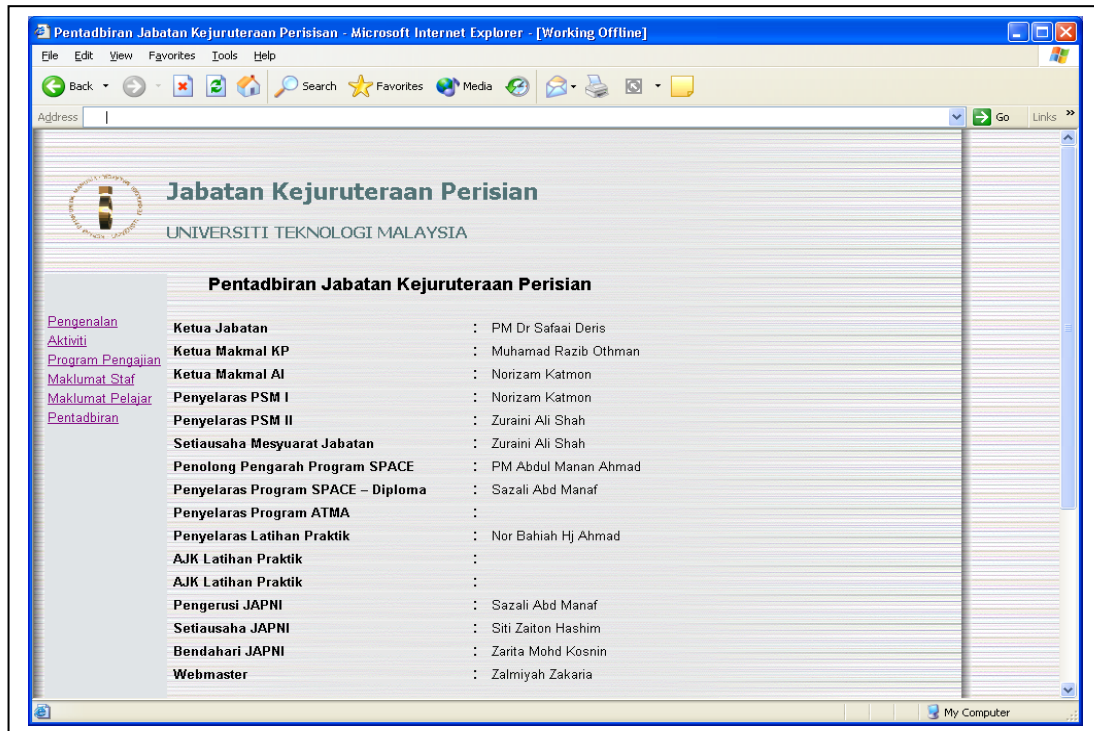
Rajah 6.4 : Petikan paparan senarai program pengajian bagi Jabatan Kejuruteraan Perisian berbentuk maklumat



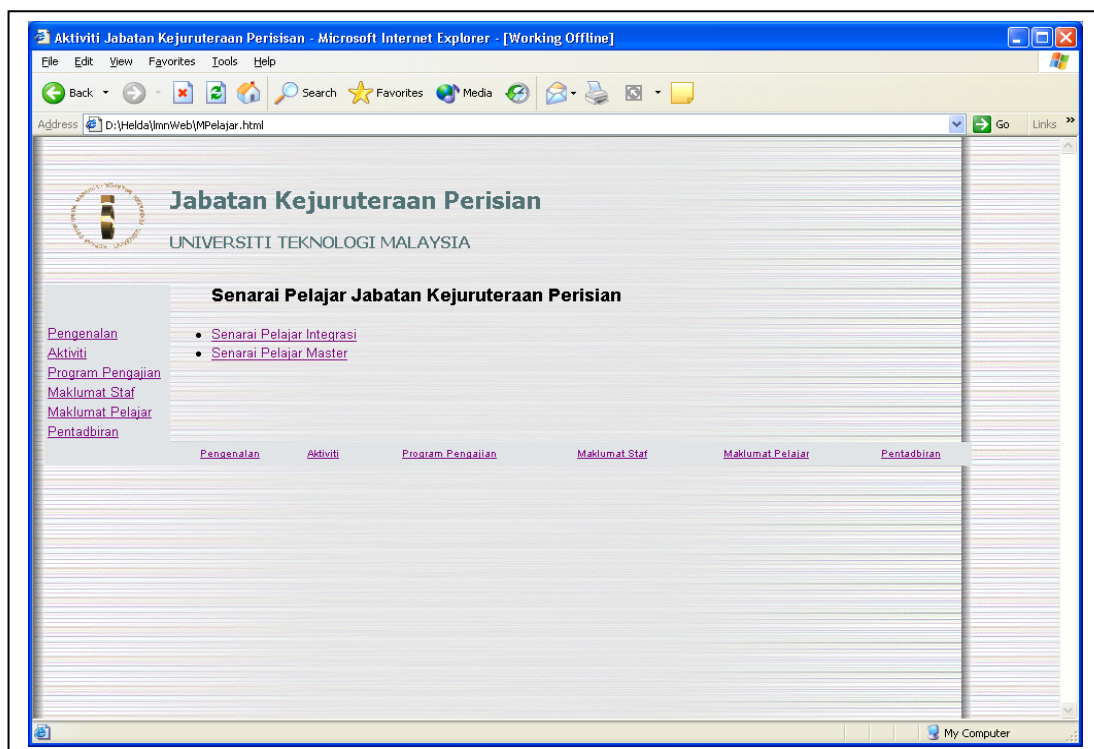
Rajah 6.5 : Petikan paparan senarai subjek yang ditawarkan bagi Jabatan Kejuruteraan Perisian berbentuk maklumat



Rajah 6.6 : Petikan paparan senarai pensyarah bagi Jabatan Kejuruteraan Perisian berbentuk maklumat

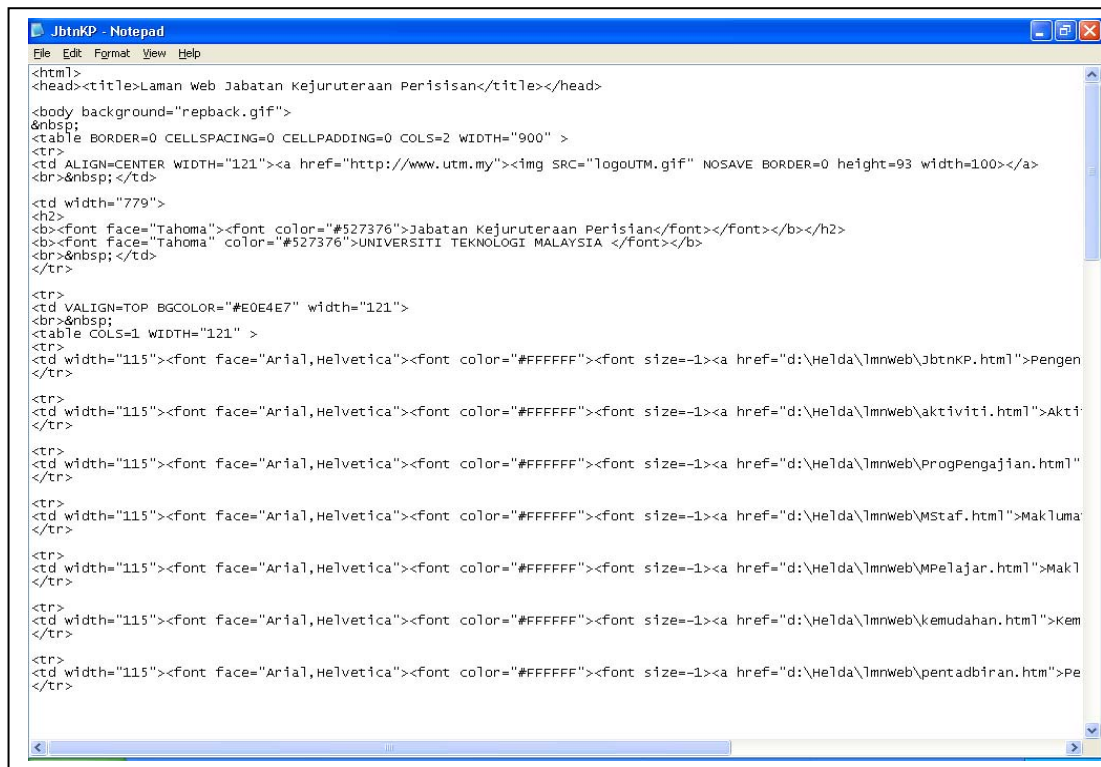


Rajah 6.7 : Petikan paparan pentadbiran bagi Jabatan Kejuruteraan Perisian berbentuk maklumat



Rajah 6.8 : Petikan paparan senarai pelajar bagi Jabatan Kejuruteraan Perisian berbentuk maklumat

Contoh kod atur cara HTML yang ditulis untuk membangunkan laman web jabatan KP berbentuk maklumat adalah seperti pada Rajah 6.9.



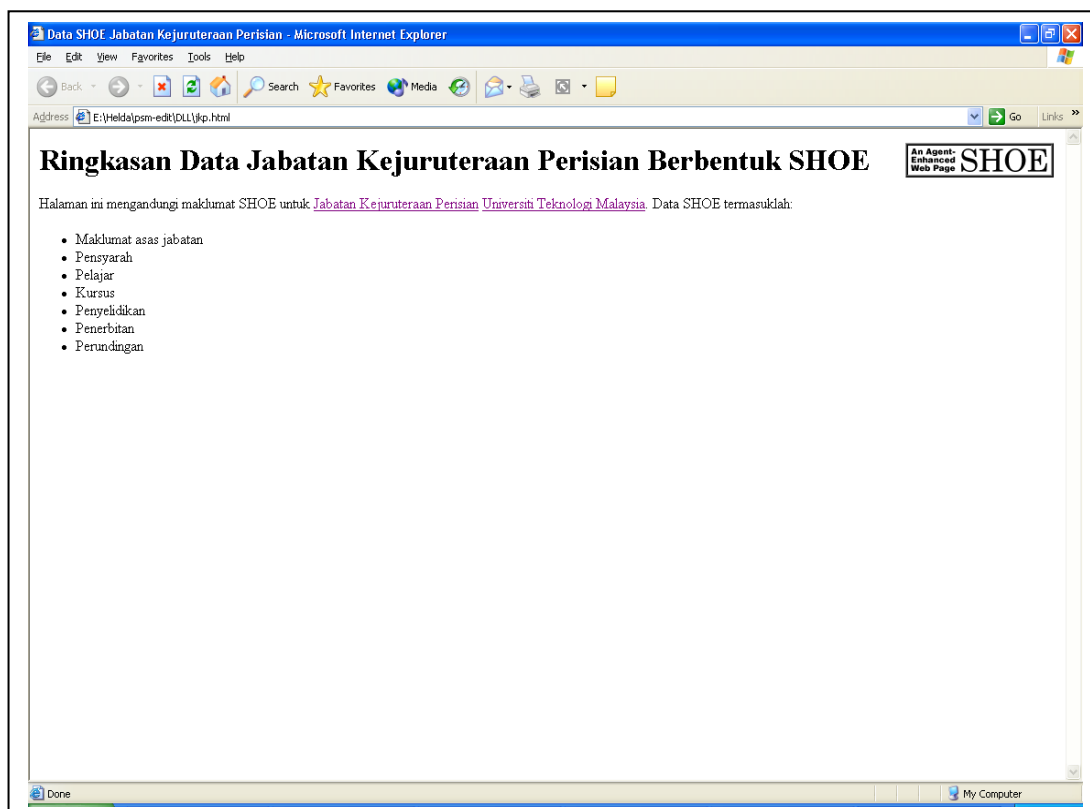
```

<html>
<head><title>Laman web Jabatan Kejuruteraan Perisisan</title></head>
<body background="repback.gif">
<table BORDER=0 CELLSPACING=0 CELLPADDING=0 COLS=2 WIDTH="900" >
<tr>
<td ALIGN=CENTER WIDTH="121"><a href="http://www.utm.my"><img SRC="logOUTM.gif" NOSAVE BORDER=0 height=93 width=100></a>
<br>&nbsp;</td>
<td width="779">
<h2>
<b><font face="Tahoma"><font color="#527376">Jabatan Kejuruteraan Perisian</font></b></h2>
<b><font face="Tahoma" color="#527376">UNIVERSITI TEKNOLOGI MALAYSIA </font></b>
<br>&nbsp;</td>
</tr>
<tr>
<td VALIGN=TOP BGCOLOR="#E0E4E7" width="121">
<br>&nbsp;</td>
<table COLS=1 WIDTH="121" >
<tr>
<td width="115"><font face="Arial,Helvetica"><font color="#FFFFFF"><font size=-1><a href="d:\Hel\da\l\mnweb\jbtnKP.html">Pengen
</tr>
<tr>
<td width="115"><font face="Arial,Helvetica"><font color="#FFFFFF"><font size=-1><a href="d:\Hel\da\l\mnweb\aktiviti.html">Aktif
</tr>
<tr>
<td width="115"><font face="Arial,Helvetica"><font color="#FFFFFF"><font size=-1><a href="d:\Hel\da\l\mnweb\ProgPENGajian.html"
</tr>
<tr>
<td width="115"><font face="Arial,Helvetica"><font color="#FFFFFF"><font size=-1><a href="d:\Hel\da\l\mnweb\Mstaf.html">Makluma
</tr>
<tr>
<td width="115"><font face="Arial,Helvetica"><font color="#FFFFFF"><font size=-1><a href="d:\Hel\da\l\mnweb\MPelajar.html">Makl
</tr>
<tr>
<td width="115"><font face="Arial,Helvetica"><font color="#FFFFFF"><font size=-1><a href="d:\Hel\da\l\mnweb\kemudahan.html">Kem
</tr>
<tr>
<td width="115"><font face="Arial,Helvetica"><font color="#FFFFFF"><font size=-1><a href="d:\Hel\da\l\mnweb\pentadbiran.htm">Pe
</tr>

```

Rajah 6.9 : Petikan kod atur cara HTML bagi laman web Jabatan Kejuruteraan Perisian berbentuk maklumat

Hasil kedua implementasi bagi projek ini adalah dengan penukaran laman web jabatan KP yang berbentuk maklumat seperti pada Rajah 6.3 sehingga Rajah 6.8 kepada laman web jabatan KP yang berbentuk pengetahuan atau semantik web seperti pada Rajah 6.10. Laman web jabatan KP yang berbentuk pengetahuan ini akan diindeks ke dalam pangkalan pengetahuan Parka oleh Expose. Sekiranya carian ke atas laman web jabatan KP dilakukan, SHOE Search akan merujuk kepada laman web yang berbentuk ontologi ini dan memaparkan URL yang mempunyai pautan ke laman web jabatan KP yang berbentuk maklumat tadi.



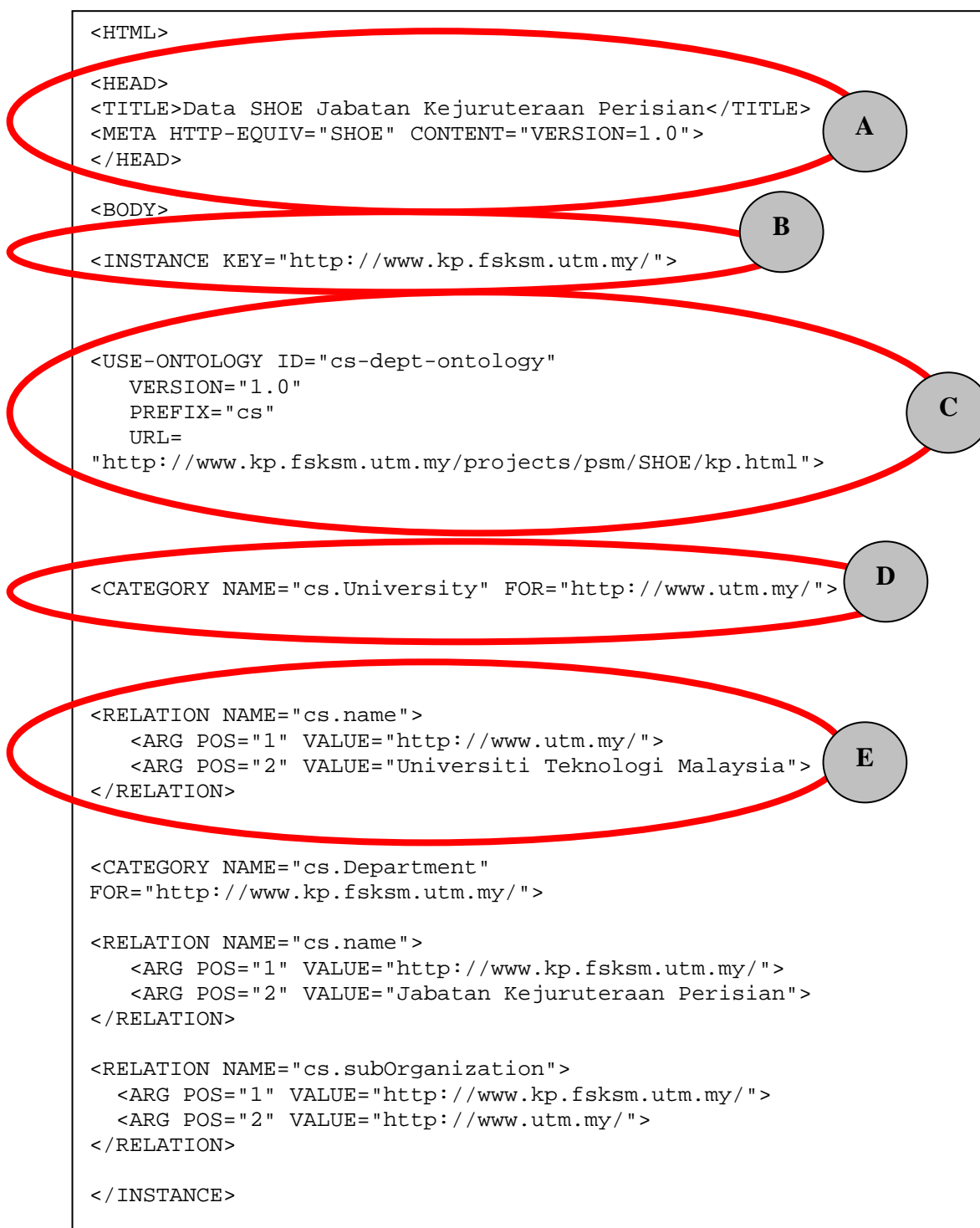
Rajah 6.10 : Petikan paparan laman web Jabatan Kejuruteraan Perisian berbentuk ontologi

Contoh kod atur cara laman web jabatan KP yang berbentuk SHOE adalah seperti pada Rajah 6.11. Bagi membolehkan Expose mengenal pasti bahawa sesuatu dokumen adalah berbentuk SHOE, pengisytiharan laman web tersebut sebagai SHOE 1.0-compliant tags perlu dibuat. Pengisytiharan ini dibuat pada seksyen <HEAD> seperti pada Rajah 6.11 (a). Sebelum maklumat semantik ditambah ke dalam laman web SHOE, satu atau lebih tika perlu ditakrifkan termasuk objek data yang manakah perlu dikelaskan atau dihubungkan antara satu sama lain. Penting bagi suatu tika menjadi unik dari tika yang lain, di mana situasi dua individu berlainan menulis tika dengan nama yang sama adalah tidak dibenarkan sama sekali. SHOE menangani masalah ini dengan mengaitkan setiap tika dengan kunci yang unik. SHOE mempunyai protokol yang piawai mengenai kunci untuk tika, iaitu setiap tika adalah berdasarkan kepada hanya satu URL untuk laman web yang berkaitan seperti yang ditunjukkan pada Rajah 6.11 (b).

Kemudian, sebelum mengisytiharkan fakta mengenai dokumen, apakah ontologi yang digunakan untuk menstrukturkan semua fakta tersebut perlulah diisytiharkan terlebih dahulu. Tanpa ontologi, agen tidak dapat memahami fakta-fakta yang terkandung di dalam sesuatu dokumen. Bagi menandakan bahawa dokumen tersebut menggunakan ontologi tertentu, pengisytiharan seperti pada Rajah 6.11 (c) perlu dibuat.

Setelah mengisytiharkan ontologi, tika di dalam dokumen yang telah diisytiharkan sebelum ini akan dikelaskan atau dikategorikan, iaitu mengisytiharkan apakah yang berkaitan dengan sesuatu tika. Di dalam SHOE, pengkategorian ini dilakukan menggunakan tag CATEGORY bersama-sama dengan satu atau lebih kategori yang telah diambil dari ontologi yang digunakan. Di bahagian badan kod atur cara ditambah pernyataan seperti yang terdapat pada Rajah 6.11 (d). Pernyataan ini menunjukkan tika tersebut dipunyai oleh kelas atau kategori “University” dan “Department” seperti yang ditakrifkan di dalam ontologi yang telah dikenal pasti sebelum ini dengan menggunakan awalan “cs”.

Hubungan di antara tika dan data yang lain adalah perlu dinyatakan. Contohnya, hubungan di antara tika yang telah dihasilkan sebelum ini dengan beberapa data yang biasa adalah seperti pada Rajah 6.11 (e). Pernyataan tersebut dapat menunjukkan bahawa terdapat hubungan “name” seperti yang ditakrifkan di dalam ontologi “cs” dengan “Universiti Teknologi Malaysia”, yang berjenis “http://www.utm.my/”. Pernyataan ini juga menunjukkan bahawa terdapat hubungan “subOrganization” seperti yang ditakrifkan di dalam ontologi “cs” dengan nilai data “http://www.kp.fsksm.utm.my/”, di mana data ini berjenis “http://www.utm.my/”.



Rajah 6.11 : Petikan kod atur cara SHOE bagi laman web Jabatan Kejuruteraan Perisian berbentuk ontologi

Beberapa contoh lain keratan kod atur cara SHOE adalah seperti pada Rajah 6.12 dan 6.13. Rajah 6.12 menunjukkan pengkategorian pensyarah. Manakala, Rajah 6.13 menunjukkan pengkategorian kumpulan penyelidikan.


```

<CATEGORY NAME="cs.Lecturer"
FOR="http://www.kp.fsksm.utm.my/~zuraini/">
<CATEGORY NAME="cs.Lecturer"
FOR="http://www.kp.fsksm.utm.my/~zalmiyah/">
<CATEGORY NAME="cs.Lecturer"
FOR="http://www.kp.fsksm.utm.my/~bahiah/">
<CATEGORY NAME="cs.Lecturer"
FOR="http://www.kp.fsksm.utm.my/~puteh/">

<RELATION NAME="cs.name">
  <ARG POS="1" VALUE="http://www.kp.fsksm.utm.my/~zuraini/">
  <ARG POS="2" VALUE="Zuraini Binti Ali Shah">
</RELATION>
<RELATION NAME="cs.name">
  <ARG POS="1" VALUE="http://www.kp.fsksm.utm.my/~zalmiyah/">
  <ARG POS="2" VALUE="Zalmiyah Binti Zakaria">
</RELATION>
<RELATION NAME="cs.name">
  <ARG POS="1" VALUE="http://www.kp.fsksm.utm.my/~bahiah/">
  <ARG POS="2" VALUE="Nor Bahiah Binti Haji Ahmad">
</RELATION>
<RELATION NAME="cs.name">
  <ARG POS="1" VALUE="http://www.kp.fsksm.utm.my/~puteh/">
  <ARG POS="2" VALUE="Puteh Binti Saad">
</RELATION>

```

Rajah 6.12 : Petikan kod atur cara SHOE bagi kategori pensyarah

```

<CATEGORY NAME="cs.ResearchGroup"
FOR="http://www.kp.fsksm.utm.my/~bioinfo/">
<CATEGORY NAME="cs.ResearchGroup"
FOR="http://www.kp.fsksm.utm.my/~speech_recog/">
<CATEGORY NAME="cs.ResearchGroup"
FOR="http://www.kp.fsksm.utm.my/~soft_comp/">

<RELATION NAME="cs.name">
  <ARG POS="1" VALUE="http://www.kp.fsksm.utm.my/~bioinfo/">
  <ARG POS="2" VALUE="BIOINFORMATICS">
</RELATION>
<RELATION NAME="cs.name">
  <ARG POS="1"
  VALUE="http://www.kp.fsksm.utm.my/~speech_recog/">
  <ARG POS="2" VALUE="SPEECH RECOGNITIONS">
</RELATION>
<RELATION NAME="cs.name">
  <ARG POS="1" VALUE="http://www.kp.fsksm.utm.my/~soft_comp/">
  <ARG POS="2" VALUE="SOFT COMPUTING">
</RELATION>

```

Rajah 6.13 : Petikan kod atur cara SHOE bagi kategori kumpulan penyelidikan

6.5 Hasil Integrasi

SHOE disokong oleh beberapa alatan bagi melengkapkan pelaksanaannya. Antara alatan yang dimaksudkan adalah Expose seperti yang ditunjukkan pada Rajah 6.14, Parka iaitu pangkalan pengetahuan yang menyimpan maklumat SHOE seperti pada Rajah 6.15 dan SHOE Search seperti pada Rajah 6.16. Bagi projek ini, setelah laman web yang berbentuk SHOE dihasilkan, maka peranan alatan seperti yang telah dinyatakan di atas iaitu Expose, Parka dan SHOE Search adalah diperlukan. Sebelum alatan ini dapat digunakan, beberapa perkara perlu dilaksanakan terlebih dahulu seperti mengkonfigurasi dan setkan persekitaran lariannya serta mengintegrasikan kesemua alatan ini bagi memastikan keberkesannya.

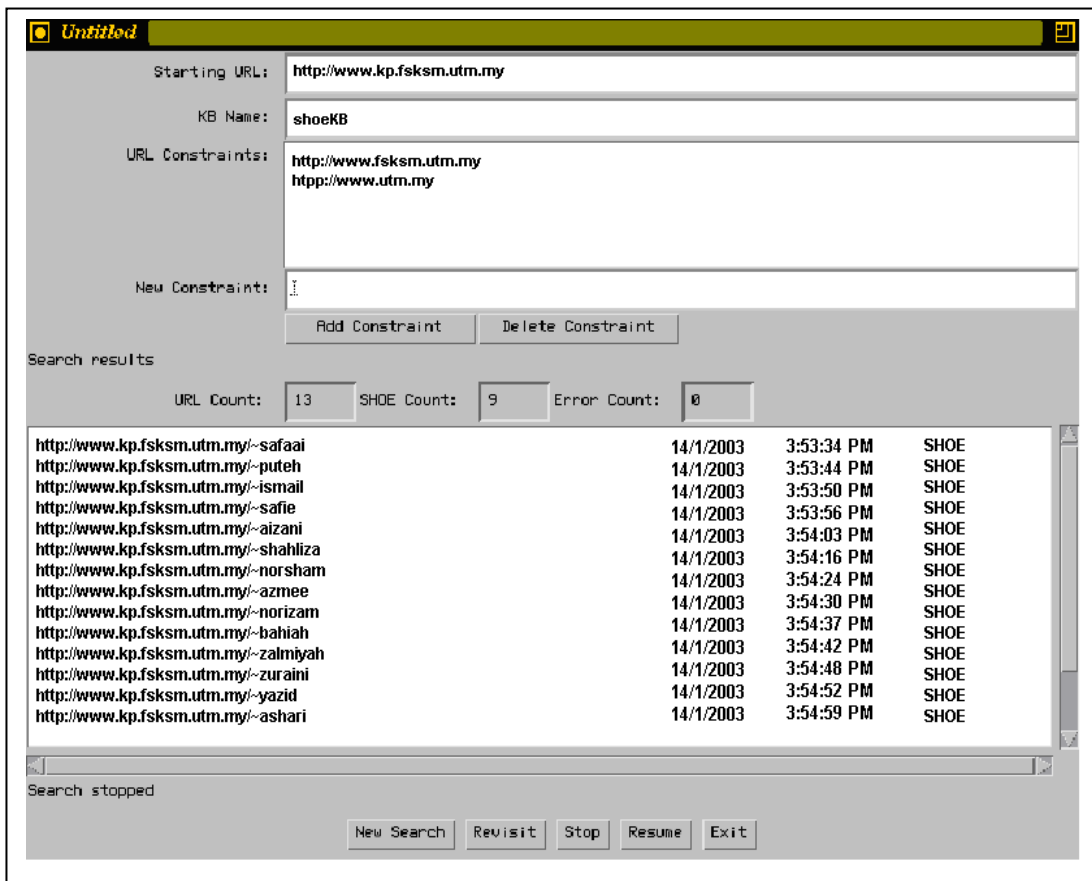
Langkah pertama adalah memuat turun alatan tersebut dari URL <http://www.cs.umd.edu/projects/plus/SHOE/downloads/>. Langkah kedua adalah memasang, mengkonfigurasi dan melarikan alatan tersebut sama ada di sistem pengoperasian Linux atau Windows. Kesemua alatan tersebut adalah multi platform kerana ia dibangunkan menggunakan bahasa Java kecuali Parka yang dicipta khusus untuk Linux. Bagi memastikan pangkalan pengetahuan Parka berfungsi dalam Windows, maka cygwin diperlukan. Ia boleh dimuat turun dari URL <http://www.cygwin.com/>. Langkah terakhir adalah mengintegrasikan kesemua alatan tersebut dengan laman web SHOE dan maklumat yang telah dibangunkan.

Expose dan SHOE Search dilarikan dengan menulis arahan berikut pada tettingkap DOS, di mana shoeKB adalah nama pangkalan pengetahuan, 161.139.67.192 adalah nama pelayan dan 8765 adalah nombor port pelayan:

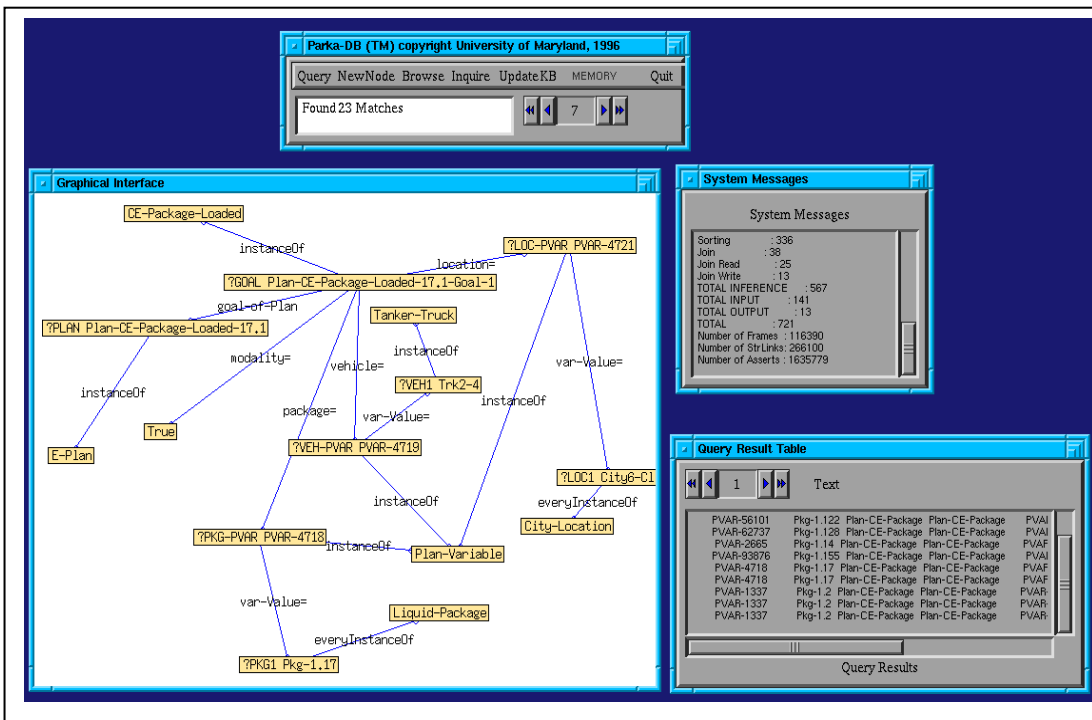
```
C:\> java Expose.ExposeApp
C:\> java SHOESearch.SSFrame shoeKB 161.139.67.192 8765
```

Manakala, Parka dilarikan dengan menulis arahan berikut, di mana 8765 adalah nombor port bagi pelayan Parka:

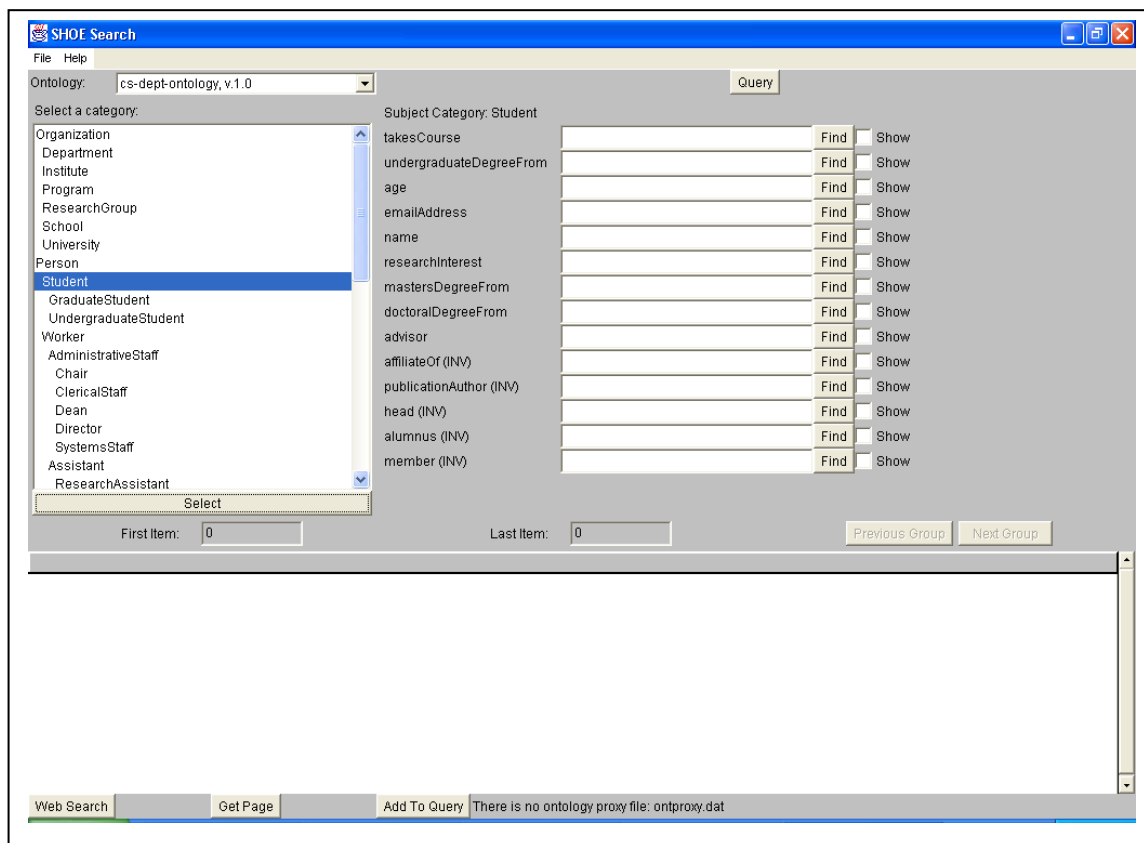
```
$ cd ClientServer
$ ./parkaServer -P8765
```



Rajah 6.14 : Petikan paparan Expose



Rajah 6.15 : Petikan paparan Parka



Rajah 6.16 : Petikan paparan SHOE Search

Setelah pengintegrasian selesai dilakukan, maka laman web berbentuk SHOE yang telah dihasilkan sebelum ini akan diindeks ke dalam Parka dengan menggunakan Expose. Ini adalah kerana Expose merupakan robot yang akan mencari laman web yang mengandungi penanda SHOE, membaca pengetahuan yang terkandung di dalam laman web berkaitan dan mengindeksnya ke dalam Parka. Expose digunakan dengan cara meletakkan permulaan URL dan meletakkan had ke atas apakah pelayan atau direktori yang dibenarkan untuk dilawati oleh Expose ini. Langkah kedua ini adalah penting kerana ia membenarkan agen untuk mengfokus terhadap bahagian web yang berkemungkinan terdapat SHOE. Selain itu, ia juga berguna dalam mengfokuskan pengumpulan pengetahuan di sepanjang topik yang khusus.

Apabila robot ini membaca sesuatu laman, ia akan mengambil maklumat SHOE dan menyimpannya ke dalam pangkalan pengetahuan Parka. Kemudian, ia akan mengenal pasti kesemua URL yang terkandung di dalam dokumen tersebut dan menggunakan *evaluation function* terhadap setiap URL tersebut untuk menentukan

susunan laluan yang perlu bagi melawatinya satu persatu. Kemudian, Expose akan memilih laman berikutnya untuk dilawati dan mengulangi proses ini. Proses ini berterusan sehingga tiada laman yang baru untuk dilawati. Pemilihan *evaluation function* ini adalah untuk mengawal susunan yang dilakukan oleh Expose. Fungsi yang digunakan ini beroperasi mengikut andaian berikut:

- (a) Laman SHOE umumnya berbentuk kluster.
- (b) Laman SHOE umumnya mempunyai pautan dengan laman SHOE yang lain.
- (c) Laman SHOE umumnya terletak dalam direktori yang sama dengan laman SHOE yang lain.

SHOE Search merupakan program yang menyemak kesemua laman yang telah diindeks ke dalam Parka oleh Expose sepertimana yang telah diterangkan sebelum ini untuk mencari padanan kepada carian dan menyenaraikan URL laman-laman web yang bersesuaian dengan apa yang dikehendaki oleh pengguna. Di dalam melakukan carian menggunakan SHOE Search, pengguna perlu melaksanakan langkah-langkah berikut:

- (a) Pilih ontologi. Kekotak senarai membenarkan pengguna memilih salah satu ontologi yang terdapat di dalam senarai. Secara lalainya, “*base ontology*” akan dipilih untuk pengguna.
- (b) Pilih kategori yang dirasakan paling relevan dengan perkara yang dicari. Senarai ini memaparkan kesemua istilah yang digunakan oleh ontologi. Subkategori terletak di bawah kategori utama (*parent category*) mereka.
- (c) Tekan butang “*Select*”. Senarai slot kepada kategori yang telah dipilih akan dipaparkan pada sebelah kanan senarai kategori.
- (d) Masukkan istilah bagi carian yang dilakukan. Masukkan nilai seperti URL, nama, nombor dan lain-lain pada senarai slot yang telah dipaparkan bagi menghadkan carian.
- (e) Tekan butang “*Query*”. Hasil bagi carian dipaparkan pada bahagian bawah senarai kategori dan senarai slot.
- (f) Paparan laman web. Dwi-klik pada URL atau pilih satu URL dan tekan “*Get Page*” untuk membuka tettingkap lain bagi memaparkan laman web.

- (g) Mengisui pertanyaan yang sama terhadap web yang umum. Tekan “*Web Search*” untuk mengisui pertanyaan yang sama kepada enjin carian yang popular.

6.6 Hasil Pengujian

Selepas kerja pengekodan selesai, maka proses pengujian pula akan dilaksanakan. Bagi projek yang dibangunkan ini, pengujian yang merangkumi beberapa peringkat yang biasanya dimulakan dengan ujian unit, ujian integrasi, ujian sistem dan ujian penerimaan tidak dapat dilakukan. Pengujian tersebut tidak dapat dilakukan kerana projek ini bukanlah melibatkan pembangunan program perisian. Namun, beberapa pengujian yang bersesuaian dengan projek ini akan dilakukan seperti pengujian pengejaan perkataan, pengujian format HTML, pengujian pautan, pengujian kesesuaian pengimbas (*browser*) dan pengujian masa muatan (*load time*).

Pengujian ke atas kod atur cara yang telah dibina adalah penting bagi memastikan alatan SHOE dapat mengesan sesuatu dokumen. Struktur dokumen yang tidak betul seperti kedudukan tag di seksyen yang salah, tag META diletakkan pada seksyen BODY dan bukannya pada seksyen HEAD dan sebagainya akan mengakibatkan Expose mengabaikan tag tersebut. Jelaslah bahawa pengujian perlu dilakukan dengan teliti dan teratur bagi memastikan perlaksanaan projek ini berjaya dan tidak menghadapi masalah yang kritikal kelak. Di dalam projek ini pengujian dijalankan pada komputer dengan spesifikasi Pentium III 800 MHz, 256 MB RAM, 30 GB HDD, JDK 1.3.1 dan Windows XP. Manakala, NetMechanic digunakan sebagai alatan pengujian. Contoh hasil laporan pengujian keseluruhan yang dijana oleh NetMechanic adalah seperti pada Rajah 6.17.

Your Results		
Page Summary		
URL: http://www.kp.fksm.utm.my/projects/psm/SHOE/kp.html		
Date Tested: Friday, Dec 8, 14:13 EDT		
Tool	Rating	Summary
Link Check	☆☆☆☆☆	0 bad links View a Detailed Report
HTML Check & Repair	☆☆☆☆☆	1 errors View a Detailed Report
Browser Compatibility	☆☆☆☆☆	0 problems View a Detailed Report
Load Time	☆☆☆☆☆	22.77 seconds, height/width problems View a Detailed Report
Spell Check	☆☆☆☆☆	3 possible errors View a Detailed Report

Rajah 6.17 : Contoh ringkasan laporan pengujian NetMechanic

6.6.1 Pengujian Pengejaan Perkataan

Pengujian pengejaan perkataan dilaksanakan bagi memastikan setiap patah perkataan sama ada yang berbentuk teks ataupun elemen di dalam tag adalah betul dari segi sintaks dan format HTML itu sendiri. Contoh hasil pengujian pengejaan perkataan adalah seperti pada Rajah 6.18.

Your Results		
Page Details: Spell Check		
URL: http://www.kp.fksm.utm.my/projects/psm/SHOE/kp.html		
Page Rating: ☆☆☆☆☆		
Word Count: 100	Misspellings: 2	
Date Tested: Friday, Dec 8, 14:13 EDT		
This report shows possible spelling errors on your page.		
Misspelled Words		
Word	Line (s)	Suggestions
DEM	26	DAM, DE, DEA, DEB, DEC, DED, DEE, DEEM, DEG, DEI, DEL, DEME, DEMI, DEMO, DEMY, DEN, DEO, DERM, DEV, DEW, DEY, DE M, DE-M, DIEM, DIM, DOM, DUM, D EM, D-EM, EM, GEM, HEM, IDEM, JEM, MEMSP , OEM, TEM, WEM
Sipse	17	Dipse, Kipse, Sips ey, Sips-ey, Sip sey, Sip-sey

Rajah 6.18 : Contoh laporan pengujian pengejaan perkataan

6.6.2 Pengujian Format HTML

Pengujian format HTML dilaksanakan bagi memastikan dokumen yang dihasilkan mengikuti peraturan bagi format pengekodan HTML yang sempurna. Langkah ini perlu dilaksanakan agar tidak timbul masalah di dalam memaparkan dokumen pada pengimbas yang berlainan. Contoh hasil pengujian format HTML adalah seperti pada Rajah 6.19.

Your Results

Page Details: HTML Check & Repair

URL: <http://www.kp.fksm.utm.my/projects/psm/SHOE/kp.html>

Rating: ★★☆☆☆

Total Errors: 1 **Total Warnings: 9** **Total Problems HTML Repair Cannot Fix: 7**

Date Checked: Friday, 08 December 2002 15:36 EST

Errors are places where your page does not follow the rules for proper HTML coding. These problems may cause your page to display incorrectly under different browsers.

Warnings are things that are not good coding practices.

Compatibility Problems are places where you use an HTML tag or attribute that is not part of the HTML 4.0 Standard and may not be supported by all browsers.

Rajah 6.19 : Contoh laporan pengujian format HTML

6.6.3 Pengujian Pautan

Kadangkala pautan kepada laman web boleh berubah atau rosak. Maka, pengujian pautan dilaksanakan untuk memastikan pautan yang terdapat di dalam dokumen yang dihasilkan adalah beroperasi seperti biasa iaitu tidak terdapat masalah. Contoh hasil pengujian pautan adalah seperti pada Rajah 6.20.

Your Results

Page Details: Link Check

URL: <http://www.kp.fksm.utm.my/projects/psm/SHOE/kp.html>

Page Rating: ★★★★★

Total Links: 20 Bad Links: 0

Date Tested: Friday, Dec 8, 14:13 EDT

This report shows the status of all links on your page. Links marked in red are broken.

Line	Tag	Links to...	Image Size	Status
45	HREF	links.htm		ok

Rajah 6.20 : Contoh laporan pengujian pautan

6.6.4 Pengujian Kesesuaian Pengimbas

Pengujian kesesuaian pengimbas dilaksanakan adalah untuk memastikan setakat mana dokumen yang dihasilkan dapat dipaparkan dengan paparan yang memuaskan pada pengimbas utama yang berlainan versi. Contohnya, sekiranya sesuatu pengimbas web itu tidak menyokong sesuatu tag, maka pengimbas web tersebut akan mengabaikan tag itu tadi. Situasi ini akan menyebabkan laman web yang dihasilkan akan dipaparkan dengan pandangan yang berlainan. Contoh hasil pengujian pautan adalah seperti pada Rajah 6.21.

Your Results

Page Details: Browser Compatibility

URL: <http://www.kp.fsksm.utm.my/projects/psm/SHOE/kp.html>

Page Rating: ★★★★★

Total Incompatibilities: 3 Number of Problems : 0

Date Tested: Friday, Dec 8, 14:13 EDT

This report shows how well your page appears under different versions of the major browsers. Tags or attributes that are incompatible with more that 10% of your visitors are highlighted as problem areas.

Tag	Attribute	Lines	Visitors Affected	Microsoft			Netscape		
				3	4	5	2	3	4
BODY	ALINK	9	4.00 %	N	Y	Y	Y	Y	Y
NOSCRIPT	--	80, 91	1.00 %	Y	Y	Y	N	Y	Y
SCRIPT	SRC	66, 90	1.00 %	Y	Y	Y	N	Y	Y

Rajah 6.21 : Contoh laporan pengujian kesesuaian pengimbas

6.6.5 Pengujian Masa Muatan

Pengujian masa muatan dilaksanakan untuk menguji berapakah masa yang diambil untuk dokumen ini dimuat oleh beberapa modem yang berkelajuan biasa. Contoh hasil pengujian masa muatan adalah seperti pada Rajah 6.22.

Your Results

Page Details: Load Time

URL: <http://www.kp.fsksm.utm.my/projects/psm/SHOE/kp.html>

Page Rating: ★★★★★

Total Page Size: 45979 bytes Total Graphics: 39770 bytes - 9 images

Server Connections: 5

Date Tested: Friday, Dec 8, 14:13 EDT

This report shows how quickly your page loads at several common modem speeds.

Load Time by Modem Speed	
Modem Speed	Download Time
14.4k	35.54 seconds
28.8k	22.77 seconds
56k	16.50 seconds
ISDN (128k)	12.87 seconds
T1 (1.44 MB)	10.26 seconds

Rajah 6.22 : Contoh laporan pengujian masa muatan

6.7 Ringkasan

Fasa implementasi dan pengujian ontologi serta fasa integrasi dan pengujian aplikasi ontologi ke atas laman web jabatan KP yang berbentuk pengetahuan dan maklumat telah dilaksanakan menggunakan SHOE dan lain-lain alatan SHOE seperti Parka, Expose dan SHOE Search. Manakala, pengujian telah dilaksanakan menggunakan NetMechanic untuk menguji beberapa kriteria seperti pengejaan perkataan, format HTML, pautan, kesesuaian pengimbas dan masa muatan. Hasil yang diperolehi membolehkan pengguna membuat carian web dengan lebih bermakna dan tepat melalui semantik web yang kemudiannya akan mempunyai pautan dengan laman web sebenar.

BAB VII

PERBINCANGAN DAN KESIMPULAN

7.1 Pendahuluan

Pada era globalisasi kini, kita adalah hampir kepada generasi baru WWW, di mana Tim Berners-Lee memanggilnya semantik web sebagaimana di dalam buku yang ditulisnya “*Weaving the Web*”. Tidak seperti WWW yang sedia ada, di mana data dikhaskan untuk penggunaan manusia semata-mata tetapi semantik web pula selain menyediakan data untuk penggunaan manusia ia juga dapat menyediakan data yang berupaya diproses oleh mesin. Kemudahan ini menjadikan penggunaan servis pintar seperti broker maklumat, agen carian, penapisan maklumat dan lain-lain dapat diperluaskan bersesuaian dengan proses yang diterangkan oleh Berners-Lee sebagai “*bringing the web to its full potential*”. Perkembangan ini memberi implikasi ke atas kajian dan penyelidikan kejuruteraan pengetahuan khususnya terhadap kegunaan dan peranan ontologi kerana ontologi merupakan teras kepada usaha tersebut. Oleh itu, selari dengan persekitaran maklumat semasa yang bergantung kepada internet dengan peredaran maklumat tanpa sempadan dan persekitaran yang dinamik telah menjadikan SHOE sebagai satu usaha dan alternatif yang perlu dipandang serius. Di mana skop dan domain SHOE perlu dikaji, diperkembangkan dan diserasikan dengan keperluan semasa. Topik-topik yang dikupas dalam bab ini adalah seperti berikut:

- (a) Sumbangan dan pencapaian projek.
- (b) Masalah dan kelemahan projek.
- (c) Cadangan pembaikan dan peningkatan projek.

7.2 Sumbangan dan Pencapaian Projek

Sumbangan primer bagi projek ini adalah penggunaan SHOE sebagai salah satu daripada produk ontologi untuk menghasilkan enjin carian yang berasaskan pengetahuan atau semantik web yang diimplementasi kepada laman web berasaskan pendidikan. Secara khususnya sumbangan projek ini boleh diperincikan seperti berikut:

- (a) Penggunaan metodologi bagi analisa dan reka bentuk ontologi yang diperkenalkan oleh Noy dan McGuinness (2001) yang digabungkan dengan model Air Terjun untuk menghasilkan model kitar hayat pembangunan perisian berasaskan pengetahuan.
- (b) Penghasilan model domain ontologi yang dianalisa menggunakan Protégé dan UML.
- (c) Penghasilan model implementasi ontologi yang direka bentuk menggunakan Protégé dan UML.
- (d) Pembangunan laman web jabatan KP yang berbentuk maklumat menggunakan HTML dan laman web jabatan KP yang berbentuk pengetahuan menggunakan SHOE sebagai kajian kes.
- (e) Pengimplementasian enjin carian berasaskan pengetahuan yang telah diintegrasikan dengan Expose, Parka dan SHOE Search untuk melaksanakan carian ke atas laman web jabatan KP yang berbentuk SHOE.

Merujuk kepada analisa dan reka bentuk berasaskan pengetahuan yang diperkenalkan oleh Noy dan McGuinness (2001), satu idea baru terhadap pembangunan ontologi telah diketengahkan. Memandangkan pembangunan ontologi merupakan aktiviti yang sukar. Maka, dengan penggunaan pendekatan yang diperkenalkan oleh Noy dan McGuinness (2001), ianya dapat memberi panduan di dalam pembangunan aplikasi berasaskan ontologi. Penggunaan alatan seperti Protégé pula memudahkan kerja-kerja di dalam menganalisa dan mereka bentuk ontologi. Disamping itu, Protégé juga membantu di dalam memahami dan menganalisa domain dan skop ontologi serta istilah yang digunakan.

Laman web jabatan KP yang diimplemen di dalam projek ini merupakan perkara yang baru di arena tempatan. Maka, projek yang dilaksanakan ini berupaya menjadi perintis bagi membentuk satu laman web pustaka ontologi atau ontolingua. Ianya membolehkan ontologi yang sama khususnya tika digunakan semula untuk membangunkan laman web ontologi bagi jabatan-jabatan yang lain di UTM. Laman web berbentuk SHOE ini akan berinteraksi dengan komponen-komponen SHOE bagi memastikan keberkesanan dan matlamat penggunaan SHOE ke atas masalah pencarian di internet tercapai. Komponen-komponen SHOE yang dimaksudkan itu adalah seperti Expose, SHOE Search dan Parka.

Projek ini dibangunkan dengan memenuhi setiap matlamat, objektif dan skop yang telah didefinisikan pada Bab I iaitu ketika awal pembangunan projek. Produk akhir projek ini mampu membuat carian terhadap laman web yang dikodkan menggunakan tag-tag yang piawai di dalam SHOE, di mana kod HTML telah ditambah dengan kod penanda semantik yang berdasarkan konsep ontologi. Kemudian, alatan SHOE seperti Expose, Parka dan SHOE Search mampu beroperasi dengan sempurna untuk memastikan carian ke atas laman web yang berbentuk ontologi dapat dilaksanakan.

Semasa melakukan carian menggunakan SHOE Search, pengguna perlu memilih ontologi yang berkaitan dengan carian yang hendak dilakukan terlebih dahulu. Kemudian, satu set kategori akan dipaparkan hasil daripada pemilihan ontologi yang telah dibuat oleh pengguna. Setelah memilih kategori yang paling relevan dengan perkara yang dicari, satu senarai slot kepada kategori yang telah dipilih akan dipaparkan. Pengguna perlu mengisi nilai bagi slot yang dipaparkan bagi menghadkan carian dan menjadikan carian lebih spesifik. Setelah pengguna mengisi nilai bagi slot, keputusan bagi carian yang dilakukan akan dijana dan hasilnya akan dipaparkan kepada pengguna.

Berdasarkan kepada hasil dan pencapaian pelaksanaan projek ini, didapati projek ini mampu menyumbang ke arah pembentukan semantik web yang salah satu daripada cirinya adalah mampu menjana hasil carian yang bermakna dan bertepatan dengan kehendak pengguna.

7.3 Masalah dan Kelemahan Projek

Semantik web merupakan usaha ke arah menjadikan data di WWW dapat diproses dan difahami oleh mesin. Pelbagai usaha yang dilakukan oleh pihak-pihak tertentu seperti ahli-ahli di W3C bagi menjadikan semantik web ini memenuhi peranannya dalam erti kata yang sebenar. Begitu juga dengan kewujudan SHOE yang dibangunkan oleh kumpulan PLUS dari Universiti Maryland merupakan usaha ke arah pembentukan semantik web, di mana usaha inilah yang telah diketengahkan di dalam projek ini. Di sebalik kejayaan di dalam mengimplementasi SHOE ke dalam domain projek ini dan kejayaan di dalam mencapai matlamat yang digariskan, terdapat beberapa masalah dan kelemahan yang dikenal pasti. Masalah dan kelemahan ini secara kasarnya tidak menjejaskan usaha ke arah pembentukan semantik web tetapi sekiranya masalah dan kelemahan ini dikaji dengan lebih mendalam dan diperbaiki ia dapat menyempurnakan keberkesanan semantik web dan sekaligus dapat memperluaskan penggunaannya.

Salah satu kelemahan yang dapat dikenal pasti adalah pembinaan kod atur cara yang berbentuk SHOE adalah agak rumit untuk dihasilkan oleh pembangun laman web biasa yang kurang arif mengenai hal-hal berkaitan dengan semantik web dan perwakilan pengetahuan. Hal ini amat berbeza jika dilihat kepada pendekatan yang dilakukan oleh Tim Berners-Lee yang telah membina sistem WWW yang mudah, di mana sistem ini senang untuk difahami, digunakan dan disenggara sehingga menjadikannya faktor yang penting di dalam perkembangan web yang pesat. Pendekatan seumpama ini adalah perlu untuk mempengaruhi pembangun laman web untuk menukarkan web mereka ke dalam bentuk semantik web yang lebih bermakna.

Selain itu, memandangkan semantik web merupakan sambungan kepada web yang sedia ada pada masa ini, maka alatan yang berkaitan dengan semantik web mestilah boleh digunakan oleh pengguna biasa dan bukan hanya untuk mereka yang pakar. Disebabkan alatan carian merupakan alatan yang paling kerap digunakan, maka perhatian yang lebih perlu diberikan terhadapnya. Secara idealnya, pengguna mahukan alatan yang mudah digunakan seperti yang diimplemen pada enjin carian berdasarkan kata kunci tetapi ia mestilah menyediakan jawapan munasabah yang

bertepatan dengan pertanyaan yang dibuat oleh pengguna. Merujuk kepada alatan SHOE iaitu SHOE Search yang digunakan untuk melakukan carian, walaupun agak mudah untuk memilih konteks carian daripada senarai pilihan ontologi yang terdapat pada SHOE Search sekiranya hanya terdapat satu dozen ontologi yang perlu dipilih daripada senarai tetapi masalah pasti akan timbul kerana semantik web secara umumnya mempunyai ontologi yang mencecah kepada ribuan jenis.

Masalah lain yang dikenal pasti adalah bagaimana pengguna dapat memilih ontologi yang betul dan bersesuaian dengan carian yang dilakukan daripada senarai yang terdapat pada SHOE Search. Kebanyakan pengguna mahukan jawapan yang segera dan tidak perlu bersusah payah untuk meneliti laman yang mengandungi dokumentasi mengenai ontologi bagi membolehkan mereka membuat carian di web.

7.4 Cadangan Pembaikan dan Peningkatan Projek

SHOE merupakan satu kaedah yang perlu diberi perhatian di dalam usaha membentuk semantik web. Tidak dinafikan bahawa terdapat beberapa kelemahan yang dapat dilihat pada SHOE tetapi memandangkan pembentukan semantik web merupakan suatu bidang yang baru dan masih berada di peringkat yang awal, maka penghasilan SHOE mampu menyumbang kepada usaha pembentukan semantik web yang mempunyai faedah yang besar kepada pengguna web.

Cadangan bagi pembaikan projek ini adalah kajian yang lebih mendalam perlu dilakukan untuk menghasilkan alatan yang dapat menukarkan laman web sedia ada yang berbentuk maklumat kepada laman web berbentuk SHOE secara semi-automatik. Ini adalah kerana tidak realistik sekiranya alatan ini mampu beroperasi secara automatik disebabkan oleh terdapat masalah yang timbul berkaitan dengan pemprosesan bahasa tabii di dalam domain yang umum. Sekiranya terdapat alatan seperti itu, maka sedikit sebanyak dapat mempengaruhi pembangun laman web untuk cuba menukarkan laman web mereka ke dalam bentuk SHOE. Seandainya semua pembangun laman web berpendapat adalah wajar bagi mereka untuk menukarkan

laman web mereka ke dalam bentuk SHOE, maka secara tidak langsung dapat memperluas usaha pembentukan semantik web.

Selain itu, pembaikan dari segi alatan carian yang digunakan bagi projek ini iaitu SHOE Search perlu dilakukan agar pengguna lebih selesa dan mudah berinteraksi dengannya tidak ubah seperti melakukan carian menggunakan enjin carian berdasarkan kata kunci tetapi menghasilkan jawapan yang tepat dan relevan. Walaupun ianya kedengaran agak sukar tetapi usaha yang berterusan perlu dilakukan agar penggunaan SHOE mendapat sambutan secara menyeluruh dari pelbagai lapisan pengguna web.

Keberkesanan projek ini dapat ditingkatkan sekiranya SHOE menggabungkan teknik carian kata kunci dan semantik web. Kajian yang teliti perlu dijalankan agar pengguna web dapat menikmati faedah hasil daripada kelebihan kedua-dua teknik ini.

7.5 Kesimpulan

Pembentukan semantik web memerlukan lebih banyak kajian yang mendalam mengenainya disamping usaha-usaha yang berterusan perlu dilakukan bagi memastikan pembentukan semantik web yang sempurna menjadi satu kenyataan. Pelbagai perkara perlu diambil kira seperti isu bahasa kerana isu ini memainkan peranan yang penting di dalam memastikan penggunaan semantik web yang meluas. Ini dapat dilihat semasa Tim Berners-Lee mencipta WWW. Beliau juga mencipta HTML, yang merupakan tunjang kepada pengformatan laman web. Disebabkan bahasa ini merupakan bahasa yang tidak begitu rumit dan dapat diimplementasi dengan mudah kepada hampir kesemua jenis sistem komputer, maka idea penggunaannya yang dijadikan sebagai asas menjadi ikutan dengan pantas dan meluas serta menyebabkan penggunaan web mengalami peningkatan yang besar.

SHOE yang telah berjaya dibina oleh kumpulan PLUS merupakan satu usaha yang baik ke arah pembentukan semantik web. SHOE dianggap sebagai berjaya di

dalam mengetengahkan teknik carian semantik web dengan menghasilkan jawapan kepada carian yang dilakukan oleh pengguna dengan tepat. Walaupun begitu, banyak lagi kajian yang perlu dilakukan oleh kumpulan pembangun SHOE itu sendiri bagi memastikan SHOE dapat diterima ramai dan mampu beroperasi sejajar dengan konsep semantik web. Sebagaimana yang diketahui umum, pembentukan semantik web merupakan satu usaha yang baru di web, maka pelbagai persoalan dan masalah yang belum dapat ditangani buat masa ini sedikit sebanyak memberi kesan kepada pembangunan SHOE itu sendiri. Tetapi masa adalah pengubat kepada segalanya kerana keadaan ini dapat digambarkan seperti pada permulaan web sedikit masa dahulu, di mana indeks global dan algoritma Google yang ada sekarang tidak pernah terlintas di fikiran pada waktu itu. Oleh kerana itu, usaha ke arah pembentukan semantik web ini perlulah diteruskan dan disokong kerana ia menawarkan pelbagai kemudahan dan kelebihan berbanding daripada web yang sedia ada.

Kepentingan semantik web dapat dilihat pada pembangunan aplikasi web seperti e-perdagangan kerana ia menyediakan keupayaan carian dan yang lebih sofistikated termasuk sokongan daripada agen pintar seperti *shopbots* iaitu "*robots*" beli-belah yang mengakses laman web penjual, membuat perbandingan harga dan lain-lain. Contoh bagi penggunaan ontologi atau taksonomi untuk menyokong carian dapat dilihat pada *Yahoo Shopping* dan *amazon.com*.

Semantik web menyediakan infrastruktur yang membolehkan bukan sahaja laman web tetapi pangkalan data, servis, program, sensor dan peralatan isi rumah untuk menggunakan dan menghasilkan data di web. Agen perisian boleh menggunakan maklumat ini untuk membuat carian, menapis dan menyediakan maklumat di dalam cara yang baru sebagai pembantu kepada pengguna web.

Pembangunan ontologi merupakan teras kepada usaha mengetengahkan semantik web. Ontologi merupakan meta data dan menyediakan pengawalan perbendaharaan kata bagi istilah, di mana setiap istilah didefinisi dengan makna yang jelas dan semantik yang berupaya diproses oleh mesin. Dengan mentakrifkan perkongsian teori domain, ontologi dapat membantu manusia dan mesin untuk berkomunikasi dengan lebih berkesan. Ontologi juga mempunyai peranan yang penting di dalam membolehkan aktiviti di web dilaksanakan berdasarkan makna dan

menyediakan servis pada tahap yang baru secara kualitatif iaitu semantik web. Selain daripada projek SHOE, contoh bagi penggunaan ontologi untuk menyokong kepada aktiviti yang berdasarkan makna ini dapat dilihat pada projek European IST iaitu ON-TO-KNOWLEDGE, di mana ontologi digunakan untuk memudahkan pengaksesan kepada rangkaian intranet yang luas.

Masa depan ontologi di dalam menyelesaikan masalah dalam persekitaran dinamik khususnya carian web bergantung kepada sejauh mana kejayaan menyatukan dan mengintegrasikan serta seterusnya berkongsi dan mengguna semula sumber-sumber yang telah tersedia ada bagi membolehkan ontologi memainkan peranan dan tanggungjawabnya dengan lebih efisien dan relevan.

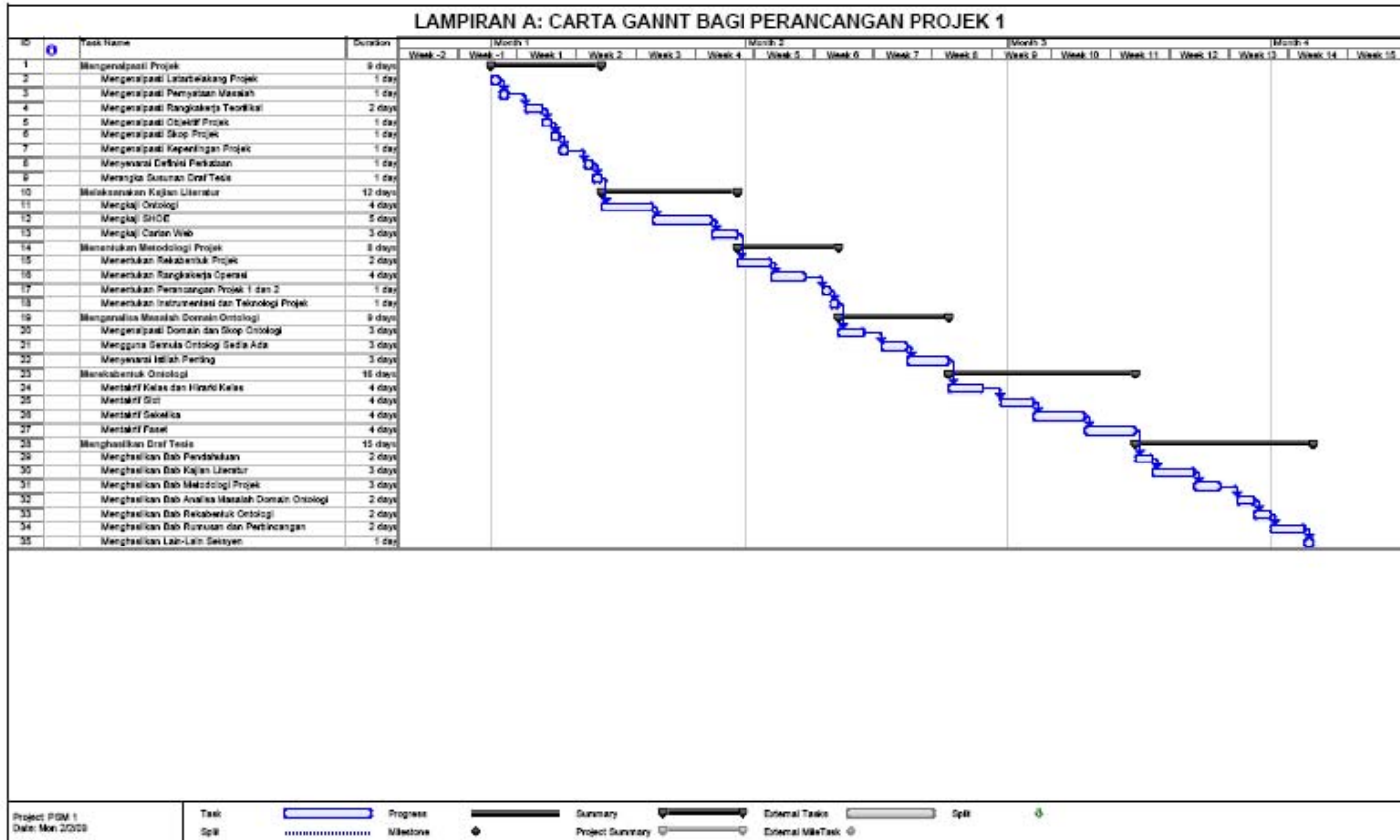
Di sepanjang pembangunan projek ini pelbagai kajian telah dilakukan ke atas semantik web, ontologi dan SHOE itu sendiri. Pelbagai cabaran yang di hadapi dalam meneruskan projek ini disebabkan bidang tersebut merupakan bidang yang baru dan kadang kala terdapat sedikit kekaburan mengenainya. Masalah ini dapat di hadapi dengan melakukan lebih banyak kajian dan pembacaan. Kesimpulannya, projek ini berjaya mencapai matlamatnya, di mana laman web berbentuk SHOE yang telah dihasilkan berjaya dikesan oleh alatan SHOE yang telah diintegrasikan dan hasilnya membolehkan pengguna dapat membuat carian terhadap laman web secara lebih bermakna dan tepat.

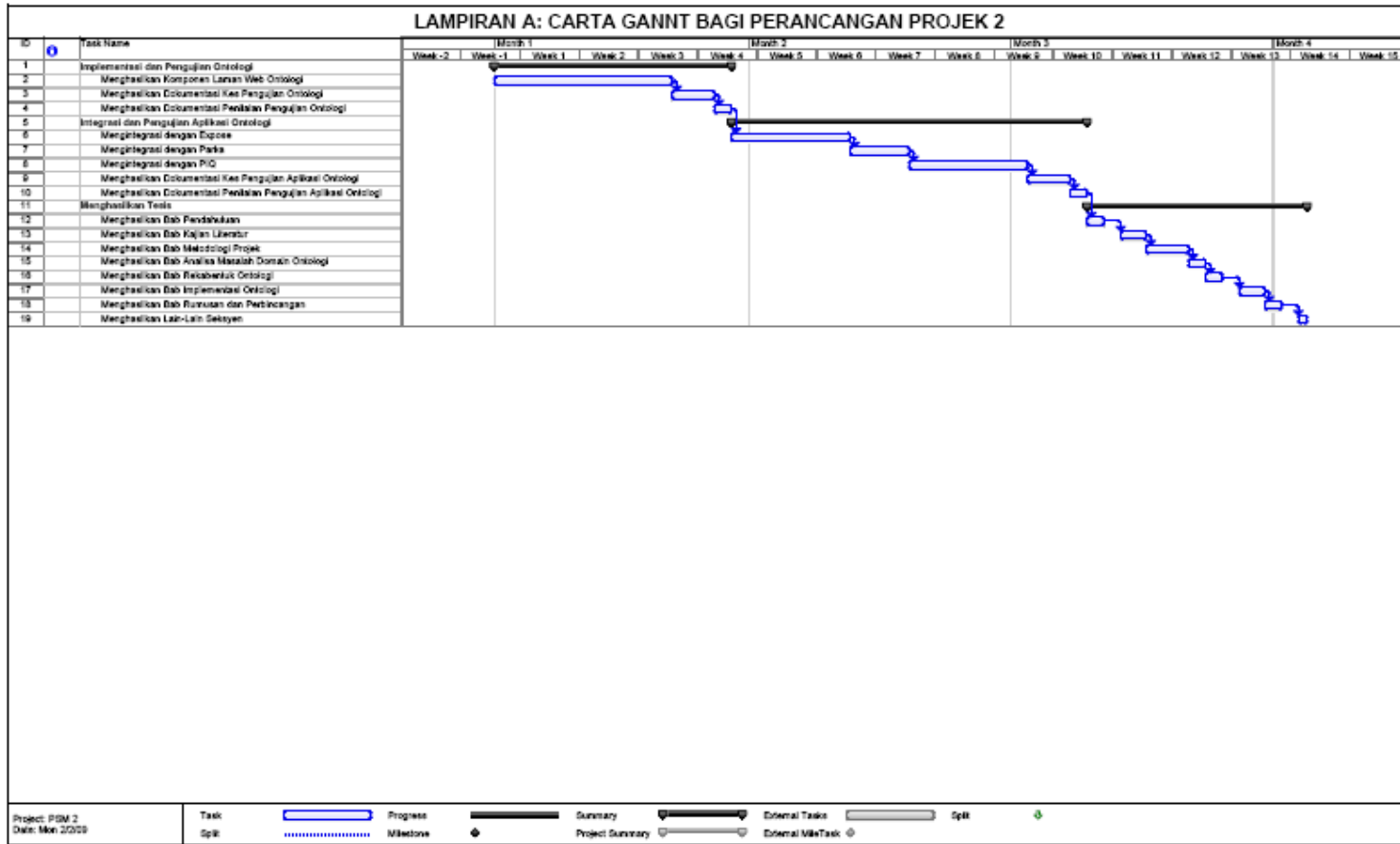
**SENARAI DOKUMEN RUJUKAN
DAN BIOLIOGRAFI**

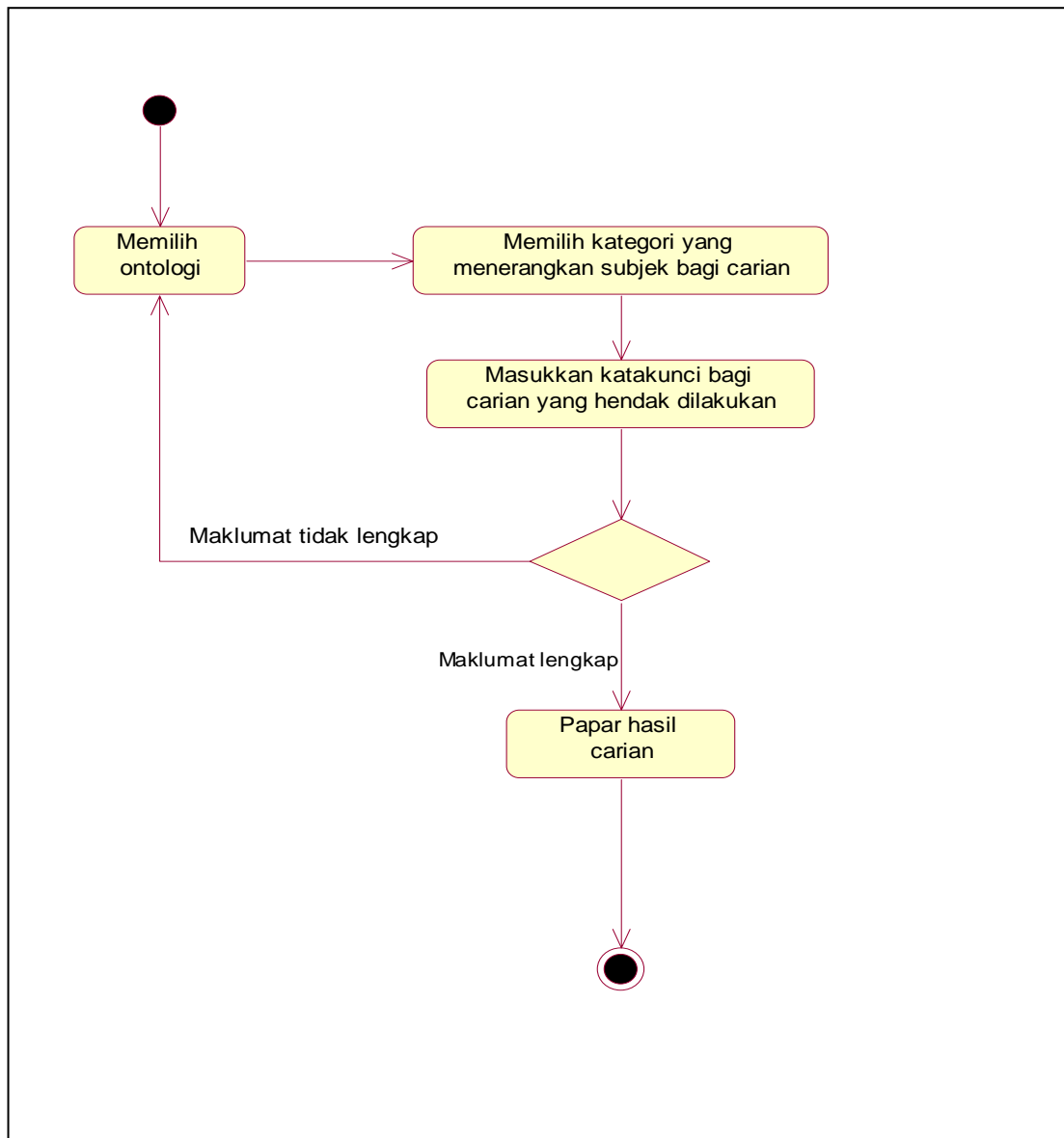
- Berners-Lee, T., Cailliau, R., Luotonen, A. and Frystyk, H. (1999). "A Secret the World-Wide-Web." *Communications of the ACM*. **37** (8). 76 – 82.
- Decker, S., Erdmann, M., Fensel, D. and Studer, R. (1999). "Ontobroker: Ontology Based Access to Distributed and Semi-Structured Information." *Proceedings of the 8th. Working Conference on Database Semantics (DS'8)*. Rotorua, New Zealand: Kluwer.
- Fensel, D., Horrocks, I., Harmelen, F. V., Decker, S., Erdmann, M. and Klein, M. (2000). "Oil in a Nutshell." *Proceedings of the 12th. European Workshop on Knowledge Acquisition, Modeling and Management (EKAW'00)*, Riviera, France: Springer-Verlag.
- Fikes, R. (1996). "Ontologies: What are They, and Where's the Research." *Proceedings of the 5th. International Conference on Principles of Knowledge Representation and Reasoning (KR'96)*. Massachusetts, USA: KR Inc.
- Gruber, T. R. (1993). "A Translation Approach to Portable Ontologies." *Knowledge Acquisition*. **5** (2). 199 – 220.
- Gruninger, M. and Fox, M. S. (1995). "Methodology for the Design and Evaluation of Ontologies." *Proceedings of the 14th. International Joint Conference on Artificial Intelligence on Basic Ontological Issues in Knowledge Sharing (IJCAI'95)*. Montreal, Canada: AAAI Press.

- Guarino, N. (1998). "Formal Ontology and Information System." Proceedings of the 1st. International Conference on Formal Ontology and Information System (FOIS'98). Trento, Canada: IOS Press.
- Heflin, J. (2001). "Towards the Semantic Web: Knowledge Representation in a Dynamic, Distributed Environment." University of Maryland: Tesis Ph.D.
- Heflin, J. and Hendler, J. (2000a). "Dynamic Ontologies on the Web." Proceedings of the 17th. National Conference on Artificial Intelligence (AAAI'00). Menlo Park, USA: AAAI/MIT Press.
- Heflin, J. and Hendler, J. (2000b). "Searching the Web with SHOE." Proceedings of the 17th. National Conference on Artificial Intelligence (AAAI'00). Menlo Park, USA: AAAI/MIT Press.
- Heflin, J. and Hendler, J. (2001). "A Portrait of the Semantic Web in Action." IEEE Intelligent Systems. **16** (2). 54 – 59.
- Heflin, J., Hendler, J. and Luke, S. (1999). "SHOE: A Knowledge Representation Language for Internet Applications." University of Maryland: Laporan teknikal.
- Hendler, J. and McGuinness, D. L. (2000). "The DARPA Agent Markup Language." IEEE Intelligent Systems. **15** (6), 67 – 73.
- Hendler, J. Berners-Lee, T. and Miller, E. (2002). "Integrating Applications on the Semantic Web." Journal of IEE Japan. **122** (10). 676 – 680.
- Lassila, O. (1998). "Web Metadata: A Matter of Semantics." IEEE Internet Computing. **2** (4). 30 – 37.
- Luke, S., Spector, L., and Rager, D. (1996). "Ontology-Based Knowledge Discovery on the World-Wide Web." Proceedings of the 13th. National Conference on Artificial Intelligence (AAAI'96). Oregon, USA: AAAI Press.

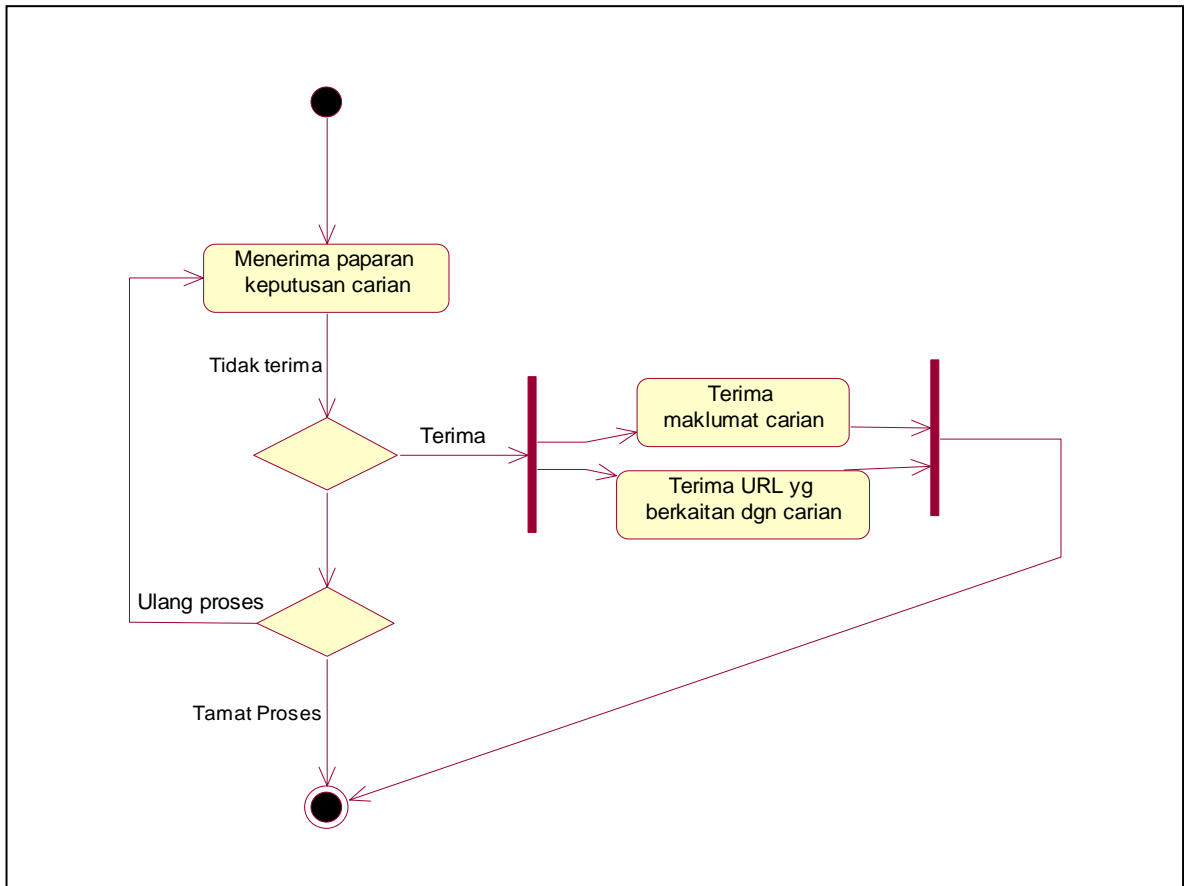
- Noy, N. F. and McGuinness, D. L. (2001). "Ontology Development 101: A Guide to Creating your First Ontology." Stanford University: Laporan teknikal.
- Noy, N. F., Sintek, M., Decker, S., Crubezy, M., Fergerson, R. W. and Musen, M. A. (2001). "Creating Semantic Web Contents with Protege-2000." IEEE Intelligent Systems. **16** (2). 60 – 71.
- Studer, R., Eriksson, H., Gennari, J., Tu, S., Fensel, D. and Musen, M. (1996). "Ontologies and the Configuration of Problem-Solving Methods." Proceedings of 10th. Knowledge Acquisition for Knowledge-Based Systems Workshop (KAW'96). Banff, Canada: Springer-Verlag.
- Uschold, M. and Gruninger, M. (1996). "Ontologies: Principles, Method and Applications." Knowledge Engineering Review. **11** (2). 93 – 137.



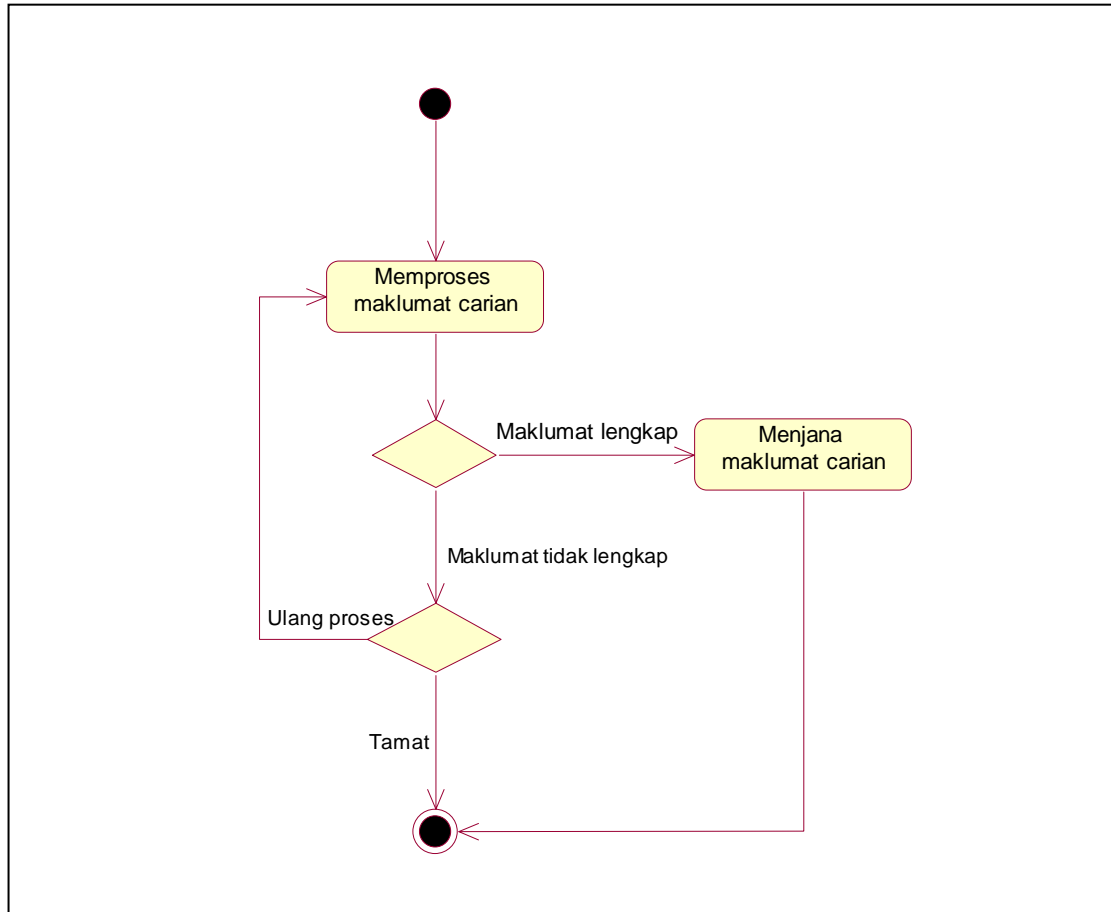




Lampiran C (a) : Gambarajah aktiviti bagi membuat carian web

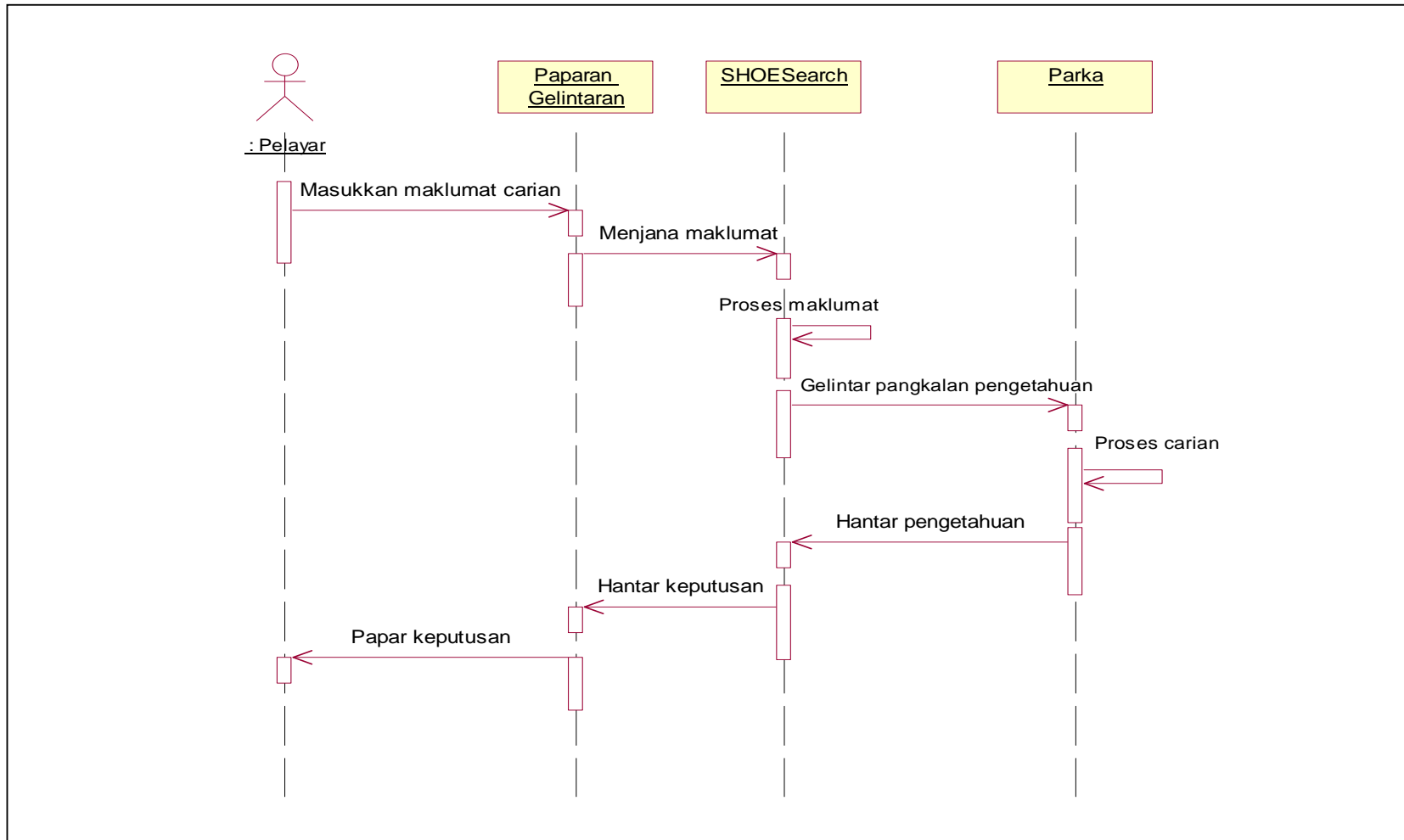


Lampiran C (ii) : Gambarajah aktiviti bagi menerima paparan keputusan carian



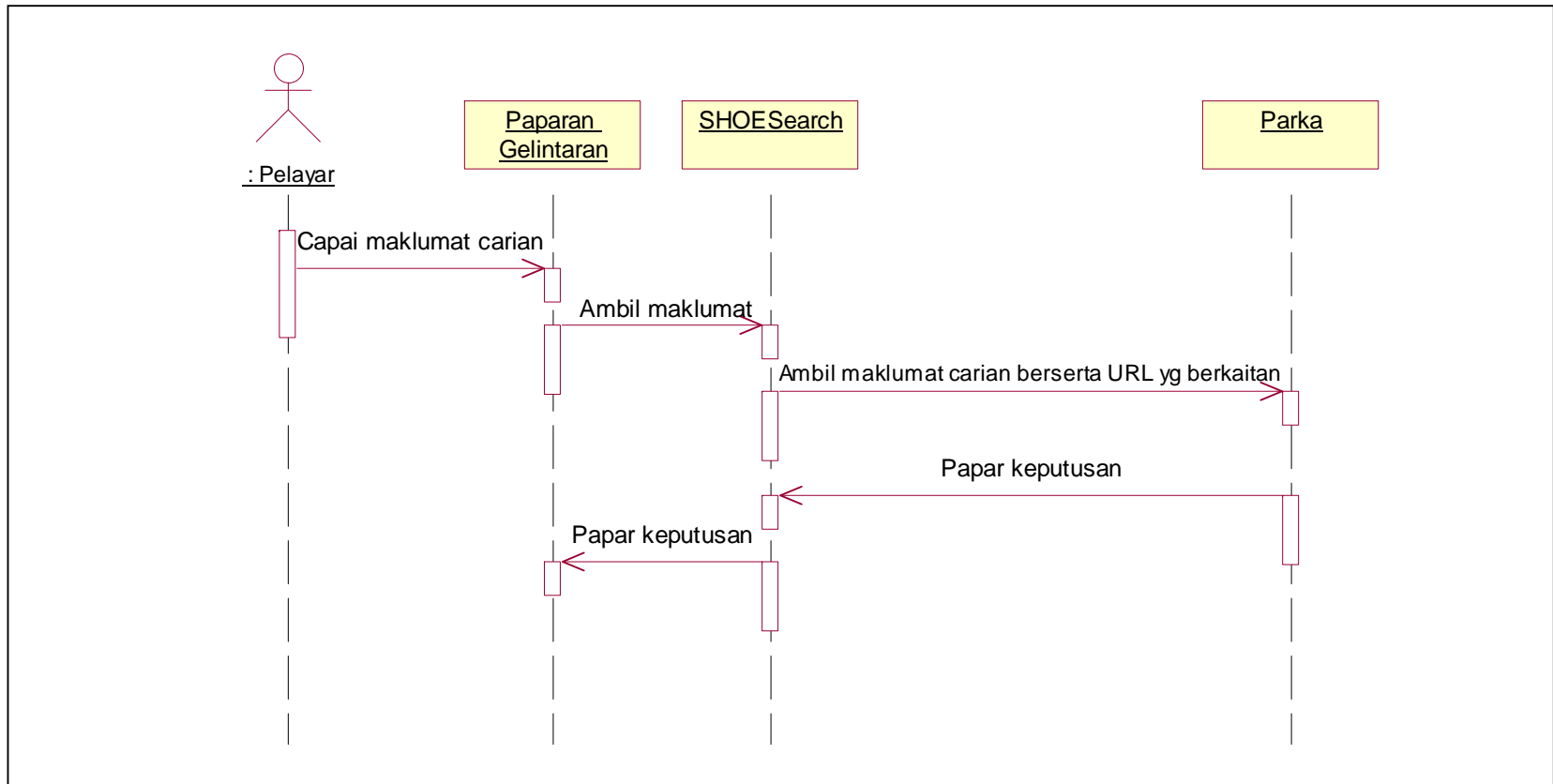
Lampiran C (iii) : Gambarajah aktiviti bagi menjana maklumat carian

LAMPIRAN D
Gambarajah Jujukan



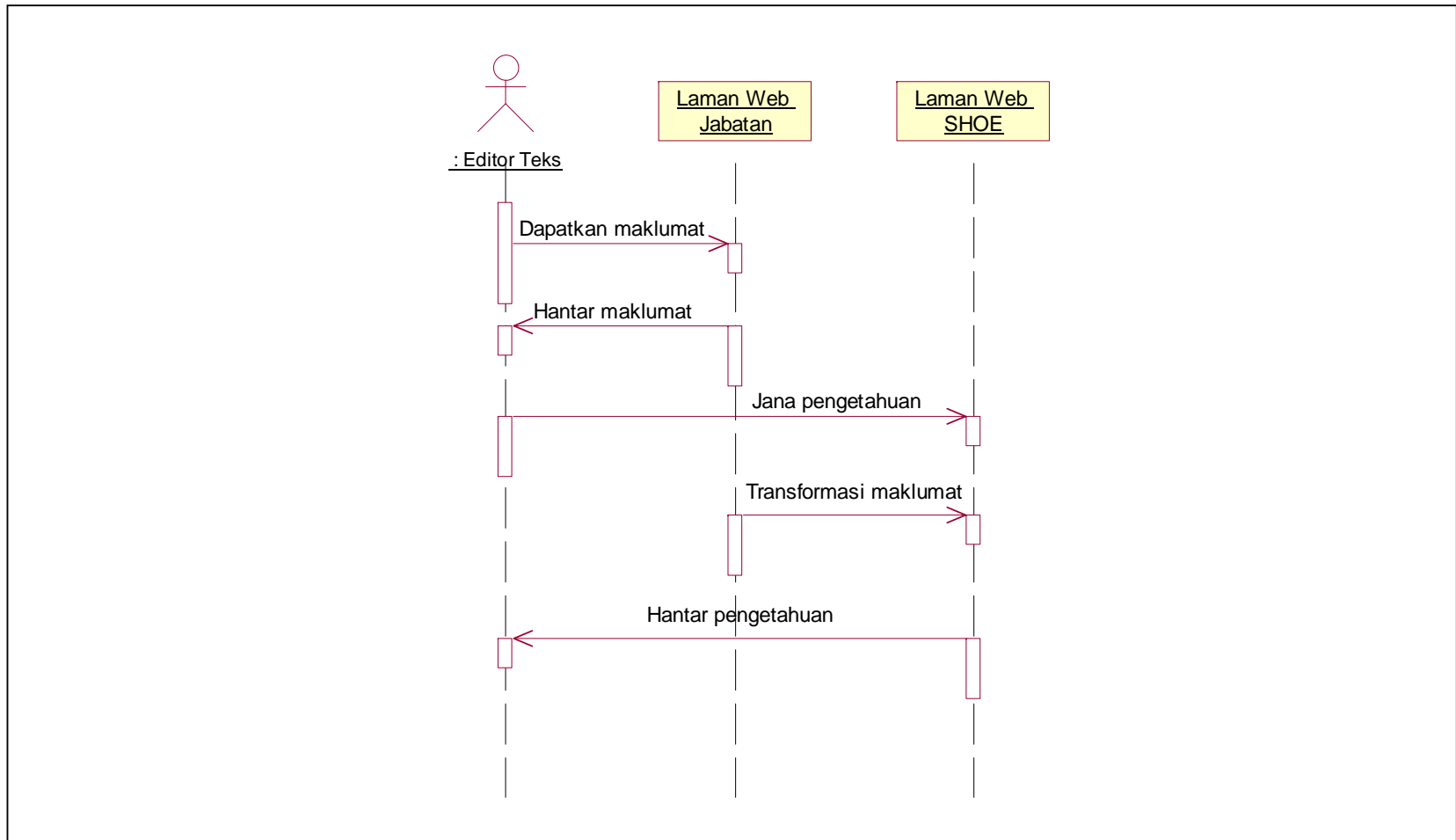
Lampiran D (i) : Gambarajah jujukan bagi memasukkan maklumat carian

LAMPIRAN D
Gambarajah Jujukan



Lampiran D (ii) : Gambarajah jujukan bagi menerima paparan keputusan carian

LAMPIRAN D
Gambarajah Jujukan



Lampiran D (iii) : Gambarajah jujukan bagi mentransformasi maklumat