

**School of Information Systems**

**Curtin Business School**

**An Employer Demand Intelligence Framework**

**Chamonix Smalberger**

**This thesis is presented for the Degree of**

**Doctor of Philosophy**

**of**

**Curtin University**

**October 2013**

## **Declaration**

To the best of my knowledge and belief, this thesis contains no material previously published by any other person except where due acknowledgement has been made.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

Chamonix Smalberger

October 2013

# Abstract

Employer demand intelligence is crucial to ensure accurate and reliable education, workforce and immigration related decisions are made. To date, current methods have been manually intensive and expensive, providing insufficient scope of the information that is required to address such important economic implications. This research has developed a state-of-the-art Employer Demand Intelligence Framework (EDIF) based on semantic technology that addressed the research goals to: 1) develop an underlying knowledge representation for employer demand intelligence; 2) populate instances that are occupation specific and linked to a wide range of employer demand intelligence variables, ensuring broad coverage of the domain; and 3) develop a semi-automated tool to gather, analyse and report on detailed employer demand intelligence concepts on a continuous basis. The EDIF can be found at [www.employerdemandintelligence.org](http://www.employerdemandintelligence.org).

# Acknowledgments

I would like to express the deepest appreciation to my main supervisor, Dr. Ponnice Clark, without whose kind guidance, persistent help and broad knowledge of the field, my research and writing of this thesis would not have been possible. To my associate supervisor, Dr. Boris Villazon-Terrazas, great thank is due for his willingness to join my thesis committee at such a late stage — albeit never having met me and being on the other side of the world — and spending hours and hours providing invaluable feedback on my thesis drafts. I also thank Professor Elizabeth Chang for her expert advice at the start of my PhD undertaking, Professor Tharam Dillon for the wise council and overall thesis direction that he provided, and my thesis chair, Dr. Omar Hussain, for his assistance in the administrative processes involved and the pre-candidacy feedback he provided me.

There were also those who assisted me in other parts and stages of my research. I would like to acknowledge Dr. John Fielder for providing initial proof reading support; Dr. Christina Houen for her editing services in the final part of my PhD tenure; Annemi Kruger for the difficult task of organising appointments for me with my supervisors amidst their busy schedules; my cousin, Johan van Rensburg, for his help in the initial data collection phase; and Zia Ur-Rehman for his research assistance during the later stages of my developmental work.

I am grateful to Professor Jeanette Hackett AM — Curtin University's Vice-Chancellor — for her invaluable advice and compassion on my path of starting my PhD degree, and for her mentorship through my many committee memberships and other higher education governance and leadership roles I held; Professor Graeme Wright (Curtin's Deputy Vice-Chancellor for Research) and Professor Kate Wright (Curtin's Associate Deputy Vice-Chancellor, Research Training) for their kind consideration of the research students' welfare and for their willingness to always consider any alternatives I raised with them on behalf of the university's postgraduate community; and to the Australian Government for the scholarships I received, allowing me to pursue my PhD full-time.

Further appreciation goes to the folk I met through the Curtin University Postgraduate Student Association— Andrew Cameron, Marion Farr-Devé and Alinta Kirke, who provided healthy interruptions during my study time so we could represent and assist other postgraduate

students to have a smooth tenure at the university; the DEBII students for supporting me through many university elections and whose friendships enriched my life; my friend, Dr. David McMeekin, for his encouragement and laughter shared over many coffees; my PhD confidant, David E Forbes, with whom I had countless conversations ranging from thesis formatting, student and federal politics, business startup fundamentals and the colour of socks to heart stories of health, family losses and family gains. We have proven that age makes no difference in friendship.

I never thought this journey would be so hard. It has been a time of extreme character testing, and has taught me to trust my instinct, and follow my heart. When you are working on the edge, you need a firm base to hold on to; and that is precisely what my family provided me. They also instilled in me the interest to know more, explore, and contribute to the betterment of mankind; they taught me that with God by my side, anything is possible! I love you all very dearly — thank you for your unconditional love.

Finally — for his unwavering believe, encouragement and trust in me — my deepest appreciation to my soul mate and husband. You are my God-sent angel.

I dedicate this thesis to my son — Monaco — who at the time of this writing, is due to be born in two weeks' time. May you always be hungry for the unknown, and never forget Michelangelo's wise words:

*“The greater danger for most of us lies not in setting our aim too high, and falling short; but in setting our aim too low, and achieving our mark.”*

**Soli Deo Gloria!**

# Publications Arising from this Research

(At time of submission)

Smalberger, C. (2011). *Meeting Employers' and Students' Expectations through the use of Employer Demand Ontology in Curriculum Development*. Paper presented at the 5th IEEE International Conference on E-Learning in Industrial Electronics, Melbourne.

Smalberger, C., Wongthongtham, P., Villazon-Terrazas, B., & Forbes, D. E. (2013 (under review)). Ontology-based Employer Demand Management. *Management Information Systems Quarterly*.

# Table of Contents

<b>Chapter 1 - Introduction.....</b>	<b>23</b>
1.1 Introduction.....	23
1.2 A Shortage of Skills .....	24
1.3 The Importance of Employer Demand Intelligence.....	25
1.3.1 Labour Market Forecasting .....	27
1.3.2 Workforce Development Planning .....	28
1.3.3 Immigration .....	29
1.3.4 Education and Training.....	29
1.4 Motivation of Study .....	30
1.4.1 No Agreed Standards or Knowledge Sharing Exist .....	30
1.4.2 A Lack of Data Depth and Breadth .....	32
1.4.3 Time Consuming and Expensive Approaches .....	33
1.5 Research Objectives and Scope of the Problem.....	35
1.6 Advantages of the Research .....	36
1.7 Thesis Structure .....	37
1.8 Conclusion.....	39
<b>Chapter 2 - Literature Review.....</b>	<b>41</b>
2.1 Introduction.....	41
2.2 Employer Demand Identification Approaches.....	42
2.2.1 Survey-Based Approaches .....	44
2.2.2 Online-Based Approaches .....	48
2.2.3 Manual Job Advertisement Analysis .....	51
2.2.4 Structured Approaches.....	54
2.3 Critical Evaluation of Approaches: an Integrated View .....	63
2.3.1 Knowledge Sharing and Different Approach Structurings .....	64
2.3.2 Data Depth and Breadth.....	66
2.3.3 Data Out of Date.....	66
2.4 Conclusion.....	68
<b>Chapter 3 - Problem Definition .....</b>	<b>70</b>
3.1 Introduction .....	70
3.2 Key Concepts and Definitions .....	71
3.2.1 Employer Demand Domain Concepts.....	71
3.2.2 Methodological Concepts.....	73
3.3 Problems Facing Employer Demand Intelligence Domain.....	76
3.3.1 No Exhaustive Employer Demand Intelligence Framework or Standards Exist .....	76
3.3.2 Limited Employer Demand Intelligence Span .....	78
3.3.3 Manual Approaches are Infrequent or not Frequent Enough.....	79
3.4 Underlying Research Issues .....	80
3.4.1 Research Issue 1: Underlying Knowledge Representation .....	80
3.4.2 Research Issue 2: Occupation Specific Data and its Querying.....	81
3.4.3 Research Issue 3: Expensive and Time Intensive Approaches .....	82
3.5 Summary of Thesis Goal .....	85

3.6	Key Research Questions.....	86
3.7	Research Methodology.....	87
	3.7.1 Overview of Research Methods.....	87
	3.7.2 Choice of Science- and Engineering-based Research Method.....	87
3.8	Conclusion.....	89
<b>Chapter 4 - Solution Proposal.....</b>		<b>91</b>
4.1	Introduction.....	91
4.2	Solution Overview.....	92
	4.2.1 Tsunami of Data.....	92
	4.2.2 Ontology Definition and Types.....	94
	4.2.3 Ontology Characteristics.....	97
4.3	Conceptual Framework.....	98
	4.3.1 Data Input.....	99
	4.3.2 Online Resources.....	99
	4.3.3 Employer Demand Intelligence Framework.....	99
	4.3.4 Data Output.....	100
4.4	Detailed Solution Requirements.....	103
	4.4.1 Requirement 1: Employer Demand Ontology.....	103
	4.4.2 Requirement 2: Employer Demand Ontology Instance Population.....	105
	4.4.3 Requirement 3: Semi-Automatic Employer Demand Intelligence Tool.....	108
4.5	Conclusion.....	112
<b>Chapter 5 - Employer Demand Ontology Engineering.....</b>		<b>114</b>
5.1	Introduction.....	114
5.2	Employer Demand Ontology Methodology.....	115
	5.2.1 Methodology Fundamentals.....	115
	5.2.2 Overview of Existing Ontology Methodologies.....	115
	5.2.3 Employer Demand Methodology Used- NeOn.....	117
5.3	Employer Demand Ontology Modelling and Design.....	132
	5.3.1 Employer Demand Ontology Notation.....	132
	5.3.2 Components of Employer Demand Ontology.....	136
	5.3.3 Ontology Design Principles.....	161
5.4	Employer Demand Ontology Conceptual Implementation.....	162
	5.4.1 Employer Demand Ontology: Tier One.....	164
	5.4.2 Employer Demand Ontology: Tier Two.....	188
5.5	Conclusion.....	189
<b>Chapter 6 - Employer Demand Ontology Instance Population.....</b>		<b>191</b>
6.1	Introduction.....	191
6.2	Instantiation Process.....	191
	6.2.1 Employer Demand Ontology: Tier One Instance Population.....	196
	6.2.2 Employer Demand Ontology: Tier Two Instance Population.....	214
6.3	Conclusion.....	217
<b>Chapter 7 - Semi-Automatic Employer Demand Intelligence Tool.....</b>		<b>218</b>
7.1	Introduction.....	218
7.2	Overview of the Employer Demand Intelligence Tool.....	218
7.3	Detailed Processes of the Employer Demand Intelligence Tool.....	223
	7.3.1 Creation of Corpus.....	223

7.3.2	Loading the Corpus and EDO into EDIT .....	226
7.3.3	Creating an EDIT Pipeline .....	228
7.4	Results and Discussion .....	238
7.4.1	A Case Study .....	238
7.4.2	SPARQL Query Outcomes .....	245
7.4.3	The Pros and Cons of Using GATE .....	248
7.4.4	Metrics and Development Time .....	249
7.4.5	Limitations of the EDIT .....	250
7.5	Conclusion.....	251
<b>Chapter 8 - Employer Demand Intelligence Framework Evaluation .....</b>		<b>253</b>
8.1	Introduction .....	253
8.2	Verification and Validation Fundamentals .....	253
8.3	The Employer Demand Ontology Logical Consistency Verification.....	255
8.4	The Employer Demand Ontology Validation .....	257
8.4.1	Use case 1: Government Official Developing Policy and Funding Protocols.....	258
8.4.2	Use case 2: Prospective Student Choosing Course of Study .....	266
8.4.3	Use case 3: Human Resource Manager at a Company.....	273
8.4.4	Use case 4: Curriculum and Career Developer at a Tertiary Institution .....	278
8.4.5	An Integrated View of Use Case Needs .....	286
8.5	Conclusion.....	288
<b>Chapter 9 - Conclusion and Future Work.....</b>		<b>290</b>
9.1	Introduction .....	290
9.2	Recapitulation of the Thesis .....	290
9.3	Significance of the Research .....	294
9.3.1	Scientific Significance .....	294
9.3.2	Social Significance.....	295
9.4	Future Work.....	296
9.4.1	Employer Demand Intelligence Framework Future Work.....	297
9.4.2	Employer Demand Ontology Future Work.....	298
9.4.3	Employer Demand Intelligence Tool Future Work.....	300
9.5	Conclusion.....	300
<b>References</b>		<b>302</b>
<b>Appendices</b>		<b>313</b>

# List of Figures

Figure 1	Simplified representation of ‘old’ and ‘new’ occupational structures (Hayton & Cheyne, 1988). .....	25
Figure 2	An overview of the various uses of employer demand intelligence. ....	27
Figure 3	Orientation of employer demand identification approaches. ....	43
Figure 4	Schematic view of Employer Demand Intelligence. ....	71
Figure 5	Schematic, Semantic and Syntactic Heterogeneity. ....	75
Figure 6	Science and engineering based research approach. ....	89
Figure 7	Information on the web covers numerous and varying employer demand related domains. ....	93
Figure 8	A small part of a company’s workers and manager, and the relationships they have with each other. ....	96
Figure 9	The three layers of ontology generalisation (Villazón-Terrazas et al., 2011, p. 13617). 97	
Figure 10	Employer Demand Intelligence Framework conceptual framework. ....	101
Figure 11	The underlying knowledge representation for employer demand intelligence. ....	104
Figure 12	Schematic overview of employer demand intelligence representation. ....	107
Figure 13	Some terminological relationships in methodologies (Asuncion Gomez-Perez, Fernandez-Lopez, & Corcho, 2003). ....	115

---

Figure 14 A graphical example of one of the methodologies- the OTK Methodology (Lau & Sure, 2002). 117

Figure 15 NeOn scenarios to build ontology networks and ontologies (M C Suarez-Figueroa, Gomez-Perez, & fernandez-Lopez, 2012). ..... 118

Figure 16 The EDO metrics for the one month trial..... 120

Figure 17 The three resource recycling options..... 124

Figure 18 *protont:Group* and *protont:Organization* class hierarchy (Terziev et al., 2005)..... 125

Figure 19 Negru and Lohmann’s ontology notation (2013)..... 133

Figure 20 The System Matrix Notation (SMN) Ontology (SMN Ontology, 2008). ..... 133

Figure 21 Employer demand ontology class notation. .... 134

Figure 22 Employer demand ontology instance notation..... 135

Figure 23 Employer demand ontology given property notation. .... 135

Figure 24 Employer demand ontology inferred property notation. .... 135

Figure 25 Employer demand ontology generalisation relation. .... 136

Figure 26 Employer demand ontology composition relation. .... 136

Figure 27 Employer demand ontology class hierarchy of the position requirements domain. .... 138

Figure 28 A schematic diagram that shows the effect of using a covering axiom to cover class A with classes B and C (Horridge, Drummond, Jupp, Moulton, & Stevens, 2009)..... 141

---

Figure 29	The effects of disjoint, decomposition and partition classes (Wongthongtham, 2006).	141
Figure 30	The intersection of classes <i>HighSalary</i> , <i>LongPeriodAdvertised</i> and <i>RemoteLocation</i> .	143
Figure 31	The union of classes <i>PositionBasis</i> and <i>WorkHours</i> ( $PositionBasis \cup WorkHours$ ) to form the class <i>EmploymentConditions</i> .	144
Figure 32	<i>InternetAdvertisement</i> conditions indicating the necessary conditions.	144
Figure 33	Illustration of Necessary and Sufficient Conditions based on the class <i>JobAdvertisement</i> .	146
Figure 34	Illustration of an inverse property: <i>hasEmployer</i> .	150
Figure 35	Illustration of a functional property: <i>hasCondition</i> .	151
Figure 36	Illustration of an inverse functional property: <i>isConditionOf</i> .	152
Figure 37	Illustration of a transitive property: <i>hasEmployer</i> .	153
Figure 38	Illustration of a symmetric property: <i>hasCoWorker</i> .	154
Figure 39	Illustration of an antisymmetric property: <i>hasBenefit</i> .	155
Figure 40	Illustration of a reflexive property: <i>hasKnowledgeOf</i> .	155
Figure 41	Illustration of an irreflexive property: <i>hasEmployer</i> .	156
Figure 42	Illustration of the Is-A relationship for a class and some of its instances.	160
Figure 43	Illustration of a composition relationship based on instances.	161
Figure 44	Employer demand ontology tiered framework.	164

---

Figure 45	Association relationship diagram of <i>Vacancy</i> and <i>JobAdvertisement</i> . .....	165
Figure 46	Typical job advertisement in a newspaper. ....	166
Figure 47	Advertisement media specialisation relationship hierarchy.....	167
Figure 48	<i>WordOfMouthAdvertisement</i> composition relationship diagram. ....	168
Figure 49	Job Advertisement class composition.....	169
Figure 50	Position Requirements Generalisation- Specialisation relationships. ....	171
Figure 51	Trichet and Leclere’s competency reference system (2003). ....	173
Figure 52	SEEMP (Villazón-Terrazas, et al., 2011) relationships for Competence.....	174
Figure 53	An example of an experience requirement in a job advertisement. ....	175
Figure 54	An example of an attribute requirement ‘compassionate’ in a job advertisement.....	176
Figure 55	The EDO indicators of a potential skill shortage based on the DEEWR definition of the phenomenon.....	178
Figure 56	ADGC Remoteness Classification (2001). ....	182
Figure 57	Location Type potential skill shortage subclasses.....	183
Figure 58	Job offer equilibrium showing example of employer’s and employee’s offers for a job vacancy.....	185
Figure 59	EmploymentConditions Class hierarchy.....	185
Figure 60	OntoGraf representation of the PositionOccupationType class’ main relationships.....	188

---

Figure 61	Major occupational grouping from ANZSCO spreadsheet. ....	188
Figure 62	Protégé modelling notations (Forbes, Wongthongtham, Singh, & Thompson, 2013 (in press)). ....	192
Figure 63	Setting up a job alert on the Seek.com.au website. ....	194
Figure 64	Annotation included for the instance <i>Adventurous</i> . ....	195
Figure 65	OntoGraf presentation of instances of the EDO class <i>JobBoard</i> . ....	197
Figure 66	Example where full title of occupation is not necessarily listed in the heading of the job advertisement. ....	201
Figure 67	OntoGraf presentation of the <i>OtherRequirements</i> class's instances. ....	202
Figure 68	Drop down salary range search criteria on Seek.com.au. ....	205
Figure 69	Salary is explicitly indicated within the body of the job advertisement. ....	206
Figure 70	Map of Australia illustrating the 2006 Remoteness Structure (Australian Bureau of Statistics, 2011). ....	208
Figure 71	OntoGraf presentation of the subclass <i>MajorCityOfAustralia</i> 's instances. ....	209
Figure 72	Snapshot for the class <i>MajorCitiesOfAustralia</i> data as it appears in the ASGC spreadsheet prior to being recycled to the EDO. ....	209
Figure 73	Some of the instances of the class <i>EmployeeBenefits</i> and their annotations. ....	210
Figure 74	A nursing advertisement where the employee benefit instances have been highlighted. ....	211
Figure 75	OntoGraf presentation of instances of the <i>EmploymentCondition</i> class. ....	212

---

Figure 76	Annotations captured for the instance <i>ReliefBasis</i> .....	212
Figure 77	A job advertisement indicating both an agent as well as an Organization where the employment will take place.....	213
Figure 78	Snapshot of the ANZSCO spreadsheet section for Midwifery and Nursing Professionals.....	214
Figure 79	OntoGraf presentation of the EDO <i>MidwifeAndNursingProfessionals</i> instances...	215
Figure 80	Seek.com.au nursing categories.....	216
Figure 81	The job advertisement with the combined occupational naming code (unique identification number) of RNMS18, still capturing the ANZSCO occupational title correctly — <i>RegisteredNurseMedicalPractice</i> in this case.....	216
Figure 82	Main processes of the Employer Demand Intelligence Tool. ....	222
Figure 83	Employer Demand Intelligence Tool flowchart.....	224
Figure 84	A pre-processed job advertisement to remove unnecessary HTML elements as links and images.....	225
Figure 85	The EDO loaded as a language resource into the EDIT (GATE application). ....	227
Figure 86	Processing resources in the EDIT pipeline (GATE application).....	228
Figure 87	Processing resources for OntoRoot Gazetteer in the GATE application.....	232
Figure 88	JAPE grammar extracting salary amounts from job advertisement text. ....	233
Figure 89	Ontology aware JAPE grammar for advertisement date. ....	234

Figure 90	Text appearing in job advertisements matched with the concept in the EDO by the EDIT through utilization of the Flexible Gazetteer Processing Resource. ....	235
Figure 91	JAPE grammar used to create instances for job advertisements. ....	236
Figure 92	Instances and property values added into EDO by EDIT (view is from the GATE application into EDO). ....	237
Figure 93	An original online job advertisement from <a href="http://www.seek.com.au">www.seek.com.au</a> . ....	239
Figure 94	The 'cleaned' job advertisement. ....	240
Figure 95	A collection of job advertisements loaded into EDIT's corpus. ....	241
Figure 96	Putting the job advertisement through the tokenization process. ....	242
Figure 97	The job advertisement is being sentence split. ....	242
Figure 98	POS and morphological checking of the job advertisement. ....	243
Figure 99	The flexible gazetteer recognises the concepts in the text. ....	244
Figure 100	Population of extracted information through the process into EDO. ....	245
Figure 101	A query to retrieve all the job advertisements and their locations that were advertised on 18 March 2013. ....	246
Figure 102	A query to retrieve all the job advertisements that were advertised in the various locations in Western Australia. ....	247
Figure 103	The SPARQL query to establish the number of job advertisements that were advertised with <i>OtherRequirements</i> . ....	248

---

Figure 104 Six dimensions that can be identified in ontology quality (Poveda-Villalon, et al., 2012).	254
Figure 105 Some of the reasoners used in Protégé.	255
Figure 106 Some of the Protégé reasoners' logs while classifying the EDO.	256
Figure 107 DL Query and result showing the number of <i>JobAdvertisements</i> advertised for the month in Western Australia's metropolitan area.	260
Figure 108 Job advertisement <i>MW1</i> .	260
Figure 109 Job advertisement <i>MW2</i> .	261
Figure 110 Job advertisement <i>RNMS5</i> .	261
Figure 111 Job advertisement <i>RNAC5</i> .	262
Figure 112 Job advertisement <i>RNMS2</i> .	262
Figure 113 DL Query and result showing the number of <i>JobAdvertisements</i> advertised for the month in Western Australia's regional areas.	263
Figure 114 DL Query and result showing the number of <i>PositionRequirements</i> advertised for the occupation class <i>Midwives</i> .	264
Figure 115 DL Query and result showing the number of <i>PositionRequirements</i> advertised for the occupation class <i>NurseManagers</i> .	264
Figure 116 DL Query and result showing the number of <i>PositionRequirements</i> advertised for the occupation class <i>NurseEducatorsAndResearchers</i> .	265

---

Figure 117 DL Query and result showing the number of <i>PositionRequirements</i> advertised for the occupation class <i>RegisteredNurses</i> .....	265
Figure 118 DL Query and result showing the <i>JobAdvertisement</i> instances that were advertised for the occupation class <i>Midwives</i> where a type of <i>AttributeRequirement</i> was specified.....	268
Figure 119 Job advertisement <i>MW9</i> has attribute requirement <i>CustomerServiceCommitment</i> . ....	268
Figure 120 Job advertisement <i>MW6</i> has attribute requirements <i>Caring</i> and <i>ConfidentialityCommitment</i> .....	269
Figure 121 Job advertisement <i>MW3</i> has attribute requirement <i>Enthusiastic</i> . ....	270
Figure 122 Detailed view of <i>MW6</i> — one of the <i>JobAdvertisement</i> instances retrieved through the query specified in the previous figure.....	271
Figure 123 DL Query and result showing the <i>JobAdvertisement</i> instances that were advertised for the occupation class <i>NurseManagers</i> where a type of <i>AttributeRequirement</i> was specified.....	272
Figure 124 DL Query and result showing the <i>JobAdvertisement</i> instances that were advertised for all nursing occupation classes where a type of <i>AttributeRequirement</i> was specified. ....	274
Figure 125 Job advertisement <i>MW1</i> with location <i>Perth</i> , and benefits on offer <i>SalaryPackaging</i> , <i>SponsorshipVisaAvailable</i> and <i>ProfessionalDevelopment</i> . ....	275
Figure 126 Job advertisement <i>MW11</i> in Mount Lawley (a suburb of <i>Perth</i> ), and benefits offered <i>OrientationProvided</i> , <i>ProfessionalDevelopment</i> , <i>FreeCarParking</i> , <i>SalaryPackaging</i> , <i>CloseToPublicTransport</i> and <i>FlexibleWorkPractices</i> . ....	275

---

Figure 127 Job advertisement *RNMS7* in *Perth*, offering benefits *WACallCentreTeam*, *UniformProvided*, *WeeklyPayCycles* and *ShiftsCloseToHome*..... 276

Figure 128 Job advertisement *RNMS2* in *Perth*, offering benefits *SalaryPackaging*, *ProfessionalDevelopment* and *SponsorshipVisaAvailable*. .... 276

Figure 129 Job advertisement *RNMS3* in *Perth*, with benefit offered *RemoteWorkAllowance*. .... 277

Figure 130 DL Query and result showing the *JobAdvertisement* instances that specified the employee benefit *SponsorshipVisaAvailable*..... 278

Figure 131 DL Query and result showing some of the *JobAdvertisement* instances that specified employee skills as a requirement. .... 280

Figure 132 Job advertisement *RNAC20* requires *GoodCommunicationSkills* and *TeamLeadingSkills*..... 281

Figure 133 Job advertisement *RNMS8* requires *TeamPlayerSkills*, *ComputerSkills* and *InterpersonalSkills*..... 281

Figure 134 Job advertisement *RNMS9* with skills *TeamPlayerSkills*, *RecallAndTestResultMaintenanceSkills*, *WoundCareSkills*, *WorkUnsupervisedSkills*, *GoodCommunicationSkills* and *ProcedureAssistanceSkills*. .... 282

Figure 135 Job advertisement *RNCE9* with skill requirement *GoodCommunicationSkills*..... 283

Figure 136 Job advertisement *RNDM1* requires *WorkUnsupervisedSkills* and *TeamPlayerSkills*. .... 284

Figure 137 DL Query and result showing the number of job advertisements the employer *BethesdaHospital* advertised during January 2011. .... 285

Figure 138 DL Query and result showing the number of job advertisements the employer *StJohnOfGodSubiaco* advertised during January 2011. .... 285

Figure 139 DL Query and result showing the number of job advertisements the employer *MercyHospital* advertised during January 2011. .... 286

# List of Tables

Table 1	Number of approaches related to employer demand identification discussed. ....	44
Table 2	A summary of structured eRecruitment and eHumanResources approaches. ....	62
Table 3	Estimated man hours required for an example short employer demand survey.....	85
Table 4	Approximate time taken to manually develop the EDO. ....	120
Table 5	A table developed by this research to indicate the classes of the PROTON ontology (Terziev et al., 2005) that was reused or reengineered in the EDO. ....	129
Table 6	A table developed by this research to indicate the classes of the SEEMP ontology (Villazón-Terrazas et al., 2011) that were reused or reengineered in the EDO.....	131
Table 7	General statistics from the one month trial during January 2011. ....	195
Table 8	Number of instances populated for the <i>PositionRequirements</i> subclasses.....	202
Table 9	Number of instances populated for the <i>LocationType</i> subclasses.....	207
Table 10	Midwifery and Nursing Professionals mapping between Seek and ANZSCO. ....	215
Table 11	The approximate time taken to develop EDIT. ....	249
Table 12	A variety of metrics relevant to the EDIT. ....	250
Table 13	Attribute requirements for <i>Midwives</i> versus <i>NurseManagers</i> . ....	272

# List of Acronyms

ABS — Australian Bureau of Statistics

ANZSCO — Australian and New Zealand Standard Classification of Occupations

ASGC — Australian Standard Geographic Classification

DEEWR — Department of Education, Employment and Workplace Relations

DOL — Department of Labour

EDD — Employer Demand Dataset

EDI — Employer Demand Intelligence

EDIF — Employer Demand Intelligence Framework

EDIT — Employer Demand Intelligence Tool

EDO — Employer Demand Ontology

SEEMP — Single European Employment Market Place

# Chapter 1 - Introduction

'If we could first know where we are, and whither we are tending, we could then better judge what to do, and how to do it'.

Abraham Lincoln, House Divided Speech, Springfield, Illinois

June 16, 1858.

## **1.1 Introduction**

Employer demand intelligence are being used across a variety of domains for important purposes such as determining labour market needs, developing immigration policy and deciding where educational funding should be directed to. As such, employer demand intelligence should be reliable, have a wide coverage of data, and be widely available.

This thesis explains the issues that have been identified in the employer demand intelligence domain that have resulted in such intelligence not being able to meet the above-mentioned criteria. To address these issues, the thesis details the innovative solutions that have been developed in this research in the form of an underlying knowledge representation, the population of data into this knowledge representation, and delivering a semi-automatic tool that can continuously extract employer demand intelligence from a range of sources on the web.

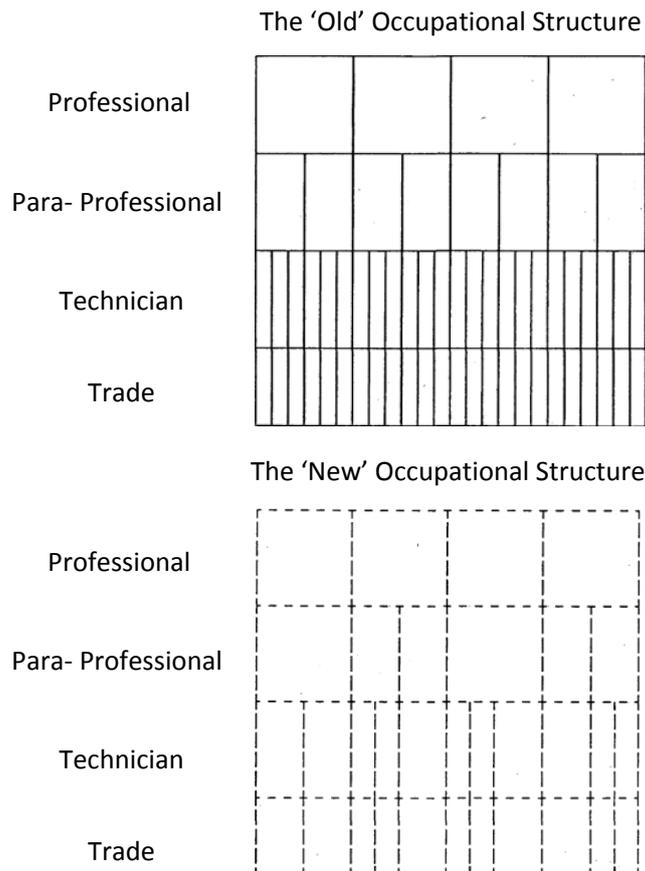
To this end, Chapter One gives a brief introduction to skills shortages and the need for an agreed methodology to be able to provide employer demand intelligence. The chapter then explains that the identification of timely and reliable data is paramount to address skill shortages, and follows with the motivation for this PhD research. The chapter then focusses on the research objectives, scope and advantages of this research and concludes with an outline of the thesis structure.

## **1.2 A Shortage of Skills**

Many economies such as Australia, New Zealand, Bulgaria and Canada, have suffered severe damage to its prospects of economic growth for decades. This damage has been due to frequent periods of skill shortages (Curtain, 1988) — especially post World War I and World War II. Skill shortages occur when employers struggle to fill vacancies for an occupation, or significant specialised skill needs within that occupation, at current levels of remuneration and conditions of employment, and in reasonably accessible locations (DEEWR, 2008-09). The lack of sufficiently skilled and experienced labour has been a topic of great concern in the media for the past decade (Access Economics, 2009; Chamber of Minerals and Energy WA, 2009; MacDonald & Klinger, 2009; Probyn, 2009; Storey, 2001).

The shortages have caused severe labour supply bottlenecks in industries that have been to the forefront of these countries' developments (Turnbull, 1980). Much of the damage is due to frequent adjustments which companies have had to make when they were unable to employ individuals who possessed the skills companies required to meet their goals (Senker, 1992).

Until recently, the major trends over the past few decades in the manufacturing industries, were greater refinements of mass production techniques, greater specialisation of machinery, specialisation of workers' skills and the segmentation of occupations (Hayton & Cheyne, 1988). The changes in the manufacturing and other industries due to these trends, as well as the major role that technology has started to play during the last thirty years, have dramatically changed the way people do their jobs. Figure 1 shows how occupations have changed structurally over the years, indicating that the trend from mass production towards flexible manufacturing has caused the need for broader multiskilled occupations, which require flexibility.



**Figure 1** Simplified representation of 'old' and 'new' occupational structures (Hayton & Cheyne, 1988).

By directing a government's funding to relevant, specific educational and learning needs for each occupation, educational offerings can be focused and shaped drastically to address the shortage of skills for that area. Employer demand should be identified on a region by region basis to ensure that each area has relevant employer demand intelligence pertaining to their specific needs. This, in turn, will indicate which educational courses should be delivered at a higher rate for specific geographical areas.

### **1.3 The Importance of Employer Demand Intelligence**

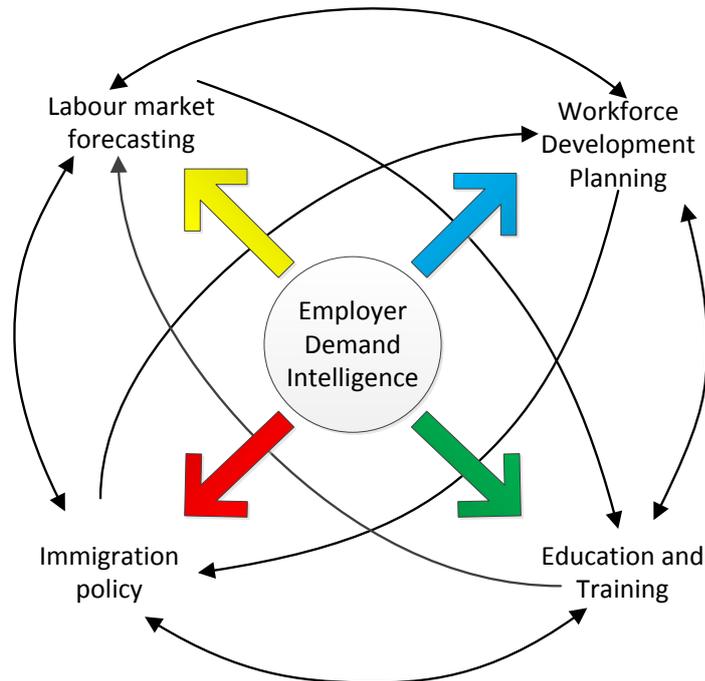
Businesses will be seriously impacted if they cannot adapt to the changing circumstances already evident in markets characterised by shortfalls. These businesses will struggle to obtain and recruit staff and it will become more imperative to retain current staff. To meet this challenge,

innovative recruitment and retention strategies are required (Department of Employment and Workplace Relations, 2007), especially to address the broader recruitment problems that exist in regional areas (Australian Government, 2005).

The delivery of detailed employer demand intelligence will assist government to direct their funding to appropriate and specific areas (Department of Employment Education Training and Youth Affairs, 1998). The funding should be delivered to educational courses that will address the problem of occupations that are in demand but not being able to be filled due to a lack of appropriately skilled employee base in the specific areas. Educational institutions will be able to ensure their courses train students with the appropriate skill sets that employers need at that point in time. Industry will be able to plan better by up skilling their current employee base with appropriate educational material that will provide employees with the needed skills. Employees and future students will be better informed when choosing a course to study to ensure they are employable post their studies.

As discussed, skill shortage periods have caused major economic damage over the centuries, resulting in production lines and services taking much longer than necessary to be delivered. This situation leads to higher cost of products and services and a nation as a whole experiences much higher costs of living. To solve skill shortages, it should be ensured that there will be enough people, with the right skills, in the identified geographical locations necessary to perform whatever job is required at a specific point in time. Without meeting all these criteria, the problem will effectively not be solved.

Accurate identification of employer demand needs and subsequent proactive addressing of skill shortages, provide many economic benefits to a nation and its people. Employer demand intelligence is used in various ways to counter skill shortages as highlighted in Figure 2, with a brief discussion of each of these efforts provided in the next section.



**Figure 2** An overview of the various uses of employer demand intelligence.

### 1.3.1 Labour Market Forecasting

To adequately plan for any changes in skill requirements at a specific point in time and place, a country's future employer demand needs have to be determined. Labour market estimations need to be done on future labour needs for each geographical area's specific requirements. These labour market forecasts provide vital data to many levels of government (e.g. where they should direct their funding to and how they should amend policies that are related to labour market needs in Australia); industry (e.g. strategic organisational planning and training requirements); and to the general public (e.g. employees considering up-skilling). For these reasons it is imperative that the data is accurate, reliable and timely.

Even though there are various labour market forecasting models in use today, all forecasting models are underpinned by current employer demand market conditions and by past trends on which the forecasting is based. Forecasting steps include the analysis of past and present

conditions, and incorporating this data when calculating as accurately as possible what the future holds.

### **1.3.2 Workforce Development Planning**

Workforce development entails planning for the skill needs of a company or region to ensure that the operational unit's labour requirements are met throughout the lifetime of the project or needs of the business. To perform workforce development planning, a company should first decide on their preferred model to use for their specific needs and then follow the chosen model's step-by-step briefings to ensure they plan effectively and proactively for their future labour needs. Even though there are various types of workforce development models in use today, they all entail analysing current employer demand intelligence as part of their plan.

The Department of Training and Workforce Development WA, for example, indicates that Western Australia's workforce development plan:

- Addresses labour market demand and supply on an industrial and regional basis;
- Detects skill and labour shortages;
- Increases the attraction and retention of skilled workers;
- Grows participation amongst workforce;
- Highlights the need for coordination amongst government departments;
- Results in targeted use of migration;
- Recognises the need for comprehensive career development; and
- Improves training and workforce development {Department of Training and Workforce Development, 2009 #7}

To ensure that the above aims are met, it is imperative that a workforce development plan is based on accurate and real data.

### **1.3.3 Immigration**

Employer demand intelligence directly influences a country's immigration policy, and yearly intake of the number and type of immigrants and temporary residents into the country. For this reason, governments undertake employer demand research on a continuous basis. Per example, the Australian Government's Department of Education, Employment and Workplace Relations (DEEWR) research forms the basis of the National Skill Needs list that is used to determine the eligibility of Australian Apprentices and their employers for a range of incentives, and to inform the Skilled Occupation List ("New List of Skilled Occupations Intended to Replace the Current Skilled Occupation List", 2010; "Skilled Occupation List (SOL) and Employer Nomination Scheme Occupation List (ENSOL)", 2009) which serves as a basis for Australia's skilled migration policy. Employer demand intelligence about skill needs in each state and territory is vital to provide data that underpin immigration policy, planning and resource allocation (2008-09).

### **1.3.4 Education and Training**

Employer demand intelligence is useful for several education and training related areas. In the first instance, education providers could use the data to determine which occupations are in shortage or high demand, by employers in their geographical area, and subsequently offer the appropriate number of places for future students to undertake courses in these industries.

Secondly, education providers can utilise this data to determine whether their course content will provide graduates with the skills required by employers, and ensure graduates are job ready.

Thirdly, employer demand intelligence will provide students with information about the current types of skills that specific occupations require, and what the demand is for specific occupations in different geographical locations. This information assists them to decide on the type of career they may like to pursue.

Lastly, employer demand intelligence provides government vital information on where they should direct their funding in order to educate and train more students within these industries that are suffering as a result of skill shortages.

## **1.4 Motivation of Study**

Many authors have indicated their concerns with current employer demand intelligence approaches. These concerns range across more than one issue, but relate mainly to the lack of data and unreliability thereof. The main concerns are briefly described in the sections below, and in more detail in Chapters Two and Three.

### **1.4.1 No Agreed Standards or Knowledge Sharing Exist**

Due to the ambiguity that exists around employer demand terminology globally (Green, Machin, & Wilkinson, 1998; Oliver & Turton, 1982), research reports and papers can sometimes be misinterpreted. This ambiguity is very apparent in the media and in industry, and authors of studies often fail to clarify their intended meaning when using terms like ‘skills shortages’. The questions, in this instance, arises whether they studied the shortage of skills (professionals) for a given area, or did they study the quality of a person’s skills sets to identify its shortcomings? Authors do not make this clear. Even though these two concepts are markedly different, authors use the same terminology — ‘skills shortages’ — to refer to different intended concepts. Handel (2003), for example refers to ‘skills shortage’ and ‘skills mismatch’ as being the same concept, where he classifies these terms to refer to a situation where workers’ skills exceed or fall short of employer requirements. The Department of Education, Employment and Workplace Relations (DEEWR), however, refers to ‘skills shortages’ as a condition that exists when employers are unable to fill, or have considerable difficulty filling vacancies for an occupation, or in meeting significant specialised skill needs within that occupation, at current conditions of employment and levels of remuneration, and in reasonably accessible locations (Department of Education,

Employment and Workplace Relations, 2008-09). These two definitions have very different meanings assigned to them by the authors, though they are both referred to as 'skills shortages'.

The inverse to this scenario is also true. Authors use *different terminologies* when they are in fact referring to the *same concept*. Green, Machin and Wilkinson (1998) found that most existing research and policy discussions tend to equate the notion of a 'skills shortage' with that of a hard-to-fill vacancy. Sloan (2010) refers to vacancies as hard-to-fill when there are qualified people to perform the job who, for some known or unknown reason, are not interested in doing this work. Reasons may include uncompetitive remuneration, working arrangements and nature of the job. Thus, if we apply the DEEWR definition, the notion of 'skills shortages' can include hard-to-fill vacancies, although hard-to-fill vacancies do not always imply that the vacancy is a skills shortage vacancy.

Situations also exist where authors combine both concepts. When there is a shortage of skilled persons for a set area or industry as well as a shortage of specific skills within a person's skill set, authors sometimes combine these two concepts into one definition for 'skills shortages'. Calderon and Stockdale (2008, p. 4) define a trilateral definition of a skills shortage, allowing for three distinct labour market situations: a) a shortage of specific skills; b) a situation where the quality or general skill level of the workforce restricts productivity growth; and c) inadequate labour market participation. Marchante, Ortega and Pagan (2006) consider skills shortages to be an obvious scarcity of suitably skilled workers in the labour market as a result of a lack of workforce, or the existence of skills and spatial mismatches.

Previous research (Green, et al., 1998; Oliver & Turton, 1982; Richardson, 2009) has also noted how ambiguities exist for the terms 'skills', and 'shortages', when analysed separately. Consequently, even more confusion is added to the already ill-used 'skills shortage' term.

In order to assist a nation's economy, or even a company, to function at its full capacity, it is important to have the required number of skilled persons to perform the job, as well as for these employees to have the necessary skill sets to perform their jobs adequately. For example, if a health clinic requires the skill sets of five nurses, but only two of these nurses have the full skill sets required to perform the job at an optimum, the clinic will still fall short of skills. If the clinic can employ five nurses with the full skill sets required, the clinic will be able to operate to its optimum, and there will be no skill shortage. If the clinic, however, needed to employ five nurses, and could only find three nurses to appoint, the clinic would also function at a lower capacity and experience skill shortages. It is important to enumerate both the required number of qualified people to perform the job, as well as the quality of the occupational skill sets required by employers. As per the example, it is clear that there needs to be agreement on different terminologies to ensure appropriate information is included in each study.

There is a large need to reform current employer demand identification methods (Skills Australia, 2010). Data cannot be accurately compared from one year to another to identify changes that have taken place in the interim, nor can the public be at ease with the quality of these outcomes, as it is entirely subjective and founded on baseline terminological differences amongst respondents.

#### **1.4.2 A Lack of Data Depth and Breadth**

The Warwick Institute for Employment Research<sup>1</sup> was a pioneer in the areas of labour market assessment and forecasting work in the United Kingdom (UK), and is currently engaged in a major project funded by the European Centre for the Development of Vocational Training (CEDEFOP) (European Centre for the Development of Vocational Training, 2010) to produce labour market forecasting for all of Europe (Warwick Institute for Employment Research, 2010a).

---

<sup>1</sup> <http://www2.warwick.ac.uk/research>

The Warwick Institute has indicated that local economic users' needs include the need for substantial local labour market detail (sector, gender and occupation).

Labour market research is usually addressed on a metropolitan versus regional level, and does not meet policy advisers and business planners' needs adequately (Warwick Institute for Employment Research, 2010b). In Australia, for example, most labour market analyses indicate current labour market conditions for the major metropolitan areas in each state, and then provide an aggregate response for the 'regional area' (as if it is one location) of each respective state (Shah & Burke, 2005). Most employer demand intelligence studies do not address specific regional and remote areas, as the available information for these areas are scarce and cannot be generalised for all states due to each area's unique characteristics. The regional and remote areas are, however, usually the areas most affected by skills shortages, and the lack of employer demand intelligence for these areas is usually the most noticeable.

There is a need for specialised occupation focused data. This is especially true for those occupations which are in dire shortage. There is insufficient detailed information about skills and competency requirements that employers are looking for in these occupations. It is critical that employers' needs are analysed to get accurate data about current employer demand intelligence on a continuous basis. As technology is increasingly being introduced into new areas, and increasingly being used in existing technology utilising areas, it changes the way employees perform their jobs, and these skill set changes should be constantly monitored and updated for employer demand intelligence stakeholders.

### **1.4.3 Time Consuming and Expensive Approaches**

Current employer demand analysis methods are time and labour intensive as they rely on manual examination of the information contained in each job advertisement. This results in researchers only analysing a small sample of data, which provide skewed intelligence that is

location or industry and occupation poor, and as such does not meet employer demand intelligence stakeholders' needs.

The Australian Chamber of Commerce and Industry<sup>2</sup> (ACCI), for example, stated that current approaches do not address the guidelines the Chamber has set. The ACCI advised there are three essentials when a National Workforce Development Strategy is determined (Australian Chamber of Commerce and Industry, 2009):

- Broad coverage of all firm sizes;
- Data pinned to regional and local levels; and
- Comprehensive information on scarce occupations where critical skills shortages exist or may develop.

Current data collection methods are mostly done manually through methods like survey distribution and membership based approaches. **Survey based** methods are known to be cost intensive, time intensive, labour intensive, to comprise great travel distances for face-to-face surveys, provide only anecdotal information, have skewed sample coverage and contain outdated information. **Membership based approaches** typically cannot collect micro level information on individual occupations; industry experts instead convey generalised information about industry trends and employer needs and do not distinguish between, and provide, detailed information for each occupation. This intelligence is also subjective, suffering from varying interpretations of terminologies, and is time constrained as it depends on the meetings taking place between data collector and company representative; some occur as infrequently as half yearly or yearly.

Due to the dire need for a detailed employer demand intelligence framework, it has always been critical from the onset of this research, that the research be developed to such an extent that it

---

<sup>2</sup> <http://www.acci.asn.au/>

can be used by industry directly after the first phase of the research has been completed. As such, it is vital that any proposed framework should be evaluated to such an extent that it is not only scientifically valid, but also that it will be accepted by the employer demand intelligence industry.

## **1.5 Research Objectives and Scope of the Problem**

This PhD research addresses the industry and researchers' concerns by proposing a conceptual framework that identifies detailed employer demand intelligence. ICT and state of the art scientific approaches enable this framework to obtain a comprehensive sample of industry, through analysing data from ALL sectors (not just the largest sectors, as is usually done in other studies). The ICT techniques also enable the analysis of a complete geographical area where vacancies exist, not only information related to vacancies that are in easily accessible metropolitan regions. Lastly, the ICT and scientific techniques employed provide the ability to identify employer demand intelligence in real time, thus the analysis will not reflect outdated information as current techniques do, as discussed in previous sections.

The objective of this thesis is to create an employer demand intelligence framework that is able to:

1. Function on an underlying knowledge representation basis;
2. Provide data instances that are occupation specific in the first place, then linked to a wide range of other employer demand intelligence variables, to ensure broad coverage of the domain;
3. Semi-automatically gather, analyse, and report on employer demand intelligence, and be evaluated for its engineering soundness, and practical readiness for industry uptake.

## **1.6 Advantages of the Research**

As this research will have a strong practical influence on the way industry currently applies employer demand intelligence, it is important to not only state the scientific related advantages that this framework provides, but also its practical advantages.

The proposed framework, therefore:

- Highlights the low accuracy rate of job advertisements found on the job board SEEK, where only 35% of the job advertisements were usable for the purposes of identifying accurate employer demand data. 25% of the job advertisements that were advertised in certain categories, were not actually related to those categories, and 41% of the 416 advertisements for the one month's trial period were found to be duplicates;
- Identifies critical educational needs for specific regional areas;
- Enables current regional workers to up-skill by accessing specific educational material relevant to the local area's needs;
- Identifies employers' needs accurately;
- Allows stakeholders to proactively address skill shortages;
- Adapts to globalisation consequences;
- Provides maximum value for public regional educational funding;
- Identifies and meets the learning needs of all occupations and not only mainstream occupations;
- Assists the expansion of a country's labour market as efficiently as possible by educating people in occupational qualification needs that are specific to that region;
- Provides for a true reflection of all employers' needs and not only employers who agree to participate in employer demand surveys;
- Provides detailed geographical location information relating to skill shortages;

- Provides intelligence about broad transferable skills required by employers for occupational, industrial and geographical levels;
- Enables detailed employer demand analysis, provides results for a wide range of occupations, and eliminates the need to collect data via surveys for each occupation;
- Provides a detailed employer demand intelligence representation framework in the form of an ontology (described in Chapter Four);
- Allows semi-automated data collection;
- Is much more cost efficient than traditional data collection methods;
- Enables faster and less cumbersome data collection than current methods;
- Provides up-to-the-minute access to real time data continuously.

## **1.7 Thesis Structure**

To assist in understanding the structure of this thesis a brief description of what each Chapter consists of, is provided below.

### **Chapter 1: Introduction (current chapter).**

### **Chapter 2: Literature Review**

This chapter starts off by discussing the various types of employer demand intelligence approaches that exist, and provides a critical evaluation of the integrated issues which are: i) knowledge sharing and differing approach structuring; ii) data depth and breadth; and iii) data out of date.

### **Chapter 3: Problem Definition**

Chapter Three discusses the key concepts and definitions applicable to both the employer demand intelligence and the information systems domains that are used in this research. The

discussion then analyses the problems that the employer demand intelligence domain faces in detail, and then discusses the underlying research issues related to these problems. The thesis goal and key research questions are provided, and the chapter concludes by supplying an overview of research methods, and the justification of why a science- and engineering- based research methodology has been chosen for this research.

#### **Chapter 4: Solution Proposal**

A solution overview of the research looks at the scientific approach of ontology on which this research is based. The chapter also details the backbone of the research, its conceptual framework, and concludes by focusing on each proposed solution requirement to meet the aims of this research.

#### **Chapter 5: Employer Demand Ontology Engineering**

Chapter Five addresses research issue one (explained in Chapter Three). This chapter first brings attention to the fundamentals of the research methodology, provides an overview of existing ontology methodologies, and then explains the NeOn methodology, the chosen methodology for this research. The focus then turns to the Employer Demand Ontology (EDO) notations, components and design principles, and concludes by elucidating on the conceptual implementation of the EDO.

#### **Chapter 6: Employer Demand Ontology Instance Population**

Chapter Six addresses research issue two (explained in Chapter Three), by focusing on the instantiation of employer demand data, and revealing the process that was implemented to populate the EDO.

#### **Chapter 7: Semi-Automatic Employer Demand Intelligence Tool**

Chapter Seven addresses research issue three (explained in Chapter Three). The chapter details the steps that were taken to develop the Employer Demand Intelligence Tool (EDIT). It starts with an overview of the EDIT, then details each process involved in its development. A results and discussion section provides an example with screenshots to illustrate the workings of the tool, and the chapter concludes with the limitations of the EDIT as it currently stands.

### **Chapter 8: Employer Demand Intelligence System Evaluation**

This chapter has two evaluative foci. First, the Employer Demand Ontology is verified for logical consistency. Secondly, the EDO is validated through an array of contemporary use case scenarios, to ensure it is ready for uptake in industry. Actual industry use cases have been listed as future work of this research.

### **Chapter 9: Conclusion and Future Work**

This thesis concludes by recapitulating the research issues in the employer demand intelligence domain, the solutions that have been provided to address each issue, and outlining the future work that is planned for this research.

## **1.8 Conclusion**

This chapter has provided an introduction to the issue of skill shortages and introduced the concept of employer demand intelligence as a way of identifying current skill shortages on a micro level of detail. The major concerns related to the employer demand intelligence domain, were briefly provided, after which the research objectives, advantages of this research, and thesis structure were given.

In Chapter Two, a literature review of the current employer demand intelligence domain is delivered. The focus of the chapter, is to elucidate the state of the employer demand intelligence

domain, with a critical evaluation of each approach that currently exists to gather employer demand intelligence.

# Chapter 2 - Literature Review

## 2.1 Introduction

Skills shortages globally pose a real and urgent need for proper investigation and workforce development planning into the future. Analysing workforce development and employer demand needs through electronic datasets allows much deeper and wider research into skill shortages. One way to gather employer demand intelligence is by analysing online job advertisements. Adoption rates for using e-based resources to advertise and recruit future employees were already 96 per cent for North -American based global 500 companies in 2003 (JobsDB Dimension, 2011), and 75 per cent for all companies in 2008 (O'Callaghan, 2008), making online job advertisement analysis a very viable option. In general, approaches to gathering information about employment needs have been manual to date, not founded on structured frameworks, and mainly ad -hoc. Current online approaches specifically used to identify economic trends are far too high level to acquire the detailed data necessary for workforce development, skill shortages and educational need analyses. Structured electronic approaches, on the other hand, do not focus on gathering data for large scale occupational and geographical workforce development needs.

In this chapter, the focus is on in-depth discussion about the various current approaches to employer demand intelligence in existence. Approaches broadly obtain their source data through one of two ways: 1) data obtained through directly surveying an organisation, or 2) data obtained through structured or unstructured analysis of job advertisements. This chapter discusses existing approaches in four main categories found: survey- based approaches, online- based approaches, manual job advertisement analyses and structured approaches. After reviewing the approaches, the focus turns to a critical evaluation of the approaches in an integrated view.

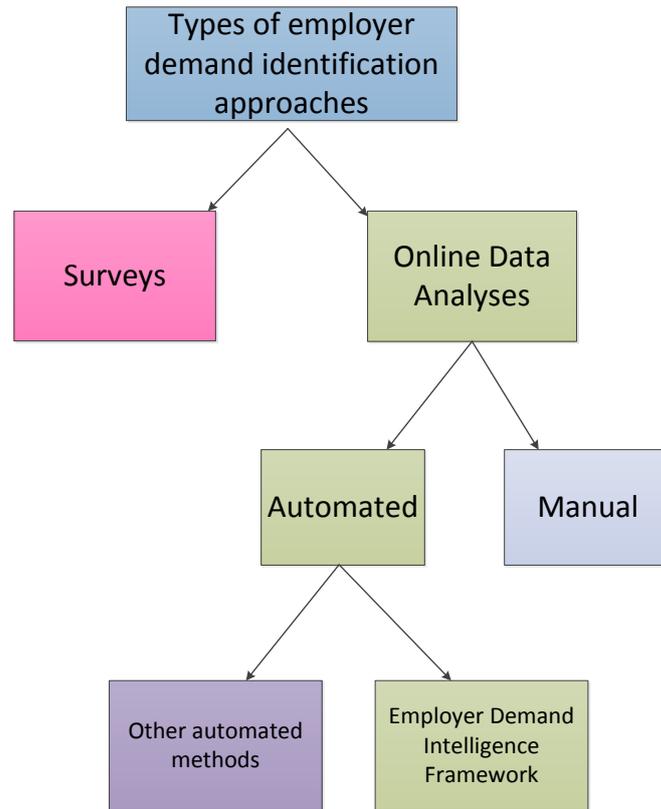
## **2.2 Employer Demand Identification Approaches**

There are various approaches being used to identify employer demand needs. Current providers of employer demand data in Australia include the Australian Bureau of Statistics<sup>3</sup>, the Department of Employment, Education and Training, employer associations, unions, state and territory employment and training bodies, industry training committees and TAFE (Hayton, 1988). No one study's outcomes should be used as the only data source for determining which skills are in shortage at present. There is a great need for studies that gather much wider, deeper and more accurate data about the needs that employers have for specific occupation types and skill sets in their organisations. As discussed in Chapter One, many stakeholders have mentioned the need to accurately gather and analyse employer demand data in order to obtain a better picture of up-to-date employment needs. Such intelligence would greatly assist workforce development, immigration, education and training policy development, and individual career and organisational human resource planning in the future.

Current approaches mainly focus on gathering employer demand data through conducting surveys with employers, or by looking at high-level online data about the number of job advertisements from one period compared with another. Figure 3 provides a graphic orientation of the different approaches. Governments mainly utilise data obtained through surveys and high level job advertisement counts. Other approaches consist mainly of job analyses done by academics to obtain one-off intelligence about a specific occupation. Existing structured job advertisement analyses do not cater for detailed workforce development, educational and immigration decision needs.

---

<sup>3</sup> <http://www.abs.gov.au/>



**Figure 3** Orientation of employer demand identification approaches.

Table 1 provides a list of the approaches discussed as a part of the Literature Review, and the respective number of studies in each approach that was analysed. A detailed explanation of each approach and study follows thereafter.

	Type of approach	# of studies discussed for each approach
1	Survey based approaches	3
2	Manual job advertisement analysis approaches	3
3	Online based approaches	4
4	Structured approaches	11
	<b>Total number of studies discussed</b>	<b>21</b>

**Table 1** Number of approaches related to employer demand identification discussed.

### 2.2.1 Survey-Based Approaches

Survey-based approaches include telephone and in-person market intelligence gathering. These are the main approaches that different government departments rely on for data to inform their policy development and funding decisions. It is helpful to consider examples of the regular efforts of both a federal government department and a state government department to gather employer demand intelligence.

The Australian Government's Department of Education, Employment and Workplace Relations (DEEWR)<sup>4</sup> continuously undertakes employer demand research. This research forms the basis of the National Skill Needs list, which is used to determine eligibility of apprentices and their employers for a range of incentives, and the Skilled Occupation List (Australia. Department of Education, Employment and Workplace Relations, 2010; Australia. Department of Education, Employment and Workplace Relations, 2009) which serves as a basis for the nation's skilled migration policy. Employer demand research also addresses the need for information about skill

<sup>4</sup> <http://deewr.gov.au/>

needs in each State and Territory to underpin policy, planning and resource allocation (DEEWR, 2008-09).

The DEEWR employer demand identification approach provides qualitative information based on analysis of the Survey of Employers (SERA), wherein employers who have recently advertised in selected skilled occupations participate in the survey by providing this detail to DEEWR. The survey is usually performed through in-person or telephone discussions with individual employers.

Collection of local and international employer demand data is usually done by means of industry consultation — either via distributed surveys or face-to-face interviews. Various institutions also gather employer demand intelligence, mainly by utilising surveying instruments. In Western Australia, the Department of Training and Workforce Development (DTWD)<sup>5</sup> has embarked on its task of creating a State Workforce Development Plan to improve the state's ability to attract and retain workers, increase workforce and participation rates, continue to improve the state's training system, and to develop strategies to better incorporate the needs of stakeholders (2010).

As part of the DTWD's consultation process to inform the development of the State's Workforce Development Plan, the department employs the ten West Australian Training Councils to identify current skills shortages. The ten Training Councils' main employer demand identification approach consists of a survey to a random sample of employers.

Another survey-based approach is that of the Australian Government's Australian Census<sup>6</sup>. Every five years, the Australian Government administers the census to gather information about Australia's people, their dwellings, and some key characteristics of each (Australian Bureau of Statistics, 2012a). Census data are widely used for many purposes by government, individuals and organisations in the public and private sectors, to make informed decisions about policy and

---

<sup>5</sup> <http://www.dtwd.wa.gov.au/>

<sup>6</sup> <http://www.abs.gov.au/census>

planning issues. Even though some skills shortage studies include census data in their analyses, it still only reflects the information of the employee and their current occupation and status, for the data of the employer as such cannot provide an accurate picture of employers' needs for specific occupations and their skill sets.

### **Issues with Survey-Based Approaches**

The following issues have been identified with survey based approaches:

1. Survey-based studies are **reliant upon the acceptance and participation of employers**. Due to surveys being extremely time-intensive, many employers do not participate in the study. This makes the study sample statistically invalid or less representative of employers across all industries and locations.
2. The survey mainly **gathers qualitative data** and does not easily enable the compilation of quantitative estimates. Due to the nature of questions asked in surveys, they usually require a sentence-like, qualitative type response which is hard to transform into statistical quantitative measures. Gathering short answers with one or two clear measurables makes it much easier to compare responses from one respondent to another and from one survey cycle to another.
3. Survey-based studies **are subjective**. Due to the nature of surveys being completed by an individual, the information collected is subject to personal opinion. Employer demand surveys, by and large, contain questions like, 'Which occupation or skill do you think is in shortage?' An individual's answer is commonly influenced by other people's opinions. For example, the individual may have watched a news bulletin the night before they completed the survey and may adopt their survey response from the bulletin's views. Unfortunately, such reports are not usually based on facts, because media information frequently relates inaccurate facts. Employers themselves, for

their own human resource needs, often lack or do not gather the data requested through the survey, and survey respondents' answers made up on the spot can be unreliable.

4. Survey-based studies **are expensive**. Due to their manual intensity, a number of staff are needed to administer a survey, making it cost inefficient; in the case of the Australian Census, 43,000 field staff were employed in the most recent census (Australian Bureau of Statistics, 2012b). In Australia, which is one of the largest countries by square kilometre, a national survey is not only labour intensive, but also requires huge distances to be travelled for the staff to perform the survey, with many employers and households situated in regional and rural areas.

5. Survey-based studies **have great time lags between rounds**. Most studies are done annually at the most, and the Australian Census is done once every five years. Due to regular and frequent changes experienced in the labour market one cannot rely on information gathered a year ago — it would be already out of date. It is imperative that policy decisions are based on data analyses that are as current and relevant as possible to ensure accurate policy development, funding distribution, syllabus and workforce development and career decisions.

6. Survey-based studies **are time intensive**. Apart from the labour intensive issue raised in point 4) above, because of some nations' vast distances, as such Australia, it can take a significant amount of time for the surveyors to complete interviewing a range of employers based in different metropolitan, regional and rural areas. To ensure an accurate and reliable sample, survey administrators have to include as wide a range as possible of different geographical areas, occupational types and industrial sectors that have been scoped for each study. Large travel distances make this task onerous and it is often skimped over.

### 2.2.2 Online-Based Approaches

Historically, the Department of Employment, Education, Training and Youth Affairs (DEETYA, now called DEEWR)<sup>7</sup> manually administered the Skilled Vacancy Survey (SVS). The SVS entailed counting newspaper vacancies advertised in the major daily papers in each State capital and the Northern Territory (Department of Employment Education Training and Youth Affairs, 1998) and is thus not truly a survey, but more a quantitative count of the number of existing vacancies at a point in time. With the introduction of technological advances, and most job advertisements now being reflected on the internet, the manual administration of this process has ceased.

DEEWR in Australia and the Department of Labour in New Zealand both analyse data pockets that are received monthly from job websites. Both these reports only calculate variances in the number of job advertisements over certain periods of time, map the online advertisements to occupational codes and aggregate the information into sector and state sets. Some of the approaches used are briefly described below.

DEEWR's online job analysis report is called the Internet Vacancy Index (IVI)<sup>8</sup> which indicates the monthly online vacancy counts received by DEEWR from four job websites (Australian Government, 2010): SEEK.com.au; MyCareer.com.au; CareerOne.com.au; and JobSearch.gov.au.

The New Zealand Department of Labour's report<sup>9</sup>, on the other hand, is called Jobs Online and receives monthly online vacancy counts from three job websites: "Jobs Online", (2009): SEEK.com.au; heraldjobs.co.nz; and TradeMe.co.nz/jobs.

---

<sup>7</sup> <http://deewr.gov.au/>

<sup>8</sup> <http://deewr.gov.au/lmip/default.aspx?LMIP/VacancyReport>

<sup>9</sup> <http://www.dol.govt.nz>

The ANZ Job Advertisement series<sup>10</sup> contains even less detailed information than the previous two reports, as it only counts the number of new online vacancies each week. The report does not aggregate any other information relating to employer demand.

### **Issues with Online-Based Approaches**

The following issues have been identified with online-based approaches:

1. **Keyword- and taxonomic-based retrieval efforts cannot ensure precision** and good recall ratios, due to online-based approaches using text-based searches and not using synonyms. This in turn leads to missing and redundant information when performing job offer searches for analysis (Lv & Zhu, 2006). For example, if one searches for all current vacancies that require the job seeker to have experience in Java and ASP, this search will not pick up those vacancies that specified VC++ and ASP.NET, where in fact these skills are regarded as equivalent.
2. **Categorisation of information** contained in the vacancies **is difficult** to do due to the large spectrum of terminology used for similar concepts related to any vacancy. In point 1) above, the issue in finding information contained in job advertisements was pointed out. Similarly, using this information to categorise it for analysis is extremely hard, due to the multiplicity of varying terminology used, and failure to sort like with like results in unreliable and fuzzy employer demand data.
3. Due to the large number of varying types of job portals, employers advertise their vacancies repeatedly in as many as possible of these boards, in an attempt to attract the best job applicants. Through posting the same vacancy on many job boards, it becomes **hard to track vacancies** for the purposes of employer demand identification and analysis. Determining whether a vacancy is a duplicate or only a similar offering found on another job board, especially when different job boards

---

<sup>10</sup> <http://www.anz.com.au/corporate.research/australian-industry-economics/job-advertisement-series/>

require employers to enter the information in varying formats or levels of detail, makes vacancy tracking extremely hard and requires manual intervention to determine whether it's a duplicate or not. Subsequently, from a labour market analysis point of view, current online approaches cause inaccurate vacancy statistics.

4. Today's web content mainly consists of **information readable by humans only**. As the current online-based employer demand approaches utilise current web functionality only, their information is only readable by the human brain, and this processing cannot be supported by computer intelligence. The Semantic Web, however, which is the web of the future, will bring structure to the meaningful content of web pages, so computers can readily carry out sophisticated tasks for humans (Berners-Lee, Hendler, & Lassila, 2001). We therefore need employer demand approaches that are based on semantic foundations in order to be future-proof and ready for the uptake in the semantic web.

5. Online-based approaches **rely on only one data source** for their employer demand information. With varying types of employer demand data in different places on the internet, we need approaches that can gather, integrate and analyse information from more than one online knowledge source. This will result in employer demand intelligence that is more effective and accurate.

6. Online-based approaches **do not have agreed structures across borders**. The approaches are developed in-house by some organisations and put into practice without scientifically validating the concepts contained in the approach as accepted by broader industry or geographical regions. This makes comparisons of study outcomes and interpretation of results hard, as each approach has its own interpretation of concepts and relationships among those concepts.

7. Online-based approaches **do not specify relationships among concepts** that make up their framework. Concepts are usually listed in a random fashion (e.g. alphabetically) and there is no

relationship mapping between all the concepts in the framework. If the relationships among the concepts contained in the study are specified, clearer, agreed understanding about what each outcome means is possible.

8. Online-based approaches are **not detailed enough**. They merely provide information about the number of vacancies from month to month and their geographical distribution based on occupational codes. This information is not detailed enough for meaningful decisions about policy, funding, career and workforce development. Detail such as required skill sets for each occupation is also imperative in these decisions. The employer demand data gathered by online-based approaches thus need to be much wider and deeper to facilitate accurate decision-making.

The information that is derived from these reports is extremely high level and does not detail employer needs sufficiently to accurately inform policy development, funding decisions and educational and workforce development needs.

To summarise, online employer demand identification approaches need more intelligent methods to: carry out automatic categorisation; achieve information sharing among heterogeneous online approaches; improve machine-understandable capabilities for better search results; be semantic and web ready; and to have agreed structures across borders that show relationships among employer demand concepts.

### **2.2.3 Manual Job Advertisement Analysis**

Manual job advertisement analysis is mainly undertaken in academic research. Much of this is done as a one-off type of study to identify employer demand skill sets required for specific occupation types at a point in time. These studies do not usually recur (e.g. monthly or annually). Three of these efforts are briefly described below.

Arcodia and Barker (2003) from the University of Queensland performed a nation-wide study of event management positions in Australia to re-evaluate educational and job training curriculums. They analysed 105 web-based job advertisements sourced from the following websites: Seek.com.au, CareerOne.gov.au, Employment.com, Monster.com and AllJobs.com.au. To ensure they received new vacancies posted for the period they were looking at, they had set up email alerts based on occupation-related keywords for event management advertisements. The study aimed to identify what skills industry required from event management staff during a period in 2002, specifically recording information related to the job title, location, industry, skills and personal attributes contained in each of the job advertisements.

Another manual job advertisement analysis study was that of Mules (2007). Mules collected online and hardcopy job advertisements published during the period May 2004 to April 2005 from The Canberra Times<sup>11</sup>. Mules focused on tourism and hospitality related occupations for his study, intending to explore empirically the feasibility of a monthly data series for job vacancies in these industries within a geographically well-defined region. Mules specifically collected data relating to the position title, requirements for skills, qualifications, experience, knowledge, and the type of employment (full-time or part-time), date listed, organisation, and closing date of each advertisement. It was concluded that such a series is feasible and adds to knowledge about skill shortages in the industry.

The final approach discussed in this section is that of Robinson, Arcodia, Tian and Charlton (2010) and Robinson and Beesley (2010). Based at the University of Queensland, Robinson et al., in collaboration with the Hospitality Training Association, tracked job advertisements from 16 job advertisement websites for cookery related employment. Their approach entailed collating data from and investigating aspects of electronically analysed job advertisements. Their aim in doing so

---

<sup>11</sup> <http://www.canberratimes.com.au/>

was to address the dearth of available information related to skills shortages in cookery. Specific aspects investigated for each vacancy were the employment status (full-time, part-time or contract), the location of the vacancy, the industry sector (government, private restaurant etc.), the organisation, skill and attribute requirements stipulated, and key benefits available to the employee. Robinson et al. utilised Microsoft Access and SPSS software to analyse the interrelationships between variables. They also performed cross tabulations with a Pearson Chi-Square test to determine statistically significant differences. Their study revealed that the increased digitalisation of information facilitates both the collection and generation of scarce labour market research.

### **Issues with Manual Job Advertisement Analyses**

The following issues have been identified with manual job advertisement analyses:

1. Manual job advertisement analyses are **resource and time intensive**, making it an expensive approach. It is an onerous task to gather job advertisements from various job portals, do the manual data entry for each job advertisement into the employer demand dataset one-by-one, and analyse such vast information without computer assistance.
2. Even though manual job advertisement analyses gather wider and deeper data types than online-based approaches, they are still **not detailed enough** due to the big task involved in manual collation and processing of this information.
3. Manual job advertisement analyses are typically **ad hoc studies**. As mentioned earlier, these types of studies are usually performed in one-off academic research efforts and not set up to be administered weekly, monthly or even yearly.
4. Current manual job advertisement analyses **are not standardised**, making comparisons between different studies and between time periods nearly impossible. If concepts and principles used in

these studies are agreed upon by the wider industry population, it would assist in better data sharing across geographical and occupational borders.

5. Manual job advertisement analyses **do not indicate the relationships between the various concepts** analysed. When relationships among concepts are established and clarified, researchers and industry specialists are more able to understand the study's findings. Furthermore, researchers can find other potential indirect influences and outcomes between the variables in each study, and support further development of employer demand research efforts.

#### **2.2.4 Structured Approaches**

eRecruitment and eHumanResources models are increasingly being developed for various purposes. In order to locate related works, continuous online search engine searches were performed throughout the life of this research for eRecruitment and eHumanResources models based on semantics. Even though it is acknowledged that other online recruitment and human resources models exist, this research focuses on structured models — those based on semantics. References to any Web 2.0-related projects were included where there was an awareness of them, although no detailed analyses of these were included because of their more restricted capacities, as described in Online-Based Approaches (section 2.2.2).

Searches performed included carefully checking the publicly available knowledge structures in the following areas:

1. Google and Swoogle (Swoogle Semantic Web Search, 2007) searches:

These included using various human resources terms and varying combinations of words related to human resources (e.g. 'competence') and existing structured approaches (e.g. 'ontology') in the same search. Searches specifying certain file types per se were also performed, e.g. searching for OWL file types by typing 'human resources filetype:owl' into the search engine.

2. Ontology/ taxonomy repositories:

- Protégé Ontology Library<sup>12</sup>
- Swoogle Semantic Web Search<sup>13</sup>
- DAML Ontology Library<sup>14</sup>
- Schema Information<sup>15</sup>
- Schema.Org<sup>16</sup>
- Watson<sup>17</sup>
- FIPA Ontology<sup>18</sup>
- Ontosearch<sup>19</sup>
- Ontolingua<sup>20</sup>
- OpenCyc<sup>21</sup>
- Taxonomy Warehouse<sup>22</sup>
- Linked Open Vocabularies<sup>23</sup>
- Vocab<sup>24</sup>

The most common purposes of existing models are either to classify and agree on structures of an organisation's internal human resource needs in order to match job seekers with job offers,

---

<sup>12</sup> [http://protegewiki.stanford.edu/wiki/Protege\\_Ontology\\_Library](http://protegewiki.stanford.edu/wiki/Protege_Ontology_Library)

<sup>13</sup> <http://swoogle.umbc.edu/>

<sup>14</sup> <http://www.daml.org/ontologies/>

<sup>15</sup> <http://www.schemaweb.net/>

<sup>16</sup> <http://schema.org/docs/schemas.html>

<sup>17</sup> <http://kmi-web05.open.ac.uk/Overview.html>

<sup>18</sup> <http://ceur-ws.org/Vol-52/oas01-suguri.pdf>

<sup>19</sup> <http://www.ontosearch.eu/>

<sup>20</sup> <http://www.ksl.stanford.edu/software/ontolingua/>

<sup>21</sup> <http://www.opencyc.org/>

<sup>22</sup> <http://www.taxonomywarehouse.com>

<sup>23</sup> <http://lov.okfn.org/dataset/lov/>

<sup>24</sup> <http://vocab.cc>

or for the purposes of competency management. In this next section, the research efforts in these three domains are discussed.

### **A. Approaches to Representing Human Resources in Organisations**

Models in this domain focus on human resources management solutions (eHumanResources), which represent a company or business perspective. Gualtieri and Ruffolo (2005) describe an organisational knowledge framework to provide design support for a knowledge management infrastructure system. The framework is two-tiered, with the first level representing a typical organisational background, and the second level representing core organisational knowledge entities such as human resources, business processes and technical resources.

Terziev et al. (2005) developed the PROTON knowledge management structure, which is a multi-layered ontology with basic concepts as standardised options. PROTON is a general-purpose, domain-independent ontology intended to be reused for consistency's sake. During the EDO development some of the PROTON concepts were considered for reusing. A detailed account of each class and why it has been reused is described in Chapter Five of this thesis.

Zemmouchi-Ghomari (2012) built an ontology to model the human resource structure of a Higher Education organisation based in Algeria. The main categories are location, research work, and role for each person related to the organisation. The ontology focuses on describing the role that each person in the organisation holds, including those of students. Unfortunately, we could not locate any research papers on the topic apart from the high-level OWL files that have been made available online.

From Zemmouchi-Ghomari's OWL files and Terziev, Gualtieri and Ruffolo's papers, no indication of the incorporation of widely accepted non-ontological classifications incorporated into these ontologies could be found. Instead, Zemmouchi-Ghomari specifically invites feedback through

her website on her ontology with the aim of validating, sharing and reusing the structure. This suggests that the structure still needs verification.

## **B. Matching Job Seeker with Job Offer Approaches**

The first job matching eRecruitment ontology to be discussed is the one developed by Gomez-Perez, Ramirez and Villazon-Terrazas (2007) and Villazon-Terrazas, Ramirez, Suarez-Figueroa and Gomez-Perez (2011) for the Single European Employment Market Place (SEEMP) project. The aim of the SEEMP project was to provide a system that was based on platforms and networks between different online recruitment websites. The system would match Information Communication and Technology (ICT) job seekers with ICT vacancies. This system bridged language barriers and provided a single port of call for job seekers to look for ICT employment in Europe. Unfortunately, due to the need to continuously update the system, and no one being available to maintain it at present, it is not available to users online anymore. The system was operational for a period of time in the late 2000s. Job seekers used the online interface to upload their CVs. The CVs were in turn matched for skill compatibility by the system against currently advertised vacancies across all the various online recruitment websites that were subscribed to SEEMP. Matches from all member websites were ultimately displayed to the job seeker.

The SEEMP project was developed on a system-of-systems framework comprising thirteen modular structures. The developers incorporated twelve standards, codes and classification structures from existing classification structures accepted worldwide and in Europe.

Garcia-Sanchez et al. (2006) also developed a job-matching ontology, but their geographical area covered rural south-east Spain. Their ontology was designed to help specific organisations look for specialised employees and, on the other hand, to assist job seekers find appropriate work. The main concepts of this ontology are 'employer' (entity offering the work), 'offer' (any information placed

online by an employer such as a news article or employment advertisement), 'profile' (characteristics of the job seeker), and 'applicant' (the job seeker).

What makes the SEEMP project unique compared to the other existing eRecruitment structures is that they have reused as many existing structures as possible. Furthermore, SEEMP, along with Garcia-Sanchez et al.'s (2006) systems, were actually developed to such a stage that the end users could utilise it in reality, further validate it, and reap its benefits. As Simperl (2009) states, most heterogeneous ontologies that are widely accessible are hardly being used beyond the boundaries of their originating context. Their limited impact leads to only minimal uptake in industry. The SEEMP system is also the one found with the most detailed explanation of its actual development stages, whereas Garcia-Sanchez et al.'s system included the most information about their concepts used in their research paper. It was not necessary to obtain the ontology files, as these structures were included in their paper as an addendum. These aspects were particularly attractive for the development of EDO; detailed explanation around which parts of Garcia-Sanchez et al.'s SEEMP structure were reused for the EDO structure is given in Chapter Five.

Another framework is that by Bizer, et al. (2005). Bizer et al.'s Knowledge Nets system for the German market was developed to assist employers to post job offers in a more efficient way. Previously, the German job portal market consisted of many portals, which cost employers who wanted to reach a broad audience a considerable amount of money to advertise on many of the portals. The Knowledge Net system provided a platform that job portals could use to trawl companies' own websites. Specific job offerings are advertised on these websites and then published through their own job portal, without much additional manual intervention. This process would result in all job portals operating on the same information, and postings reaching a wider audience, thus providing a more transparent market. Furthermore, the system assists job seekers to standardise their job seeking profiles through the Knowledge Net, which in turn would apply algorithms that calculate the similarity of job requirements with applicants' profiles.

Lastly, Lv and Zhu (2006) also developed a job matching (eRecruitment) framework. Their framework had a Chinese setting, where they assigned weightings to employee requirements and matched these with the weightings assigned to job seekers for their skill sets. The only information found in the paper relating to the actual ontology classes and their approach was the mention of their main skills being classified into two broad categories: soft and hard skills. They do not seem to have used any existing widespread standards or methodology.

### **C. Competency Management Approaches**

Dorn, Naz and Pichlmair (2007) developed a human resources structure to serve two purposes, thus effectively falling in the previous category as well as this category. Their first aim was to be able to manage a university's competence in terms of their students' abilities, and their second aim to provide a Job Meta-search function. The University Competence Management System focused on gathering data about each employee's skill sets and their individual level of attainment of skills, which would then be presented as a 'total grade of a competence' after an algorithm was applied to these two factors, offering the university a snap-shot of their total skill sets available. The Job Meta-search system would match the skills of job seekers with offers from employers through analysing and aligning the data available on several specific job portals, eliminating the need to search many portals one-by-one.

Dorn et al. define a general competency structure consisting of functional and behavioural competencies, where functional relates to knowledge skills and abilities of an individual, and behavioural refers to soft competencies, which influence a person's behaviour or attitude. However, it is concerning that they refer to providing one grade for each applicant overall, where there could be varying levels of knowledge and experience for different types of skill sets of one person. This single grade could provide inaccurate information on candidates. They may receive a lower type grade due, perhaps, to one knowledge unit that they do not have many years'

experience in; however, that specific skill may not hold the same level of importance to employers and may be more relevant to some occupations than others, placing the student at a disadvantage.

Furthermore, Dorn et al. infer that the level of competency for each person's grade should be assigned by some higher level of authority, making this study only valid to people who would be able to have their skills constantly reassessed by someone (e.g. students). The only existing widespread human resource standard classification that these authors incorporated into their work was that of computer- and business-related occupations from the Standard Occupations Classification (SOC) list, and skills from International Co-operation Europe Ltd.

A related structured approach, but not an eRecruitment one as such, is that developed by Biesalski and Abecker (2005). These authors developed an ontological Human Resource Management system that would facilitate data analysis about a company's human resources situation at any given point in time (eHumanResources). The system provides for job requirement skills to be entered and matched against those of current employees, effectively doing a gap analysis of the company's human resources. This is based on a competency catalogue in which data from both the job and the incumbent are entered, allowing a manager to compare the differences and to do strategic planning for future training and recruitment needs. The system's skills structure has once again been based on the German KOWIEN project's skills structure, tried in real life at the Daimler-Chrysler AG Würth Plant.

Lau and Sure (2002) also developed a similar competency management system at Swiss Life, a large insurance company based in Switzerland. The system's aim was also to facilitate gap analyses of the company's existing human resources, and their skills structure was developed through in-house focus groups with their own employees. They did not reuse any other existing structures or widespread standard classifications to develop their structure, nor did they make their detailed list

of skills or any other facet of their structure publicly available to be analysed by other researchers globally.

The final human resources project reviewed was by Trichet and Leclere (2003) based in France, and focused on the Information and Telecommunication Technology (ITT) domain. They built a system that provides job matching capabilities and measures an organisation's human resources needs. They did not incorporate any widespread standard classification systems; rather, they built their framework's skill set template from a report provided by an industry group that represents major ITT corporations in France.

To summarise, Table 2 lists the existing structured data systems and their respective purposes.

Source	Ontology purpose	Existing standards used?	Country	Occupation/ Industry focus
(Lau & Sure, 2002)	Competency management	No	Switzerland	Private insurance, human resources, IT
(Trichet & Leclère, 2003)	Job match and competency management	No	France	IT
(Biesalski & Abecker, 2005)	Competency management	KOWIEN	Germany	Automotive industry
(Gualtieri & Ruffolo, 2005)	Generic Organisational HR representation	No	Italy	NA
(Terziev et al., 2005)	Generic Organisational HR representation	No	European	NA
(Bizer, 2005)	Job matching	HR-XML, KOWIEN American & German occupational classification systems, industry classification systems.	Germany	All

(García-Sánchez et al., 2006)	Job matching	No	Spain	Rural sector occupations
(Lv & Zhu, 2006)	Job matching	No	China	Unclear but seems to be related to ICT.
(Dorn et al., 2007)	Job matching and competency management	HR-XML	Austria	University student skills
(Gómez-Pérez et al., 2007)	Job matching	KOWIEN and 11 other standard classification systems	Europe	ICT
(Zemmouchi-Ghomari, 2012)	Generic Organisational HR representation	No information available to determine this.	Algeria	University staff

**Table 2 A summary of structured eRecruitment and eHumanResources approaches.**

### Issues with Existing Structured Approaches

The following issues have been identified with existing structured approaches:

1. Current structured approaches have been **designed for other purposes**. As discussed above, eHRM and eRecruitment structures, to date, have been designed for three distinct purposes: i) to provide human resources templates for the development of organisational structures; ii) to match job offers with job seekers; and iii) to compare an organisation's current human resource set of competencies with its ideal set of competencies. Whereas point i) above does not aim to gather any data around its structure, but is merely provided to assist the development of future structures related to human resources, points ii) and iii) involve temporarily gathering data for the purposes of an ad hoc matching or comparison. Structures designed for the purposes of job matching, for example, would record the occupational requirement that the job seeker has provided and match it with any job offers that have the same occupational vacancy type as that which the job seeker has indicated; the structure allows the matching to take place, and thus the process has been completed. Structures built for the purposes of competency management (point iii) above) record

the skill set information about an organisation's current human resources pool and then compares it with the organisation's ideal pool of human resources skill sets. This is another matching process to find the same type of information, generally involving either one individual or one company only. A detailed employer demand intelligence structure, on the other hand, requires a structural design that can gather, continuously analyse and permanently keep various types of information at different points in time, across geographical, industrial, occupational and sectoral borders.

2. **None of the structured approaches** described in this thesis **has been maintained or is still operational**. Some of the approaches merely made it to a 'pen on paper' stage; others were utilised by industry at some point in time, but only for a short period and have not been used again. Technological and other scientific advances have resulted in labour market changes happening at a much more frequent pace these days. Quicker travelling times, electronic communication, and machinery advances have drastically changed the ways we perform our jobs and continue to modify what we do. In order to reflect these significant changes and accurately identify current trends, it is imperative that human resource structures are kept up-to-date and in line with these changes.

3. Current structured approaches **do not include the ability to record the period of time** that an advertisement was published for. The period a vacancy remains vacant is a significant indicator of a possible skill shortage (Green et al., 1998) and hence imperative to be recorded alongside other vital skills shortage and employer demand indicators.

## **2.3 Critical Evaluation of Approaches: an Integrated View**

This section builds on the critical evaluation of existing approaches to employer demand identification that are reviewed in the first part of this chapter. Above, respective issues were discussed on an approach-by-approach basis; this section provides an integrated view of all the issues. To date, no approach has been developed on a basis that fully meets the requirements for

a detailed approach to employer demand intelligence. Some structured approach researchers have described their approaches' designs in a high level fashion but, mostly, they do not explicitly refer to the detailed approach taken to develop and refine underlying theories behind their design decisions. These structured employer demand approaches have been highly successful in the job matching and competency management communities, but they fail to address the detailed employer demand intelligence requirements discussed in this work.

The main shortcomings of the above approaches and methodologies are summarised in three areas:

- Current studies **do not share standards** and **do not meet the needs** of detailed employer demand intelligence requirements;
- The **data depth** is shallow **and the breadth** narrow;
- **Data** gathering and **analyses** are done **infrequently** and in an ad-hoc manner.

### **2.3.1 Knowledge Sharing and Different Approach Structurings**

In an employer demand identification approach, differences like various perspectives, procedures and perceptions cause misunderstanding. This increases the risk that employer demand identification outputs diverge from output stakeholders' requirements. Amalgamated knowledge sharing is needed for efficiency of discussion, to overcome cross-over settings and ultimately to eliminate misinterpretations and misunderstandings.

The only existing approaches to employer demand that were developed on the basis of sharing and reusing existing standards, and sharing the knowledge, are the structured approaches described in this chapter. By sharing structures and thus the foundations that an approach was built on with industry experts and other knowledgeable developers, an agreed view is formed between individuals and groups of people about the domain being investigated.

A study's purpose or aim is the main driver for the inclusion or exclusion of certain concepts in a study. A study's purpose therefore also influences the way its architecture is designed — in other words, each concept's positioning and the specific and varying relationships it holds with other concepts in the study. Human resource availability relating to this part of the development does not have much of an influence on the way the study's concepts and relationships are positioned (as opposed to the big role it plays in the inclusion and exclusion decisions of certain concepts).

As an example, for the existing structured approaches to job matching, it is beneficial to record the location of a specific vacancy in order to feed this information back to the job seeker on his or her search outcomes, so that they may know where they might potentially have to move to if they are successful in applying for that position. For a detailed employer demand intelligence framework, the location of a specific position would be recorded for a different purpose, to establish how many vacancies of the same type exist in a specific geographical region. More specifically, in terms of establishing whether that occupational type might be in shortage, it is imperative that the study also records whether this location falls into one of the following categories: metropolitan, regional or rural areas. Thus the relationship that 'location' has with other concepts in the study, and its own attributes and properties in each of these studies, will be different to serve different needs.

Of the four related approaches to employer demand intelligence, eRecruitment and eHumanResource structured approaches are the only ones with internal, clearly defined concept architecture. Unfortunately, due to these approaches' purposes of serving eRecruitment and eHumanResources needs, as mentioned earlier, their architectural designs do not meet the needs of a detailed employer demand intelligence framework specified in this study, and they cannot be readily adopted in such a setting.

### **2.3.2 Data Depth and Breadth**

The approaches discussed do not have the required depth and breadth of data categories needed to inform policy, funding, workforce and career development decisions. Data depth and breadth included in each study is generally determined by two factors: i) to serve the purpose that the study has been developed for; and (ii) the quantity and quality of resources available for the development of the study's framework (and ultimately funding at hand). In most cases, the project's funding overrules the wide scope of a study, and resource frameworks are only developed to meet the minimum requirements.

Most of the job-matching studies developed to date have only been designed for one type of occupation matching, mostly in ICT related fields. Efforts related to competency management structures have mostly been designed with standard larger organisational human resource needs in mind. A detailed employer demand intelligence structure needs to cater for all types of occupations, across all areas, sectors and industries, and needs to be able to drill down deep into each of these employer demand needs. In other words, an extensive slicing and dicing of a diverse range of occupational factors needs to be included in an employer demand intelligence study to fully serve the needs of informing policy, funding, workforce development, and career decisions.

### **2.3.3 Data Out of Date**

Of all the approaches studied, only a handful was developed to a stage where they could be brought into real life and used by end-user stakeholders in different fields. The manual job advertisement analyses were mainly developed for a one-off type of academic study to determine employer demand needs at a specific point in time, and they never went back to analyse the same data types at a following event.

The structured approaches, on the other hand, were developed with the intention of being available to end-users for a longer period of time. Unfortunately, maintenance requirements on

the one hand, and human resource limitations on the other hand, ultimately did not match up.

The result is that none of these approaches are available and still being used in practice.

Due to some countries' vast geographical expanses (like Australia where the country is regarded as one of the largest countries in the world per square metre), it is an extremely time-consuming process to physically visit target participants for in-person, survey-based studies. Many employers would not complete electronic or posted hardcopy surveys if these were to be mailed out to them, and hence survey administrators need to physically visit these employers in order to obtain their inputs. Keeping in mind the vast distances and number of industries and sectors involved in such a detailed employer demand study, it is nearly impossible to gather and analyse a truly representative number of participants' inputs within a one year timeframe. Due to the time and manual intensity of conducting such surveys, the data of survey-based approaches are almost always out of date by the time the study has been completed. Furthermore, the share scale of such studies means they typically do not happen on a regular basis anyway, and thus do not meet this study's requirements.

In order to ensure accurate and reliable intelligence to inform policy, funding, workforce and career development decisions, it is imperative that these studies are performed on a regular and continuous basis.

There are currently no structured approaches that are suitable for the purposes of providing detailed occupational and regional intelligence. None of the employer demand intelligence approaches provide precise and explicit processes that explain how domain specialists can use the approaches to develop a knowledge representation structure for detailed employer demand intelligence purposes. Even though some knowledge representation structures are available publicly, most do not provide detailed explanations (if any) about their choice of concepts and classifications thereof. Job matching and competency management structures have been highly successful in the eRecruitment and eHumanResources communities, but they fail to address the

need for an economic outlook on all existing occupational and employer demand needs, in either a snapshot or detailed analytical view. Unstructured approaches generally provide even less information about their choice of data gathering categories. Where these do exist, they are extremely basic. Thus, no suitable structured approaches exist that can be directly adopted for an employer demand intelligence setting.

## **2.4 Conclusion**

In this chapter, twenty-one existing employer demand identification approaches were reviewed in four categories: the first category entailed a discussion of the survey based approach, in which three current studies utilising this type of approach were discussed. The second approach discussed looked at manual job advertisement analyses, which are mainly performed in a piecemeal way in academic settings, and three specific studies were detailed. Thirdly, online-based approaches where only very high level detail is being gathered from online sources to inform employer demand intelligence were discussed; another four studies were specifically analysed and discussed in that section. The final approach category discussed was that of existing structured approaches (eRecruitment and eHumanResources). Eleven studies in total under a further three subdivisions (organisational human resource representation, job offer and job seeker matching, and competency management) completed the literature review for this final type of approach. A summary table (Table 2) was provided to compare existing structured approaches' scopes. To conclude, a detailed critical analysis of all the approaches' shortcomings was provided, demonstrating no standards and no knowledge sharing, insufficient data coverage, infrequent data gathering and dissimilar concept and relationship structures.

In summary, the current structured approaches are clearly significantly improving the eRecruitment and eHRM domains; however, none of them address the needs of a detailed employer demand intelligence framework identified by this study. Methodology for a wide scale employer demand

intelligence system is therefore required. In the next chapter, the issues as they relate to employer demand intelligence needs will be presented.

# Chapter 3 - Problem Definition

## 3.1 Introduction

In the previous chapter, a review of the work done in the fields of eRecruitment and eHumanResources was presented and discussed. It was established that while a significant amount of work has already been done in these areas, the work does not, unfortunately, meet the needs of a system that can provide detailed employer demand intelligence for policy and workforce development needs, updated career advice or curriculum development purposes.

This chapter highlights the specific goals of this research, discusses the issues pertaining to eRecruitment and eHumanResources systems and also issues that need to be addressed in order to develop a system that can provide detailed employer demand intelligence. Although similar structured systems have been developed by researchers in the eRecruitment and eHumanResources fields, as discussed in Chapter Two, they address the need to either match up vacancies with applicants or determine the human resources gap within an organisation. This research, on the other hand, attempts to provide a tool capable of delivering employer demand intelligence that can inform stakeholders of employers' current needs on both a broader and deeper scale.

The chapter concludes with an overview of the types of research methodologies that exist, and a justification for the science- and engineering-based methodology that was chosen for this research.

## 3.2 Key Concepts and Definitions

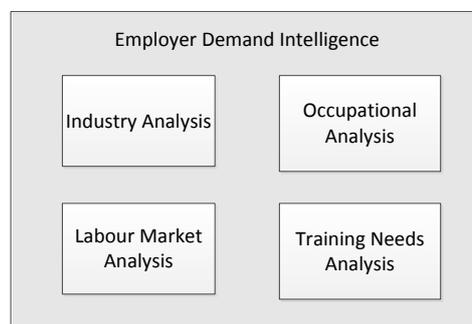
The section below will provide definitions of the concepts used throughout the thesis, in order to clarify them. It will be divided into two subgroups of clarification — concepts associated with the employer demand domain, and those associated with the methodological aspects of this research.

Although it is acknowledged that more than one definition may exist for the concepts mentioned, those most appropriate and accurate for this research's needs have been provided.

### 3.2.1 Employer Demand Domain Concepts

#### A. Employer Demand Intelligence

This research has developed a novel definition of Employer Demand Intelligence (EDI) as this term has not been coined in the literature previously. EDI refers to the analysis, understanding and sharing of information that relate to employers' needs for elements such as skills and number of employees required in a specific location to perform a task. It contains intelligence that can be used for industry, labour market, occupational and training needs analysis (Figure 4).



**Figure 4** Schematic view of Employer Demand Intelligence.

## **B. Industry Analysis**

Industry analysis pertains to determining the boundaries of the industry in focus, inspecting the determinants of the industry's activity levels and ascertaining the likely future activity and employment levels of that industry based on expected movement in the said determinants (Hayton, 1988).

## **C. Job Advertisement**

A job advertisement is an advertisement pertaining to a specific vacancy that has been advertised via some type of advertisement media to attract applications for the position (Baker, 2013).

## **D. Labour Market Analysis**

Labour market analysis comprises the assessment of occupation or job numbers in relation to the predicted and current needs for staff in comparison with the current and predicted availability of qualified persons to fill such positions (Hayton, 1988).

## **E. Occupational Analysis**

An occupational analysis examines the types of skills necessary to hold a specific occupation, the aptitude of such persons and the lifelong training and upskilling requirements necessary to stay appropriately skilled in a specific occupation (Hayton, 1988).

## **F. Skill Shortages**

Skill shortages occur in situations where employers are struggling to fill vacancies in an occupation with current level of remuneration, conditions of employment and in a realistic location (Australian Department of Employment Education Training and Youth Affairs, 1998).

## **G. Vacancy**

A vacancy is a job opening that exists at an organisation (“Vacancy”, 2013). A vacancy usually exist to perform specific duties and tasks in return for some form of remuneration.

### **3.2.2 Methodological Concepts**

#### **A. Activity**

An activity is a defined body of work that needs to be performed and includes the required input information and output information (Institute of Electrical and Electronics Engineers (IEEE), 2006, p. 6).

#### **B. Design**

A design is the process where architecture, interfaces, components and other characteristics of a system or component is defined, often linked to the development of a conceptual schema to meet a user’s requirements (Institute of Electrical and Electronics Engineers (IEEE), 1990).

#### **C. Method**

A method is the orderly process used to perform a service or build a product (Institute of Electrical and Electronics Engineers (IEEE), 1987).

#### **D. Ontology Lifecycle**

An ontology life cycle refers to the different stages through which an ontology goes during its life time (Fernandez-Lopez, 1999).

### **E. Process**

A process is a sequence of steps that are performed for a given purpose and are composed of a range of activities (Institute of Electrical and Electronics Engineers (IEEE), 1990, p. 57).

### **F. Prototype**

A prototype is a product that is developed through scientific research in order to test the product's foundations and purpose that it was built for. The product may either be some type of technology, tool or methodology that is iteratively evaluated for its verification and validity (Wongthongtham, 2006).

### **G. Schematic Heterogeneity**

Schematic heterogeneity encompasses the schematic or structural differences between various information systems (Kashyap & Sheth, 1996) (Figure 5).

### **H. Semantic Heterogeneity**

Semantic heterogeneity is caused by the lack of well-defined semantics or meaning of information items (Kashyap & Sheth, 1996) (Figure 5).

### **I. Syntactic Heterogeneity**

Syntactic heterogeneity involves using varying representation languages or models (Kashyap & Sheth, 1996) (Figure 5).

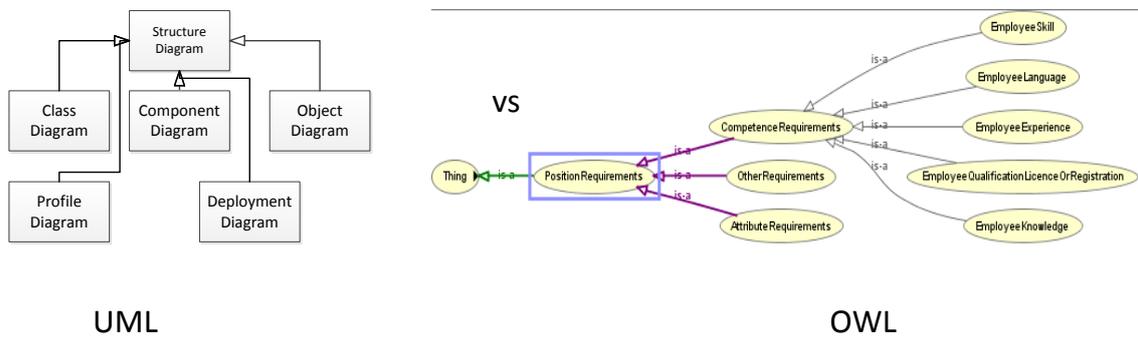
### Schematic Heterogeneity



### Semantic Heterogeneity



### Syntactic Heterogeneity



**Figure 5 Schematic, Semantic and Syntactic Heterogeneity.**

## J. Task

A task is a range of instructions that are treated as a basic unit of work (Institute of Electrical and Electronics Engineers (IEEE), 1990, p. 74).

## **K. Taxonomy**

A scheme that divides a body of knowledge and clarifies the relationships among the pieces. It is used for classifying and understanding the body of knowledge (Institute of Electrical and Electronics Engineers (IEEE), 1987).

### **3.3 Problems Facing Employer Demand Intelligence Domain**

The employer demand intelligence domain has produced a vast amount of information to date, with wide-ranging implications. Despite the great deal of employer demand intelligence available, many areas still suffer from scenarios of skill shortages, deficiency of proper training to meet employers' needs, and challenged career, funding and policy decisions. The causal factors for this overarching concern are reviewed below.

#### **3.3.1 No Exhaustive Employer Demand Intelligence Framework or Standards**

##### **Exist**

Two sub-issues exist in this category: i) the various types of employer demand related approaches (such as labour market, skill shortage, and training analysis approaches) are not standardised; and ii) employer demand-related approaches do not properly meet the needs of the detailed employer demand intelligence outcomes required for policy, training, and funding decisions.

Employer demand statistics are very poor (Williams, 1988), methods to gather such statistics notoriously unreliable (Fuller & Oxley, 1987), and approaches have not been designed appropriately to meet specific regions' needs for employer demand studies (Jennings, 1980). It is extremely difficult to compare the outcomes of any employer demand-related studies with each other. Studies have been done in silos and with different purposes or aims in mind, which

causes extreme difficulty in using those results again at a later stage, to measure how they fit with other or more recent studies' outcomes. The differences in studies' aims and unsystematic approaches used cause challenges when comparing the outcomes of studies with each other. A common frame of reference and a common type of data that are much more accessible need to be collected (Sweet, 1988).

As an example, if a study has been performed with the aim of establishing skill shortage numbers, this study could have very different results to a study which may have focused on occupational gaps and why there is a lack of uptake in training positions for certain occupational fields. Even though these two studies may utilise the same type of factors, the way the studies are formulated and analysed may cause very different and incomparable outcomes. Furthermore, even studies that have been designed with the same purpose (such as identifying skill shortages), do not have the same outcomes, as these studies have been drafted with varying perceptions of the terminology and required factors involved in establishing skill shortages.

Technological advances in industry have also caused key changes in occupational tasks. Major amendments to the training system and the way employer demand data is collected are needed (Hayton, 1988). Models need to be developed that are much more responsive to changes in the ways occupational skills are utilised — in particular industry and market settings (Curtain, 1988). The way these changes are incorporated on a continuous basis needs to happen systematically and rigorously across all employer demand intelligence studies, to ensure standardisation of this research industry.

The results of employer demand-related studies have not informed universities and training institutions sufficiently to train students with skills that employers want (Williams, 1988). It is essential to have closer partnerships between university career services, employers, charities,

and ultimately, young people looking for work, to properly establish what employers want and what makes a young person employable (Ballinger, 2012).

### **3.3.2 Limited Employer Demand Intelligence Span**

Current employer demand intelligence related studies do not have a satisfactory research span to address labour market, career, training, policy and funding decisions efficiently. The employer demand domain involves many divisions and sub divisions; for instance: skill shortage studies tend to focus on the level of remuneration and geographical aspects of vacancies on offer; industry analyses focus on broad economic factors that could influence the job market; and career-based analyses focus on personality traits and skill sets required to be employed in a certain occupational field. The differences in granularity and span of employer demand research can make the meaningful representation of employer demand intelligence very challenging.

Where some studies have attempted to cover a broad range of factors, these studies lack the required depth of data to properly examine such areas. Simple macro studies of employer demand needs can offer little assistance to the process of addressing things like skill shortages (Curtain, 1988).

There is a tremendous shortage of quality employer demand intelligence data that can properly assist decision-making and policy development (Gordon & Gordon, 1980). The type of employer demand data available to some training organisations has been identified as inadequate, especially in regards to capital planning purposes, and it has been argued that that their systems would be far better off if the quality of the employer demand data was improved (Sweet, 1988).

As an example, the Job Vacancies and Overtime survey is conducted on a very small number of data items — one for vacancies and three for overtime (Wilson, 1988). The number of data items being studied would have to be increased significantly to obtain more vacancy related information. As a result, a substantial increase in sample size would be needed in order to try

and eliminate standard errors in improved detail, but will create a considerable increase in cost (Wilson, 1988). In order to make the census and other employer demand data outputs more useful, the small range of employer demand related data collected, would need to be increased drastically — both in amount of data type and geographical distribution (Bray, 1988). Current surveys on, for instance, skill shortages, can only yield limited information and need to be supplemented by more detailed analysis about what is happening in a particular industry (Senker, 1992).

### **3.3.3 Manual Approaches are Infrequent or not Frequent Enough**

Employer demand studies do not happen on a regular or sufficiently frequent basis. The usefulness of such approaches' outcomes, such as the census data, is restricted by the infrequency and lack of timeliness associated with it (Bray, 1988). Even though censuses are historically done to estimate the number of people in each country, this tool has evolved to include a whole range of other information, such as household incomes, occupations, and unemployment numbers. In some countries, their census is undertaken only once every ten years. In Australia, the census is done every five years. It is typical for most recurring employer demand studies to be performed with great time lags in-between studies, causing the data to be grossly outdated soon after the study has been completed.

Other manual approaches to obtaining employer demand intelligence include those done on an ad hoc basis such as for individual academic studies. Typically, a researcher from a university or organisation may identify the need to perform a study into, for example, a specific occupation's skill sets or a geographical area's unemployment scenario, after which the study's results will be used for that one-off purpose, and the study will not be performed again. Even though these studies' results may be useful immediately after they have been completed, there is not much use for the data thereafter, as it becomes outdated very quickly.

## **3.4 Underlying Research Issues**

In the previous section, the main problems associated with employer demand intelligence have been identified. To address these matters, the underlying research issues will be described below.

### **3.4.1 Research Issue 1: Underlying Knowledge Representation**

A comprehensive employer demand intelligence framework requires development of an extensive model that encompasses the needs of all of the following types of possible employer demand data related studies: training analyses, labour market analyses, occupational analyses, and industry analyses. It would seem that there does not currently exist a formal framework for the classification and conceptualisation of this type of exhaustive employer demand information.

The principle of employer demand intelligence largely comprises identifying employer demand components and examining the interaction between such components. A wide variety of studies is possible, such as focusing on the geographical scale of the study (e.g. national, state wide or regionally), or the setting of the study (e.g. an industry or a single enterprise). The relations between these concepts are very complex and interrelated, involving many components and different aspects of relationships. As an example, sometimes the producer or end user of employer demand data fails to take into account important aspects, such as: distinctions between skill quantity (the number of people possessing a particular set of skills) and quality (whether the occupationally qualified incumbent matches the job skill specifications of that position); variability across industries and regions (analysis on an industry by industry or region by region basis is more meaningful than averaging over the whole workforce or nation); and

distinctions between past and expected future trends (Sweet, 1988). It is also important to ensure that emerging 'state of the art' trends are identified (Curtain, 1988).

As technology evolves even further, and employer demand information becomes more and more available to researchers through the internet, the risk of comparing and utilising study outcomes that are based on vast differences of terminology and conceptual perceptions (perceived to be the same ideas) is increasing. It is almost impossible for a single employer demand researcher to manually represent the information contained in all the different sources found online within a structured framework of concepts and relations. As an example, a Google search on the phrase 'employer demand studies' returned almost 21 million results in August 2013. As the amount of data available online grows at an alarming rate each day, it becomes harder for researchers to integrate and analyse newly presented employer demand data that are based on agreed terminology, standards, perceptions and similar underlying aims of studies.

The creation of new employer demand data in non-standardised ways also adds to the issue of the heterogeneity of information available to employer demand researchers. The autonomy of individual employer demand studies allows creation of data outputs that have different content and dissimilar formats, such as natural language or Extensible Markup Language (XML). These varying information formats allow semantic, schematic and syntactic heterogeneity (Kashyap & Sheth, 1996) (definitions provided in section 3.2.2). There is a vast need to promote the interoperability of employer demand resources and provide a platform where information can be shared and utilised efficiently.

### **3.4.2 Research Issue 2: Occupation Specific Data and its Querying**

To populate employer demand studies with sufficient breadth and depth of information so that it is useful for stakeholders of these studies, the data collection instrument needs to have substantially more data points, data storing capacity, and a structured hierarchy. The common

denominator for employer demand-related studies, is based on the occupational field. As such, for data to be meaningful to any employer demand domain stakeholder, a schematic representation of an employer demand data collection tool needs to be centered on occupational clusters. Major investigations of occupational skill needs in each industry have to be undertaken to ensure proper assessments are achieved (Curtain, 1988). This will allow the occupation template to be populated with real-time, occupation-specific data after which specific instances and scenarios can be queried to exceptional levels of granularity.

Providing this level of granularity, structured with standardized occupational concepts on the top level of each dataset, will also provide employers the ability to view an industry in a conglomerated view (e.g., looking at the state of all nursing vacancies with an eagle-eye view, or assessing all engineers' skill requirements for a specific region). The development of a structured hierarchy, where lower level, more specific concepts can be generalised to broader, high level concepts, will allow the continuous and systematic population of data for different time periods into one overarching dataset, covering all necessary data points required for a detailed employer demand intelligence resource.

### **3.4.3 Research Issue 3: Expensive and Time Intensive Approaches**

Most current labour market, economic outlook, training and skill shortage analysis studies are performed via some type of manual intensive option. The most common approach used by those wishing to study these domains consists of surveying the target market. Employer demand surveys can take the form of: i) paper-based postal or electronically emailed surveys (self-administered surveys); or ii) telephone surveys or in-person surveys (interview-based surveys) with the self-administered option having been the favourite among most study conductors to date.

The advantages of self-administered employer demand surveys are: i) self-administered surveys can be completed in the respondent's own time; and ii) self-administered surveys allow for a larger set of questions to be asked of the respondent.

The advantage of an interview-based employer demand survey is that the survey administrator can clarify the survey questions to the respondents, and a further advantage of employer demand telephone interviews is that survey administrators do not need to travel to respondents in order to hold the interview with them.

There are, however, disadvantages in utilising employer demand survey-based approaches. Firstly, survey-based approaches are resource intensive, and as such, expensive. Whether the survey is self-administered by the respondent, or administered by staff, it requires at least one, usually two persons to participate in the process of completing the questionnaire. This could mean that a survey consisting of, for example, 30 short survey-based questions could potentially take at least 30 minutes per questionnaire, totaling one man hour per survey, if both a respondent and a survey administrator are involved. For a study that would consist of targeting at least 50 completed surveys, this would mean that there are at least 50 man hours spent only on the completion of the survey.

Secondly, where in-person interviews are conducted in order to try and ensure the highest return of completed questionnaires, survey staff need to travel to each respondent. If a typical survey sample of employers is targeted, it would mean that staff would have to either travel to companies in metropolitan areas, where traffic and parking cause a time-intensive process, or to regional companies, where distance would impact extensively on the journey time. For a survey of the same dimensions as described above, that is, where 50 completed surveys would be required for a study, with an additional two hours travelling to and from each respondent is taken into consideration, it could add an additional 100 man hours to the execution of the study,

bringing the sub-total of required man hours for this short survey with a small sample size to 150. In a country such as Australia, where travel distances are vast, typical travel distances to reach a company within the same state could mean at least two full days' air travel to and from the organisation's location.

Lastly, employer demand survey-based studies require substantial human intervention to set up the questionnaire, enter, analyse and report on the results of the survey. It can be estimated that a short answer questionnaire with 30 questions could take five hours to draft (bringing the sub total to 155 man hours). Furthermore, if the survey was conducted via a paper-based questionnaire, either through self-administration or where a survey administrator noted down the answers of respondents, this step would involve yet another time commitment. This could require an additional one hour per questionnaire simply to enter the answer sheet data into the study's electronic database, bringing the employer demand survey man hour total to 205 hours in this example. The actual analysis and reporting of the survey data would also need manual intervention, as responses need to be collated into reporting measurements such as medians of question answers, reading and analysing qualitative type responses, and providing intelligence as an output of the data entered. If another hour per questionnaire is allocated to this task as an estimate for the example employer demand survey, it would mean that the small survey would need at least a total of 255 man hours to be conducted (Table 3).

Activity		Estimated man hours
Completing questionnaire		50
Travel time		100
Survey administration	Survey drafting	5
	Data entry	50
	Data analysis	50
<b>Total</b>		<b>255</b>

**Table 3** Estimated man hours required for an example short employer demand survey.

Over and above the manual and time intensive (and as such very costly) aspect of employer demand survey approaches, these tactics could potentially also incur large expenses due to the travel that may need to be done by survey administrators (for in-person surveys), and to fund the postage, printing and paper requirements for each survey.

The research issue identified here highlights the fact that a more efficient time- and resource-based approach needs to be developed so that analysts can quickly and easily collect employer demand data and process it to provide employer demand intelligence.

Once the employer demand intelligence tool has been developed, it will need to be evaluated to demonstrate its feasibility, in order to allow proof of the claims made that it will address the issues discussed above.

### **3.5 Summary of Thesis Goal**

Recapping the main points of the problem definition section, the purpose of this thesis is to create an employer demand intelligence framework that is able to:

1. Function on a basis that is supported by underlying knowledge representation;
2. Provide data instances that are occupation specific in the first place, then linked to a wide range of other employer demand intelligence variables to ensure broad coverage of the domain;
3. Semi-automatically gather, analyse and report on employer demand intelligence and ensure its engineering soundness and practical readiness for industry uptake.

Although there will always be further issues that still need to be met and resolved, the issues mentioned in this chapter form the basis upon which the proposed model will be built, as a foundation for future development in this domain. Therefore, the next section will focus on the key questions that need to be answered through this research.

### **3.6 Key Research Questions**

Three key issues have been defined in the sections above. In order to address these three issues, the following research questions have been developed:

1. How can employer demand concepts be modelled in order to represent the employer demand domain? Specifically, how can complex employer demand intelligence concepts be represented in terms of concepts and relationships?
2. How can instances be populated that are occupation-specific, linking them to a wide range of other employer demand intelligence variables?
3. How can a semi-automatic framework be developed that gathers, analyses and reports on employer demand intelligence? How can the employer demand intelligence framework be evaluated for industry readiness so that it can serve as the foundation for future employer demand intelligence expansion?

## **3.7 Research Methodology**

The overarching objective of this thesis is to develop a framework for employer demand research which ensures that ample breadth and depth of employer demand intelligence concepts are captured. This vast framework will require a semi-automated system to ensure continuous population of up-to-date employer demand intelligence. In order to proceed with such task, a systematic and logical approach needs to be followed. This will ensure that the framework's development is based on a quality scientific method, which will be discussed in the below section.

### **3.7.1 Overview of Research Methods**

Research methods can be grouped under two categories: the social sciences- and science- and engineering-based methods. A social science approach consists of researching social issues — concepts and ideas. Where social science research can identify how well methodology is accepted or not, science and engineering approaches involve developing practical solutions for identified issues. From an engineering perspective, science is viewed as knowledge accumulated from use of a scientific method to further understanding, where in contrast, engineering not only seeks understanding, but also generates systems and devices capable of meeting human needs (Kiplagat & Rotich, 2007).

### **3.7.2 Choice of Science- and Engineering-based Research Method**

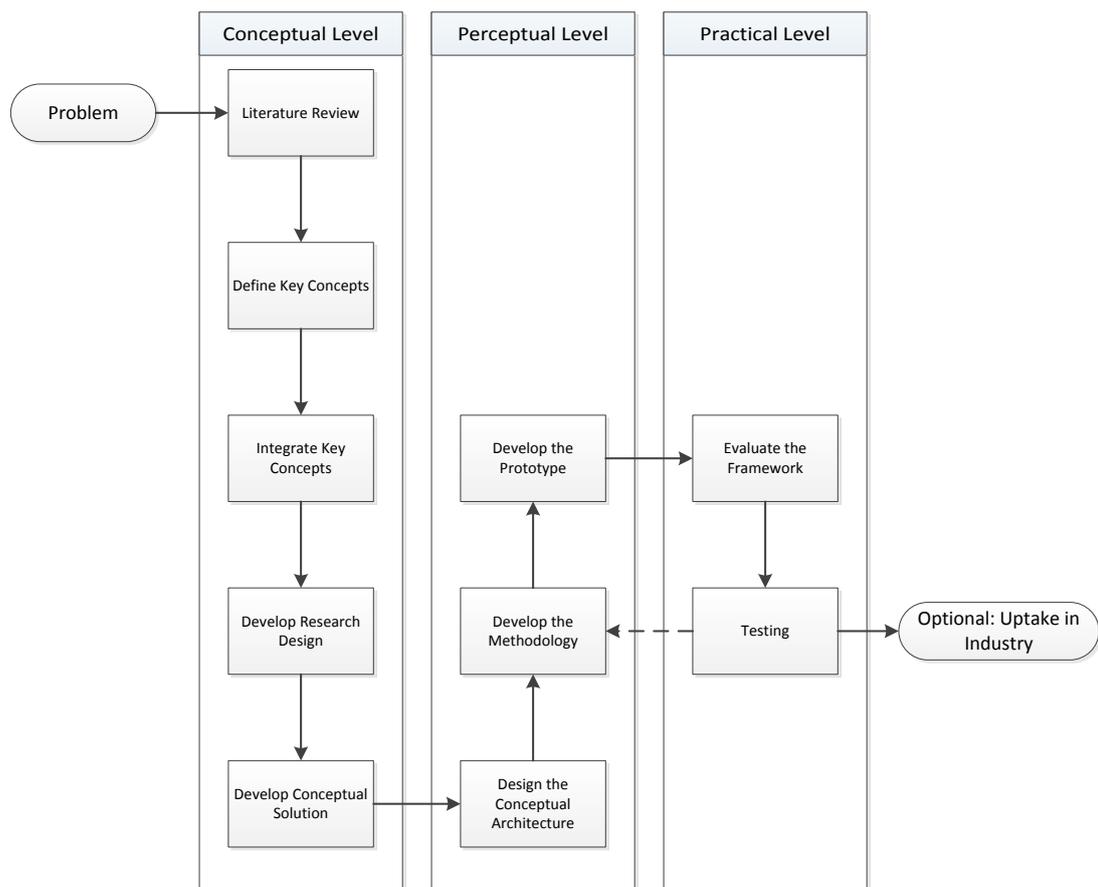
Research methods in both social science and science and engineering are concerned with improving the quality of life; however, it is technology, as an outgrowth of science, that fuels the industrial engine; and engineers, not scientists, who make technology happen (Kiplagat & Rotich, 2007).

The aim of this research is to develop an employer demand intelligence framework by utilising the science- and engineering-based approach, which aims to develop a scientific solution to the identified problems in this chapter by developing a theoretical framework, and later confirming it (Galliers, 1992; Hevner, March, Park, & Ram, 2004; Nunamaker, Chen, & Purdin, 1991). Of the various research methodologies available under this approach, it was decided to use a Systems Development approach, which is a combination of the following research steps (Nunamaker et al., 1991):

1. Envisaging a system;
2. Developing key concepts and theories;
3. Developing a conceptual framework;
4. Building the system;
5. Testing and validating the system.

Each of the above-mentioned steps has various sub-steps in it, which can be classified into three broad categories, as listed below (Galliers, 1992) and depicted in Figure 6:

1. **Conceptual** (identification of research objective, research issues and challenges);
2. **Perceptual** (development of conceptual framework to underpin system analysis, modelling, designing and implementation);
3. **Practical** (implementation, validation and assessment of the developed framework) phase of the research.



**Figure 6 Science and engineering based research approach.**

### 3.8 Conclusion

In this chapter, the main concepts involved in the employer demand intelligence research are provided as an introduction to the defining terminology used. The chapter then looks at the main problems facing the employer demand intelligence domain from a practical and user-based perspective, after which it describes the research issues underlying each of these problems.

In conclusion, a science- and engineering-based research methodology is presented as the preferred option to address these issues and allow the development of a practical, systems-based approach for the employer demand intelligence framework.

In the next chapter, a detailed solution proposal and solution requirements for each of the issues identified in this chapter are offered.

# Chapter 4 - Solution Proposal

## 4.1 Introduction

In Chapter Three, the problems facing the employer demand intelligence domain, and their underlying research issues, were discussed: i) underlying knowledge representation; ii) occupation specific data and its querying; iii) expensive and time intensive approaches; and iv) employer demand intelligence adoption.

This chapter addresses the identified issues and suggests solution proposals to each of the identified research issues: i) employer demand ontology; ii) employer demand ontology instance population; iii) semi-automatic employer demand intelligence processing; and iv) evaluation of the employer demand intelligence tool.

The main technique that enables the development of an overarching solution to these underlying research issues is a knowledge representation which enables employer demand intelligence to be presented formally. Therefore, this chapter will justify the decision to use a knowledge representation as the basis for the overarching solution in this thesis, then provide the conceptual framework of the research, and finally provide a detailed explanation of each individual solution requirement proposed in the research.

Chapter Four has been dedicated to detail the specific functionality that ontology provides, and as such, why it has been chosen as the framework to the issues of this research and thus being a superior choice to that of data mining to address the needs raised in this research. I.e. the approach has been based on ontology to provide a standardised model that allows for a shared understanding of the domain. Data analytics and data mining algorithms do not provide this function. Chapter Three specifically discussed the fact that there is a need for an underlying

knowledge representation in this domain due as other approaches (like data analytics) do not provide an exhaustive employer demand intelligence framework based on an agreed understanding of the knowledge domain.

## **4.2 Solution Overview**

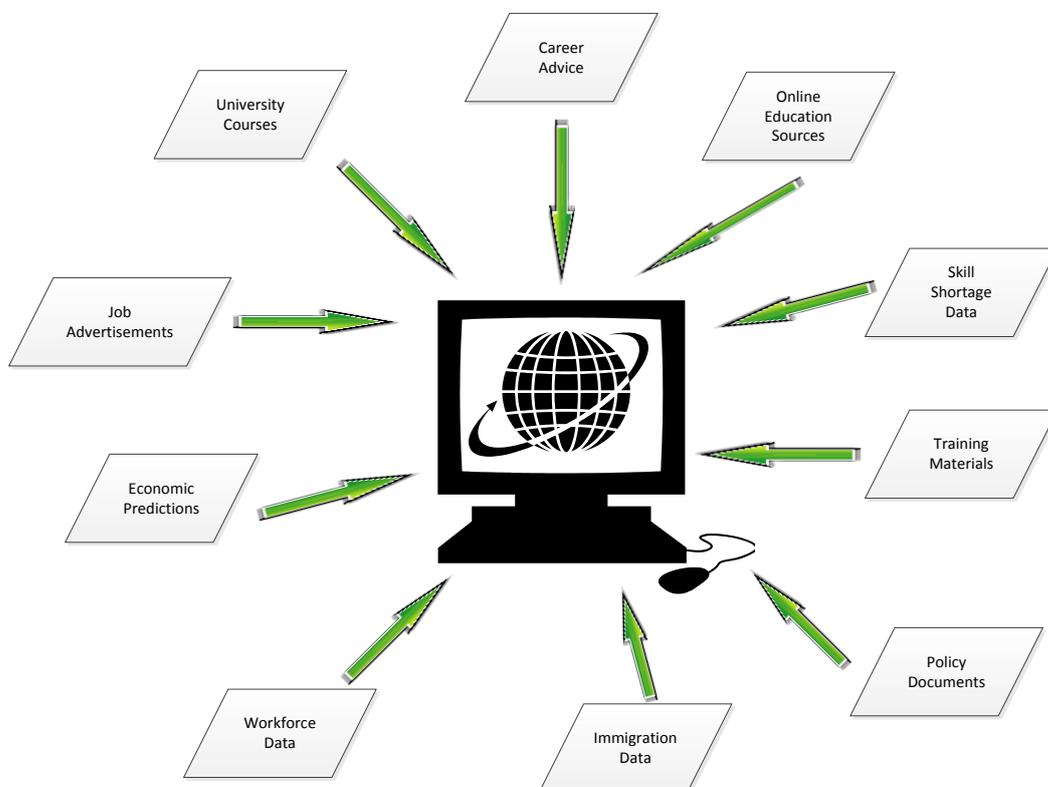
The aim of this research is to develop a detailed employer demand intelligence tool (EDIT) by using a science- and engineering-based research approach, as discussed in Chapter Three. A science- and engineering-based methodology aims to develop a scientific solution to the identified problem through first developing a theoretical framework and later confirming it (Galliers, 1992; Hevner, et al., 2004; Nunamaker, et al., 1991). The specific approach chosen for this research is a Systems Development-based approach (Nunamaker, et al., 1991) as discussed in Chapter Three, where a knowledge representation is the underlying basis upon which the software tool will be built (discussed in this chapter). To sufficiently discuss the proposed conceptual framework, it is necessary to first reflect on a knowledge representation as a scientific solution. As such, the next section is dedicated to knowledge representation fundamentals. Following this, the conceptual framework of the proposed solution is covered to conclude this section of the chapter.

### **4.2.1 Tsunami of Data**

The body of existing knowledge in the world is rapidly increasing every day. In the past, this knowledge was filed in physical paper-based databases such as academic journals, books, reports, record keeping journals, company files and newspapers. These days, most information is stored in some kind of electronic domain and shared via various digital channels such as emails, on a company's website, in e-journal repositories, through sharing of USB drives, and network sharing options. The vast increases in the number of people and domains contributing to knowledge discovery and the advanced methods and technological tools used in knowledge

discovery, retrieval and sharing has significantly influenced the size, availability and quality of this body of knowledge. Figure 7 shows how there is an array of information on the web that is generated from multiple sources, for instance, in the employer demand-related domains, information can be generated by web users such as career advisors, university personnel and policy developers, to name but a few.

The availability of this body of knowledge through the e-world, has provided for many more people being able to access it in a timely, cost efficient, environmentally friendly and reliable manner. However, the mostly unstructured nature of the information available on the World Wide Web makes the searching for and collection of relevant information for a specific goal an onerous and almost unmanageable task. The value of available information needs to be maximised. As such, the management and organisation of the available information online and in other repositories are important factors.



**Figure 7** Information on the web covers numerous and varying employer demand related domains.

In order to successfully retrieve, access and share information from a collage of resources, the following difficulties need to be overcome (Wongthongtham, 2006):

- A daily increase in distributed and heterogeneous types of information;
- There are no underlying knowledge bases available to enable shared understanding of concepts;
- Databases and other information sources are autonomous, diverse and dynamic.

To this extent, ontology is proposed as the overarching scientific basis upon which to build the proposed employer demand intelligence framework as ontology's uses are (Gruninger & Lee, 2002; Uschold & Gruninger, 1996):

- For communication between implemented computational systems, humans, and a combination of humans and computational systems;
- For computational inference, to internally represent and manipulate plans and analyse aspects such as internal structures and algorithms; and
- To reuse and organise new and existing knowledge.

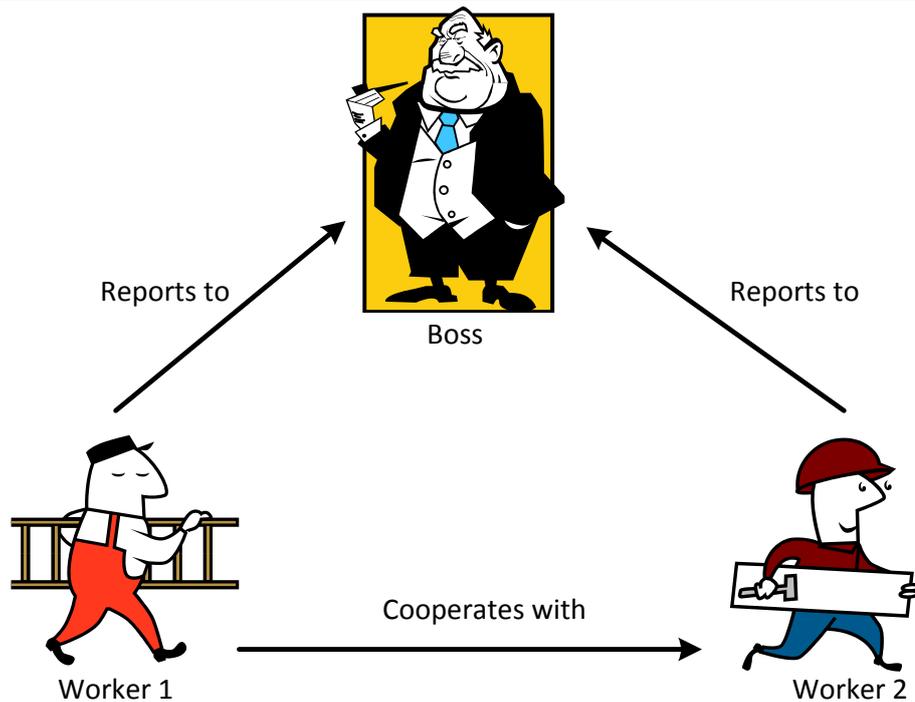
To understand more about ontology's fundamentals, the next section is dedicated to the origin, definition and types of ontology in existence.

#### **4.2.2 Ontology Definition and Types**

Ontology originates from the metaphysics and philosophical sciences, when Aristotle first introduced it as a philosophical domain to investigate the existence of being (Corazzon, 2013). However, the concept of ontology has been introduced to the computer science and artificial intelligence domains during the latter part of the previous century, and grown significantly in its usefulness and uptake in this domain since then.

Among the computer science and artificial intelligence circles, an ontology is most commonly known as ‘a formal and explicit specification of a conceptualisation’ (T. R. Gruber, 1993), which allows information owners to annotate their material (Berners-Lee, Hendler, Lassila, 2001). Several definitions of ontology have been provided to date (Maedche, 2003; Neches, Fikes, Gruber, Senator, & Swarout, 1991; Studer, Benjamins, & Fensel, 1998); however the core of them all relate to the usefulness of ontology in technology for the representation and sharing of knowledge about a domain, by modelling concepts and the relationships among those concepts within the domain (T. Gruber, n.d.). Figure 8 shows an example of the relationships in a small business that can be modelled by using ontology: ‘Worker one’, ‘Worker two’ and ‘Boss’ are all concepts from this domain. ‘Worker one’s’ relationship with ‘Worker two’ is ‘cooperates with’, i.e. ‘Worker one’ cooperates with ‘Worker two’ in this business. Both ‘Worker one’ and ‘Worker two’ report to ‘Boss’, and as such have relationships of ‘reports to’ with ‘Boss’. The four main concepts of ontology, however, are:

- A domain conceptualisation which is an abstract model of a specific view of a domain’s contents and the relationships among them;
- Formal, which relates to a form of knowledge representation that is machine readable and mathematically formulated;
- Explicit, which provides a precise and clear definition of concepts and their respective relationships utilised in the ontology;
- Shared, meaning that the knowledge represented through an ontology is a form on consensual knowledge and that there is agreement among the users about the concepts and their relationships’ intended meanings.



**Figure 8** A small part of a company's workers and manager, and the relationships they have with each other.

With the acceptance that an ontology is based on the notion of a shared conceptualisation, users, designers and domain experts need to agree on the specific knowledge contained in any ontology to deem the ontology usable. To assist in this task, different layers of ontology have been introduced to eliminate disagreement and foster the re-use of ontologies as opposed to reinventing new ontologies each time. The three layers of ontology knowledge, based on their generality, are (Chang, Dillon, & Hussain, 2006; Guarino, Oberle, & Staab, 2009):

- Top level ontologies, which relates to general concepts that are domain independent (e.g. space, time etc.)
- Domain ontologies, which capture the general knowledge of a specific domain or task. In this research, it is the employer demand domain where general concepts such as competence and geographic concepts are captured.
- Application ontologies, which represent concepts that capture the knowledge necessary for a specific task or application within a domain. In this research, it is the occupation employer

demand domain relating to specific job offers that are included as concepts of this layer.

Figure 9 shows these different layers of ontologies graphically as described here.



**Figure 9** The three layers of ontology generalisation (Villazón-Terrazas et al., 2011, p. 13617).

With the intended definition of an ontology defined as it is used in this work, it is timely to reflect on the specific characteristics that ontologies hold in the next section.

### 4.2.3 Ontology Characteristics

Ontologies are characterised by a range of factors, with the most important aspects provided below (Sidhu, 2008):

- Ontologies are used in order to describe a certain domain;
- Concepts are regarded as abstract or general ideas;
- Concepts and their relationships are formally and explicitly defined through using axioms and definitions, presented in a formal language that is computer understandable;
- Concepts are linked through their relationships with each other, and modelled in a hierarchical or non-hierarchical manner as a semantic network;
- Ontology users have agreement about the terminology and concepts included in an ontology;
- A concept is any unit of thought;

- A term is a lexical representation of a concept;
- Terms can exist in a variety of relationships with each other;
- The knowledge classification can evolve over time as the domain adapts in everyday life;
- Different terms may be combined to form new meanings.

It is clear that an ontology provides a window into a specific area of interest, represented in such fashion that it fulfills the purpose set out for it in the way the concepts and their relationships are modelled with each other. The ontology users agree that the content of the ontology, the ontology's positioning in the hierarchy, and the relationships that the concepts have with each other, are a true representation of that specific domain. As such, the ontology can be reused in various other domains, and extended to be adapted into or extend other examinations of a domain. Due to these characteristics, ontology has been chosen as the preferred scientific basis upon which the employer demand intelligence tool will be founded, as explained in the next section — conceptual framework.

### **4.3 Conceptual Framework**

Most information resources exist autonomously and independently from one another (Goble, 2003). For the most part, they are developed for a single purpose, where each independent knowledge domain produces its own data and information sources.

This thesis serves to introduce an Employer Demand Intelligence Framework (EDIF) to assist the identification and analysis of real-time employer demand data from a range of sources. Information to be collected include elements of employer demand intelligence such as skill shortage indicators, workforce development information, skill set requirements for specific occupations, benefits offered by employers and the types of occupations which are in demand in specific geographical areas. Figure 10 provides a graphical view of the conceptual framework proposed in this research and is described under the corresponding sections below.

### **4.3.1 Data Input**

The data is sourced from a variety of employer demand intelligence related websites, and are contained in some type of textual format depending on the source the data originates from, such as online job board advertisements in html format.

### **4.3.2 Online Resources**

The EDIF proposes to analyse employer demand data gathered online from a multitude of sources such as job boards, company websites, blogs, social media (like Facebook, LinkedIn and Monster) and electronic academic papers. The EDIF will crawl online sources, identify, gather and store relevant original data, analyse and process this data into intelligence, and subsequently provide this up-to-date and comprehensive intelligence to users of the EDIF when requested.

### **4.3.3 Employer Demand Intelligence Framework**

#### **A. Data Collection Process**

The EDIT will prompt the process to access the online resources in order to collect the relevant employer demand data from those employer demand data sources that have been identified as valuable for the EDIF stakeholders. The process can happen as little or as often as required.

#### **B. Employer Demand Intelligence Tool**

The EDIT instigates the collection of the data from the websites (as described above), processes it through a range of text identification and language processing steps (detailed in Chapter 7), and then populates the information into the EDO Repository described below.

#### **C. Employer Demand Ontology Repository**

The EDO Repository consists of both the EDO and the Employer Demand Dataset (EDD).

### **Employer Demand Ontology**

The EDO holds the employer demand intelligence concepts and some non-occupation specific records that allow the EDIT to look for data in the online resources that match the EDO concepts.

### **Employer Demand Dataset**

The EDD holds the data relating to each time period when information is collected from the sources, and subsequently populated into this dataset for analysis by the employer demand intelligence stakeholders.

### **D. Interface for Employer Demand Intelligence Framework**

The EDIF interface is hosted on the web, where stakeholders can access the EDIF from anywhere in the world to obtain results on numerous types of queries they may have for a given Employer Demand Dataset.

#### **4.3.4 Data Output**

Overarching, the EDIF will be a one-stop online user interface for employer demand intelligence stakeholders such as policy developers, career advisors, students, curriculum developers, and employers. Users will be able to enter a query into the EDIF interface online based on their needs, and be presented with a result that provides current, reliable and comprehensive information that was located from a diverse range of information sources in a matter of seconds.

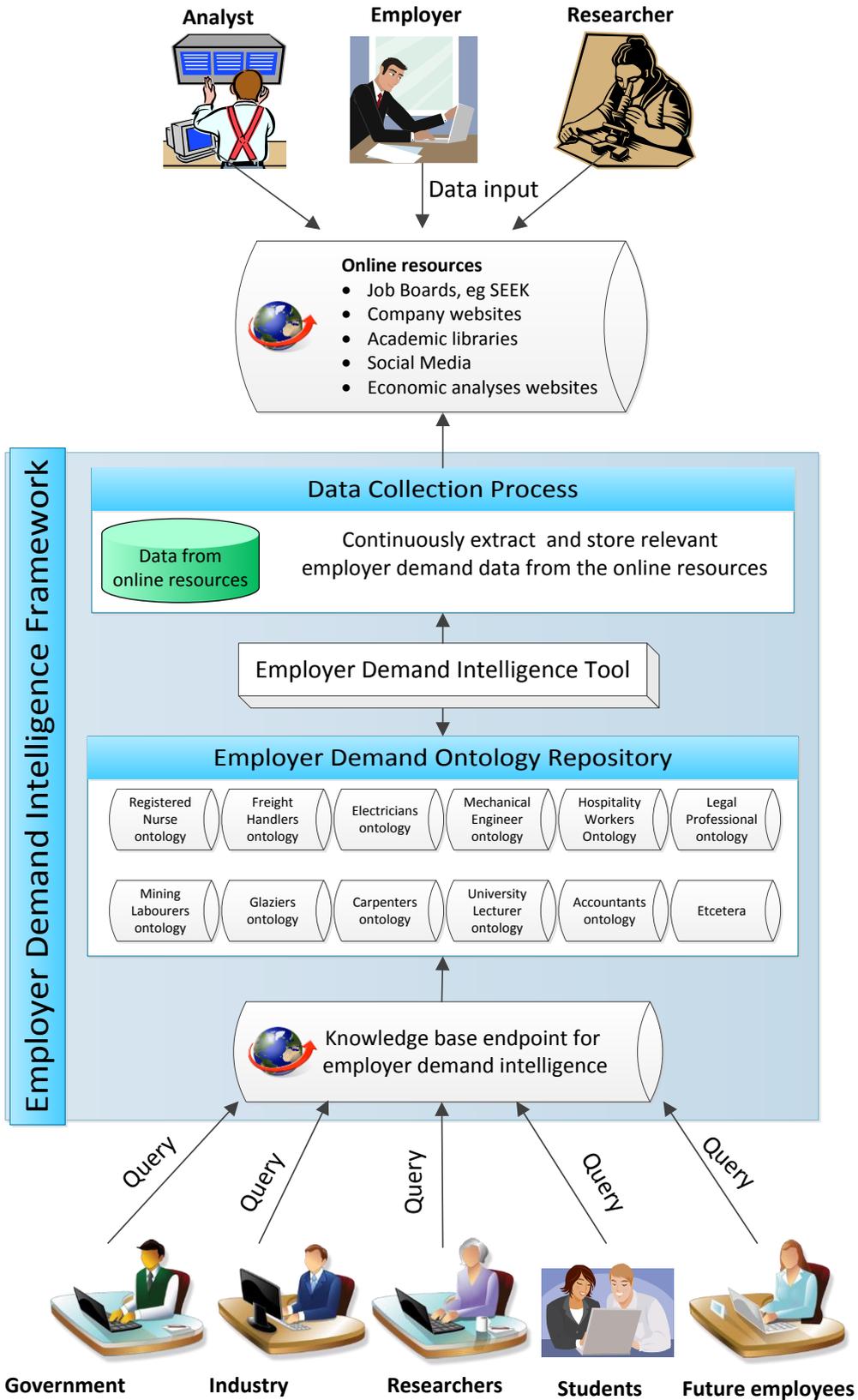


Figure 10 Employer Demand Intelligence Framework conceptual framework.

For the purposes of this PhD research (stage one of the EDIF's development), due to limited time and resources, the EDIF will only be developed to prototype with a limited range of sample features as first listed and then described in more detail below; a greater scope is intended as future work (detailed in Chapter Nine). It will be up to the research team who progress this research, to decide which development of the EDIF they will deem as stage two of the EDIF's development.

- Occupation type: Midwifery and Nursing Professionals
- Geographical location: The state of Western Australia
- Employer demand data source: SEEK<sup>25</sup> job board
- Data collection period: For EDO development- one month (January 2011), for EDIT development, another month (March 2013).
- EDIT to be semi-automatic for this research, full automisation intended as future work.
- EDIF interface is a SPARQL (a recursive acronym for **SPARQL Protocol and RDF Query Language**) endpoint at this stage of the research.

It is well known that job advertisements contain rich data pertaining to employer needs (Rawling, 1988). Due to the increased use of the internet and online advertising strategies, job advertisements have been increasingly advertised electronically through mediums such as company websites, online newspapers and job boards. The EDIF prototype proposed for this research has been developed to analyse data collected from the job board SEEK, for the occupations under the category Midwifery and Nursing Professionals in Western Australia (refer to the Registered Nurse Ontology that is indicated in Figure 9 in the Employer Demand Ontology Repository). As discussed, other occupational types are intended for future work.

---

<sup>25</sup> <http://www.seek.com.au>

The conceptual framework as described above and represented in Figure 10, is the overarching solution to the research issues that were identified in Chapter Three. Firstly, a comprehensive employer demand ontology (EDO) has been manually developed as the underlying knowledge representation for the EDIF. Secondly, the EDO has been manually populated with instances of Midwifery and Nursing job advertisements found on the job board SEEK, for a one month period. Thirdly, a semi-automated employer demand intelligence tool (EDIT) has been developed. It is intended that the EDIT's automation will be further enhanced in future work. Lastly, all three facets of the EDIF (the EDO framework, the EDO repository (which includes both the EDO and the EDD) and the EDIT) were evaluated for usability, quality and usefulness. It is worthwhile noting that the two manual steps (EDO development and population) happened before development of the EDIT, because it provided the EDIT with the ontological reference framework, as well as a point of reference for incremental measurement and evaluation during the EDIT developmental process.

Each of these proposed solutions will be described in more detail in the next section.

## **4.4 Detailed Solution Requirements**

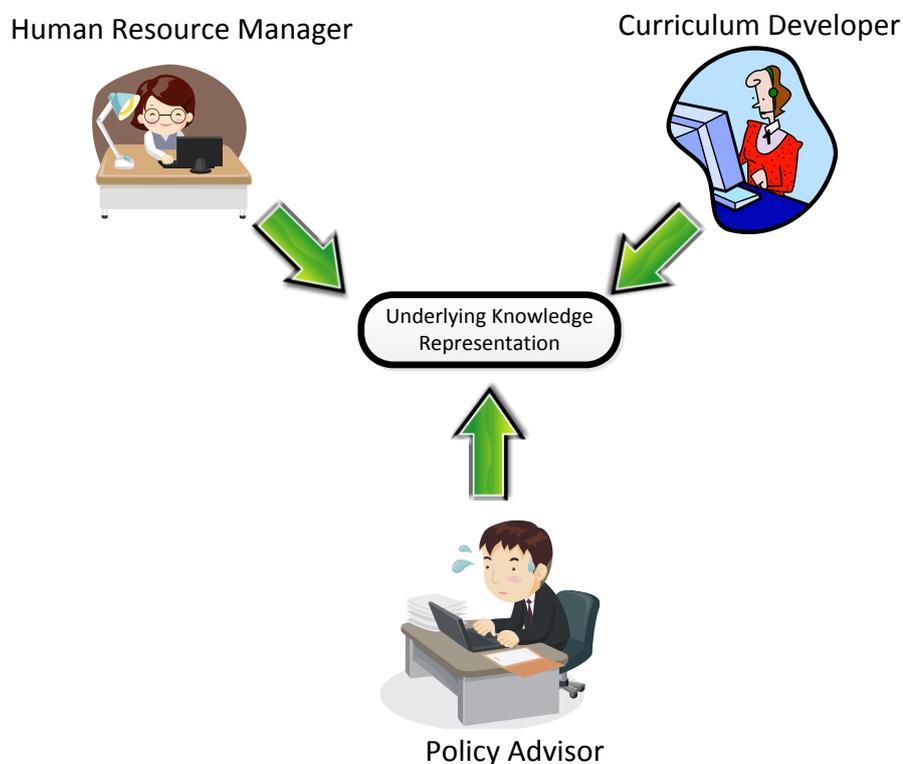
### **4.4.1 Requirement 1: Employer Demand Ontology**

*This solution addresses the research issue of needing an underlying knowledge representation framework for detailed employer demand intelligence.*

Technological advancement and a growing number of interconnections between resources and organisations are increasing the importance of knowledge integration methodologies (Hadzic, Wongthongtham, Dillon, & Chang, 2009). As such, data needs to be standardised and structured to allow efficient storage, acquisition and organisation of data from various types of employer demand resources found online.

In order to do this, a unifying ontological framework that represents a shared understanding of the employer demand domain knowledge will enable efficient and intelligent management of all its existing information sources across the domain. Figure 11 provides an example of different types of information stakeholders entering and extracting data from an underlying knowledge representation. The ontological framework will overcome the issue of information heterogeneity to a large extent by ensuring syntactic, schematic and semantic agreement among employer demand intelligence stakeholders (definitions provided in Chapter Three).

Another key issue of current employer demand related studies has been highlighted as the need to integrate all labour market, industry, occupational, and training needs analyses into a single point of reference (Hayton, 1988). Portraying a domain in ontological terms facilitates intelligent access to the information (Benjamins, Fensel, & Gomez-Perez, 1998) and assists in clearing up ambiguities in the terms used within the domain (Wongthongtham, 2006).



**Figure 11** The underlying knowledge representation for employer demand intelligence.

The requirement of an underlying representation framework has been addressed through the development of a comprehensive, two-tiered, underlying Australian knowledge representation template for employer demand intelligence consistency, comparison, breadth and depth. The employer demand ontology model consists of two tiers:

- Tier one: High level non-specific occupation type template for general concepts that are applicable to all occupation types; and
- Tier two: Detailed occupation-specific subsection template relevant to each separate occupation as defined under the ANZSCO.

The reason the employer demand ontology framework has been tiered is to allow for the EDO to be developed and populated incrementally; that is, whilst one occupation's specific ontology concepts and relationships are still being developed, the EDO can be used in parallel to it for the occupations on which development has already been completed. This is because of the current expensive and time intensive process involved in developing comprehensive domain specific ontologies manually. Automatic and semi-automatic ontology development in Information Systems is not yet mature and will take some time before it will be confidently used in ontology creation. Until such time, ontologies are still being created manually from scratch or through reusing other existing ontologies.

#### **4.4.2 Requirement 2: Employer Demand Ontology Instance Population**

*This solution addresses the research issue of providing occupation-specific data for detailed employer demand intelligence datasets.*

Ontologies help both people and machines to communicate succinctly by exchanging semantics as well as syntax (García-Sánchez et al., 2006). Populating an ontology with instances is one of the evaluative measures of an ontology (Chen, 2013). More specifically, for this research, this method:

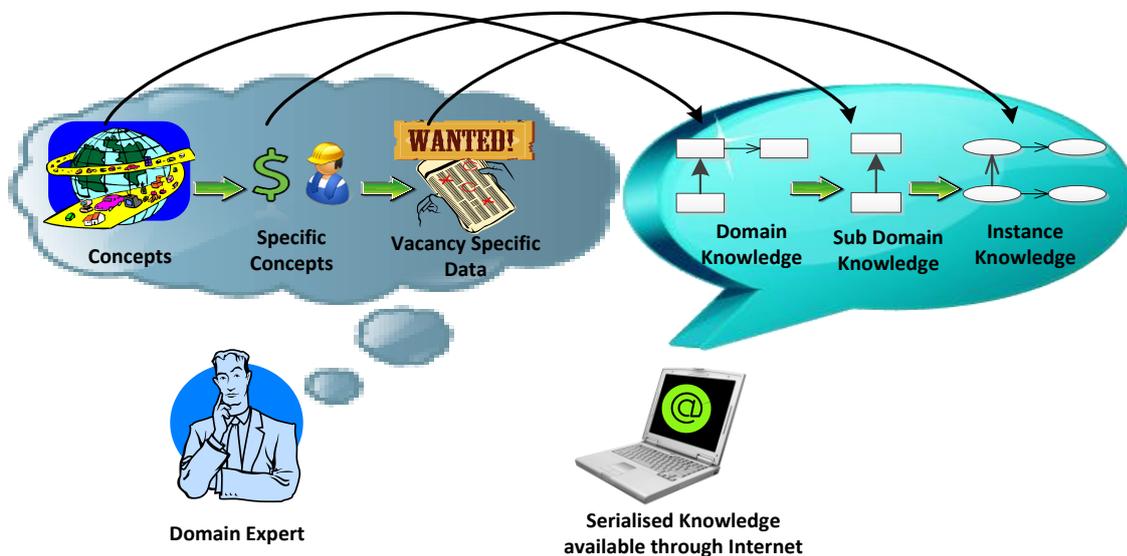
1. Proves the viability of the solution proposed in solution 1 above by providing an outcome that was built based on the EDO framework;
2. Provides the detailed occupation-specific instances, addressing the need for a more comprehensive source of employer demand intelligence for industry, academia and others, as opposed to having only the scant and silo-driven existing datasets available for analysis;
3. Delivers the point of reference that is necessary for the development of EDIT (proposed in solution 3 below) and provides a measurement against which EDIT can be tested.

Figure 12 shows an abstract view of the employer demand intelligence representation. The domain knowledge represents all employment concepts; the sub-domain knowledge represents specific concepts related to occupational employer demand data, while the instance employer demand knowledge represents concepts from the occupation specific employment domain for particular job advertisements. For each job advertisement, there is information or actual data, such as geographical locations and requirements for the specific job advertisement. The job advertisement information is needed to define instance knowledge in the EDO. The instance population will happen through collecting one month's job advertisements from the job board SEEK. The specific occupations that the data will be collected for, will be for the sixteen Nursing and Midwifery occupations from the ANZSCO:

- Midwives
- Nurse Educators
- Nurse Researchers
- Nurse Managers
- Nurse Practitioners
- Registered Nurses (Aged Care)
- Registered Nurse (Child and Family Health)

- Registered Nurse (Community Health)
- Registered Nurse (Critical Care and Emergency)
- Registered Nurse (Developmental Disability)
- Registered Nurse (Medical)
- Registered Nurse (Medical Practice)
- Registered Nurse (Mental Health)
- Registered Nurse (Perioperative)
- Registered Nurse (Surgical)
- Registered Nurse not elsewhere classified

The geographical area that will be covered will be that of Western Australia, Australia's largest geographical state, where Perth is the capital city. An extension of this geographical parameter is intended as future work to also include other Australian states, and eventually other countries too.



**Figure 12** Schematic overview of employer demand intelligence representation.

### **4.4.3 Requirement 3: Semi-Automatic Employer Demand Intelligence Tool**

*This solution addresses the research issue of eliminating time and manual intensive approaches to obtain detailed and reliable employer demand intelligence.*

A main issue of efficient information sharing is the dynamic aspect of the internet — what was available for scrutiny one minute ago, may not be available the next. Data are continuously being added and removed without warning, creating potential gaps in knowledge gathering attempts. The proposed employer demand intelligence prototype will address this issues by regularly and automatically retrieving relevant information from the web and storing it in the EDIT dataset for on-the-spot and later analysis. This will ensure a record of original data sources is kept. Where, when and how this information was sourced will also be available, guaranteeing a record trail of all sources of information used and allowing cross-checking to happen at any stage into the future.

The EDIT will also allow sharing and retrieval of employer demand data across diverse knowledge platforms by collecting, analysing and then re-storing the knowledge in the EDIT's underlying EDO; this is made possible by the nature of ontology's underlying representative framework that unifies heterogeneous information. The EDIT will have the capacity to cross varying structures and content, and querying languages through its unique data-crawling capacity to provide information that is 'understood' by computers, overcoming the issues that existing analysis tools have, where they work only on a limited subset of the available information (Hadzic, et al., 2009).

Ontologies can enhance the accuracy of web searches, as the search program can look for only those pages referring to specific concepts rather than those that use ambiguous terms (García-Sánchez, Martínez-Bejar, Contreras, Fernández-Breis, Castellanos-Nieves, 2006). Currently, this task requires a human to sift through all the content of various pages returned by a search engine. Due to the semi-automatic nature of EDIT, it will not only permit regular collection of

employer demand-related data, but also allow a much broader range of information gathering to feed into its knowledge base compared to using manual methods. The fact that the EDO that will be developed in solution 1, formalises the employer demand intelligence concepts will enable software applications such as EDIF to analyse and elicit the desired information embedded within various sources in an accurate and integrative manner. This makes EDIF much less resource intensive as well as a much more comprehensive information gathering alternative to traditional employer demand knowledge gathering approaches.

EDIF has three main processes involved in its development:

- Data collection

The data collection process crawls the job board SEEK for the same occupational and geographical categories as in solution two above. The aim is to identify, remove the noise of and store text documents in the corpus resources.

- Annotation

The annotation process analyses each word in the document to identify matching terminology with that provided in the EDO.

- Ontology population

The EDIT populates the EDO instance base with the relevant records from that document.

To validate and logically verify the proposed methodology and knowledge representation prototype for consistency, two steps are taken, as discussed below.

Firstly, its structure is tested for logical consistency, proving that the philosophy behind the methodology underpins the proposed structure. Secondly, use cases (also known as case studies) are applied to the prototype as proof of concept experiments. These experiments simulate a real life scenario that puts the prototype to use in a variety of settings, confirming the prototype's practical uptake readiness.

As discussed, it is important that employer demand intelligence case studies are able to be compared with each other and be related to the broader picture (Vaughan, 1988). Evaluation of the proposed employer demand intelligence prototype will prove that the use of common classification variables and a common framework will allow for this to happen, ensuring that each variable can be linked to the broader picture, and show each study's outcome significance.

Ontology evaluation is still an emerging research field (Vrandecic, 2009). As such, evaluating a new software prototype based on an ontological framework requires novel ideas. To formulate such ideas, it is important to reflect on what a quality system would deliver, and to reflect on the concept of quality in the first instance.

The notion of quality could refer to a variety of options for software. Four of these examples illustrate that quality means:

1. Conforming to specific requirements (Crosby, 1979);
2. A system is fit to use (Juran, Gryna, & Bingham, 1974);
3. The sum of a software product's characteristics and features that bear on its ability to satisfy defined or implied needs ("ISO/IEC 9126", 2013) are relevant to software's use in commercial environments;
4. Making a technical judgement of the ontologies and their associated software environment, and documenting this in regard to a frame of reference (requirements, specifications, competency questions, or the real world) (A. Gomez-Perez, Jursito, & Pazos, 1995).

When the notion of a quality software system is discussed, the different types of criteria involved in the evaluation of ontology should be deliberated. There have been several different criteria posed in the literature to this end:

- **Accuracy** (Obrst, Ceusters, Mani, Ray, & Smith, 2007): whether the axioms comply to the expertise of the users;
- **Adaptability** (Obrst et al., 2007): whether the ontology offers a conceptual foundation for a range of expected tasks;
- **Clarity** (Gruber, 1995): whether the definitions are objective and independent of context;
- **Completeness/ Competency** (Chen, 2013; Gomez-Perez, 2001; Gruninger, 1995): whether the ontology includes all relevant concepts and their lexical representations;
- **Computational efficiency** (Gangemi, Catenacci, Ciaramita, & Lehmann, 2006; Obrst et al., 2007): whether reasoners can easily and successfully reason the ontology in an acceptable time;
- **Conciseness** (Chen, 2013; A. Gomez-Perez, 2001): whether the ontology includes relevant axioms in regards to the domain being covered;
- **Consistency/ coherence** (Chen, 2013; A. Gomez-Perez, 2001; T. R. Gruber, 1995): whether the ontology's definitions are consistent and excludes contradictory information;
- **Extensibility** (Chen, 2013): whether the ontology can have new definitions added to it by users without changing the well-defined properties of the ontology;
- **Minimal encoding bias** (Chen, 2013): whether conceptualisations have been specified at knowledge level and not at notation or system level;
- **Organisational fitness** (Gangemi et al., 2006): whether the ontology can be easily deployed within an organisation or commercial environment.

The EDIF prototype's usability and usefulness will be measured with the above qualitative and evaluative criteria in mind, ready for end user uptake. As no exhaustive Employer Demand Intelligence Framework exists yet, and adoption by end-users of such a tool has not been achieved as such, the following steps are proposed to validate and verify the EDIF prototype:

1. Perform reasoning checks on EDO.

The general aim of the reasoning is to derive facts that are not expressed in the ontology or in the knowledge base explicitly. More specifically, it assesses the satisfiability and subsumption of concepts, consistency of the ontology, and the checking, retrieval and realization of individuals (Obitko, 2007).

## 2. Apply use cases to EDO through DL Querying

Use cases will be applied to EDO to check that the ontology conforms to its claims and resolves key employer demand intelligence questions. The queries will be performed through a DL reasoner that annotates semantics and formalises the knowledge to retrieve and combine information in the ontology gathered from multiple sources (Stevens, Goble, & Bechhofer, 2000).

## 3. Execute validator and SPARQL on EDIF

SPARQL is defined in terms of International Resource Identifiers (IRIs) (“W3C SPARQL Query Language for RDF”, 2013) and the queries are performed to retrieve and manipulate data stored in the Resource Description Framework (RDF) format (“SPARQL”, 2013). Since EDIF will be built on an ontological foundation (EDO), the employer demand data will be exposed in RDF format. This will allow queries to be run on EDIF that consist of triple patterns, conjunctions, disjunctions and optional patterns to check general conformity to the qualitative evaluation measures described above. SPARQL allows queries to be run on EDIF that are based on a Boolean type search (“W3C SPARQL Query Language for RDF”, 2013) which DL Query cannot do. As an example, a query relating to the term ‘citizen’ can be performed to allow other words relating to ‘citizen’, such as ‘citizenship’, and ‘citizens’, to be allocated and identified in EDD.

## **4.5 Conclusion**

This chapter has presented the three solutions that are proposed to the research issues described in Chapter Three, namely: i) developing an employer demand ontology; ii) populating

the employer demand ontology with instances; and iii) developing a semi-automatic employer demand intelligence tool.

All three of these solutions are founded on the scientific foundation of ontology, which has been extensively justified in this chapter as the overarching framework for this research. To this extent, the ontology definitions, types, and characteristics are provided, as well as the conceptual framework of the proposed framework of this research.

The next three chapters are designated to extensively discuss each solution's implementation and outcomes, in the order summarised. These solution chapters are followed by the evaluation chapter, which details the validation and verification of the complete framework.

Chapter 5- Solution 1: Employer Demand Ontology Engineering

Chapter 6- Solution 2: Employer Demand Ontology Instance Population

Chapter 7- Solution 3: Semi-Automatic Employer Demand Intelligence Tool

Chapter 8- Employer Demand Intelligence Framework Evaluation

# Chapter 5 - Employer Demand Ontology Engineering

## 5.1 Introduction

The lifecycle of employer demand ontology development spans everything from methodology, modelling and design to implementation, validation and verification. The employer demand ontology development lifecycle defines the range of problems that the Employer Demand Intelligence Tool (EDIT) is intended to solve. At the time of writing, such modelling does not seem to have been published for employer demand intelligence in Australia or anywhere else in the world yet.

This chapter focuses on addressing research issue one — underlying knowledge representation — as discussed in Chapter Three. It starts by discussing the type of methodology that has been employed to guide the development of EDO and the reasons for choosing the NeOn methodology. It then details the modelling notations and design principles that have been employed to present the ontology, and concludes by providing an in-depth explanation of the conceptual implementation of the various EDO classes and subclasses in a two-tiered framework. The ontology editor and knowledge acquisition system, Protégé<sup>26</sup>, was used to develop EDO throughout this research.

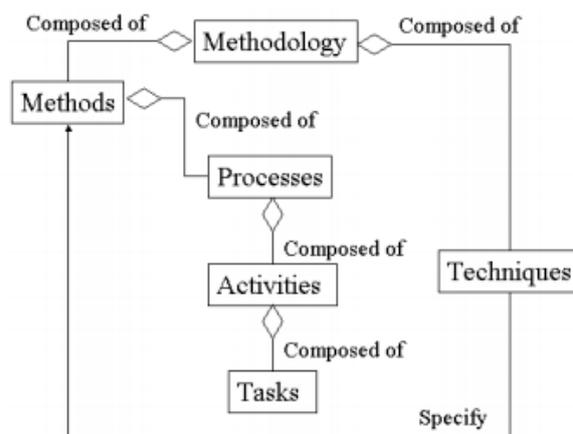
---

<sup>26</sup> <http://protege.stanford.edu/>

## 5.2 Employer Demand Ontology Methodology

### 5.2.1 Methodology Fundamentals

Terms like method, methodology, design, process, task and activity are used haphazardly in the literature (de Hoog, 1998). The Institute of Electrical and Electronics Engineers (IEEE) provides descriptions for each of these terms which, for consistency's sake, have been detailed in Chapter Three and utilised throughout this thesis. Figure 13 shows a flowchart of how some methodology relationships described in Chapter 3 relate to each other. From the figure it is clear that techniques and methods form part of methodology. Methods comprise processes; processes comprise activities; and activities comprise a number of tasks.



**Figure 13** Some terminological relationships in methodologies (Asuncion Gomez-Perez, Fernandez-Lopez, & Corcho, 2003).

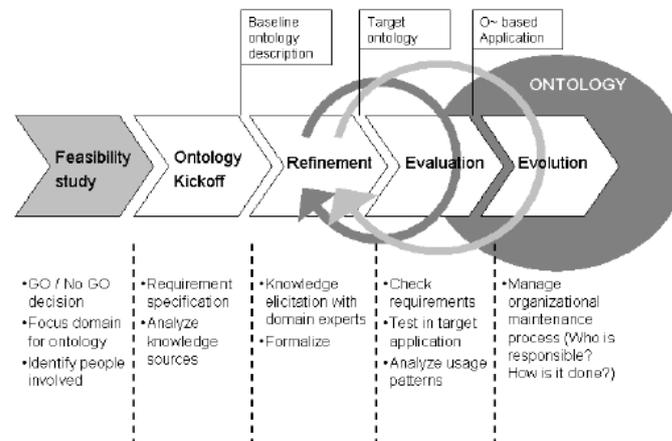
### 5.2.2 Overview of Existing Ontology Methodologies

Many methods have been put forward since ontology development has taken off in information systems. Yet, there is no one correct methodology to use (Noy & McGuinness, 2001), and it is up to the relevant ontology architect to establish the best method to develop that specific domain's ontology. This should be done based on the application in mind. In other words, what the ontology is going to be used for and how general or detailed it has to be will guide many of

the ontology modelling decisions. Some ontology methodology and knowledge-base examples below are listed from old to new, with a brief description of each:

- Toronto Virtual Enterprise (TOVE) (Gruninger & Fox, 1994): this project has been created to build a set of integrated ontologies that can model public as well as commercial enterprises. It describes an enterprise's activities, and states, time, resources, and cost.
- METHONTOLOGY (Fernandez, Gomez-Perez, & Juristo, 1997): METHONTOLOGY has been developed to provide the detail relating to the definition and standardisation of the ontology life cycle. Fernandez et al. claim that prior to their METHONTOLOGY, this process has been a craft instead of an engineering act due to a lack of ontology methodology and other standards.
- The Simple Knowledge-Engineering Methodology (Noy & McGuinness, 2001): this is a declarative frame-based system to develop ontologies. Noy and McGuinness make the important point that there is no single correct ontology for any domain, and that ontology design is a creative process which, if designed by different people, would not be exactly the same.
- Cyc (Reed & Lenat, 2002): Cyc is a large commonsense knowledge base that relies on an interactive clarification-dialog-based tool set. It provides upper level general knowledge concepts that can be plugged into lower level, domain specific ontologies.
- On-To-Knowledge (OTK) (Lau & Sure, 2002; Sure, Staab, & Studer, 2009; Sure, Tempich, & Vrandecic, 2006): this distinguishes itself from other methodologies by focusing on the application-driven development of ontologies rather than focusing on the ontology itself. On-To-Knowledge's methodology is based on five steps: a) feasibility study, b) kickoff, c) refinement, d) evaluation, and e) application and evolution (Figure 14).
- NeOn (Gomez-Perez, Motta, & Suarez-Figueroa; Suarez-Figueroa, 2010): The NeOn methodology's aim is to provide pathways and activities for a range of situations instead of stipulating a rigid workflow when developing ontologies. The methodology details nine

scenarios, which include common tasks such as aligning and integrating existing ontological resources.



**Figure 14** A graphical example of one of the methodologies- the OTK Methodology (Lau & Sure, 2002).

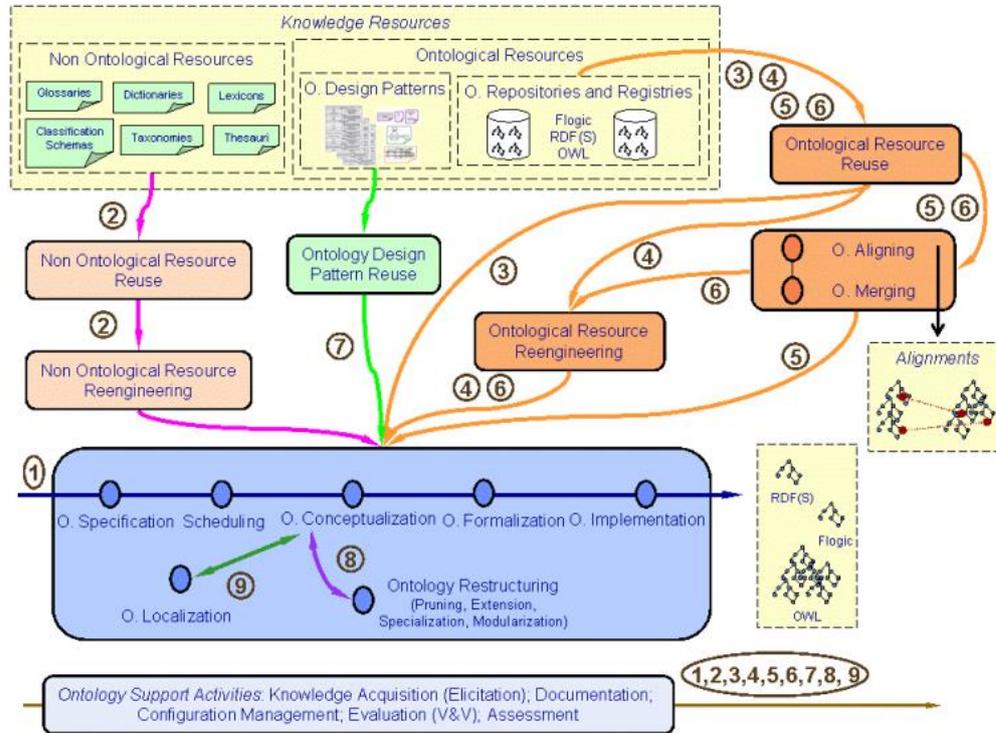
It would seem that across the board, most ontology methodologies, albeit different, agree that the design objectives should allow ontologies to be collaborative, lightweight, domain-oriented, integrated and incremental (Bergman, 2010). The next section will elaborate on the NeOn methodology which was chosen for the development of the Employer Demand Ontology.

### 5.2.3 Employer Demand Methodology Used- NeOn

The novel facet of the NeOn methodology is the creation of an Ontology Requirements Specification Document (ORSD) as the spine of the ontology development process. The Ontology Requirements Specification Document aims to facilitate 1) the search and reclaim of obtainable knowledge-aware supplies with the idea to reuse them; 2) the search and reclaim of obtainable ontological supplies such as ontology modules, statements and design patterns); and 3) the verification of the ontology during the ontology development (Suarez-Figueroa, Gomez-Perez, & Villazon-Terrazas, 2009).

The NeOn methodology was seen as the most comprehensive methodology available at the time of the Employer Demand Ontology development process. The nine scenarios available

from the NeOn methodology are illustrated in Figure 15 and described in the next section, as they were applicable to the development of the Employer Demand Ontology.



**Figure 15 NeOn scenarios to build ontology networks and ontologies (M C Suarez-Figueroa, Gomez-Perez, & fernandez-Lopez, 2012).**

### A. NeOn’s Nine Scenarios

The NeOn methodology consists of a possible nine scenarios that ontology developers can use, depending on their specific domain, existing relevant and reliable resources and intended eventual use of the ontology. The scenarios are: 1) from specification to Implementation; 2) reusing and re-engineering non-ontological resources; 3) reusing ontological resources; 4) reusing and re-engineering ontological resources; 5) reusing and merging ontological resources; 6) reusing, merging, and re-engineering ontological resources; 7) reusing ontology design patterns (ODPs); 8) restructuring ontological resources; and 9) localizing ontological resources.

Of the nine NeOn scenarios, only scenarios one, two and three (to some extent) applied for the development of EDO as discussed below.

*Scenario 1: From Specification to Implementation*

Scenario 1 is made up of the core activities that have to be included in any ontology development process. This scenario is used to guide ontology development from scratch. The first step in this scenario is to produce the ontology requirements specification document (ORSO). The ORSO details the ontology's purpose, scope, implementation language, target group, intended uses, requirements to fulfil and pre-glossary of terms. The second step in this scenario deals with checking existing resources that can be reused in the development. The third step is to schedule the activity to ensure resources are available and to take time commitments into consideration. The fourth and final step in this scenario is to carry out the research by completing: 1) the conceptualisation activity; 2) the ontology formalisation activity; and 3) the ontology implementation activity.

Table 4 indicates the time it took one researcher to develop the employer demand ontology. There was consultation with employer demand as well as ontology experts during this process. Figure 16 presents the metrics of the EDO after all concepts have been included for the one month trial period.

Activity		Uninterrupted time taken in days to complete activity
ORSD production		40 days
Check existing resources		15 days
Activity scheduling		10 days
Ontology development	Conceptualisation	60 days
	Formalisation	70 days
	Implementation	80 days
<b>Total</b>		<b>275 days</b>

**Table 4** Approximate time taken to manually develop the EDO.

Metrics	
Axiom	5831
Logical axiom count	3616
Class count	197
Object property count	31
Data property count	7
Individual count	1729
DL expressivity	SRIF(D)

**Figure 16** The EDO metrics for the one month trial.

*Scenario 2: Reusing and re-engineering non-ontological resources*

This scenario deals with the identification and possible reuse of existing resources that have not yet been engineered into an ontology. Examples of such non-ontological sources include

classification schemes, thesauri, folksonomies and lexicons. One aim in doing this is to speed up the ontology development process (Villazón-Terrazas et al., 2011). Another aim in doing so for the development of the Employer Demand Ontology was to allow cross-usage comparisons and to utilise existing agreed standards for each of the relevant sources that were used (Smalberger, Wongthongtham, Villazón-Terrazas, & Forbes, 2013 (under review)). The main difference between this scenario and scenario three below is that the non-ontological resources identified to be reused still need to be engineered into ontologies.

Following the three actions involved in carrying out the non-ontological resource reuse process (search non-ontological resources, assess the set of candidate non-ontological resources and select the most appropriate non-ontological resources), three non-ontological resources were identified for reuse in the EDO:

1. Australian and New Zealand Standard Classification of Occupations (ANZSCO spreadsheet)<sup>27</sup> (Original source from the Australian Bureau of Statistics)<sup>28</sup>
2. Australian Standard Geographical Classification<sup>29</sup> for locations (Original source from the Australian Bureau of Statistics), and
3. ISO 639 classification system for language<sup>30</sup>.

Following the selection of these resources, they were manually forward-engineered into ontological resources.

### *Scenario 3: Reusing ontological resources*

Reusability is an intrinsic property of ontology (Simperl, 2009). The terms ‘reusing’ and ‘re-engineering’ are generally used when other ontology concepts are screened for relevance and

---

<sup>27</sup> <http://www.abs.gov.au/ausstats/abs@.nsf/mf/1220.0>

<sup>28</sup> <http://www.abs.gov.au>

<sup>29</sup> [http://www.abs.gov.au/websitedbs/D3310114.nsf/home/Australian+Standard+Geographical+Classification+\(ASGC\)](http://www.abs.gov.au/websitedbs/D3310114.nsf/home/Australian+Standard+Geographical+Classification+(ASGC))

<sup>30</sup> [http://www.loc.gov/standards/iso639-2/php/code\\_list.php](http://www.loc.gov/standards/iso639-2/php/code_list.php)

possible inclusion into an ontology under development (Simperl, 2009). There are several advantages for sharing, reusing and re-engineering ontologies: a) it increases the quality of the applications that utilise them; b) it reduces the costs related to ontology development; and c) it has the potential to improve the quality of the reused ontologies through the revising and refining of the original concepts (Simperl, 2009; Villazón-Terrazas, et al., 2011).

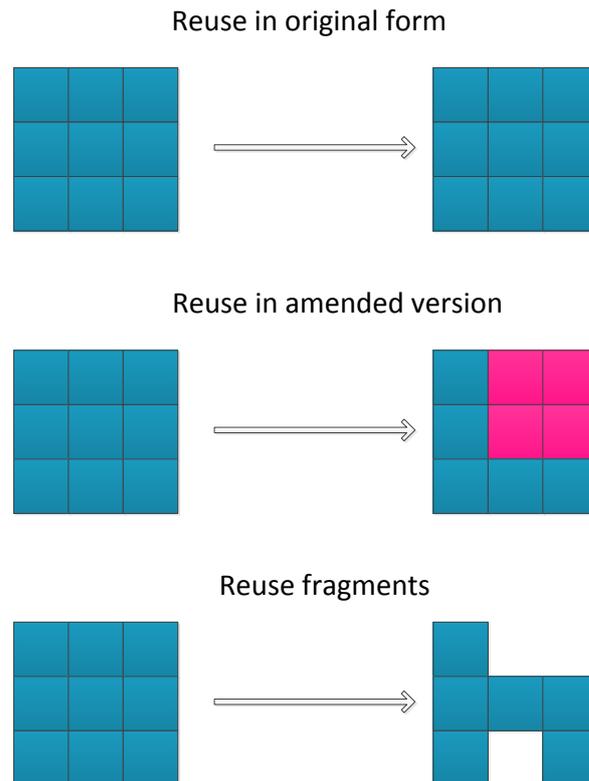
Even though there are many terms used in the engineering, sharing and reusing of ontologies, the term 'ontology recycling' has never been used as a collective term for this process. Almeida & Simões (2006) used the term 'recycling taxonomies' as a header of a section in which they described the process of converting existing taxonomies into ontologies. Apart from this one mention and by including the word 'recycling' into the title of the paper 'T2O — Recycling Thesauri into a Multilingual Ontology', they do not specifically discuss the intended meaning of the term anywhere nor make any reference to its purpose. In the paper they do however mention the existence of several verbs related to ontology construction from existing structured information: 'transform', 'conciliate', 'merge', 'translate', 'complete', 'reuse', 'invert' and 'join' but they only clarify the terms 'translation', 'conciliation'/'merging' and 'completion'. The only other recycling reference found was that of Kohn et al. (2010) where they used the term 'recycling' to describe their process of mining unstructured text from documents in order to classify them in some shape or form. Even though Kohn et al. thoroughly described their employed method for extracting this information, they did not clarify their specific definition of recycling and was it only used in relation to changing existing data into an ontological format.

Simperl (2009) analyse the various methodologies, methods and tools utilised in reuse of existing ontologies in the most prominent ontology case studies published between the early nineties and 2009. The paper argues that there are four steps involved in ontology reuse: 1) ontology discovery, 2) ontology selection, 3) ontology customisation, and 4) ontology integration. Based on their case study analysis and development of two domain-specific

ontologies to further analyse the processes involved in reusing ontologies when developing a new ontology, the author supports their argument that there is a need for practical guidelines and best practices in the selection and evaluation step. They have found this step to be difficult to perform due to the lack of flexible and fine-grained evaluation frameworks and because of the difficulties humans have in assessing extremely heterogeneous sources of information.

This thesis formally introduces the term ‘ontology recycling’ (OR) to refer to the collective processes of reusing, in any shape or format, existing information **and** ontologies in either their original form or to amend it for the same or another purpose. This means that information can be changed from one type of information (e.g. taxonomy) into another type of information (e.g. ontology) for the purposes of reusing some or all aspects of that ontology. As well as the NeOn methodology, aspects of this process are also described in Simperl’s (2009) step 2, ‘ontology selection’. Once the relevant ontologies to be reused have been selected, the developer needs to determine two things: whether this information needs to be converted from one file type to another, and which part of the selected source is going to be reused. The developer can either first convert the source file from its original format to an OWL format, then customise the content of the source to serve the new ontology’s purpose, or the developer can customise the content of the source file and then convert the concepts into the OWL format afterwards.

OR also includes the process of utilising the same type of information in either its original form or an amended version of that information. Furthermore, OR can include the partial or fragmented use of an existing concept (Figure 17).

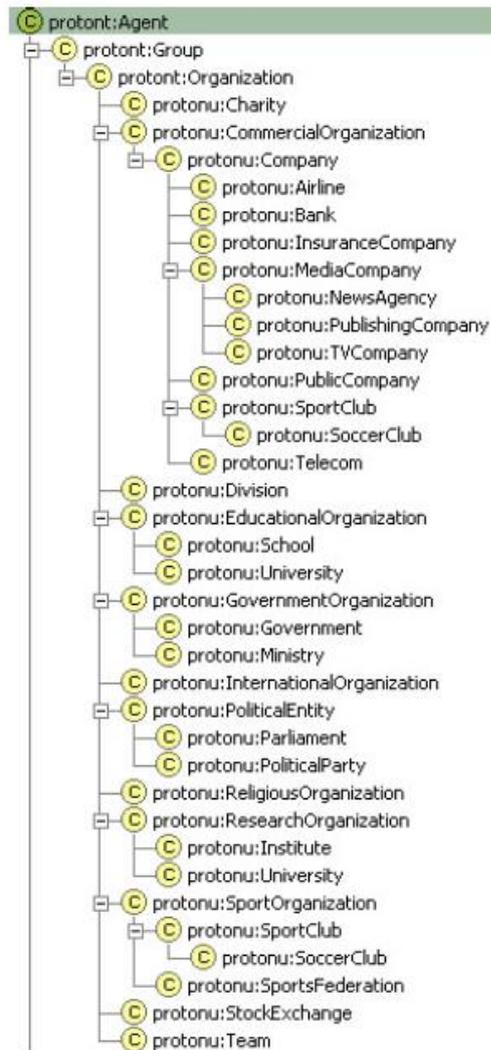


**Figure 17** The three resource recycling options.

To further explain, an OR instance is described where some concepts from the SEEMP (Villazón-Terrazas et al., 2011) were reused in their original state, but only to some extent. The challenge in recycling existing ontologies is to achieve a balance between specification and overgeneralisation, whilst still meeting the aims of the application. Reusable ontologies can easily be too general for a specific application, or they can be too specific for what is trying to be achieved (Advani, Tu, & Musen, 1998; Chandrasekaran & Johnson, 1993; Jarrar & Meersman, 2002).

NeOn's third scenario consists of five activities to reuse ontological resources: i) Ontology search, ii) ontology assessment, iii) ontology comparison, iv) ontology selection, and v) ontology integration; these are described below as they pertain to the development of EDO.

After analysis of the PROTON (Terziev et al., 2005) and SEEMP ontologies it was decided to label the concept of employer in EDO as 'Organization' (see Figure 18 for PROTON hierarchy).



**Figure 18** *protont:Group* and *protont:Organization* class hierarchy (Terziev et al., 2005).

This was to match the existing ontologies and to aid future sharing of data, enable machine understanding, link up existing ontologies in future and so on. The PROTON ontology however includes subclasses for ‘Organization’<sup>31</sup> that are not relevant to the EDO ontology (types of organisation, size of organisation etc.), and hence those parts of the concept were not reused. Only the label ‘organization’ applies for EDO and not its subclasses as it is not currently necessary for the scope of the EDO research. For ‘Organization’, instances of different organisations in Australia that have job vacancies, such as Chevron, Curtin, and Royal Perth Hospital, were

<sup>31</sup> Where the term ‘organization’ has been adopted for EDO, it uses the American-English spelling of the word, i.e., ‘organization’. Where the term is used generically, the Australian-English spelling is used for the word, i.e., ‘organisation’. Explanation for the differing use of these terms follows later in this chapter.

instead added (further detail about the specific concept in section 5.4.1). Thus the EDO meaning of Organization is the same as for existing ontologies, but the specific purpose of each ontology dictates which aspects of that concept are relevant for the specific task at hand to be included/excluded.

As described in Chapter Two, Literature Review, PROTON (Terziev et al., 2005) is a lightweight, general purpose ontology designed with a basic subsumption hierarchy and only a few axioms. This design makes it an ontology that is easy to modify and extend.

Table 5 details the high level classes of the PROTON ontology, specifically indicating the classes could be recycled during the development of the EDO. The reason why it was chosen to recycle one of the classes contained in the PROTON ontology is outlined below, along with the reasons for not reusing those classes which, at the first glance, seemed reusable in EDO.

One of the PROTON Top module super classes is *protont:Person* with properties *hasPosition*, *hasProfession*, *hasRelative* and *isBossOf*. However, as EDO does not entail describing or investigating actual persons as such, it was not possible to reuse any of the Person sub-classes for EDO. Another top module super class of the PROTON ontology is *protont:Location*. This class could seem related to EDO, however, due to EDO's requirement of specifically differentiating locations by their rural, regional or metropolitan statuses, the PROTON class *protont:Location* and its subclasses did not fit the EDO's purpose. Terziev et al. (2005) specifically do not include specialisations in *protont:Location* that could be classified directly as instances of *Location* or for any territorial or tribal area, as those would specify exact geographical locations contradicting the generalised purpose of PROTON. Section 5.4.1 further details the location needs and ontological development for EDO. The *protont:JobPosition* in PROTON refers to a person who actually holds a specific position at a specific point in time (related to *protont:Person* described above), and is thus also not related to the EDO's purpose of describing a vacancy. The

*protont:Role* class refers to the role of an entity at a particular event (e.g. defendant in a trial) and would thus not be related to the EDO purposes either.

The final layer of PROTON — the Knowledge Management Module — in its entirety is not applicable to our ontology either. The concepts in this layer refer to information resources that do not fall within the current EDO scope and could be considered as future work. If taken on board in future work it would relate to the sources aspects of the Employer Demand Information System (e.g. websites where the data can be mined from).

The only class that was able to be reused in EDO to some extent, was that of *protont:Organization* (as described above) to specifically align it with other eRecruitment options.

PROTON modules and classes		Recycled? (Y/ N)	
<b>System</b>	Entity	N	
<b>Module</b>	EntitySource	N	
<b>Coverage</b>	LexicalResource	N	
<b>Top module definitions and upper module branches</b>	Object branch	Agent:	
		<i>Group(Organization)</i>	Y
		<i>Person</i>	N
		Account	N
		Brand	N
		Currency	N
		Location	N
		PieceOfArt	N
		Product	N
		Service	N
		Statement	N
		Vehicle	N
	Happening branch	Event class:	N
	<i>Accident</i>	N	

		<i>ArtPerformance</i>	N
		<i>Meeting</i>	N
		<i>MilitaryConflict</i>	N
		<i>Project</i>	N
		<i>SportEvent</i>	N
		Situation class:	
		<i>JobPosition</i>	N
		<i>Role</i>	N
		TimeInterval:	
		<i>CalendarYear</i>	N
		<i>Date</i>	N
		<i>Month</i>	N
		<i>Quarter</i>	N
		<i>Week</i>	N
		Abstract branch	BusinessAbstraction
	ConctactInformation		N
	GeneralTerm		N
	Language		N
	NaturalPhenomenon		N
Number	N		
SocialAbstraction	N		
TemporalAbstraction	N		
Topic	N		
Knowledge Management Module	InformationSpace	N	
	SoftwareAgent	N	
	User	N	
	Profile	N	
	InformationSpaceProfile	N	
	UserProfile	N	
	Mention	N	
	WeightedTerm	N	
	Device	N	

**Table 5 A table developed by this research to indicate the classes of the PROTON ontology (Terziev et al., 2005) that was reused or reengineered in the EDO.**

The second existing ontology that the developers of EDO could consider recycling concepts from was the SEEMP ontology (Villazón-Terrazas, et al., 2011). As described in Chapter Two the purpose of the SEEMP ontology was to match ICT job seekers' CVs with available vacancies across different eRecruitment websites, languages and skills terminology.

The SEEMP ontology has been developed around two main concepts (the application ontologies): job seeker and job offer. Feeding into these two main concepts are eight sub concepts: language, competence, education, labour regulation, compensation, geography, economic activity and occupation. All eight sub concepts feed into both application ontologies (job seeker and job offer) except for language that only has a relationship with job seeker.

To determine which of the SEEMP classes could be recycled for the EDO ontology, the SEEMP OWL files were obtained from the SEEMP developers directly. Unfortunately, no annotations for any of the SEEMP classes could be found in either Villazon-Terrazas et al. (2011) or the received SEEMP OWL files, and a best guess approach was made about classes' definitions based on their relationships and properties.

Table 6 indicates the classes that were recycled from the SEEMP ontology. This section only discusses the classes that were **not** chosen to be recycled, however the classes that **were** recycled are discussed in the relevant subsections in this chapter as indicated in Table 6.

The first seven classes listed in Table 6 relate to the first application ontology of the SEEMP project which is the Job Seeker Ontology. Since EDO does not entail investigating any job seeker details at this point in time, these seven classes were not applicable to EDO development. It is worth clarifying though that the SEEMP class *Country* relates to the job seeker's residency or nationality. As EDO has been developed for an Australian geographical setting, at this stage, it is

not applicable for class recycling. The only EDO class *Country* could remotely relate to is the *WorkingVisa* class in EDO under *PositionRequirements* but even if so, Australian job advertisements do not state which nationality/ residency candidates should have but rather that they should have working rights in Australia.

The class *Sector* does not have any usage in EDO at this point in time either; rather, it was decided that the classification of the position via its occupation title is suffice for the purposes of EDO. If necessary, however, sector could be derived from the organisation that the vacancy is linked to if this information is found to be required later on.

Another class that was not used directly is that of *Education*. EDO instead has the class *EmployeeQualificationLicenceOrRegistration* which is a more accurate description of the instances that were included under this grouping. Not all licenses or registrations are necessarily education per se.

The class *ComputingProfessionals* are referred to in the ANZSCO as *ICTProfessionals* and as such, the ANZSCO version was used instead to match the Australian market needs.

Another class that was not used from the SEEMP ontology is *JobVacancy*. This was deemed too confusing and close to the class *Vacancy* which was in fact used and is described in the Ontology based Employer Demand Knowledge: Tier One section later in this chapter. EDO instead used the class *JobAdvertisement* to eliminate any confusion between a vacancy that exists at an organisation, versus a vacancy that has actually been advertised and thus referred to as job advertisement instead.

SEEMP also used the class *ICTVacancy*, which was irrelevant to EDO as EDO classified the type of vacancy under the class *PositionOccupationType* instead, using the ANZSCO.

SEEMP class	Recycled? (Y/ N)	Recycling decision detailed in which thesis section?
Job Seeker	N	5.2.3
Candidacy	N	5.2.3
Objective	N	5.2.3
ICT Objective	N	5.2.3
OfferedWorkExperience	N	5.2.3
Language	N	5.2.3
Country	N	5.2.3
Competence	Y	D
Work/Condition	Y	K
ContractType	Y	K
Compensation	Y	G, J
Location	Y	H
Sector	N	5.2.3
Education	N	5.2.3
ComputingProfessionals	N	5.2.3
Occupation	Y	M
Organization	Y	L
JobVacancy	N	5.2.3
RequestedWorkExperience	Y	D
Vacancy	Y	C
ICTVacancy	N	5.2.3

**Table 6** A table developed by this research to indicate the classes of the SEEMP ontology (Villazón-Terrazas et al., 2011) that were reused or reengineered in the EDO.

Unfortunately, the Knowledge Net (Bizer, 2005; Mochol & Simperl, 2006; Simperl, 2009) ontology could not be analysed in detail because it is not available publicly online. These papers only provide high level information on the structure of the system, and through analysing that information we established that we could not recycle any of their concepts in any way. More specifically, the Knowledge Net system reused the skill set from KOWIEN which, because it is in

German, is based on German occupations and job settings, is outdated by now (2005), and cannot be located anywhere online, could not be compared with the EDO skill and competency classes.

### **5.3 Employer Demand Ontology Modelling and Design**

As discussed in the previous section, the employer demand ontology lifecycle consists of five elements: methodology, modelling, design, implementation, and validation and verification.

This section discusses the second element namely the modelling of employer demand ontology.

The section firstly discusses the notation of fundamental ontology parts with graphical depictions of each notation type, and then turns to employer demand ontology fundamentals such as classes, instances and properties — the basic building blocks of ontology.

#### **5.3.1 Employer Demand Ontology Notation**

Ontology notation is a matter of taste and has been the cause of many arguments (Sowa, 2009).

As a result, many notations exist; examples are provided in Figure 19 and Figure 20. There is no 'right' or 'wrong' notation, but some notations are more easily understood than others. As a basic rule, most notations have some kind of symbol (e.g. a triangle, square or circle) that represents a class, and another type that represents an instance, and properties are usually indicated by a line connecting related concepts to each other.

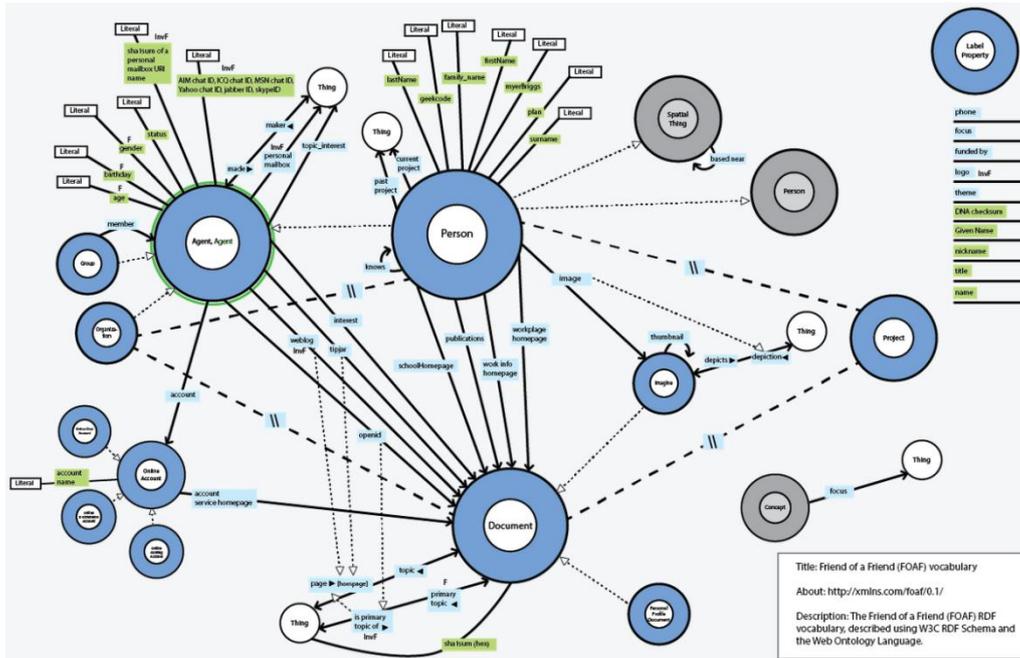


Figure 19 Negru and Lohmann's ontology notation (2013).

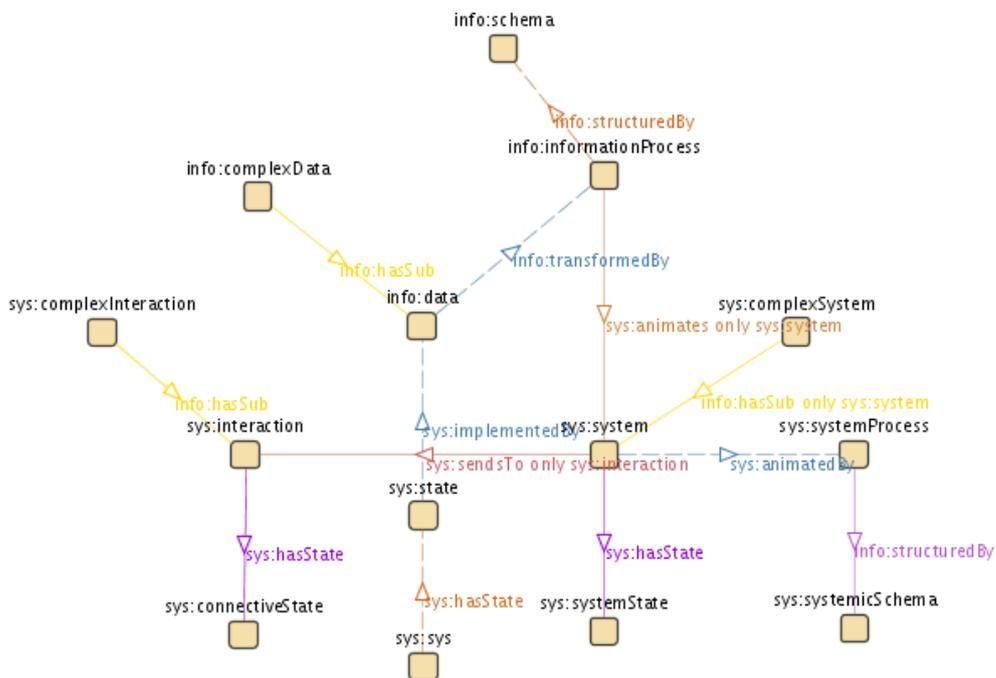


Figure 20 The System Matrix Notation (SMN) Ontology (SMN Ontology, 2008).

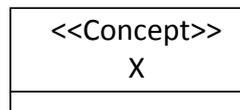
The Employer Demand Ontology uses a variation of the Software Engineering Ontology developed by Wongthongtham (Wongthongtham, 2006). The basic graphical notations for

classes, instances and properties are briefly described below. More information on all of these concepts can be found in section 5.3.2. This section only deals with the graphical notations of each of these important ontological concepts.

### A. Class notation

A double field box represents employer demand ontology classes (Figure 21). In the top compartment, the mandatory term <<Concept>> is accompanied by the actual class label or concept's name, e.g. JobAdvertisement. The bottom compartment is provided to indicate the properties relating to that class.

A class will typically be connected to other classes via some kind of property and it will usually also consist of a set of instances that are specific occurrences of that class (and thus share certain properties which make them a set).



**Figure 21** Employer demand ontology class notation.

### B. Instance notation

In employer demand ontology, an instance is represented as an oval shape containing the instance label or name (Figure 22). The oval represents both class instances and property instances. The oval should always be attached to either a class or a property and this relation is indicated via an arrow. The specific type or arrow depends on whether it is a class instance (generalisation relation), or a property instance (composition relation) which are described in the relevant sections below.



**Figure 22** Employer demand ontology instance notation.

### C. Given and inferred property notation

Properties can indicate the existence of one or more relations among concepts. When property characteristics have more than one relation with another concept or concepts, there are often a given (or true) scenario, and an inferred (or uncertain) scenario which is clarified by specifying the property characteristic type. Some property characteristics can be described with the following question: if this is true, will that be true too? Or, via the following statement: if A is true, will B be true too? The given property represents the true statement. The inferred property represents the indirect action that occurs based on the given property's characteristics. Property characteristics are described in detail in section 5.3.2.

Given properties in employer demand ontology are indicated via an arrow line where the line is solid and the arrow head is closed and filled (Figure 23). Inferred properties in employer demand ontology are indicated via an arrow line where the line is dotted and the arrow head is closed and filled (Figure 24).



**Figure 23** Employer demand ontology given property notation.



**Figure 24** Employer demand ontology inferred property notation.

#### D. Generalisation relation notation

Where relationships exist between different classes or properties that represent either the generalisation or specialisation of a concept, the relation is indicated via a solid arrow line with an unfilled, closed arrow head (Figure 25). The direction of the arrow is dependent on whether the creator wishes to detail the specialisation relation (arrow pointing from the superclass/property towards the subclass/property) or the generalisation relation (arrow head pointing from the sub class/property towards the superclass/property).



**Figure 25** Employer demand ontology generalisation relation.

#### E. Composition relation notation

Relationships that indicate the constituents of specific concepts are graphically indicated via a solid line with an open, unfilled arrow head (Figure 26). The direction of the arrow is dependent on whether the creator wishes to highlight that an instance or instances are part of an upper ontology concept, or to highlight the composition of an upper level concept — it is a very fine nuance.



**Figure 26** Employer demand ontology composition relation.

### 5.3.2 Components of Employer Demand Ontology

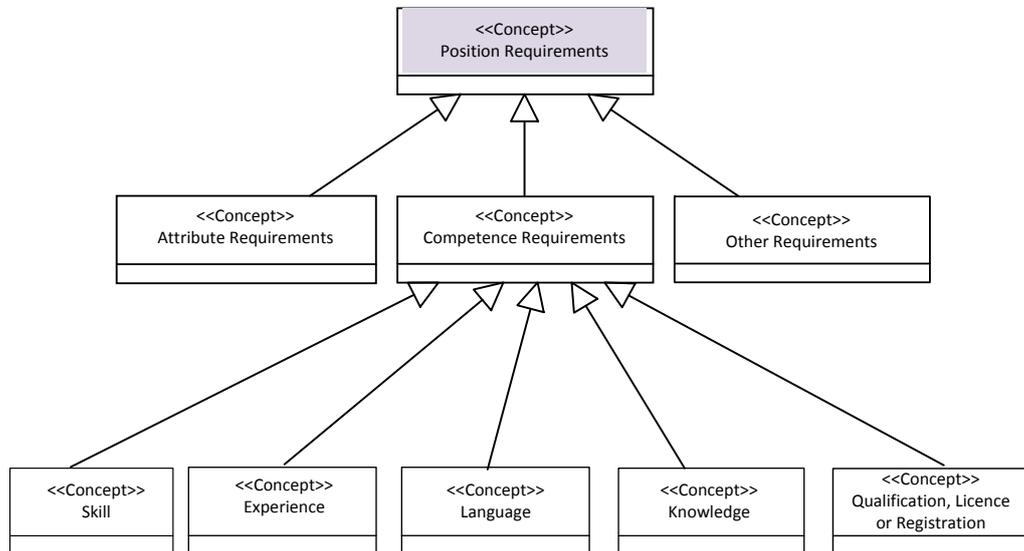
The employer demand ontology is similar to ontologies in other domains in that it consists of classes, properties and instances. Employer demand ontology consists of: instances that

---

represent specific occupational classifications and locations; properties that represent relationships among employer demand concepts and instances; and classes that represent the employer demand concepts. Class, property and instance names are written in *italics* throughout this thesis. For class labels, the first letter of each word is capitalised with no alphanumeric symbols or intervals between words. For relation and attribute labels, the same principle applies except for the first letter of the label that is not capitalised (see section 5.3.3).

**Employer demand ontology instances** represent specific occupational employer demand data, also known as employer demand ontology individuals — referred to in this thesis only as ‘instances’. **Employer demand ontology properties**, referred to only as ‘properties’ in this thesis, are binary relations between employer demand classes or employer demand instance; that is, properties can link either two employer demand concepts or two employer demand occupational instances together. For example, the property *hasEmployer* links the class *JobAdvertisement* with the class *Organization* but also links the instance *MW1* (an instance of the class *JobAdvertisement*) to the instance *IPAHealthcare* (an instance of the class *Organization*). Properties can have inverses; for example, the inverse of *hasRequirement* is *isRequirementOf*. Properties can be limited to comprising of a single value, classifying them as being functional. Additionally, properties can be symmetric, transitive, inverse, inverse functional, antisymmetric, and reflexive or irreflexive. All properties, however, are used to set restrictions that limit the instances belonging to a specific class. **Employer demand ontology classes** are real representations of employer demand concepts, also referred to as sets that contain specific instances or examples of employer demand data. For the remainder of this thesis, employer demand ontology classes are simply referred to as ‘classes’. Classes are described by formal descriptions stating precisely the conditions for membership of each class. Employer demand ontology classes are organised into a super class-subclass hierarchy, where subclasses are subsumed into their respective super classes, also known as a taxonomy. As an example, the type of applicant requirements indicated as necessary for a vacancy in employer

demand ontology are *AttributeRequirements*, *CompetenceRequirements* and *OtherRequirements*. Figure 27 shows the class hierarchy of position requirements contained in job advertisements.



**Figure 27** Employer demand ontology class hierarchy of the position requirements domain.

The classes *EmployeeExperience*, *EmployeeKnowledge*, *EmployeeLanguage*, *EmployeeSkill* and *EmployeeQualificationLicenceOrRegistration* are all subsumed by the super class *CompetenceRequirements* in the first place and ultimately *PositionRequirements*. This implies that when reference is made to an employee's experience, it includes referring to the requirements that are expected of an employee's competence. Furthermore, the *PositionRequirement* class structure supposes that reference to any employee experience, knowledge, language, qualification, licence or registration and skills are all part of competence requirements of the employee. The class *EmployeeExperience* also includes all the instances that reflect an employee's experience. Classes *CompetenceRequirements*, *AttributeRequirements* and *OtherRequirements* are all subclasses of the class *PositionRequirements*. Correspondingly, this hierarchy infers that when reference is made to attribute requirements contained in a job advertisement, it implies inclusion of attribute, competence and other employee requirements.

In conclusion, employer demand ontology classes consist of employer demand concept descriptions that specify the conditions to be satisfied by employer demand data in order for it to be a member of those classes. A more detailed explanation of the characteristics and restrictions of instances, properties and classes are provided in the following sections.

### **A. Employer Demand Ontology Classes**

The class *Thing* is the super class of all classes; it subsumes all subclasses and represents all individuals in an ontology. Employer demand ontology classes are sets of specific employer demand data — the main building blocks of an OWL ontology. These employer demand data sets have common characteristics in order to satisfy limited expressions of the classes. Classes can be defined in three manners: a) by a named super class (e.g. *EmploymentCondition*); b) by anonymous super classes (i.e. restrictions that are imposed on the super class); or c) by enumerated classes where individuals belonging to that class are all listed.

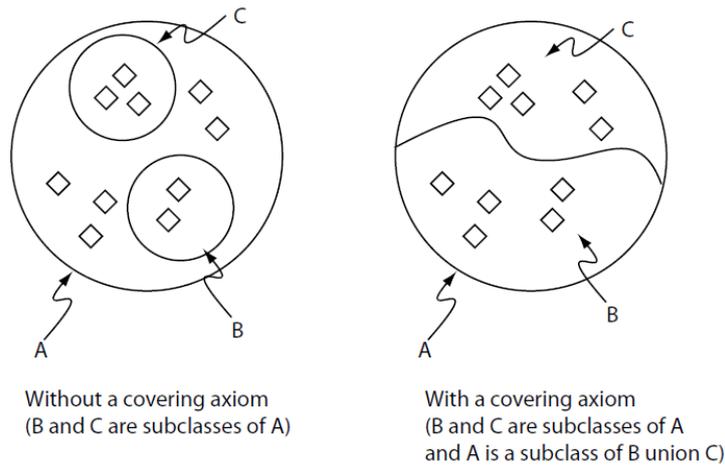
‘Job advertisement’ in the employer demand domain is defined by using anonymous super classes where several concepts are combined to form a single class definition. The concepts of *EmployeeBenefit*, *EmploymentCondition*, *PeriodAdvertisedValuePartition*, *PositionRequirement*, *PositionResponsibility*, *JobAdvertisementIdentificationNumber* and *SalaryLevelValuePartition* all assist in the formation of the class *JobAdvertisement*. An instance of the class *JobAdvertisement*, can thus be expected to include any of the instances associated with any of the mentioned seven restrictions that make up the *JobAdvertisement* super class. Even though it is not a prerequisite, most of the employer demand ontology classes usually hold one or more employer demand ontology properties. That is, an employer demand ontology property may or may not be linked to one or more employer demand ontology classes. It is also possible for a property not to be linked to any classes.

### 1. *Employer Demand Ontology Classes and Subclasses*

Employer demand ontology classes are hierarchical — each subclass is subsumed by its super class. Subclasses can be classified into three sections based on their characteristics: disjoint, partition and decomposition classes. It is important to enumerate the type of subclasses based on their characteristics, as they provide fundamental bases for the decisions to model EDO in the particular way described in these sections.

**Disjoint classes** infers that the instance of one subclass can only be a member of that one class and no other classes. In other words, the class extension of a class description cannot have any common members with those of another class — it can only belong to one class at a time. For example, in the domain of a vacancy's position requirements, the classes *AttributeRequirements*, *CompetenceRequirements* and *OtherRequirements* have been made disjoint as they are very different in nature. It would not make sense for a requirement to be both an attribute requirement (i.e. relating to the employee's personality such as having a can-do attitude) as well as an instance of the domain of other requirements (where it does not refer to an employee's character, such as having their own vehicle or holding working rights for Australia).

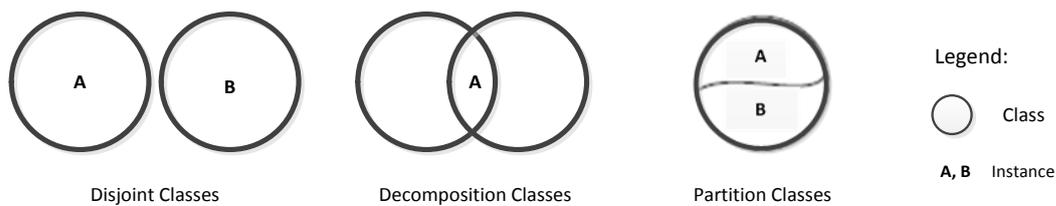
**Partition classes** have covering axioms consisting of two parts: the classes that form the covering, and the class that is being covered. This means that an instance of the super class must be an instance of **at least one** of its subclasses. However, *if the subclasses are made disjoint, an instance of the super class has to be an instance of only one of the subclasses* (Figure 28). For example, in the domain of a vacancy's location type, a location can be either from inner regional Australia, from outer regional Australia, from remote Australia, from very remote Australia or from a major city of Australia. This implies that the ontology classes *InnerRegionalAustralia*, *MajorCityOfAustralia*, *OuterRegionalAustralia*, *RemoteAustralia* and *VeryRemoteAustralia* form the covering to the ontology class *LocationTypeValuePartition*.



**Figure 28** A schematic diagram that shows the effect of using a covering axiom to cover class A with classes B and C (Horridge, Drummond, Jupp, Moulton, & Stevens, 2009).

**Decomposition classes** are classes that overlap at some point, and an instance can be a member of several subclasses in the group. For example, in the employer demand domain for printed media advertisements, a vacancy in Western Australia can be advertised in hard copy in any or all of the following papers: the local community newspaper, The Australian Newspaper or The West Australian newspaper. It means that the class *PrintedMediaAdvertisement* is decomposed into the classes *CommunityNewspaper*, *TheAustralianNewspaper* and *TheWestAustralianNewspaper* and that an instance can be considered as a member of any one or all three of these subclasses.

Figure 29 illustrates the conditions of disjoint, decomposition and partition classes as an aid for EDO examples provided in the text above.



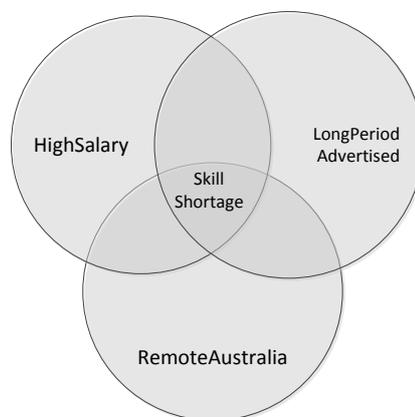
**Figure 29** The effects of disjoint, decomposition and partition classes (Wongthongtham, 2006).

A hierarchical class can also be classified into two categories based on its reason for existence: those that are a type of the parent class also known as Is-A relationships, and those that are a part of the parent class, also known as composition relationships. In **Is-A relationships**, the focus is on generalisation and specialisation of a concept. For example, in the employer demand ontology domain, the class *CompetenceRequirements* is a generalisation of its subclasses *EmployeeExperience*, *EmployeeKnowledge*, *EmployeeLanguage*, *EmployeeSkill* and *EmployeeQualificationLicenceOrRegistration*. Similarly, these five subclasses are specialisations of their super class *CompetenceRequirements*. In **composition relationships**, the focus is on an aggregation of concepts — the association to or composition of a class based on behavioural as well as structural concepts. For example, in the employer demand ontology domain, the class *JobAdvertisement* is an aggregation of the classes *EmployeeBenefit*, *EmploymentCondition*, *PeriodAdvertisedValuePartition*, *PositionRequirements*, *Organization*, *PositionResponsibilities*, *UniqueJobAdvertisementIdentificationNumber*, *LocationTypeValuePartition* and *SalaryLevelValuePartition*.

The employer demand ontology allows super classes to have multiple inheritances, that is, it allows a collection of subclasses to be combined into one super class. As such, every property that the super class holds will also be a property of the subclasses. For example, the class *PositionRequirements* that holds its own properties has the subclasses *AttributeRequirements*, *CompetenceRequirements* and *OtherRequirements*. These subclasses consequently inherit the properties of the super class *PositionRequirements* as well as their own unique properties. The class *CompetenceRequirements* has further subclasses *EmployeeExperience*, *EmployeeKnowledge*, *EmployeeLanguage*, *EmployeeSkill* and *EmployeeQualificationLicenceOrRegistration*. The properties of the class *EmployeeKnowledge*, furthermore, are an aggregation of properties inherited from the super class of its direct super class (*PositionRequirements*), of those inherited from its direct super class (*CompetenceRequirements*), and of its own properties (*EmployeeKnowledge*).

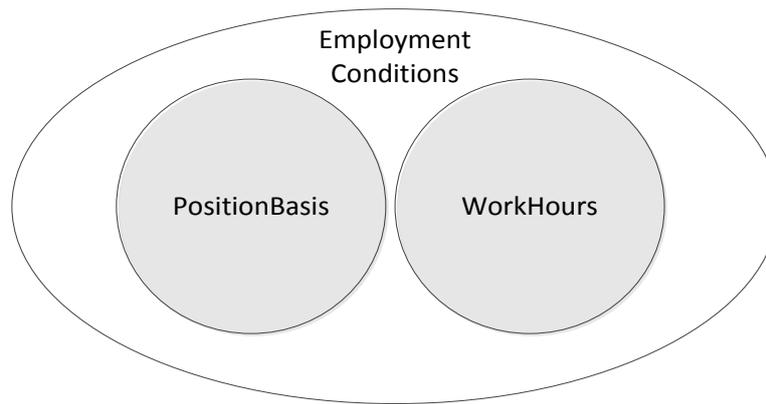
## 2. Complex Employer Demand Ontology Class Descriptions

An employer demand ontology class can consist of either a complex or a primitive description. A number of simple class descriptions combined result in a complex class description. To this end, the Boolean logical operators OR ( $\cup$ ), and AND ( $\cap$ ) are used to form these statements. An intersection class is created through the binding of two or more classes with a partial overlap of the classes. To do this, the AND operator ( $\cap$ ) is used. For example, consider the concept of a skill shortage. As described in earlier chapters, a skill shortage exists if three conditions are fulfilled — a high salary is offered, the location of the vacancy is in a remote location, and it takes a long time to fill this vacancy. This concept can thus be described as the instance where a high salary intersects with a remote location and a long time to fill the position, making it equivalent to the class *SkillShortage* (Figure 30). Intersection classes are never disjointed.



**Figure 30** The intersection of classes *HighSalary*, *LongPeriodAdvertised* and *RemoteLocation*.

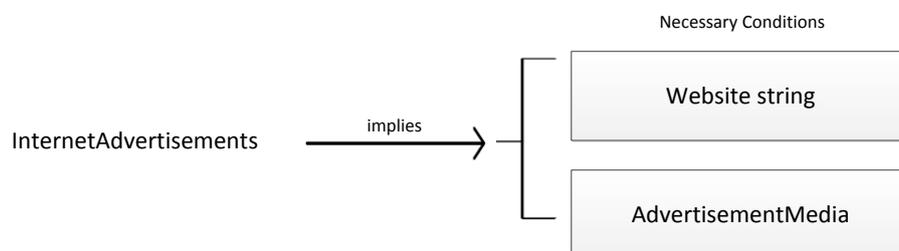
A union class, on the other hand, is created by combining two or more classes but without having an overlap of the classes. All the classes together, then, form the super class. To do this, the OR operator ( $\cup$ ) is used. For example, consider the concept of *EmploymentConditions* in employer demand ontology. The classes *PositionBasis* and *WorkHours* are united (*PositionBasis*  $\cup$  *WorkHours*) to form the class *EmploymentConditions* (Figure 31).



**Figure 31** The union of classes *PositionBasis* and *WorkHours* (*PositionBasis*  $\cup$  *WorkHours*) to form the class *EmploymentConditions*.

### 3. Primitive and Defined Employer Demand Classes

Two conditions describe employer demand ontology classes: 1) necessary conditions, and 2) necessary and sufficient conditions. In an ontology only using necessary conditions, if an instance or a class is part of a super class, it has to fulfill the super class's one or more requirements or conditions as a result (a one-way condition as illustrated in Figure 32). When a class has only necessary conditions in the employer demand ontology, it is known as a primitive class.



**Figure 32** *InternetAdvertisement* conditions indicating the necessary conditions.

For example, in the domain of employer demand ontology, vacancies on the internet can be done via one of three avenues at present: on a company's own website, on a job board website, or on a social media website of sorts. It means that when a class is a subclass of *InternetAdvertisement*, it has to have the same condition imposed on itself as that of *InternetAdvertisement*.

An ontology using both necessary and sufficient conditions is one where instances and classes that are part of a super class will not only fulfill the conditions of the super class but they also **have** to be members of that super class if they satisfy the same conditions (a two-way relationship as illustrated in Figure 33). An employer demand ontology class that has both necessary and sufficient conditions is referred to as a defined class.



**Figure 33 Illustration of Necessary and Sufficient Conditions based on the class JobAdvertisement.**

An instance of any of the conditions of *JobAdvertisement* (*EmployeeBenefits*, *AdvertisementMedia*, *EmploymentConditions*, *PeriodAdvertisedForValuePartition*, *PositionRequirements*, *LocationTypeValuePartition*, *SalaryLevelValuePartition*, *UniqueJobAdvertisementIdentificationNumber*, *PositionOccupationType* and *Organization*) also satisfies the conditions that define the class *JobAdvertisement*.

## B. Employer Demand Ontology Properties

Employer demand ontology properties indicate the relations that exist between two or more employer demand ontology classes or instances. For example, in the domain of word of mouth advertisements, the class *Employee* has a relation called *hasEmployer* with class *Organization*. Another example is for the class *InternetAdvertisements*, which has a relation called *hasString* with the data type *String*. From these two examples it is clear that a property does not necessarily have to be a class but rather that a property is a twofold relation between two things. The difference between the two example properties, *hasEmployer* and *hasString* is discussed below.

### 1. Employer Demand Ontology Property Types

There are three types of OWL properties in employer demand ontology: object properties, datatype properties and annotation properties. The first type, **object properties**, relates two objects with each other, that is, a concept to a concept or an individual to an individual through some kind of relationship. The second type — **datatype properties** — describes the relationship between an individual or class and its data value. That is, it relates a class or individual to an XML Schema Datatype or RDF literals. The third type, which is an OWL specific property, is the **annotation property**, which allows the ontology creator to provide descriptive text for any class, instance or property to clarify contextual meaning.

For example, in the domain of job advertisements, the class *JobAdvertisement* has a relation with the class *AdvertisementMedia* that is called *isAdvertisedVia*. This relation is presented as an **object property**. The object property *isAdvertisedVia* also links the instance *RNAC1* (an instance of *JobAdvertisement*) to the instance *Seek* (an instance of *JobBoard* which is a sub class of *InternetAdvertisement* with *AdvertisementMedia* as its super class).

In the domain of vacancies advertised via the internet, for example, the class *InternetAdvertisements* has a relation called *hasWebsite* with a type of `xsd:string`. This relation is presented as a **datatype property**. The datatype property *hasWebsite* furthermore also links the instance *Seek* (an instance of the class *JobBoard*) to the `xsd:string` <http://www.seek.com.au/>, or the instance *RNAC1* (an instance of the class *JobAdvertisement*) to the `xsd:string` <http://www.seek.com.au/Job/registered-nurse-aged-care/in/perth-fremantle-southern-suburbs/24126888598>.

In the domain of position requirements for a vacancy, for example, the class *OtherRequirements* has an **annotation property** `rdfs:comment` that is used to store the definition for that class as a string: ‘Additional qualities or abilities that the employer expects their employee to have’. The annotation property for *OtherRequirements* also links its instance *WorkingRights* to its own specific data literal (string) ‘Includes Working Visa, Permanent Residency, Australian Citizenship and Visa 475’ that provides a more detailed definition applicable to that specific instance.

## 2. Employer Demand Ontology Property Domains and Ranges

Employer demand ontology properties can link a class (or its individuals) from one domain to a range of classes (or their individuals) of a related domain. It is a binary relation with distinct beginnings and ends. For example, in the employer demand ontology, the property *hasLocationType* would link individuals belonging to the class *JobAdvertisement* to individuals belonging to the class of *LocationTypeValuePartition*. In this case the domain of the *hasLocationType* property is *JobAdvertisement* and the range is *LocationTypeValuePartition*.

Multiple classes can also be specified as the range of a domain’s property. The range of the property is interpreted as the intersection or union of the classes when multiple classes are specified as the range of a property’s domain. For example, if the range of a property has the classes inner regional Australia, major city of Australia, outer regional Australia, remote Australia

or very remote Australia, the range of the property will be interpreted as one of two options: 1)  $InnerRegionalAustralia \cap MajorCityOfAustralia \cap OuterRegionalAustralia \cap RemoteAustralia \cap VeryRemoteAustralia$ , or 2)  $InnerRegionalAustralia \cup MajorCityOfAustralia \cup OuterRegionalAustralia \cup RemoteAustralia \cup VeryRemoteAustralia$ . The difference between option 1) and 2) is indicated through making the classes contained in the ranges either **not disjoint** with each other (indicating the intersection of the classes with each other) or **disjoint** with each other (indicating the union of the classes with each other (as explained earlier, Figure 30 and Figure 31). In the employer demand ontology, the ranges for *LocationTypeValuePartition* are indeed made disjoint from each other as a location can only be categorised into one of the above ranges at any point in time —that is, the location types *InnerRegionalAustralia*, *MajorCityOfAustralia*, *OuterRegionalAustralia*, *RemoteAustralia* and *VeryRemoteAustralia* cannot be overlapped.

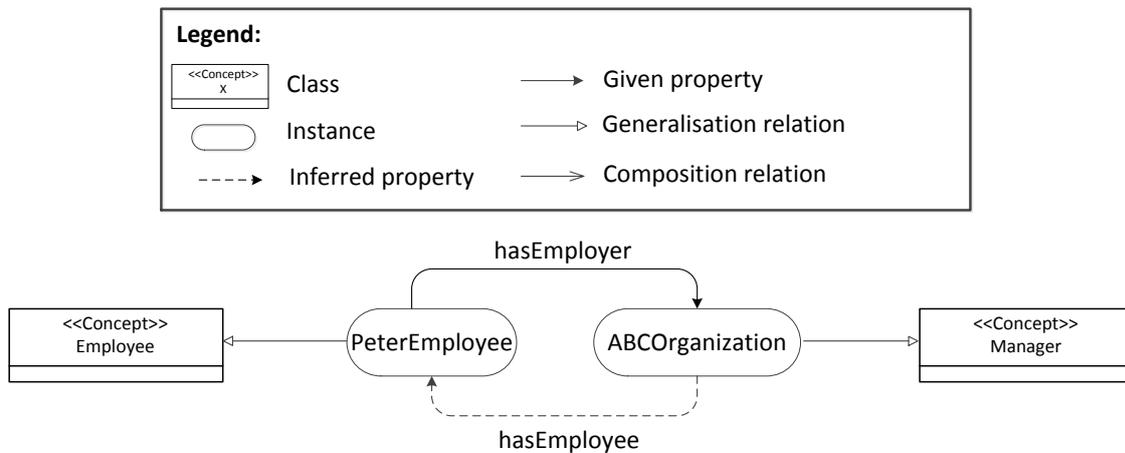
As discussed earlier, a class can be defined by three means, one of which is through the enumeration of the class: either through the exact listing of its subclasses or instances, or through the provision of its exact data values for that range when using datatype properties. In the employer demand ontology, for example, the datatype property *SalaryLevelValuePartition* has been deliberately related to the set of data values: high salary level, average salary level and low salary level. Each item in this range has been made disjoint from each other, inferring that the datatype property *SalaryLevelValuePartition* relates to either high salary level or average salary level or low salary level. In other words, a salary contained within a job advertisement can only fall into one of the categories mentioned at any particular point in time (the salary ranges do not overlap with each other).

### 3. Employer Demand Ontology Property Characteristics

Employer demand ontology utilises property characteristics to enrich the meaning of properties. There are several types of property characteristics (inverse, functional, inverse functional, transitive, symmetric, antisymmetric, reflexive, and irreflexive) which are described in their respective sections below.

#### i) Inverse properties

An inverse property implies that it has two opposite values that are related to each other in both ways. In other words, in the employer demand domain where an employee has an employer and the property is an inverse property, it implies that an employer will also have an employee (Figure 34).

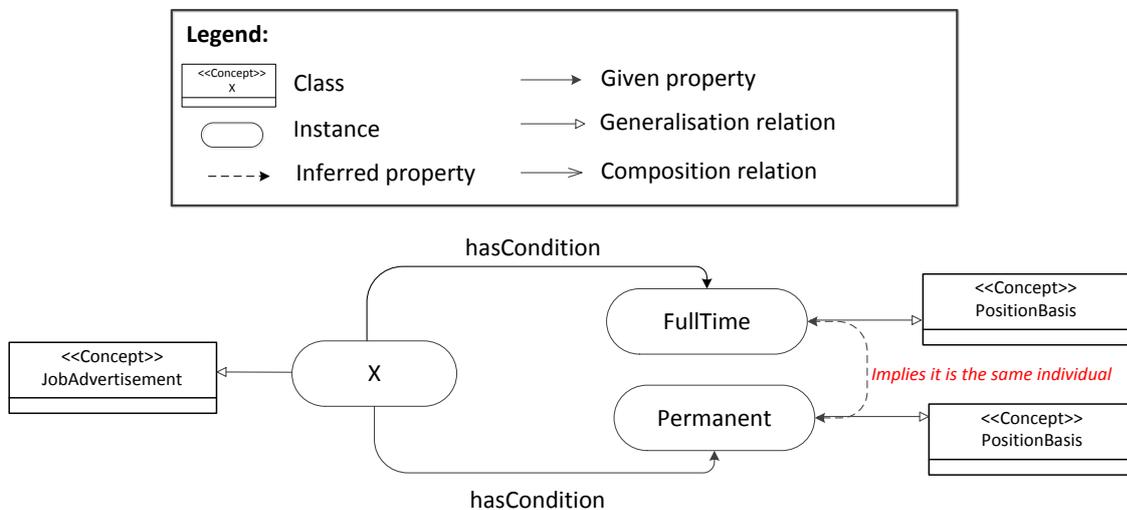


**Figure 34** Illustration of an inverse property: *hasEmployer*.

#### ii) Functional properties

A functional property implies that there can be only one relationship of an individual with another individual. In other words, the individual has a single valued property relationship with one other individual. In the employer demand domain, this would be the case for a job advertisement as it can only have one type of position basis specified — one of either casual,

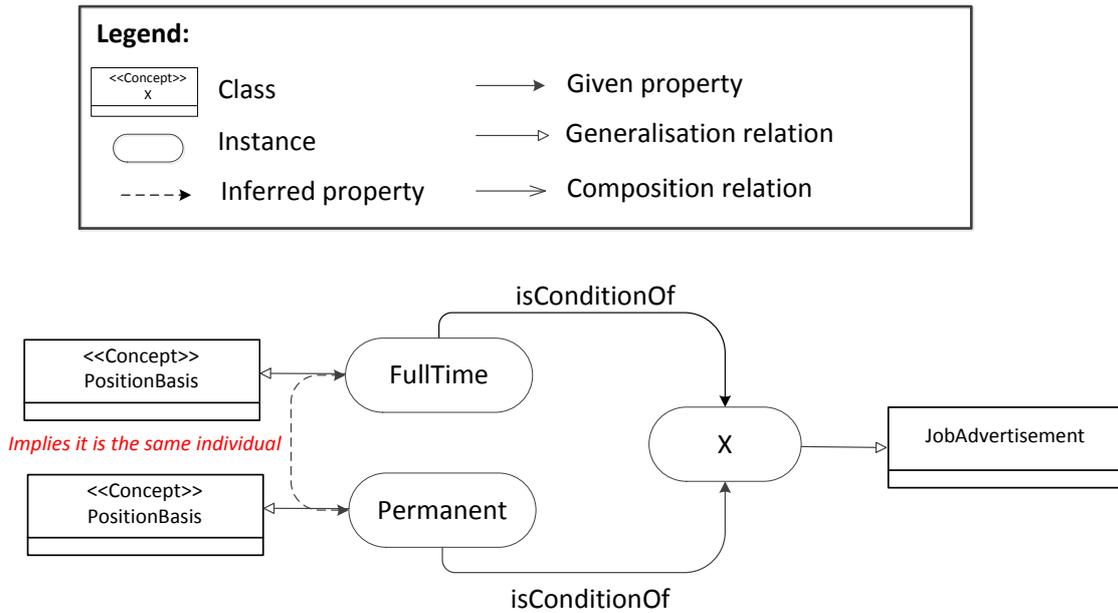
full-time, part-time or relief. A specific vacancy instance, for example *RNCE11* (Registered Nurse Critical Care and Emergency Advertisement Number 11), thus has only the one *PositionBasis* condition that is specified in the vacant position’s advertisement (*JobAdvertisement hasCondition FullTimeBasis*). This is because you can be employed as a full-time, a part-time, a casual or a relief person depending on the needs of the position (Figure 35). The functional property characteristic is also the only characteristic that can be applied to datatype properties.



**Figure 35** Illustration of a functional property: *hasCondition*.

iii) Inverse functional properties

Inverse functional properties, as the name would suggest, are properties that hold the characteristics of both a functional and an inverse property. In other words, it means that the inverse property is functional and that there can only be one instance related to the instance via the property. Figure 36 shows an example of an employer demand ontology inverse functional property *isConditionOf*.

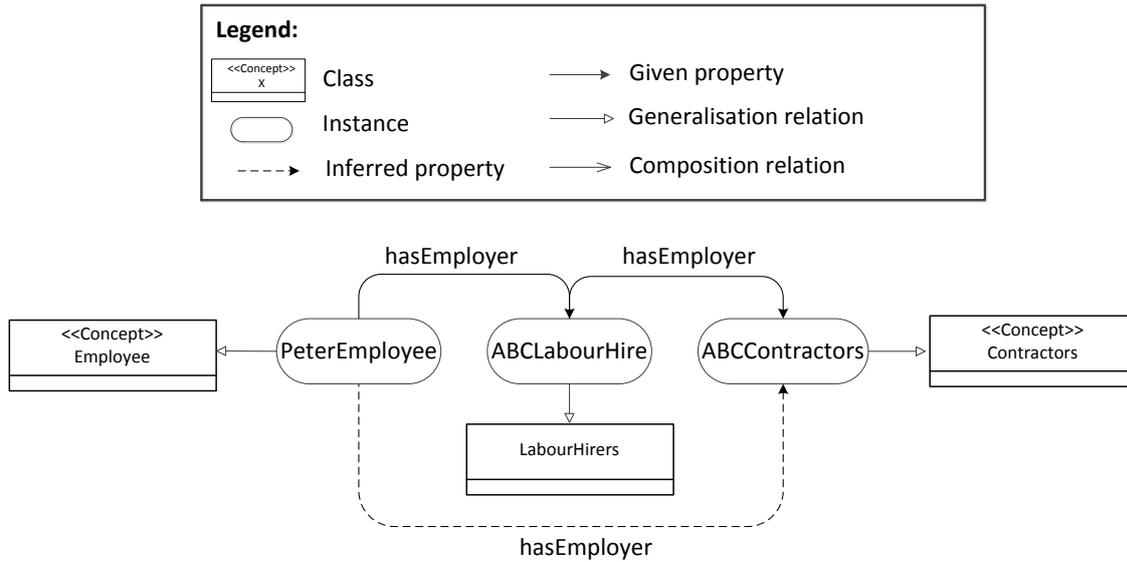


**Figure 36** Illustration of an inverse functional property: *isConditionOf*.

This is the inverse property of *hasCondition*, and since *hasCondition* is functional, *isConditionOf* is inverse functional. If it is indicated that *FullTime* is a condition of *JobAdvertisement*, and also that *Permanent* is a condition of *JobAdvertisement* while this property is functional, it can be inferred that *FullTime* and *Permanent* has to be the same instance.

#### iv) Transitive properties

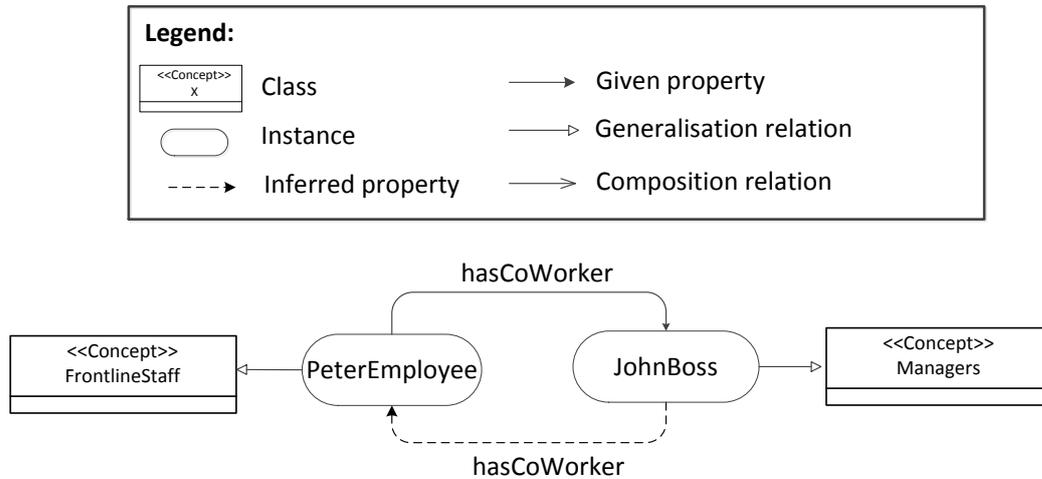
Another type of property characteristic allowed in employer demand ontology is transitive properties. If a property is transitive and it relates property x to property y, and property y to property z, it can be inferred that property x is related to property z through the transitive property specified – the property can thus not be functional. For example, in the employer demand ontology where *PeterEmployee* is related to *ABCLabourHire*, and *ABCLabourHire* is in turn related to *ABCCContractors* through the property *hasEmployer*, it can be accepted that *PeterEmployee* is also related to *ABCCContractors* via the transitive property *hasEmployer* (Figure 37).



**Figure 37** Illustration of a transitive property: *hasEmployer*.

v) Symmetric properties

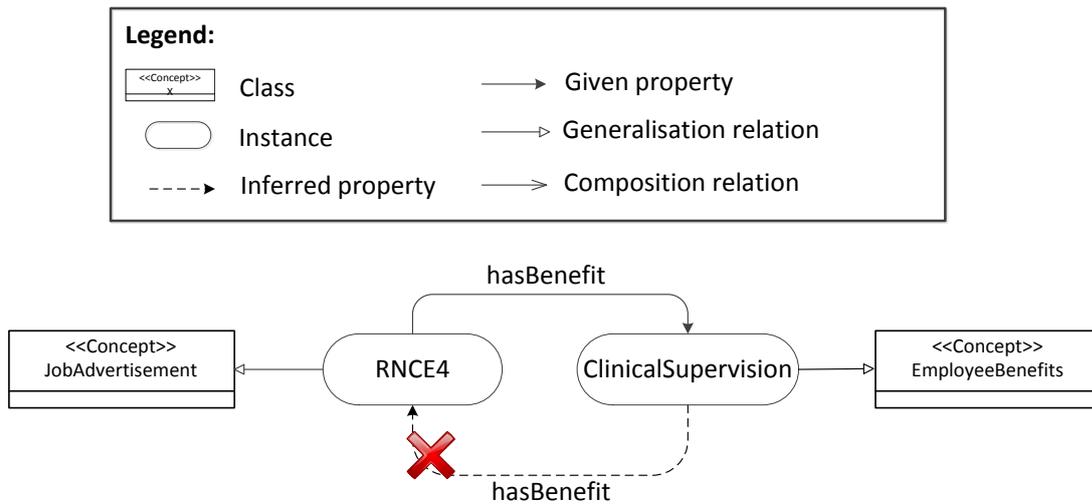
As the name suggests, symmetric properties relate two individuals to each other via the same property. The property relation is valid both ways between the individuals — thus the domain and range for each individual are the same. Figure 38 shows an employer demand domain example of a symmetric property. If the instance *PeterEmployee* is related to the instance *JohnBoss* via the *hasCoWorker* property, then it can be inferred that *JohnBoss* must also be related to *PeterEmployee* via the *hasCoWorker* property. In other words, if *PeterEmployee* has a coworker that is *JohnBoss*, then *JohnBoss* must have a coworker that is *PeterEmployee* — the property becomes its own inverse property.



**Figure 38** Illustration of a symmetric property: *hasCoWorker*.

vi) Antisymmetric properties

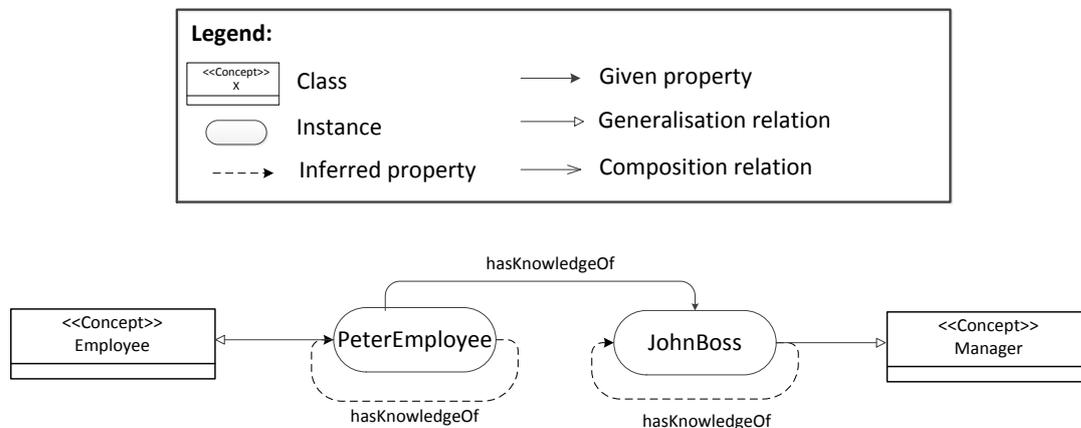
Property characteristics opposite to those of symmetric properties also exist in ontology. In other words, if a property is asymmetric or antisymmetric and relates two instances to each other, the relationship is only valid in one direction between the two instances. For example, in the employer demand ontology domain, the *JobAdvertisement RNCE4* is related via the *hasBenefit* property to the instance *ClinicalSupervision* (Figure 39). Due to the property being asymmetric, the instance *ClinicalSupervision* cannot be related to the instance *RNCE4* via the same *hasBenefit* property. In other words, *RNCE4* has the benefit of clinical supervision; however clinical supervision cannot have a benefit called *RNCE4*.



**Figure 39** Illustration of an antisymmetric property: *hasBenefit*.

vii) Reflexive properties

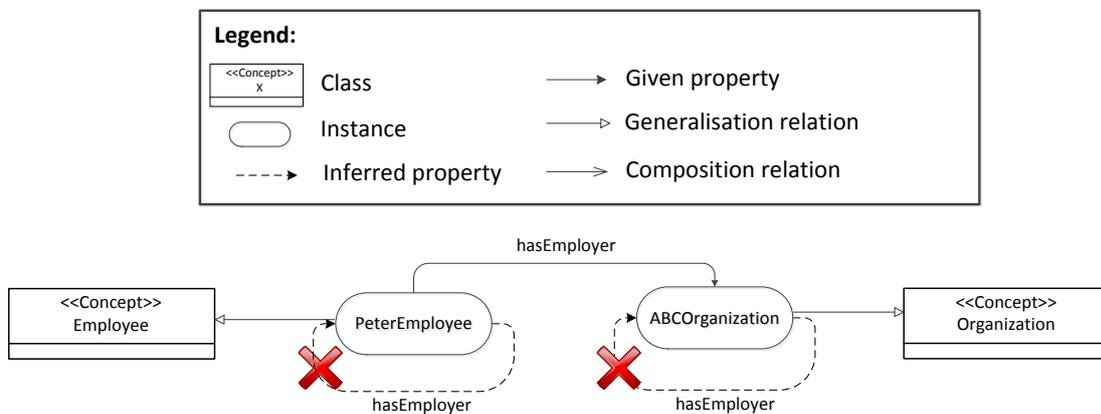
When an instance can be related to itself as well as to another instance via a property, that property is set to be reflexive. In the employer demand domain, for example, using the property *hasKnowledgeOf*, an instance *PeterEmployee* must have a relationship with itself via the property *hasKnowledgeOf* if the property is set to be reflexive. Thus *PeterEmployee* must have knowledge of himself and can also have knowledge of others such as his boss John via this reflexive property *hasKnowledgeOf* (Figure 40).



**Figure 40** Illustration of a reflexive property: *hasKnowledgeOf*.

## viii) Irreflexive properties

An irreflexive property is the opposite of a reflexive property in that the property can relate an instance to another instance, but the property cannot relate an instance to itself via that same property. For example, in the employer demand domain, the property *hasEmployer* relates the instance *PeterEmployee* to the instance *ABCOrganization*; however, *PeterEmployee* does not have an employer of itself called *PeterEmployee*, and likewise, *ABCOrganization* does not have itself as its employer. The property *hasEmployer* is thus irreflexive (Figure 41).



**Figure 41** Illustration of an irreflexive property: *hasEmployer*.

#### 4. Employer Demand Ontology Property Restrictions

Employer demand ontology properties are used to create restrictions for child (or sub) classes and instances belonging to a parent (or super) class. Employer demand property restrictions can define a class of instances based on the relationships that they participate in — the restriction is seen as its own kind of class. Restrictions are categorised into three areas: *hasValue*, *quantifier* and *cardinality* restrictions, which are described in the next section.

## i) hasValue restrictions

The hasValue property restriction (denoted by the  $\exists$  symbol) specifies that class  $A$  will have at least one relation along a stated property  $P$  with instance  $b$  —  $A \exists b$ . For example, in the employer demand domain in Australia, the hasValue restriction *hasEmployeeLanguageRequirement*  $\exists$  *English* describes the group of instances (an anonymous super class) that have at least one relationship along the *hasEmployeeLanguageRequirement* property to the specific instance *English*. In the employer demand ontology the *hasEmployeeLanguageRequirement* property relates the instance *English* to the class *JobAdvertisement* through the following restriction imposed on it: *JobAdvertisement* *hasEmployeeLanguageRequirement*  $\exists$  *English*. This is because the primary language in Australia is English, and to enable an employee to do his or her work effectively in Australia, employees need to have mastered at least some level of English. The class *JobAdvertisement* could, however, also be linked via the same property to another instance (e.g. *Spanish*) if need be, as it is not a mutually exclusive property. Similarly, *hasEmployeeLanguageRequirement*  $\exists$  *English* could also be linked to another class e.g. *DiplomaCourse* as its domain.

## ii) Quantifier restrictions

Another type of property restriction is the quantifier restriction. A quantifier restriction defines the quantity type whereby a class can be linked to an instance value via the quantifier property. In other words, a quantifier property restricts the range of a domain by either allowing all instances (called **universal restrictions** and denoted by the symbol  $\forall$ ) or only some instances (called **existential restrictions** and denoted by the symbol  $\exists$ ) of the specified class to be linked to it. Thus, in the case of a universal restriction, the range is specifically related to a whole class whereas for existential restrictions the range infers that only certain parts of that class are linked to it. For example, for the employer demand ontology, the existential property restriction *hasLocationType* imposed on the class *PotentialSkillShortage* can only relate to the following

four members of the class *hasLocationTypeValuePartition*: *InnerRegionalAustralia*, *OuterRegionalAustralia*, *RemoteAustralia* and *VeryRemoteAustralia*. If the vacancy's location is in one of the major cities in Australia — *MajorCityOfAustralia*, the fifth subclass of *hasLocationTypeValuePartition* — it cannot be an indication of a potential skills shortage according to its definition (as explained in earlier chapters). The universal property restriction *hasEmployer*  $\forall$  *Organization* contains those instances that are connected via the *hasEmployer* property only with instances that are members of the class *Organization*; in other words, it contains instances that have only organisations as their employers. It is worth noting here that 'organizations' in the employer demand ontology refers to any institution that legally employs individuals — whether they are partnerships, companies or small businesses.

### iii) Cardinality restrictions

Another type of property restriction is cardinality restrictions, classified under three types of relation restrictions: maximum, minimum and exact cardinality.

**Minimum cardinality** connects an instance via a property to, at most, the given number of instances specified in the cardinality restriction, and is indicated by the 'equal to or greater than' ( $\geq$ ) symbol. For example, in the employer demand domain of potential skill shortages, there needs to be at least two or more potential skill shortage indicators present in order to classify it as potentially being a skills shortage occupation (*PotentialSkillShortageIndicator*  $\geq$  2).

**Maximum cardinality** connects an instance via a property to at least the given number of instances specified in the cardinality restriction and is indicated by the 'equal to or less than' ( $\leq$ ) symbol. For example, in the employer demand domain of location type, there cannot be more than one vacancy main location specified in a job advertisement (*LocationTypeValuePartition*  $\leq$  1). Even though it is ideal that there should be at least one location specified in the job advertisement for the position, it is possible that some employers may not indicate the specific

location where the vacancy is situated — in such a case the employer demand ontology caters for those advertisements where no exact location may be specified.

**Exact cardinality** property restrictions provide the precise number of relations that an instance has with instances from another class. The restriction is indicated by the (=) symbol. For example, in the employer demand domain, a job advertisement can only have one unique job advertisement identification number (*UniqueJobAdvertisementIdentificationNumber = 1*). If a specific job advertisement has more than one unique job advertisement identification number, the system has generated an error and it needs to be fixed. If an advertisement has not been assigned a unique job advertisement identification number by the system yet, it should be done as soon as possible.

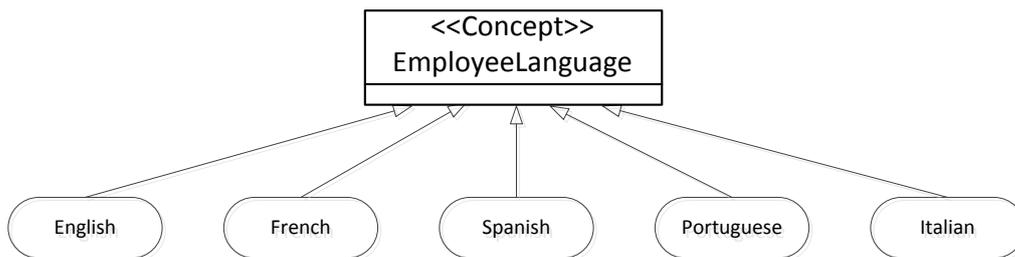
### C. Employer Demand Ontology Instances

Employer demand instances can generally be referred to as individual members of specific classes that represent objects in the employer demand domain. These instances can have properties related to them that further define their existence. For example, in the employer demand domain, the class *JobAdvertisement* holds many instances that are specific occurrences of a job advertisement in the employer demand domain. RNCE3, for example, stands for ‘Registered Nurse Critical Care and Emergency Advertisement Number 3’. This means that it is the third advertisement in our set of advertisements gathered during the one month period in 2010 that related to Registered Nurses in the Western Australia location, who were specialising in critical care and emergency departments.

It is important to mention that employer demand ontology does utilise the Unique Name Assumption (UNA) phenomenon. Where two different names of instances could in fact refer to the same instance, it must be explicitly stated in the employer demand ontology that they are the same instances. If not, they might be the same or they might also be different to each other.

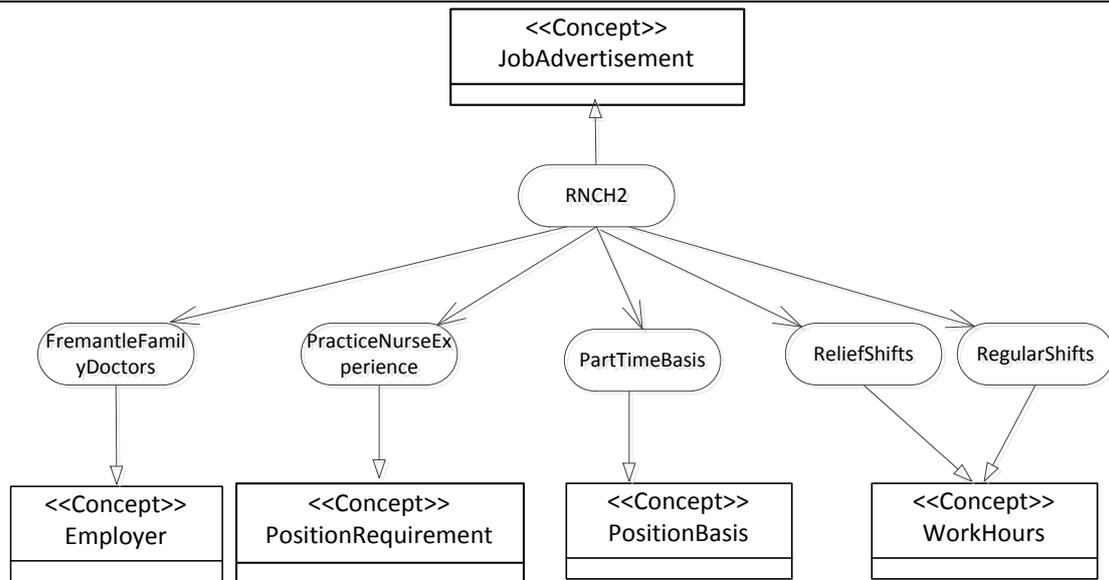
In cases in EDO where similar instance descriptions may exist, and they have been identified as the same instance, the additional names have been listed as an annotation of the instance being investigated. For example, ‘patient education skills’, ‘family education’ and ‘client education’ may, or may not, all refer to the same instance and have been correctly identified in EDO to indicate this where appropriate.

As mentioned previously, employer demand ontology classes (and their relating instances) can have either Is-A relationships or composition relationships. In Figure 42 the class *EmployeeLanguage*, containing several types of languages illustrates the Is-A relationship. In other words, it can be said that, for example, English or French are types of employee languages.



**Figure 42** Illustration of the Is-A relationship for a class and some of its instances.

In Figure 43 a composition relationship for a specific job advertisement and its instances is illustrated. The illustration shows that the job advertisement instance RNCH2 (Registered Nurse Community Health Advertisement Number 2) comprises the instances Fremantle Family Doctors as the employer, practice nurse experience as the position requirement, part time basis as the position basis for the job, and relief and regular shifts as its work hours.



**Figure 43** Illustration of a composition relationship based on instances.

### 5.3.3 Ontology Design Principles

When an ontology is developed, there are several design decisions that need to be considered. Design criteria such as clarity, coherence, extendibility, minimal encoding bias and minimal ontological commitment should be considered (Gruber, 1995); however, as with most design problems, tradeoffs among the criteria may be required.

This section describes the design principles, class and relation label decisions made during the development of EDO to ensure an agile ontology design that is simple, sustainable and highly effective. The following set of basic principles were followed during the development of the EDO.

It was ensured that:

- The ontology is an OWL ontology (OWL is derived from DAML+OIL) which is built on top of RDF and Extensible Mark-up Language (XML). OWL is more expressive and has greater machine interpretability than RDF, provides more facilities and has stronger syntax RDF, and

has become the W3C recommendation in February 2004 to use (Wongthongtham, 2006). As such, it is understood by the industry and the web community as a web standard.

- OWL-Full (the extension of OWL—DL and OWL—Lite) was chosen for its maximum expressiveness which was deemed more important for this research than the option to guarantee the decidability or computational completeness of the formal language.
- The ontology represented and correctly adhered to the strict semantics of the Web Ontology Language (OWL Working Group, 2012).
- Resources were uniquely and persistently identified by International Resource Identifiers (Duerst, 2005) and OWL documents were versioned. The manual intensive job of making the IRIs dereferenceable, however, has been included as future work of this thesis.
- Resources were described with accurate and brief English (human-readable) definitions. This was done to avoid circularity and figurative or obscure language without adding redundant information.
- Resources were described through axioms that match their human readable descriptions. For class labels, the first letter of each word was capitalised with no alphanumeric symbols or intervals between words (example: *PrintedMediaAdvertisement*). For relation and attribute labels, the same principle applied except that the first letter of the label that was not capitalised (example: *hasAdvertisementStartDate*). These naming conventions were similar to those used by Terziev et al. (2005).

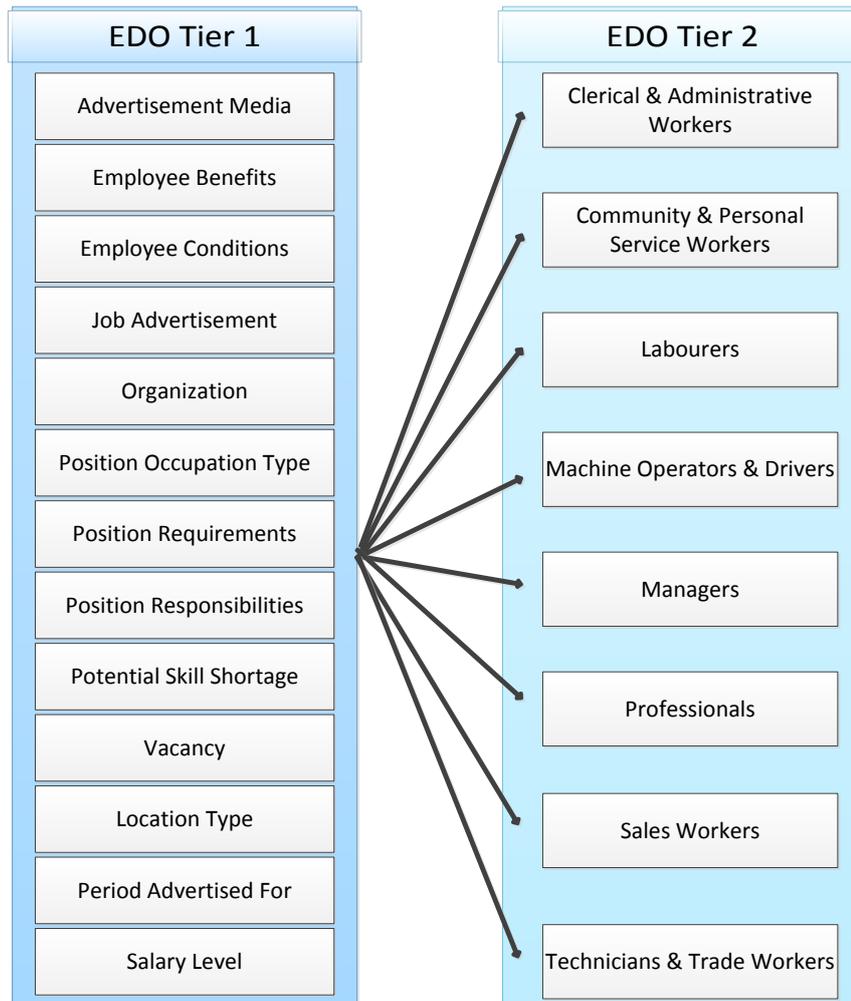
## **5.4 Employer Demand Ontology Conceptual Implementation**

This section is intended as an analysis of what employer demand ontology is, what it consists of, and the various relations the concepts have with each other. Use case scenarios of EDO are presented in Chapter Eight and highlight characteristics of EDO that represent important decisions taken into account with the development of EDIT. Policy makers, education

providers, career consultants, employers and employees should use employer demand ontology to assist in identifying employer demand to aid their respective future decisions.

As discussed, EDIT proposes an additional method that collects information from employers where employers' direct input into the data gathering process is not needed; however the information collected will be derived directly from other information that employers share in the public domain. One of the sources that contain rich employer demand data is vacancies, and these are frequently shared in the public domain by employers in the form of job advertisements. Job advertisement analyses allow data collection about what characteristics employers require in an ideal candidate for a position; that is, the required skills, experience, licences, registration etc. that a candidate should already hold when they apply for the position.

The main components of employer demand ontology are covered in a two-tiered structure (Figure 44): tier one defines the employer demand ontology as it relates to any occupation classification type in Australia, that is, general concepts that are applicable to all occupation types. Tier two defines employer demand ontology for each specific occupation as listed in the ANZSCO, and has been specifically developed further in this thesis for the Minor Group 254 *Midwifery and Nursing Professionals*, in a Western Australian setting.



**Figure 44** Employer demand ontology tiered framework.

The complete employer demand ontology structure can be found on the following website:

[www.EmployerDemandIntelligence.Org](http://www.EmployerDemandIntelligence.Org). Screenshot examples of the major occupational

classes are provided of the ontology as <http://115.146.86.127/sparql> addendum one to this

thesis. The implemented and queried employer demand ontology for Midwifery and Nursing

Professionals in a Western Australian setting is provided in Chapter Six.

### 5.4.1 Employer Demand Ontology: Tier One

This section discusses all the Tier One EDO classes and related subclasses: *Vacancy*,

*AdvertisementMedia*, *JobAdvertisement*, *PositionRequirements*, *PotentialSkillShortage*,

*PeriodAdvertised*, *SalaryLevel*, *Location*, *PositionResponsibility*, *EmployeeBenefit*, *EmploymentCondition*, *Organization* and *Occupation* (only the top level class).

### A. Vacancy Class

Employer demand can be identified through various means. Some efforts entail analysing census data, other efforts consist of surveying individual companies, and some analyse economic indicators such as the unemployment rate.

Another rich source of data that indicates employer demand is vacancies. However, locating information that is freely available in the public domain, entails having to look at job advertisements — it is only when a job is advertised that this information enters the public sphere and can be used by researchers to do data analysis. EDO makes a clear distinction between the concept *Vacancy* that indicates that a job opening exists with an employer, and the concept *JobAdvertisement* which is a vacancy that has been advertised via some type of advertisement media to attract applications for the position (Figure 45). Not all vacancies are necessarily advertised, and some can remain vacant for significant periods of time, depending on an organisation's need at that point in time.



**Figure 45** Association relationship diagram of *Vacancy* and *JobAdvertisement*.

When an employer drafts a job advertisement for an existing vacancy, they specify qualities of the ideal employee that will be capable of performing the specific job to help the company meet its operational and strategic goals (Figure 46). Employers also define the role description, work conditions and employee benefits that pertain to each vacancy and its geographical location.



**Figure 46** Typical job advertisement in a newspaper.

Vacancy was included in EDO to allow for future research direction. At this stage, the EDO instantiations have only been implemented using job advertisement data. It is anticipated that future work will entail capturing vacancy data from other sources too, such as directly from company sources, and via other domains such as LinkedIn<sup>32</sup> and Facebook<sup>33</sup>. As such, the class *AdvertisementMedia* has also been included in EDO (discussed in section B below).

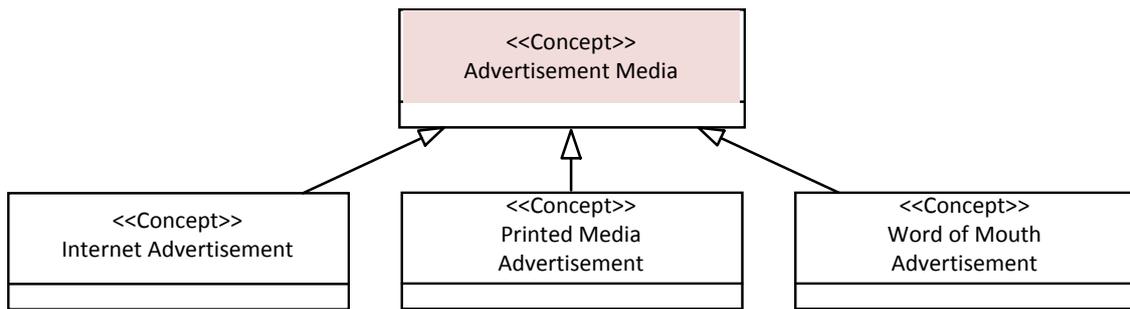
The SEEMP ontology also used the class *Vacancy*, but as described in section 5.2.3 above, the SEEMP ontology also used the class *JobVacancy*, and it was decided to call this class *JobAdvertisement* in EDO, to eliminate any confusion between vacancy and job vacancy.

## **B. Advertisement Media Class**

When an employer decides that they wish to fill a vacancy, they usually advertise the vacancy to certain stakeholders to invite applications for the vacant position. Advertising the vacancy could happen through various means, and the most common ways utilised by employers include placing a printed advertisement in a newspaper, posting it on the internet, or with Word of Mouth Advertising (Figure 47). Each of these methods will be briefly discussed below.

<sup>32</sup> <http://www.linkedin.com/>

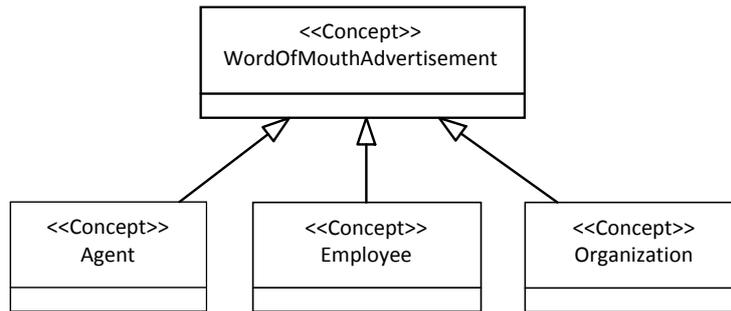
<sup>33</sup> <https://www.facebook.com/>



**Figure 47** Advertisement media specialisation relationship hierarchy.

#### *Word of Mouth Advertisements Class*

Some vacancies are not advertised via printed media (hardcopy or online). In these circumstances, employers rely on 'word of mouth' information spreading about the vacancy and inform others that applications are being invited for the position. Employers may wish to spread this information to stakeholders themselves, or inform current employees that they may do so, or employ recruitment agents to scout for possible employee matches (Figure 48).



**Figure 48** *WordOfMouthAdvertisement* composition relationship diagram.

#### *Printed Media Advertisements Class*

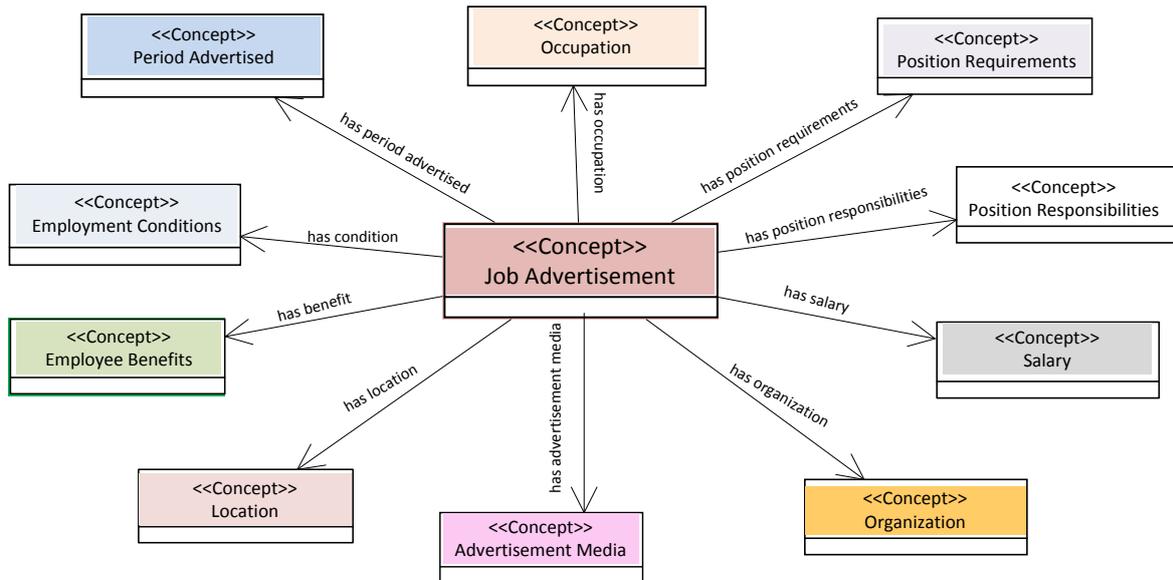
Printed media advertisements are job advertisements that are published in material where advertising space is sold, such as newspapers, newsletters, magazines and pamphlets.

#### *Internet Advertisements Class*

Website advertisements refer to vacancies being promoted on the internet, either through employers' own websites, through job boards or through social media such as Facebook or LinkedIn.

### **C. Job Advertisement Class**

Figure 49 shows the overall ontology composition for *JobAdvertisement* in EDO. Concepts that have been colour coded in the figure denote that they have subclasses; each subclass composition has been further detailed in the relevant sections below.



**Figure 49 Job Advertisement class composition.**

The generic model provides the framework found in most job advertisements. All job advertisements include as a minimum the vacancy's geographical location, the occupation required, advertisement media (discussed above) and the employer's identification (organisation). Additionally, most job advertisements also include a number of other components that are decided by each employer, to indicate:

1. Advertisement Media
2. Period Advertised
3. Position Requirements
4. Position Responsibilities
5. Salary
6. Organization
7. Occupation
8. Geographical Location
9. Employee Benefits
10. Employment Conditions

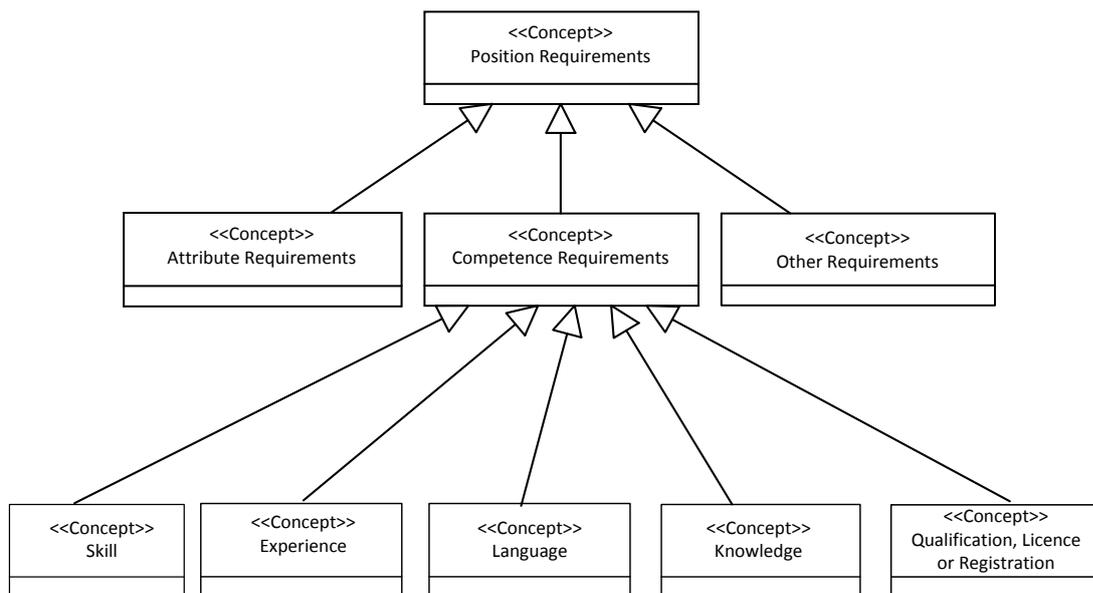
Classes two to ten have been further discussed in the following sections of this chapter.

#### **D. Position Requirements Class**

Position requirements refer to the requirements that the employer expects their future employee to bring to the position, and can be divided into three categories: attribute requirements, competence requirements and other requirements (Figure 50). A job advertisement can have multiple position requirements, thus an instance of *JobAdvertisement* can have N-ary relationships with instances of *PositionRequirement*.

For both *AttributeRequirements* and *OtherRequirements*, this is the second last specialisation level. Their next level down is the actual instance of the requirement, such as compassion, flexibility, and holding a work visa. *CompetenceRequirements*, however, are further subdivided into five subclasses that are discussed under the relevant headings to follow; Experience, Knowledge, Language, Skill, and Qualification, Licence Or Registration.

SEEMP labelled their class *RequestedWorkExperience*, however due to the EDO class *EmployeeExperience* falling under the *PositionRequirements* class, it was decided that the term 'requested' was superfluous in this instance and that *EmployeeExperience* would suffice. It was also preferred to refer to the words 'employee', 'employer' and 'employment' throughout as opposed to using the term 'work' in the EDO.



**Figure 50** Position Requirements Generalisation- Specialisation relationships.

#### *Competence Requirement Class*

In the search to locate possible existing skill ontologies, references were found relating to KOWIEN, a skill ontology supposedly hosted by the University of Essen that classifies competencies required to perform specific jobs (Bizer, 2005; Simperl, 2009), the SEEMP (Celino et al., 2010) ontology and the ontology of Trichet and Leclere (2003).

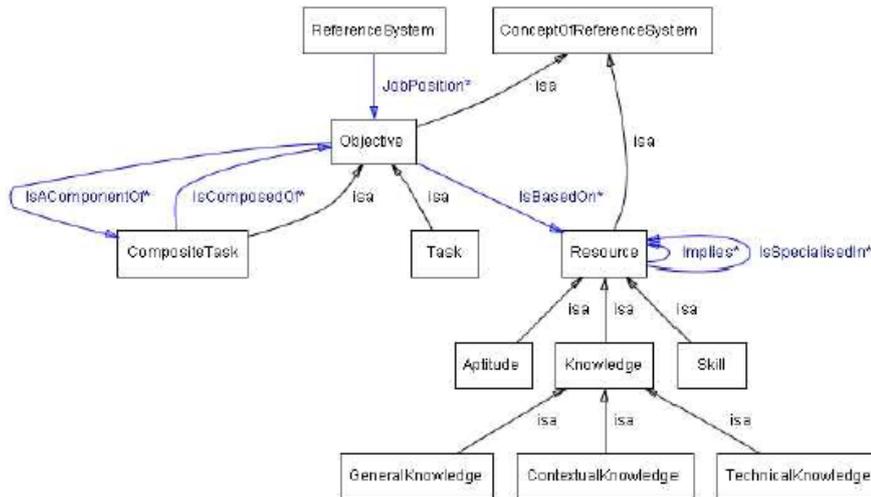
Unfortunately, the website provided in Simperl's (2009) paper for the KOWIEN ontology did not work at the time of this search, neither could it be located through various Google searches using different combinations of keywords. Even though it was anticipated that it would not be possible to use this ontology because it was in German, it still would have been preferable to investigate it for its composition for possible tips in the development of EDO. That said, EDO has specifically been constructed through the information found in online job advertisements and not from some preconceived idea of the type of skills that each job or occupation should consist of. The advent of technology and its exponential development and uptake in industry is constantly changing the nature of tasks that have been required in the past of employees. This is true in some cases for whole tasks that are not being performed by people anymore, in other

cases where only certain aspects of tasks are still performed by humans, and in other cases where tasks performed by humans may have changed. Thus, due to the constantly evolving nature of occupation-specific tasks, the developers of EDO specifically decided not to reuse an existing classification of occupation tasks or related skills, but instead to develop and continuously adapt these as time goes on. This was one of the data gathering outcomes that EDO was specifically developed for to deliver information to policy, education funding curriculum developers and immigration officials on the type of skills needed at any point in time, and to ensure their decisions are based on up-to-date, reliable data. It is acknowledged that this process is more time-consuming and costly; however, with the consideration to recycle certain aspects of existing skill ontologies, it became apparent that this process would be even more manually intensive, as each existing ontology would have to be carefully checked and matched against each job advertisement for each occupation type, to see if those skills are in fact part of that occupation's requirements at present. It was established that this process — recycling existing skill ontologies on a detailed level — would require more time and effort than to manually develop a skills ontology from scratch in order to fulfill the requirements of EDO.

That said, existing ontologies were assessed at the top level of their skill or competency hierarchy to ascertain their grouping of classes and subclasses, without becoming specific with regards to a certain industry or geographical region.

Trichet and Leclere (2003) classified their main employee ability class — *Resource* — into the following subclasses: *aptitude*, *knowledge* and *skill*, with a further subdivision of the *knowledge* class — *GeneralKnowledge*, *ContextualKnowledge* and *TechnicalKnowledge* (Figure 51). In EDO's development, it was decided that the knowledge class need not be further divided as in Trichet and Leclere's ontology (2003); however, EDO's requirement to be a detailed employer demand

intelligence tool, meant that EDO needed to have additional classes on the same level as knowledge and skill: *Experience, Language, Qualification, Licence or Registration*.



**Figure 51** Trichet and Leclere's competency reference system (2003).

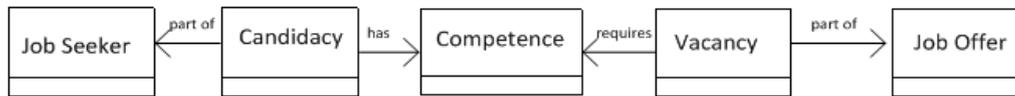
### Skill Class

Skills refer to having the ability to perform a certain task or being capable of executing specific jobs. Examples include coaching skills, good communication skills, problem solving skills and time management skills.

Where a requirement in the job advertisement was listed as 'ability or skill', it was included under the *Skill* class. Where the employer specifically indicated experience, the requirement was listed under the *Experience* class (below) instead. These two requirements are not exactly the same as a skill could be learned in a course or at home whereas when employers indicate they want experience in a specific field, it means the ideal candidate would need to have professional level experience for that specific action.

The SEEMP ontology (Villazón-Terrazas et al., 2011) also includes a *Competence* class that relates on the one side to the job seeker through a relationship with *Candidacy*, and on the other side to job offer through a relationship with *Vacancy* (see Figure 52). The SEEMP ontology assigned

the subclasses *Skill* and *Language* to the *Competence* class according to the SEEMP reference ontology. We decided to follow the SEEMP design for EDO.



**Figure 52 SEEMP (Villazón-Terrazas, et al., 2011) relationships for Competence.**

### Language Class

In EDO, *Language* refers to ability that an employee has to communicate sufficiently in a specific language, to able to work in an environment where that is the sole form of communication used. Some positions may require employees to be able to communicate fluently in more than one language, such as both English and Spanish.

As discussed, *Language* was manually recycled from the SEEMP (Villazón-Terrazas et al., 2011) ontology. SEEMP reused the widely accepted ISO 639 classification system (“Browse or Search the ISO Language Codes”, 2010) for *Language* which was also adopted in EDO and includes 559 different languages. The 559 concepts were manually entered into OWL format using Protégé during the development of EDO; this was done from the text-based classification standard found in an online table for ISO 639<sup>34</sup>. This took approximately four man hours. The manual evaluation of this standard classification system revealed that four concepts found in the ISO639 classification would be eliminated. These concepts were offered in ISO 639 for instances where an entry did not relate to a specific language, and in order to enter some kind of code for these instances, these fillers or descriptors were provided e.g. NA for Not Applicable.

<sup>34</sup> [http://www.loc.gov/standards/iso639-2/php/code\\_list.php](http://www.loc.gov/standards/iso639-2/php/code_list.php)

## Experience Class

Experience requirements refer to a period of time that an employee has spent in a certain role, doing a specific task, working in a specific industry or at a specific level (Figure 53). Examples include having care planning experience, accreditation audit experience and management experience.

### Community Mental Health Nurse

As a member of the community mental health crisis team you will be delivering a high standard of mental health and crisis care in the community.

Being part of the multidisciplinary team you will be providing...

Crisis assessments during a client's acute episode.

Provide specialist assessment, triage and treatment in the emergency departments covering all age groups from child adolescents through to aged care.

Facilitation of transfer from Emergency care to the Mental Health inpatient unit.

Undertake Medication administration.

Be involved in daily hand over of patient care plans.

Essential: **2 years post graduate experience in community mental Health**, current Australia registration, ability to work within a team and also autonomously.

**Start date** : Open start date.  
**Position Type** : This is a permanent position contract also available for the right applicant.  
**How to apply** : Submit your C.V now to find out how you can secure a great opportunity.

APPLY TO WORK WITH SUGARMAN TODAY.  
PLEASE SEND YOUR C.V TO  
leigh@sugarmangroup.com.au  
or call us on  
02 9549 5700

**Figure 53** An example of an experience requirement in a job advertisement.

## Knowledge Class

Knowledge requirements relate to an employee's understanding of or familiarity with a specific topic, and can include any topic that may assist the employee in performing their duties. Examples include clinical knowledge, legislative care knowledge and regulatory advice.

## Qualification, Licence or Registration Class

Qualification requirements refer to the attainment of an educational course, workshop or training where the person was assessed as capable in the specific discipline. Examples include a Bachelor of Commerce degree, Certificate III in Aged Care and a Doctor of Philosophy degree.

Holding a license or registration means that an authoritative institution has issued the employee with the rights to execute a certain duty or task or be able to obtain and keep certain instruments. It usually entails paying an annual fee to keep the registration valid or having to be reassessed after a certain period of time to be able to renew the license or registration. Examples include medical board registrations, firearm licenses and forklift licenses.

## Attribute Requirements Class

Attribute requirements signify a characteristic or quality that is inherent to an employee and is often referred to as personality traits. Examples include being motivated, compassionate, reliable or well presented (Figure 54).



**SURELINE**  
CARE SERVICES

Providing Tailored Solutions

### Aged Carers Wanted for PT Work!

- Join a small and close knit team
- Weekly pay and good conditions
- Enjoy the satisfaction of a variety of work

Sureline have Facility based and Community roles available for experienced Aged Care Workers. We have positions vacant both North and South of the river for those individuals looking for 20+ hours per week.

We pay top \$\$ to our Aged carers in recognition of their hard work and contribution. Our company has a great reputation for providing quality services in a variety of areas.

To be successful in this role and to join our small close knit team you will need at least 6 months experience in a similar position. You will need a police clearance, a first aid certificate and your own reliable vehicle.

We're looking for **compassionate** people who are looking for long term employment.

To apply please contact our friendly team on the details provided:  
Ph: 9330 7500  
Email: sam@surelinecares.com.au  
Email: Please click the 'Apply Now' button below.

**Figure 54** An example of an attribute requirement 'compassionate' in a job advertisement.

### *Other Requirements Class*

Other requirements include specific necessities that the employer has identified as necessary for the ideal candidate to successfully perform their job. These requirements usually relate to things that form part of the employee's direct environment or situation and are not directly embedded in the employee's persona. Examples include owning your own vehicle, having professional references and holding a work visa.

### **E. Potential Skill Shortage Class**

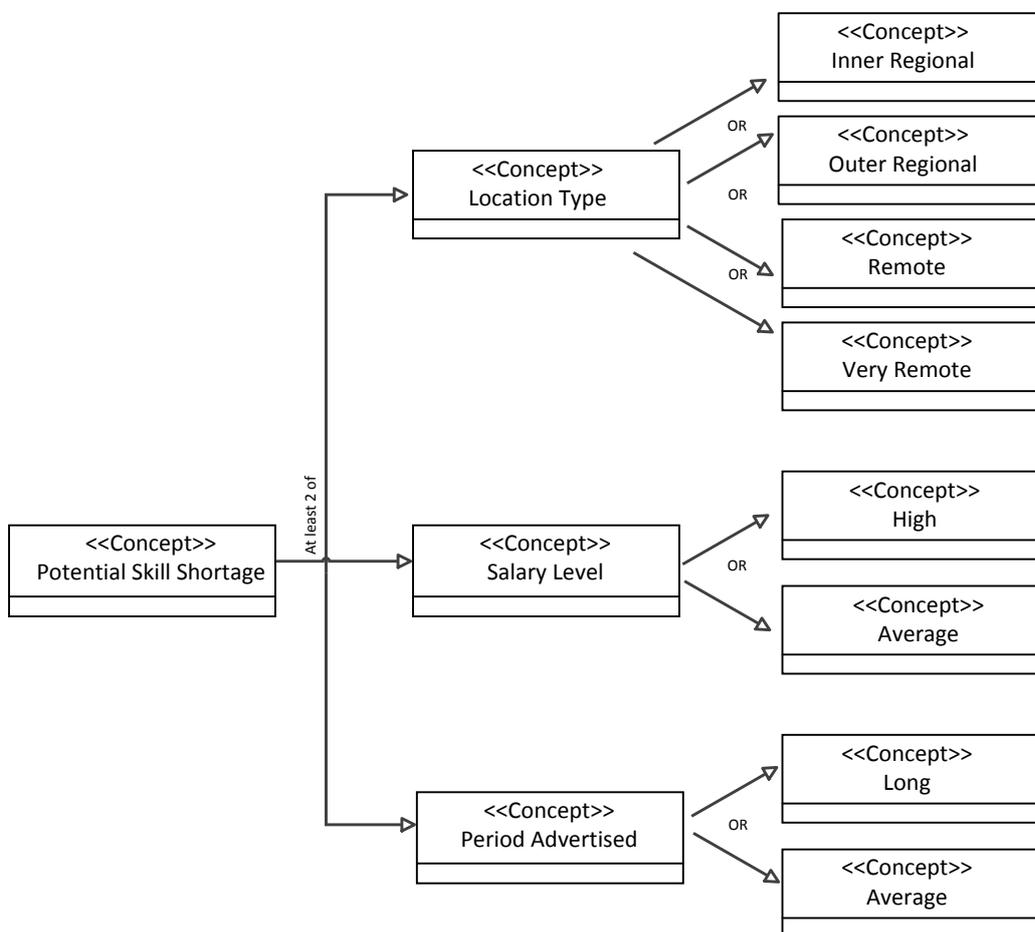
As discussed in a previous chapter, the Department of Education, Employment and Workplace Relations (DEEWR) ("Skill Shortages", 2007) defined the existence of a skill shortage as a condition 'when employers are unable to fill or have considerable difficulty filling vacancies for an occupation, or significant specialised skill needs within that occupation, at current levels of remuneration and conditions of employment, and in reasonably accessible locations' (DEEWR, 2008-09, n.p.).

The DEEWR skill shortage definition refers in basic terms to the difficulty of finding somebody to do the job. This difficulty can be measured by analysing the geographical location of the vacancy ('reasonable accessible locations'), the benefits that the employer is offering to the employee ('at current levels of remuneration and conditions of employment') and the time it takes to fill the position.

The EDO's *PotentialSkillShortage* class has been designed to capture vacancies that include the skill shortage indicators mentioned in the above definition (Figure 55). When a vacancy is identified as meeting two of the mentioned three conditions in the definition, the EDO will flag it as a **potential** skill shortage for investigation.

To assist in identifying occupations that could potentially qualify as being a skill shortage, it is necessary to analyse job advertisements by gathering as much information as possible relating

to these three skill shortage indicators. Unfortunately, not all job advertisements display the salary that is on offer for the specific vacancy; however, it is still helpful to collect, analyse and identify vacancies where at least two of the three indicators are displayed, in order to identify a *potential* skill shortage. It will be up to the workforce development specialist or skill shortage expert to investigate all cases where vacancies have been flagged as a potential skill shortage, to establish whether it is indeed an actual skill shortage. EDO thus provides the first step of identifying or flagging potential problem areas to be investigated in this regard.



**Figure 55** The EDO indicators of a potential skill shortage based on the DEEWR definition of the phenomenon.

The three indicators (Period Advertised, Salary Level and Location Type) have been further discussed below under the relevant headings.

## **F. Period Advertised Class**

As soon as a job advertisement is released to public attention, regardless of the manner in which this happens (i.e. whether it is published in a newspaper, online or outsourced to the recruitment agency), the job vacancy has started to be advertised and this date is recorded as the job vacancy's period advertised start date. Alternatively, the job vacancy period advertised end date is effective on the day the job advertisement is removed from the market. Counting the days from the start to the end date for which the job was effectively advertised provides the period (number of days/ weeks) over which stakeholders were invited to apply for the position. A vacancy needs to be advertised for a long period to be regarded as a high risk indicator of a skill shortage. If a vacancy was advertised for an average period, and this vacancy has an additional potential skill shortage indicator from the classes *SalaryLevel* or *LocationType* that also qualifies as a potential skill shortage, the vacancy will be flagged as a potential skill shortage to be further investigated by the domain expert.

It is also possible that employers may decide to cancel or retract their job advertisement after a certain period of time that it has been advertised without having employed a suitable candidate in the position. This decision could happen due to reasons such as financial or organisational changes, not needing an employee in that domain at the time anymore, or not being able to find a candidate who meets the employer's requirements (Van Ours & Ridder, 1992); however, this situation occurs very rarely and in most cases the vacancy will continue to be advertised until the position has been filled.

## **G. Salary Level Ontology**

A salary is the form of periodic payment (e.g. monthly) from an employer to an employee in exchange for the services that the employee provides to the organisation that has employed the person.

Salary level concepts are included in the Tier One EDO as either High, Average, or Low salaries. These classifications are merely that- classifications or indications of where the salary level is deemed to be at a particular point in time when information obtained for a specific instance is analysed. The top level (Tier One) ontology will not be populated with specific salaries as such, but only the classification that the instance belong to (e.g. average salary). By only using these classifications at the Tier One ontology, it allows for changes in practise which will not prompt an updated version of the Tier One ontology. However, the Tier Two ontology, which is specific to each type of occupation, will have to be versioned for specific market periods to reflect the changing landscapes in employee remuneration. Updating salary levels for each occupation periodically, will be part of the normal maintenance that the ontology will have to undergo in order to reflect current industry practices. The pros of this approach, is that the ontology allows for each occupation type to be classified into a specific salary level without having to change the Tier One ontology used by all the occupation types. The con of the approach is that Tier Two would need periodic updating.

There is more than one form of compensation that an organisation may offer an employee in exchange for his or her work; however, most employment usually has a form of financial or monetary benefit in the first instance, with other benefits viewed as 'in addition' to the monetary benefit provided (e.g. housing subsidy or salary packaging). Depending on the nature of the position (e.g. contract or permanent), an employer may decide to offer the monetary benefit in either the form of a salary (periodic basis), or in the form of piece wages where each job, hour or unit, is paid separately. For the purposes of EDO, it was decided to only use the term salary to encompass both these monetary concepts that are on offer for a vacancy.

The SEEMP ontology (Villazón-Terrazas et al., 2011) used the terminology of 'compensation'. Even though compensation includes the concept of a salary as well as other forms of receiving a benefit for the work that has been performed, it was decided to separate the concept

'compensation' into two distinct classes in the EDO: that of *Salary* and that of *EmployeeBenefits* (discussed in section J below).

One of the main purposes of EDO is to be able to flag potential skill shortage vacancies. In order to do this, the concept of salary needs to be subdivided into high, average, or low salary levels. Provided that these salary level subclasses allow the EDO to classify a vacancy's stance in terms of its potential skill shortage risk, that is, if a vacancy's salary that is on offer is classified as either a high or an average level of salary for that type of vacancy (occupation, experience and responsibility considered) — it could indicate that the vacancy may be a potential skill shortage. As discussed, this indicator needs to be linked to at least one other *PotentialSkillShortage* indicator in order for it to be flagged as a possible problem area for the domain expert to investigate further.

To allow the EDO to classify a salary as either a Low, Average or High salary, certain parameters need to be established for each occupation so the vacancy's salary can be assigned to the correct category. At this stage, EDO has not yet been developed to automatically assign salary instances to specific salary level types, due to the lack of data available for salary scales in the public domain. It is anticipated that tools such as Payscale<sup>35</sup> could be used in EDO's further development stages to assist in such automatic categorisation of salaries (see Chapter Nine).

## H. Geographical Location Ontology

Each vacancy relates to a geographical location where the vacancy exists. In Australia, the Australian Bureau of Statistics produces and publishes the Australian Standard Geographical Classification (ASGC)<sup>36</sup> to enable researchers to utilise this information in order to inform their own research, or to provide greater clarification on other research which includes these

---

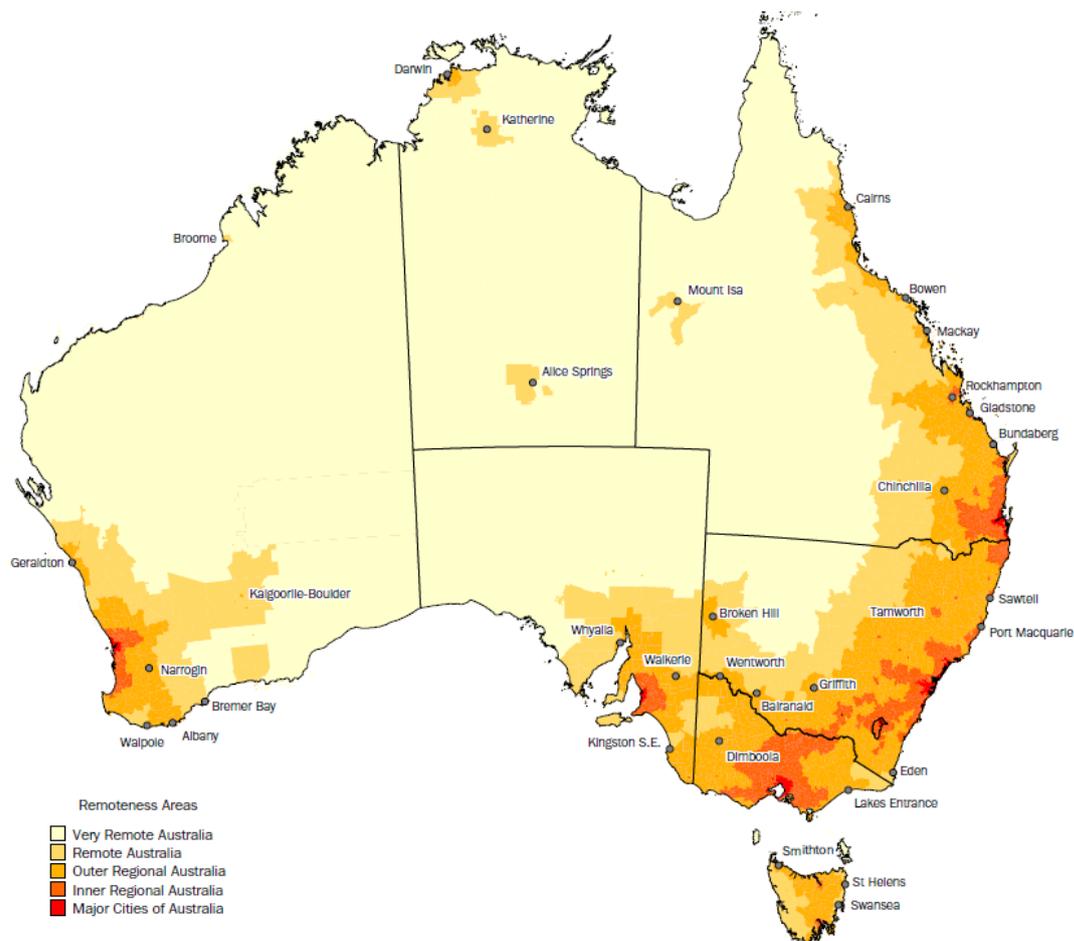
<sup>35</sup> <http://www.payscale.com/about.asp>

<sup>36</sup>

[http://www.abs.gov.au/websitedbs/D3310114.nsf/home/Australian+Standard+Geographical+Classification+\(ASGC\)](http://www.abs.gov.au/websitedbs/D3310114.nsf/home/Australian+Standard+Geographical+Classification+(ASGC))

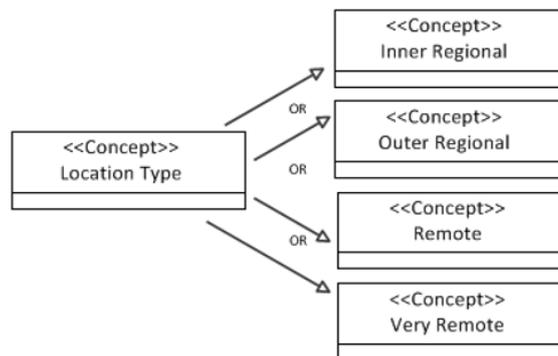
geographical classifications. The ABS has also developed a Remoteness Classification Structure for utilization in studies and policy development relating to the remoteness of a location as part of the ASGC. The remoteness of a location is very important in Australia due to the very long distances that typically need to be travelled from major cities to other areas to provide essential services for these remote locations.

The ABS classifies remoteness in the Accessibility/Remoteness Index of Australia (ARIA) which comprises six Remoteness Areas (RAs): 1) Major Cities of Australia; 2) Inner Regional Australia; 3) Outer Regional Australia; 4) Remote Australia; 5) Very Remote Australia, and 6) Migratory. These areas have been classified based on Australia's Census 2006 data (2011) and grouped into areas that share common characteristics or remoteness (Figure 56).



**Figure 56 ADGC Remoteness Classification (2001).**

As discussed, the DEEWR definition of a skill shortage includes reference to the physical or geographical proximity of a vacancy to goods and services — ‘reasonably accessible locations’. To match the DEEWR concept of a location being ‘reasonably accessible’ to concepts of the ASGC, the EDO developers decided to include only the *MajorCityOfAustralia* ASGC location parameter as a location type that is ‘reasonably accessible’ for Australia. Even though *InnerRegionalAustralia* could also potentially be viewed as being ‘reasonably accessible’ from goods and services, this class was perceived to have been borderline and as such was included instead to allow domain experts to make an informed judgement about the vacancy’s potential skill shortage risk when coupled with the other flagged indicators for that vacancy. As such, the Location Type classes included in the *PotentialSkillShortage* hierarchy are shown in Figure 57.



**Figure 57** Location Type potential skill shortage subclasses.

The SEEMP ontology (Villazón-Terrazas et al., 2011) also used location. However the EDO requirements are different in the sense that locations are based in Australia, and they are based on measurement of how far they are from metropolitan areas which influences the probability of a vacancy being filled.

Based on the EDO’s potential skill shortage indicators, some of the scenarios that could be captured by the EDO as a *PotentialSkillShortage* are listed below. Clearly, many more scenarios are possible.

1. Rural location  $\cap$  high salary level  $\cup$  long period advertised;
2. High salary level  $\cap$  very remote  $\cup$  average period advertised;
3. Remote location  $\cap$  high salary level  $\cup$  long period advertised;
4. Remote location  $\cap$  high salary level  $\cup$  average period advertised;
5. Long period advertised  $\cap$  average salary level  $\cup$  outer regional location;
6. Rural location  $\cap$  averages salary level  $\cup$  average period advertised;
7. Inner regional location  $\cap$  average salary level  $\cup$  long period advertised;
8. Average salary level  $\cap$  very Remote location  $\cup$  average period advertised.

#### **I. Position responsibilities**

As per the title, position responsibilities refer to those tasks an employee will be expected to perform as part of the position that the person has been employed in.

The developers of EDO view position responsibilities as a secondary factor in regards to establishing workforce development needs and potential skill shortages. As such and due to time and resource constraints, this class has been scheduled for further development in future stages of EDO research.

#### **J. Employee benefits**

It is important to identify the kinds of skills that employers require for specific occupations and to establish what the employer is offering in return to an employee who may hold these skills. As discussed, services performed can be remunerated through two means — monetary outputs and other benefits. This class describes those other benefits (over and above the salary) that may be offered to an employee for the job that they will be performing (Figure 58). Examples include free car parking, ongoing training that is provided, and being able to receive a

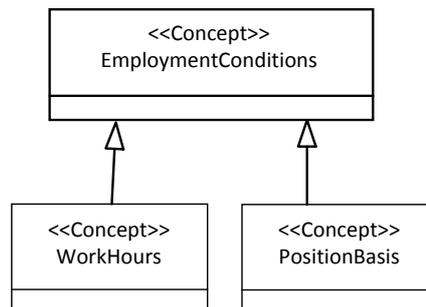
sponsorship visa to work in Australia. These types of benefits can significantly increase the attractiveness of a position for a potential candidate.



**Figure 58** Job offer equilibrium showing example of employer's and employee's offers for a job vacancy.

### K. Employment conditions

Employment conditions are conditions that indicate the work hours and position basis of the vacancy (Figure 59). Options for position bases include being employed on a casual basis, a full time basis or a relief basis. The concept of work hours cover things like having to work weekend shifts, being required to travel or only being able to work relief shifts.



**Figure 59** EmploymentConditions Class hierarchy.

The SEEMP ontology used 'work condition' as a class, however during the EDO development it was decided to call this class *EmploymentConditions* instead, to fit the naming convention of EDO throughout, where the words 'employee', 'employer' and 'employment' were the

preferred vocabulary. In terms of the SEEMP class 'contract type', the EDO's preferred terminology for this concept was *PositionBasis* instead.

#### **L. Organization**

Organization refers to the organisation that holds the vacancy and will ultimately employ the ideal candidate. Inclusion of the concept of organisation allows the identification of changes that were specific to particular companies versus those changes that are more widespread (Vaughan, 1988).

It was decided to reuse the class title *Organization* as per SEEMP (Villazón-Terrazas, et al., 2011) and PROTON (Terziev et al., 2005) ontologies. Even though EDO has been developed for an Australian environment, and as such would typically be spelled 'organisation' — with an 's' — the EDO developers decided to globalise the EDO through the usage of the American spelling convention for future potential developments of the ontology.

#### **M. Occupation**

Various occupation classification systems have been developed worldwide, and are widely used and welcomed as building blocks in eRecruitment ontology development (Simperl, 2009). Examples include O\*Net (Occupational Net), ISIC (International Standard Industrial Classification of Economic Activities), SOC (Standard Occupational Classification) and NAICS (North American Industry Classification System). Even though many classifications exist, they have been developed to different stages and with varying levels of detail. Where some occupational classifications have been developed to include descriptions of occupation titles, associated skills as well as qualification, others have only been developed as a hierarchical system of occupational groups and subgroups.

The Australian and New Zealand Standard Classification of Occupations (now called ANZSCO but previously known as ASCO) was developed as a skill-based classification of occupations to

provide a broad overview of the technical division of labour in skill terms (Vaughan, 1988). As the name suggests, this classification system was built for the Australian and New Zealand job market and is specific to these countries' occupational structures.

Using an occupational classification structure as basis of the EDO enables workforce development and skill shortages to be accurately identified and related to specific occupation titles, and translated into training and educational needs. Using this standard classification that has been developed jointly by the ABS and DEEWR enables outcomes of this framework to be compared to previous and future studies, and to relate it to the broader labour market picture.

The SEEMP ontology (Villazón-Terrazas, et al., 2011) used the Class '*Occupation*'. It was decided that a more detailed description of '*PositionOccupationType*' would be used in EDO to eliminate possible confusion in the future as to whether the class refers to the occupation that the job seeker holds, or that the job position relates to.

As discussed, EDO was developed in a two-tiered structure to allow for EDO to be used in practice for those EDO occupation types that have been finalised, while other occupation types are still being developed. The manual intensity of the development of an extensive ontology such as EDO means that each stage can take a significant time to develop. If EDO was not staged or tiered in its development, implementation and further developments of those occupational types that are ready for uptake would have been stalled.

The *PositionOccupationType* class that is included in Tier One of EDO is merely the top level hierarchy label to indicate that this class has a relationship with the class *JobAdvertisement*, however all subclasses of *PositionOccupationType* have been included under EDO Tier Two, discussed in the section below. Figure 60 shows the top level hierarchical composition of the class *PositionOccupationType*.

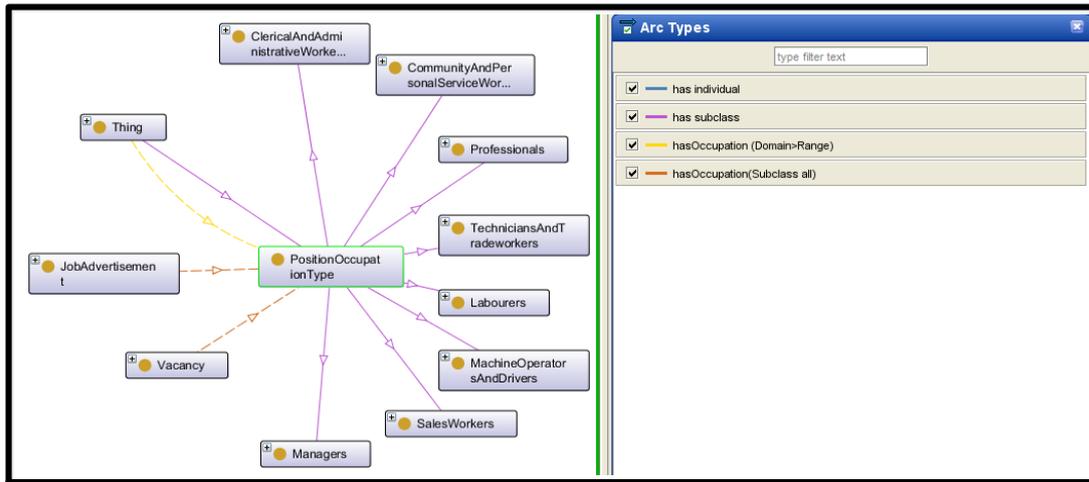


Figure 60 OntoGraf representation of the PositionOccupationType class' main relationships.

### 5.4.2 Employer Demand Ontology: Tier Two

ANZSCO consists of approximately 520 occupations. Figure 61 shows the top level (major groups) occupation types from ANZSCO's spreadsheet. The subclasses for each of these major groups have been provided as OntoGraf presentations in the appendix of this thesis, with the exception of the *MidwiferyAndNursingProfessionals* class, which is further described below.

**Australian Bureau of Statistics / Statistics New Zealand**  
 ABS Cat. no. 1220.0 ANZSCO -- Australian and New Zealand Standard Classification of Occupations, First Edition  
 Structure  
 Table 1: Major Groups

Major Group	Predominant Skill Levels
1 Managers	1, 2
2 Professionals	1
3 Technicians and Trades Workers	2, 3
4 Community and Personal Service Workers	2, 3, 4, 5
5 Clerical and Administrative Workers	2, 3, 4, 5
6 Sales Workers	3, 4, 5
7 Machinery Operators and Drivers	4
8 Labourers	4, 5

© Commonwealth of Australia, 2006  
 © Crown Copyright New Zealand, 2006

Figure 61 Major occupational grouping from ANZSCO spreadsheet.

Due to budget and resource constraints, EDO's second tier has so far only been manually developed for the 15 occupations found under the ANZSCO occupational group Midwifery and Nursing Professionals from the approximate 520 existing occupations. The Midwifery and Nursing Professionals group is a part of the greater ANZSCO Professionals group and represents the following occupations:

- Midwife
- Nurse Educator
- Nurse Researcher
- Nurse Manager
- Nurse Practitioner
- Registered Nurse in Aged Care
- Registered Nurse in Child and Family Health
- Registered Nurse in Community Health
- Registered Nurse in Critical Care and Emergency
- Registered Nurse in Developmental Disability
- Registered Nurse in Medical
- Registered Nurse in Medical Practice
- Registered Nurse in Mental Health
- Registered Nurse Not Elsewhere Classified
- Registered Nurse in Perioperative
- Registered Nurse in Surgical

## **5.5 Conclusion**

This chapter is an integral part of EDIT, as it provides the backbone to EDIT's structure — the EDO and its detailed developmental stages. The chapter has addressed research issue one — underlying knowledge representation — and provides a brief overview of some main ontology

development methodologies. The methodology that was chosen to develop the EDO (NeOn) is subsequently discussed as well as the steps involved in EDO's development, the notation of the EDO and the design principles that were employed. The chapter concludes with a detailed description of all the EDO classes and subclasses in a two-tiered format.

The following chapter will provide a solution to research issue two — occupation specific data querying.

# Chapter 6 - Employer Demand Ontology Instance Population

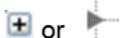
## 6.1 Introduction

The previous chapter provided the detailed solution for research issue one — underlying knowledge representation. This chapter explains the steps that were taken to address research issue two — occupation specific data querying. The chapter details the manual population of the Employer Demand Ontology instances (justification for manual process provided in Section 6.2 below). In summary, there are three classes (occupational type, geographical location and language) that were instantiated based on existing hierarchical information from respected sources, as detailed in the relevant sections in this chapter. The other instances were all populated from the job advertisements found on the job board SEEK for the month of January 2011.

As in the previous chapter, the EDO Tier One instances are first clarified, after which the EDO Tier Two instance population is explained. As discussed, due to resource constraints, only the occupational group of Midwifery and Nursing Professionals were further developed for the appropriate classes of the EDO, with a focus on job advertisements for the state of Western Australia.

## 6.2 Instantiation Process

Information technology tools have not yet been developed to such an extent that ontology instances can be confidently populated with semi-automated processes. Therefore, it was decided to populate the initial instances in this research manually. This research utilised the ontology software Protégé to instantiate the EDO — Figure 62 provides a list of modelling notations to aid graphical depiction understanding for this thesis.

Notation	Terminology	Semantics
	Primitive Class	If something is a member of this class then it is necessary to fulfil the conditions [Horridge 2011].
	Defined Class	If something fulfils the conditions then it must be a member of this class [Horridge 2011].
	Subclass Relationship (pink colour)	Superclass-subclass hierarchy. Note that it has the same notation with instance relationship but in different colour i.e. in pink colour.
	More Subclasses	More subclasses in a hierarchy.
	Object Property Relationship (different colours represent different relationships)	Link two instances or classes together.
	Instance / Individual	Objects in the domain in which we are interested.
	Instance / Individual Relationship (blue colour)	Instances of the classes. Note that it has the same notation with subclass relationship but in different colour i.e. in blue colour.

**Figure 62** Protégé modelling notations (Forbes, Wongthongtham, Singh, & Thompson, 2013 (in press)).

The EDO instance population was done through a trial run of one month's job advertisements gathered during January 2011. The advertisements were only collected on the job board Seek for the geographical location of Western Australia, and only for the ANZSCO occupational group Midwifery and Nursing Professionals. Except for those classes where existing categorisation resources were recycled (e.g. ANZSCO for *OccupationType* and ISO 639 for *Language*), the specific steps taken to populate the remaining classes with instances were as follows:

1. Setup job alerts in Seek to email all new job advertisements being advertised for a one month period with criteria Nursing, Western Australia (Figure 63).
2. Opened job alert emails daily to retrieve new postings and saved to a folder under the specific occupation title.

3. After the one month period, printed the total of 416 job advertisements under their respective job categories.
4. Read through each advertisement to establish whether it was indeed a new advertisement, or whether it was just a duplicate/ reposting of an already existing advertisement among the 416 advertisements in the Employer Demand Dataset (EDD). There were a total of 169 (41%) duplicates found among the total number of 416 job advertisements in the trial period (Table 7). This is consistent with the findings by Robinson et al. (2010) who found that 30-40% of their job advertisements analysed were duplicates.
5. Eliminated all advertisements that were posted in the wrong categories, e.g. those that were not actually for Registered Nurses or based in Western Australia. 102 advertisements were found to be posted in the wrong categories in terms of what they were listed under. For example, the category the advertisement was listed in might have been related to a job opening in Brisbane according to the free text contained in the job advertisement, when the listing (according to the drill down category definition to search for job advertisements) was for vacancies existing in Western Australia.
6. Assigned a unique job advertisement identification number for each job advertisement that was applicable in the dataset, according to the coding protocol explained in section 6.2.2A.
7. Identified and marked each job advertisement's applicable instances with a highlighter that related to the EDO classes discussed in the previous chapter.
8. Manually created each new instance into the EDD (in Protégé), cross-checking so as not to duplicate previously entered instances in perhaps a slightly different format. Where similar instances were found with minor varying grammatical versions, the most appropriate/ common instance label would be entered and other formats of the same instance would be listed under the 'comments' section of the annotations tab for each instance (Figure 64). Firstly, all parent-child instances were recorded (for example, all employment conditions, employee benefits and position requirements). Secondly, all composition relationship

instances were recorded (that is, the relationships that each specific job advertisement had with any other type of instance recorded in the previous step).

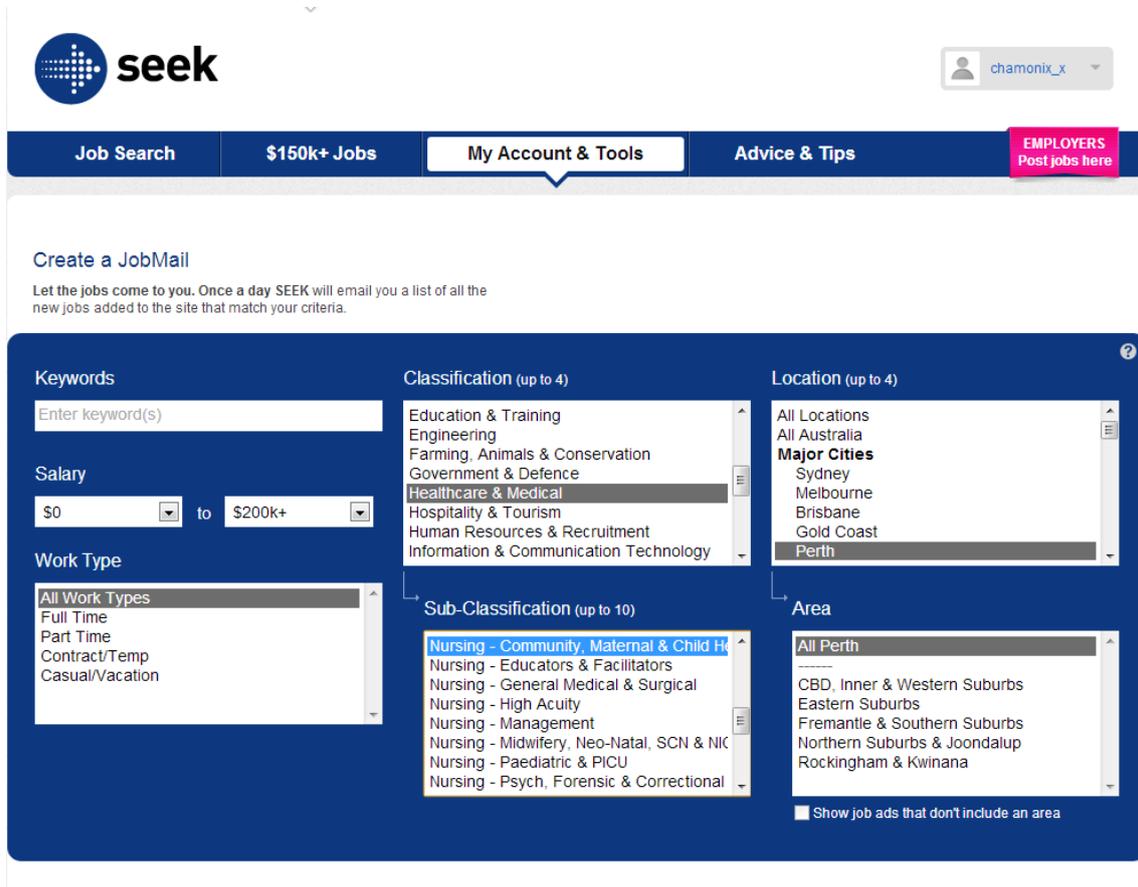
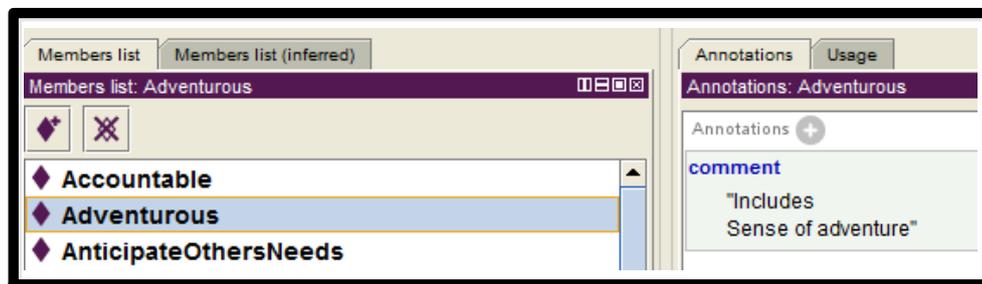


Figure 63 Setting up a job alert on the Seek.com.au website.

Parameter	Number
Total number of advertisements collected for the 1 month period	416
Total number of advertisements that were usable	145
Total number of duplicate advertisements found	169
Total number of advertisements not related to <i>Registered Nurse</i> vacancies in <i>Western Australia</i> (even though they were advertised under these categories on the job board)	102
Accuracy rate of 145/416	35%

**Table 7** General statistics from the one month trial during January 2011.



**Figure 64** Annotation included for the instance *Adventurous*.

Populating EDO or any other eHumanResourcesManagement and eRecruitment system is still an extremely onerous task at present. The scope of this thesis only allows a person to populate one occupation type's data as an example for the workings of the EDO. The EDO Registered Nurse ontology generated over 5,000 ontological concepts — a very large dataset which highlights the manual intensity required to perform such work. This is not an uncommon number for competency ontologies. Biesalski and Abecker (2005) (automotive industry) and Lau

and Sure (2002) (insurance industry) coincidentally reported having approximately 700 single *competencies* in their respective skill ontologies.

From Table 7, it is clear that the accuracy rate (number of job advertisements that are relevant to the classifications for which it was posted, and which were not duplicates of each other), and for job advertisements found with the Seek job board is very low. Only approximately one-third of the advertisements were usable in the ontology development, as only these advertisements were statistically and categorically correctly listed. As discussed in Chapter 2, many economic forecasts are done on the tallying of the number of job advertisements listed each week/ month/ quarter, however as proved in the EDO trial period, those forecasts are by and large based on skewed data. Unfortunately, due to the complicated nature of eliminating some duplicates and identifying erroneously listed job advertisements per category, those studies' outcomes are still used to inform policy development, funding and workforce planning decisions.

The following two sections detail the decisions taken during the population of each EDO class's instances in the one month trial period.

### **6.2.1 Employer Demand Ontology: Tier One Instance Population**

#### **A. Vacancy**

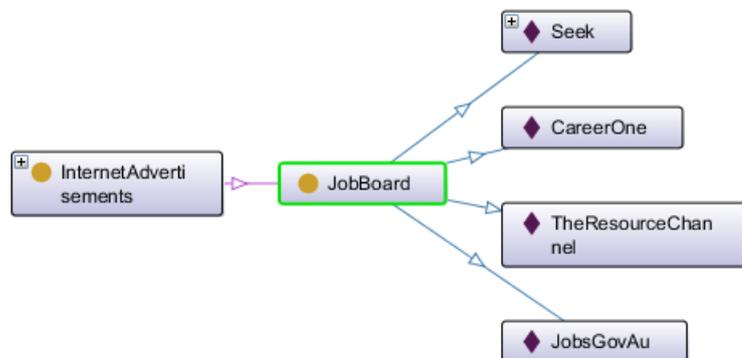
As discussed, the *vacancy* refers to the actual vacancy that exists at an employer for a vacant position related to a specific occupation. At this point in time, EDO development has only been done for actual job advertisements found on job boards as opposed to being able to access employers' private data to record all vacancies they may hold.

Future work for EDO includes the potential to develop it to such an extent that it can record actual vacancies that exist in employers' records. While job advertisements strongly suggest that a vacancy exist with an employer for a specific position, the EDO will not be populated for this class's instances until such time that data is received directly from employers confirming their

number and type of vacancies at any point in time; an employer may have several vacancies but may decide not to advertise at a point in time. Thus, at this stage it is more accurate to record job advertisements just as that — job advertisements — and not vacancies. Vacancies could also be advertised via more than one avenue (see the following section) and until such time that the EDO has been further developed to include instantiation population of these too, it was decided to only record job advertisement instances as *JobAdvertisement* instances and not as vacancies as well.

## B. Advertisement Media

As discussed, a vacancy can be advertised via many means, e.g. online, in newspapers or other printed material, and by word of mouth. At this stage, the only *AdvertisementMedia* instances that have been recorded are those relating to four popular job boards in Australia: CareerOne, Jobs.Gov.au, Seek.com and The Resource Channel (Figure 65).



**Figure 65** OntoGraf presentation of instances of the EDO class *JobBoard*.

Because EDO has been developed from information found on the job board Seek, and the other job boards being well known, these have been included. The class *SocialMedia* has also been populated with the most famous instances of this class at this stage in time: *Facebook*, *LinkedIn* and *VirtualReality*. Other advertisement media instances have not yet been included in the EDO and will be done as part of future work.

## **C. Job Advertisement**

A total of 145 job advertisement instances were recorded for the trial period. The tasks listed below were performed to instantiate the job advertisements:

### **1. Remove job advertisements unrelated to Midwifery and Nursing Professionals**

Many job advertisements advertised for other types of medical professionals such as enrolled nurses or carers and not Registered Nurses as such. EDO specifically excluded these as the prototype focus is only on Registered Nurses and any subspeciality of this occupation. This meant that only job advertisements that included the occupation title of Registered Nurse were considered. As an example, the following titles were listed under the Seek Aged Care nursing category but were excluded from the EDO Registered Nurse dataset: Care Aide, Enrolled Nurses, Clinical Manager, Community Support Workers, Community Care Workers, Facility Manager, Director of Nursing, Aged, Disability and Home and Community Care Worker, Clinical Nurse, Carer, Aged and Disability Carers, Assistant Carer, Night Duty Carer, Clinical & Lifestyle Coordinator, Aged Care Management Professionals, Multi Skilled Carer, Professional Aged Carers, Clinical Care Coordinator, Live-in Carer, Home Care Manager, Practice Nurse, Senior Care Coordinator, Care Coordinator, Assistants in Nursing and Occupational Therapist.

In the Nurses — Managers Seek category, only those vacancies that specifically stated the ideal candidate needs to be registered at AHPRA or needs to have a Nursing Degree were included. Several advertisements did not clearly state whether the advertised manager position specifically required the future employee to be a Registered Nurse or not. Where it could comfortably be assumed the position was specifically for a qualified Registered Nurse, even if it was not specifically stated, the advertisement was instantiated and populated into the EDD.

Where no specific Midwifery or Nursing Professionals occupation was indicated, there was no occupation instance recorded for that advertisement. Many advertisements merely state they

need Registered Nurses for many or a variety of specialisations, and thus for these a specific occupational title could not be recorded. However, these advertisements exist under the specific category indicated in Seek and as such provide an occupation specialisation to some degree. For example, an advertisement found in the 'General' category on Seek for Registered Nurses, was listed under the RNMS job advertisement title code which maps to the Registered Nurse (Medical), Registered Nurse (Medical Practice), Nurse Practitioner and Registered Nurse (Surgical) categories.

## **2. Consolidate multiple vacancies advertised together into one job advertisement instance**

Where job advertisements contained more than one vacancy described within the advertisement, it was still recorded as one advertisement, but the various types of Registered Nurse positions described in the advertisement would all be included as relationships of that specific job advertisement. Where this does not necessarily provide an accurate picture of the number of job advertisements to the number of vacancies as such, it was decided that it would better reflect the employer's needs to record the advertisements in this manner, rather than to enter a new job advertisement record for each occupation specified in the advertisements, and merely duplicate the rest of the information for all the other instances. Furthermore, for these job advertisement instances that listed more than one type of occupation's vacancies within the confines of one job advertisement (e.g. enrolled nurses, registered nurses and carers), only those relating to Registered Nurses were recorded in EDO (thus 'enrolled nurses' and 'carers' would not be recorded in this example as they are not registered nurses according to ANZSCO).

There were no advertisements for High Acuity (one of the ANZSCO Critical Care and Emergency categories) or Paediatric & PICU (one of the ANZSCO Child and Family Health categories) vacancies in either Perth or WA location categories over the whole trial month.

### **3. Perform manual intervention and judgement on occupational titles**

It is acknowledged that manual intervention and judgement need to be made when classifying individual advertisements into or under each occupational title or type. The persons who write and place job advertisements do not necessarily stick to the same occupation titles as the ANZSCO list, or the exact occupation specialisation might be embedded in the body text of the job advertisement. For example, the author of the job advertisement could put only the title 'Registered Nurses' in the heading of the job advertisement without also providing the sub-classification of the type of registered nurse that they are looking to employ. The example provided in Figure 66 shows that the heading only reads 'Registered Nurse'; it is the next line that indicates the type of nurse that is required, in this case a Practice Nurse. Human intervention is then further required to be able to know that the advertisement occupation is mapped to the ANZSCO title called 'Registered Nurse (Medical Practice)'. It is up to the researcher at this stage to match descriptions of job advertisements to occupation types which they think are most appropriate.

The naming convention used to assign each job advertisement's unique identification number has been detailed in section 6.2.2A.



## Registered Nurse

### Are you looking for a career in Practice Nursing?

The Independent Practitioner Network is currently recruiting for Registered Nurses to work in clinics throughout WA.

#### Full time, Part time and casual roles available

The Practice Nurse's roles include but are not limited to;

- Oversee and manage the day to day operations
- Demonstrate a high level of communication with our patients
- Provide quality service and work as part of a small team
- Identify & contribute to the development of patient educational needs
- Implement and practice Infection Control Policies
- Good understanding of OH&S
- General administration & software experience
- Multi tasking, problem solving & time management
- Confidentiality and management of patient records
- Mature, flexible and caring disposition

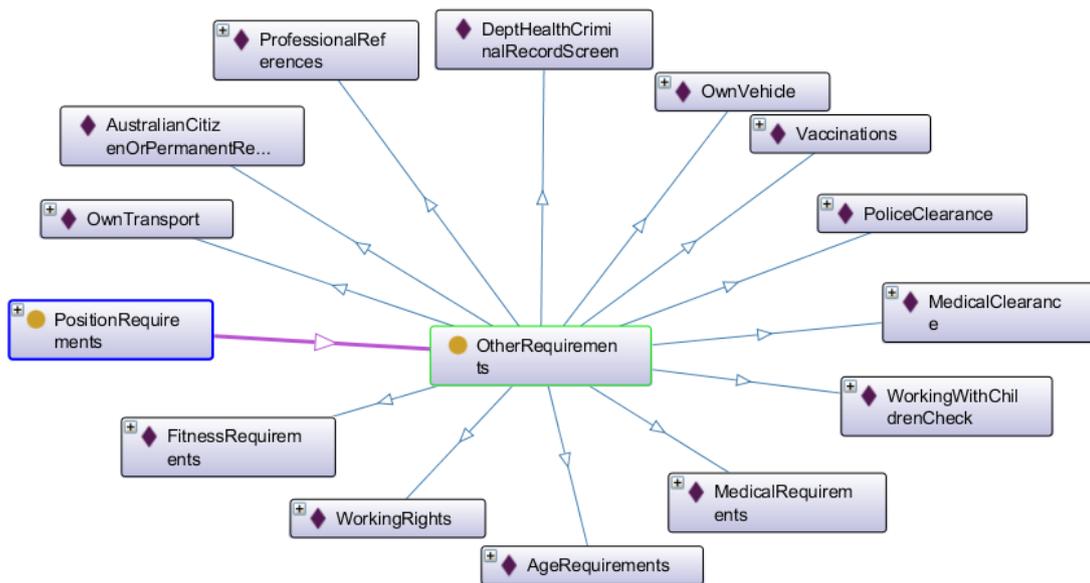
**Figure 66** Example where full title of occupation is not necessarily listed in the heading of the job advertisement.

### D. Position Requirements

A total of 820 instances were recorded during the one month trial period. Included in this number is the standard ISO 639 language classification that has been incorporated into EDO, and as such was not derived from the 416 job advertisements but through resource recycling instead (as discussed in Chapter Five). Table 8 lists the number of instances recorded for each of the *PositionRequirements* subclasses and Figure 67 shows the instances for the subclass *OtherRequirements*.

Position requirement subclass		Number of instances
Attribute requirements		51
Competence requirements	Employee experience	81
	Employee knowledge	26
	Employee language	557
	Employee qualification, licence or registration	32
	Employee skill	60
Other requirements		13
Total position requirements		820

**Table 8** Number of instances populated for the *PositionRequirements* subclasses.



**Figure 67** OntoGraf presentation of the *OtherRequirements* class's instances.

Decisions were made on the relevancy of certain benefits, conditions and requirements that were listed in the job advertisements. It was necessary to use judgement in relation to different uses of the same type of skill requirements to minimise duplicating too many of the same type of requirement. As an example, the essential criterion 'Genuine desire to work with frail aged people' was interpreted as similar enough to the criterion 'Relate to Elderly' and this criteria was instead listed for consistency purposes. Another example includes listing the criteria 'Eligible for Registration in WA as a Registered Nurse' as 'Australian Nursing Registration'.

Furthermore, the number of years' experience that some advertisements indicated were needed from potential candidates, were not included and recorded in the EDO — simply the requirement of the type of experience requested by the employer were included. This quantitative requirement measure may be included as future work for the EDO.

#### **E. Potential Skill Shortage**

As discussed, to identify potential skill shortages, information is needed about at least two of the three determining factors of a skill shortage: salary level, geographical location and period taken to fill the position. Unfortunately two of these three factors' instances were unable to be determined during the one month trial period. As the trial period ran only for one month, it was not possible to record the start and end dates for the job advertisements, as an advertisement would need to be monitored daily to know when the advertisement is no longer being run on the job portal and as such, when this is established, an end date can be recorded for the advertisement.

For salaries, we found that only two job advertisements included the salary that is on offer in the text of the job advertisement and as such, these instances could not be recorded either (see Salary and Period Advertised sections for more information below). As a result, no instances could be recorded for *PotentialSkillShortages* at this time.

## **F. Period Advertised**

One way to establish a job advertisement's 'active' period is for the system to record the first day that the job has started being advertised, and continuously checking to see when this job has been removed from the job board. As soon as it is found to have been removed, it can be concluded that a suitable candidate has been appointed to the position. Another way could be to obtain this information directly from the employer; however, it was decided not to utilise the 'application close date' that is sometimes stipulated in the job advertisement free text, as this are more than often not included in the advertisement, and it would not mean that a suitable candidate has been found by the closing of applications. Often the employer may decide to continue running the advertisement and just amend the closing date because the ideal candidate has not yet been found.

As the trial period was only performed for one month, it was not possible to track the period that a position has been advertised, as these advertisements usually run for more than four weeks at a time. Once ongoing tracking of job advertisements is being done, the period the job has been on the market should be determinable by recording the first day the new job advertisements were detected, and the last day they were recorded in the system before being taken off the market.

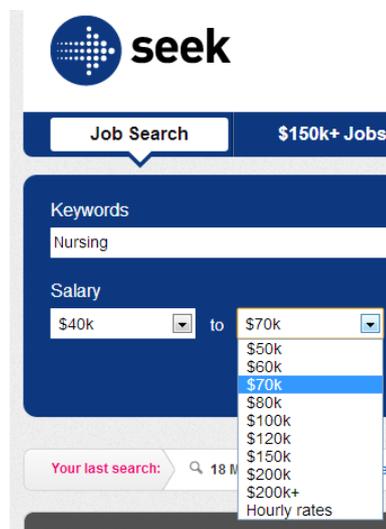
## **G. Salary Level**

It would seem that employers are less likely to advertise a position's salary than they used to do in previous years. From the 145 usable job advertisements analysed during the trial period, only three job advertisements had the salary level explicitly indicated in the body text of the job advertisement as follows:

- Job advertisement NM9= \$99,485 total value package on offer.
- Job advertisement RNP11= \$50,000- \$65,500 salary range indicated.

- Job advertisement MW8= starting package of \$75,000 specified.

It was unexpected that so few advertisements would have the salary level specified in the text of the job advertisement. It would seem that salaries are instead specified through an indirect means such as the drop-down search criteria salary ranges (Figure 68) to search for a job on the job board, as opposed to the employer putting this in the free text of the job advertisement itself (Figure 69). As a result, the EDO records populated for the trial month only had three records recorded for salary as indicated above.



**Figure 68** Drop down salary range search criteria on Seek.com.au.

The job alerts for the trial period were not set to retrieve job advertisements for specific salary ranges; this field was left blank when the job alert was setup so as to capture all salary ranges. It would seem that future trial periods would need to set up, several job advertisement alerts so as to distinguish between the different salary ranges on offer. For instance, where this trial period had one alert setup to capture all Midwifery positions being advertised in Western Australia, this would need to be changed to several alerts to cover all the options available in the drop-down salary range boxes on the website: \$0-\$30k, \$30-\$40k, \$40-\$50 etc., until the \$200k plus category. This would clearly entail much more initial setup work and ensuring that each job alert is saved under the correct salary category when the job advertisements are received via email and saved for later analysis each day.

Furthermore, once a salary level has been identified or captured for a specific job advertisement, a measure needs to be incorporated to determine whether the salary on offer is a high, average, or low salary for that specific type of occupation and similar conditions. To do this, an existing salary analysis group who provide this type of information has been suggested in the future work section to be considered in this regard.

## Nurse

As a Nurse in the Army you're a vital member of the medical team, working in both static hospitals and mobile field units on deployment around Australia and overseas.

You'll be exposed to a diverse range of opportunities, such as supporting humanitarian and disaster relief operations and you'll work with some of the most sophisticated medical technology. Most of your time will be spent with patients, but you'll also assist in general surgery; ear, nose and throat surgery and a large amount of orthopaedic surgery. With a career in the Army, you'll enjoy the sort of variety, challenges and travel that you rarely get as a civilian nurse.

You'll be paid while you undertake your training, and receive free healthcare and subsidised accommodation. Upon completion of training, you'll enjoy a **starting salary package of approximately \$75,000p.a**

You may be posted to any base within Australia or possibly overseas, although overseas postings are limited.

You must be an Australian citizen or hold Permanent Residency status. Additional age, medical and fitness guidelines may also apply.

You'll need to have successfully completed a Bachelor of Nursing degree and your application will be considered during your compulsory two-year post-graduate experience in general nursing duties.

<http://www.defencejobs.gov.au/army/jobs/NursingOfficer/?availability=fulltime>

**Entry Requirements**

You must be an Australian citizen or hold Permanent Residency status. Additional age, medical and fitness guidelines may also apply.

**Education Requirements**

You'll need to have successfully completed a Bachelor of Nursing degree and your application will be considered during your compulsory two-year post-graduate experience in general nursing duties.

**Figure 69 Salary is explicitly indicated within the body of the job advertisement.**

## H. Geographical Location

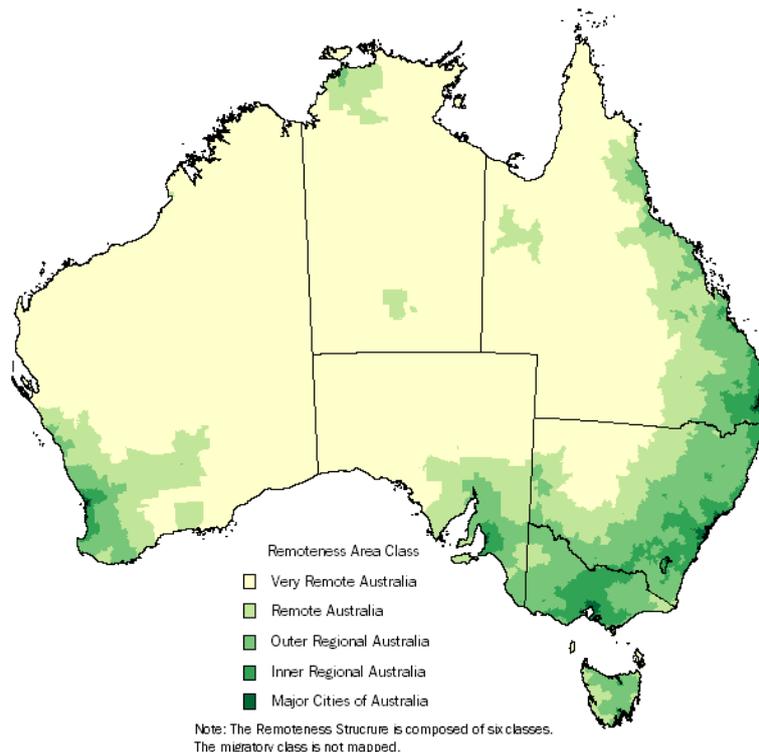
The number of instances recorded for each EDO location type in Western Australia has been specified in Table 9; Figure 70 shows the actual geographical clusters of these subclasses in Australia and Figure 71 the instances of one of the location type subclasses. These instances have been manually recycled from the ASGC tables provided by the ABS. The ABS classification data was in a text-based Excel spreadsheet which was downloaded from the ABS website; instances were entered one-by-one from the Excel spreadsheet into the EDD in Protégé. A

snapshot of the same location type example provided in Figure 71 has been provided in Figure 72 to indicate the raw data in the ASGC spreadsheet.

Due to time and resource constraints, only Western Australian locations have been recorded in EDO to date; no other state in Australia has been surveyed as this was not necessary for the trial period which focused only on Western Australian job advertisements.

Location type subclass	Number of instances
Inner regional Australia	43
Major city of Australia	11
Outer regional Australia	48
Remote Australia	37
VeryRemoteAustralia	37
Total location type instances	176

**Table 9** Number of instances populated for the *LocationType* subclasses.



**Figure 70** Map of Australia illustrating the 2006 Remoteness Structure (Australian Bureau of Statistics, 2011).

Once all the Western Australian locations were recorded in the EDO, the relationship of each job advertisement with its respective geographical location had to be recorded. Some job advertisements did not specify the geographical location of the position; in these cases the organisation was searched online to find where they were physically located and this location was recorded in the EDD instead. As an example, here is the location information stated by some of the job advertisements in the category Registered Nurse — Medical, Medical Practice, Nurse Practitioner and Surgical (RNMS):

- RNMS3 stated the vacancy is mainly Perth-based but will involve some travel to mine sites in the Pilbara and Goldfields as Fly in Fly out (FIFO).
- RNMS12 stated they have vacancies all over Australia.
- RNMS14 stated they do rural and remote placements in Australia.
- RNMS15 only mentioned the location as Peel *region* (not a location as such on the ASGC).
- RNMS16 only mentioned 'north west coast of Australia'.

- RNMS19 stated the location as ‘a mine site in WA’.

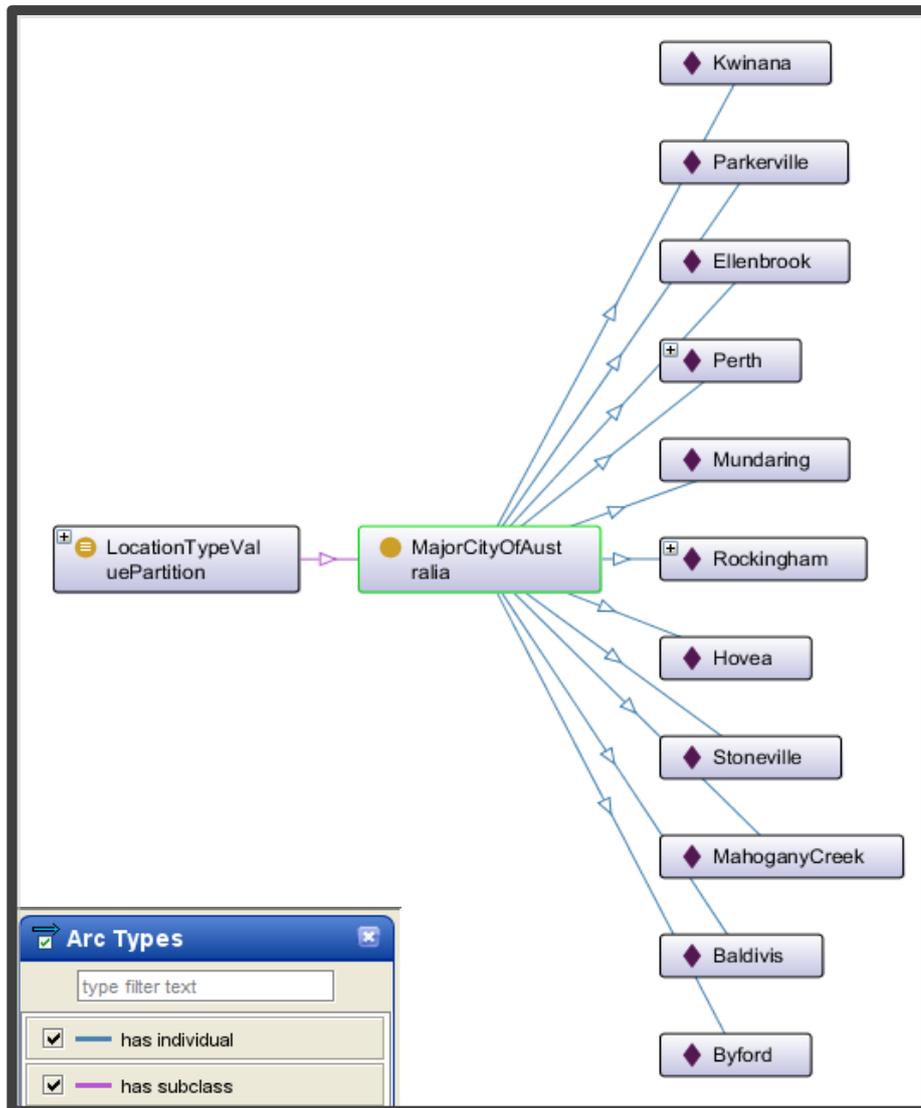


Figure 71 OntoGraf presentation of the subclass *MajorCityOfAustralia*'s instances.

RA_2006_CODE	RA_2006_NAME	UCL_2006	UCL_2006_NAME	RATIO	PERCENT
50	Major Cities of Australia	500880	Baldivis	1	100
50	Major Cities of Australia	503400	Byford	1	100
50	Major Cities of Australia	508500	Ellenbrook	1	100
50	Major Cities of Australia	511500	Hovea (L)	1	100
50	Major Cities of Australia	514400	Kwinana	1	100
50	Major Cities of Australia	516720	Mahogany Creek	1	100
50	Major Cities of Australia	519800	Mundaring	1	100
50	Major Cities of Australia	522280	Parkerville	1	100
50	Major Cities of Australia	523000	Perth	1	100
50	Major Cities of Australia	524600	Rockingham	1	100
50	Major Cities of Australia	525500	Stoneville	1	100

Figure 72 Snapshot for the class *MajorCitiesOfAustralia* data as it appears in the ASGC spreadsheet prior to being recycled to the EDO.

## I. Position Responsibility

As discussed previously, the position responsibilities class has not yet been developed and has been included as future work of the EDO. As such, it currently holds no instances.

## J. Employee Benefit

During the trial period, a total of 56 employee benefits were recorded. Examples include *AccommodationSupport*, *Bonusses*, and *FlexibleWorkPractices* (Figure 73, Figure 74).

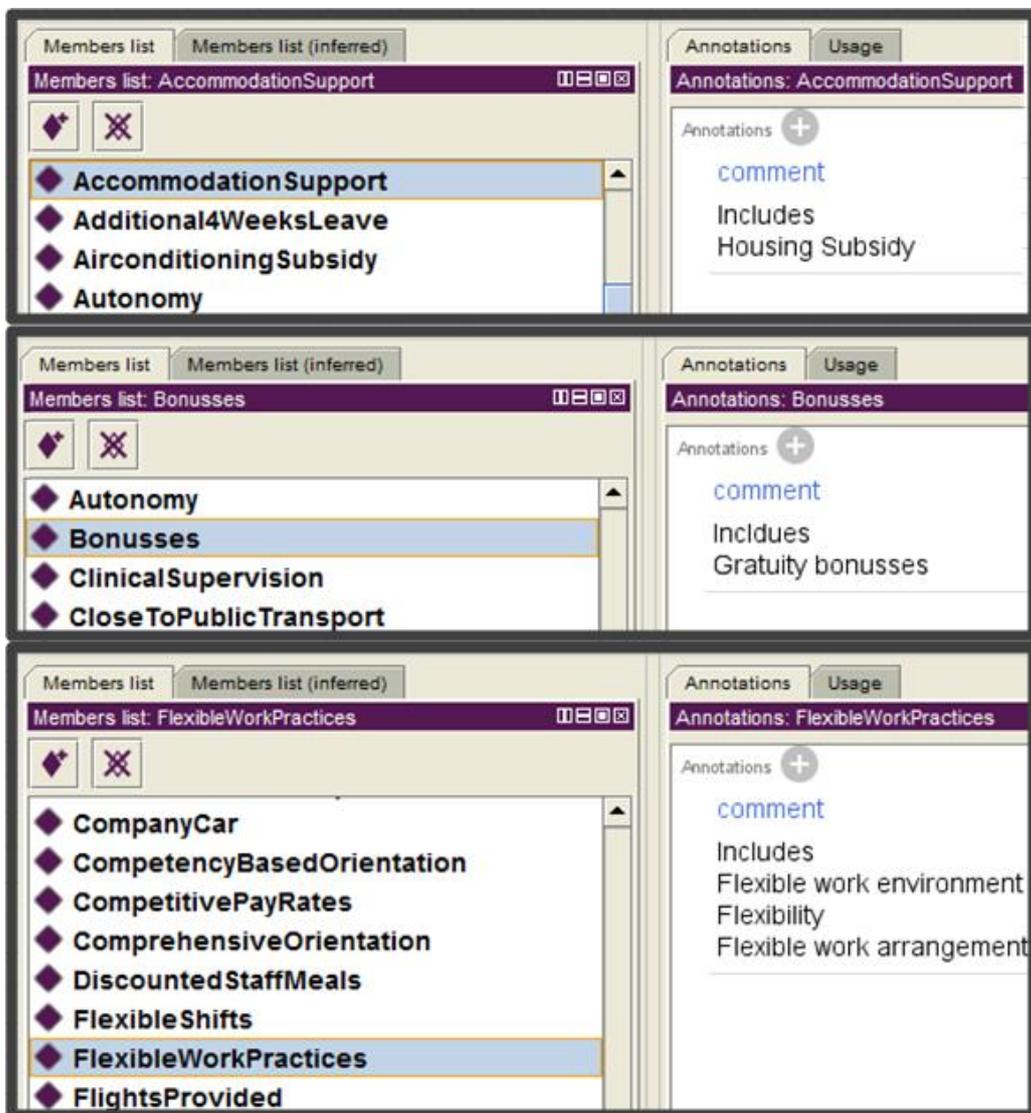


Figure 73 Some of the instances of the class *EmployeeBenefits* and their annotations.



the best health opportunities

## Registered Nurse & Midwifery - Rural and Remote Contracts - WA

- Agency rates of pay!
- Free or subsidised Accommodation + Travel Reimbursed
- Experience Real Australia

Multiple short term and on-going contracts are now available throughout regional Western Australia within Hospitals, Multi Purpose Services and Community Health Centres.  
Applicants must be experienced in one or more of the following:

- Emergency
- Acute Rural and Remote
- Mental Health inpatient
- Child and Family Health
- Aged Care
- Midwifery

These contracts are ideal for people who are keen to experience living and working in the real Australia! AHPRA registration and valid work permit is required.  
Benefits include:

- Rural and remote allowance
- Travel reimbursement
- Excellent pay rates
- Unique work experience
- Subsidized or free accommodation

If you are an experienced Nurse looking for a change and interested in gaining valuable skills in a different setting then we want to hear from you!

To confidentially apply for this role or to find out more information, please contact Alex Hall on +61 02 9994 2100 or email to [nc@nchealth.com.au](mailto:nc@nchealth.com.au), quoting reference number **AH22870/SK**

To view other exciting opportunities available with Naiman Clarke, please visit [www.nchealth.com.au](http://www.nchealth.com.au) or call one of our experienced consultants.

**Figure 74** A nursing advertisement where the employee benefit instances have been highlighted.

## K. Employment Condition

In total, 21 employment conditions were recorded from the trial month dataset for both subclasses of *EmploymentConditions*: *WorkHours* and *PositionBasis*. Figure 75 provides an OntoGraf presentation for each of the two employment condition subclasses' instances. The class *WorkHours* had a total of 14 instances recorded, and the subclass *PositionBasis* had seven records captured.

Data entry decisions for *EmploymentConditions* included scenarios where employers indicated 'permanent part-time' as the employment condition, and it was listed as *PartTime* as well as *Permanent* — two separate entries for that job advertisement record. Examples of annotations captured for the *WorkHours* instance — *ReliefBasis* — are presented in Figure 76.

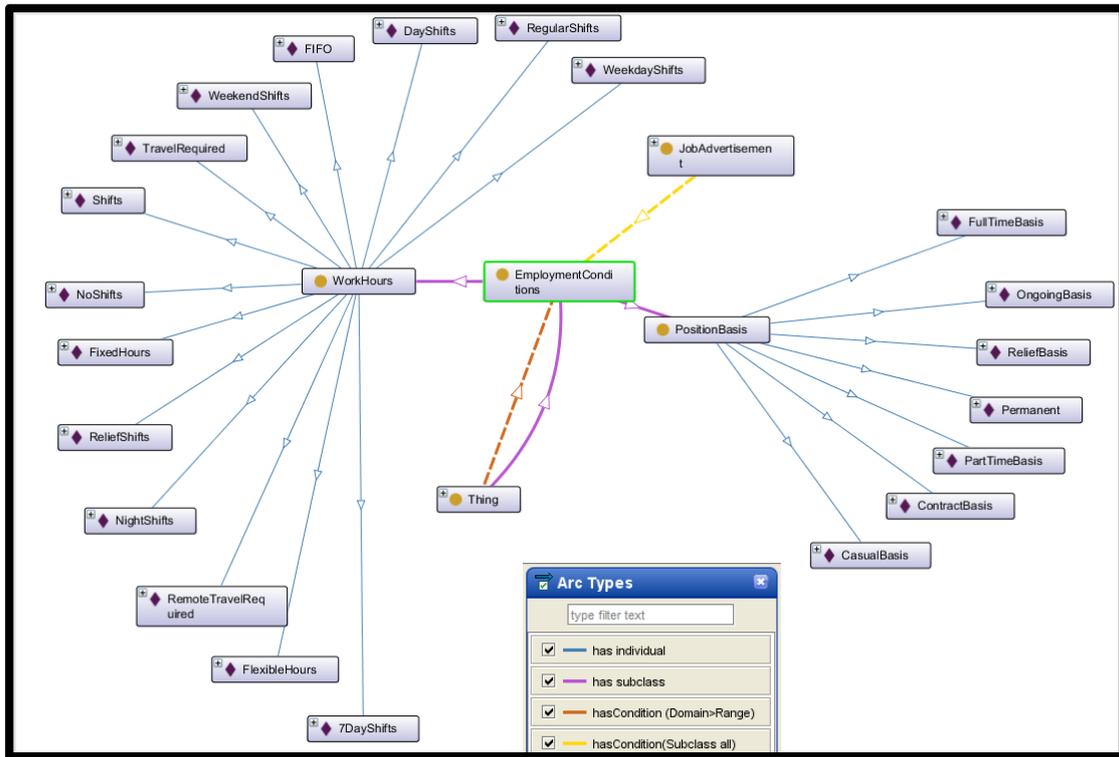


Figure 75 OntoGraf presentation of instances of the *EmploymentCondition* class.

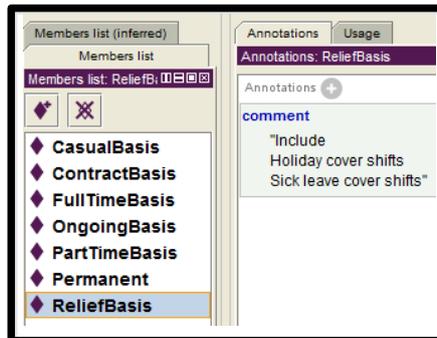


Figure 76 Annotations captured for the instance *ReliefBasis*.

## L. Organization

There were 69 different organisations who placed job advertisements for the trial period in January 2011. Where there were both a recruitment agency listed on the advertisement and the organisation where the employment would take place, the latter was listed as the organisation

related to the vacancy under question. E.g. on *RNMS3*, Kannon Recruitment was the advertiser of the vacancy, however the place of work was listed as Capstone Health, an occupational and travel medicine company operating in WA and beyond (Figure 77).

There were also instances where some job advertisements (e.g. *RNMS18*) didn't specify the employer of the vacancy, however it stated that the vacancy is for a neurological practice in Subiaco. In these instances, no *Organization* relationship was able to be recorded for such job advertisements.



  
**KANNON  
RECRUITMENT**

## Registered Nurse – FIFO Positions

**Capstone Health** is a dynamic Perth based occupational and travel medicine company based around the provision of high quality and attentive service to individuals and companies operating in Western Australia and beyond.

We are looking for enthusiastic and motivated registered nurses, exercise physiologists and other health professionals to assist with a growing demand for occupational health screening and surveillance; as well as a role in injury management and return to work co-ordination, health surveillance, travel medicine and vaccinations.

The role will involve work in the Perth based clinic; metropolitan sites as well as involvement in rostered short term FIFO clinic work in mine sites including through the Pilbara and Goldfields.

Competitive remuneration packages are available incorporating generous loading for remote work.

- Relevant and current professional qualifications with relevant professional accreditation (RN, OT, Exercise Physiologist or Physiotherapy) and 3 years industry experience in a relevant field.
- Strong interpersonal and communication skills in written and verbal contexts
- Sound computer skills including experience with Outlook, Word, Excel and Power Point software applications.
- The ability to travel regionally Monday – Friday with sufficient notice
- Excellent planning, organisational and time management skills.
- A high level of initiative and the ability to work unsupervised and within a team environment.

**Desirable**

- Previously experience in occupational health and safety fields will be highly advantageous.
- Mineworkers Health Surveillance, Drug Screen Collection and WorkCover Audiometric/Approved NIHL certification
- Current Immunisation Certificate
- Previous track record in
  - Design, Implementation, Promotion and Management of health programs including injury management and return to work co-ordination, health promotion, ergonomics and functional capacity assessments
  - Management of small teams of colleagues and staff to meet project goals and performance indicators

Flexible work arrangements including part time and casual work are available.  
Only shortlisted candidates will be contacted.  
Email: Please click the 'Apply Now' button below.

**Figure 77** A job advertisement indicating both an agent as well as an Organization where the employment will take place.

## 6.2.2 Employer Demand Ontology: Tier Two Instance Population

### A. Occupation

The EDO *OccupationalType* has one instance for the subclass *Midwives*, one instance for *NurseManagers*, two instances for *NurseEducatorsAndResearchers* and twelve instances recorded for *RegisteredNurses*. Figure 78 shows the Midwifery and Nursing Professionals section of the ANZSCO spreadsheet, which was used to populate the EDO occupational type instances shown in Figure 79. As for the ASGC spreadsheet that was used to populate the instances in the Location Type class, the occupation title instances were also manually entered from the spreadsheet into the Protégé without using a script. At this stage, the mapping of the ANZSCO and SEEK categories are only stored in a spreadsheet, and published through this thesis. Future work will include enhancing the annotation of the current *OccupationalType* class in Protégé to formalise the mapping in that EDO too (also discussed in future work in Chapter 9).

Australian Bureau of Statistics / Statistics New Zealand			
ABS Cat. no. 1220.0 ANZSCO -- Australian and New Zealand Standard Classification of Occupations, First Edition Structure			
Table 5: Occupations			
Major Group	Sub-Major Group	Minor Group	Unit Group
			Occupation
	254	Midwifery and Nursing Professionals	
		2541	Midwives
			254111 Midwife
		2542	Nurse Educators and Researchers
			254211 Nurse Educator
			254212 Nurse Researcher
		2543	Nurse Managers
			254311 Nurse Manager
		2544	Registered Nurses
			254411 Nurse Practitioner
			254412 Registered Nurse (Aged Care)
			254413 Registered Nurse (Child and Family Health)
			254414 Registered Nurse (Community Health)
			254415 Registered Nurse (Critical Care and Emergency)
			254416 Registered Nurse (Developmental Disability)
			254417 Registered Nurse (Disability and Rehabilitation)
			254418 Registered Nurse (Medical)
			254421 Registered Nurse (Medical Practice)
			254422 Registered Nurse (Mental Health)
			254423 Registered Nurse (Perioperative)
			254424 Registered Nurse (Surgical)
			254499 Registered Nurses nec

Figure 78 Snapshot of the ANZSCO spreadsheet section for Midwifery and Nursing Professionals.

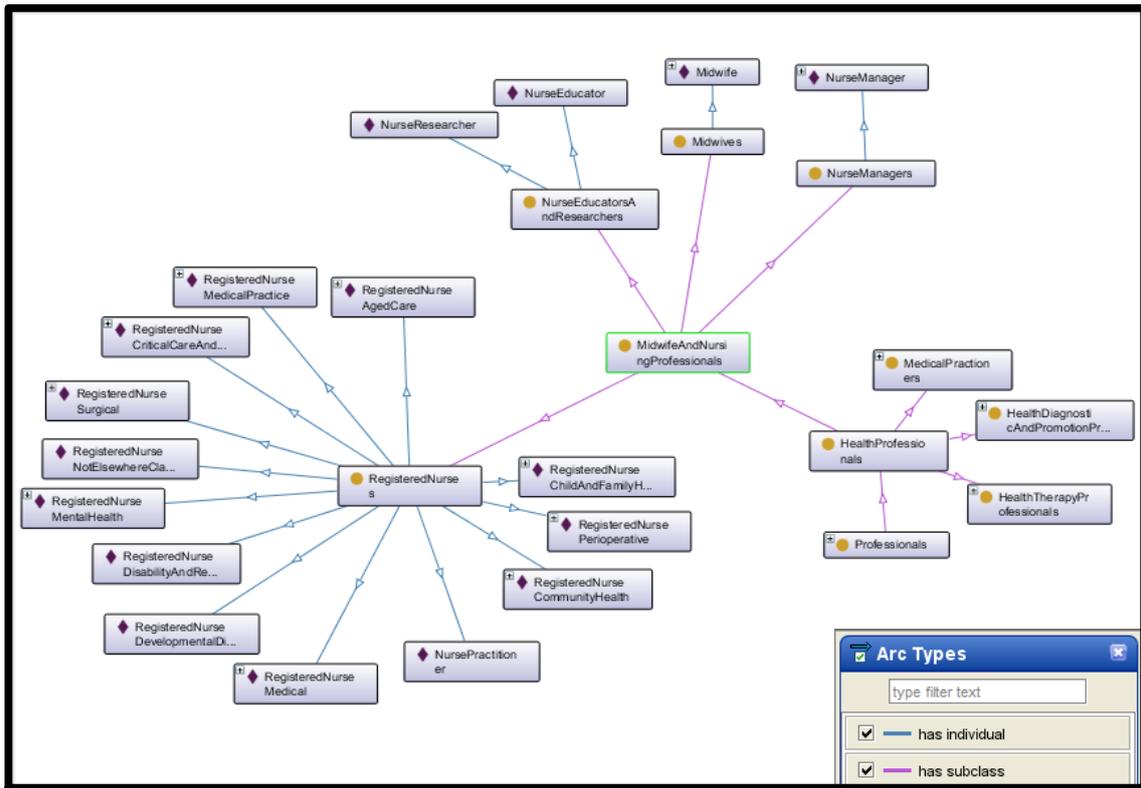


Figure 79 OntoGraf presentation of the EDO *MidwifeAndNursingProfessionals* instances.

In order to capture accurate occupational type data in the EDO, it was necessary to develop a mapping system from the ANZSCO and Seek occupation categories, because the job board Seek containing the job advertisements analysed for the trial period (Table 10). Figure 80 shows the Seek nursing categories.

seek	ANZSCO title and EDO abbreviation
Nursing- Aged Care	Registered Nurse - Aged Care (RNAC)
Nursing- Midwifery, Neo-Natal, SCN & NICU	Midwife (MW)
Nursing- Paediatric & PICU	Registered Nurse- Child and Family Health (RNCF)
Nursing- Community, Maternal & Child Health	Registered Nurse- Community Health (RNCH)
Nursing- A&E, Critical Care & ICU	Registered Nurse- Critical Care and Emergency (RNCE)
Nursing- High Acuity	Registered Nurse- Critical Care and Emergency (RNCE)
Nursing- Psych, Forensic & Correctional Health	Registered Nurse- Developmental Disability, Disability and and Rehabilitation and Mental Health combined (RNDM)
Nursing General Medical & Surgical	Registered Nurse- Medical, Medical Practice, Nurse Practitioner and Surgical combined (RNMS)
Nursing- Theatre & Recovery	Registered Nurse- Perioperative (RNP)
Nursing- Management	Nurse Manager (NM)
Nursing- Educators and Facilitators	Nurse Educators and Researchers (NER)

Table 10 Midwifery and Nursing Professionals mapping between Seek and ANZSCO.



**Figure 80** Seek.com.au nursing categories.

ANZSCO has 17 Midwifery and Nursing Professionals occupations, while Seek only has 11. Because Seek combines different occupational types into combinations that are different to the list as presented in the ANZSCO spreadsheet, it was necessary to develop a mapping system between the ANZSCO occupations and SEEK's occupations, ultimately bringing it down to 10 classes for the job advertisement naming convention only. The relationship with between each job advertisement and its subsequent ANZSCO occupational title was still captured accurately according to the correct ANZSCO and EDO occupational type classes (Figure 81).

It is acknowledged that the ideal scenario would be to keep the ANZSCO categories exactly as they have been presented in the ANZSCO spreadsheet for the naming convention to assign each job advertisement a unique identification number, however practical limitations as described above, with the SEEK classification being different, necessitated this change.



**Figure 81** The job advertisement with the combined occupational naming code (unique identification number) of RNMS18, still capturing the ANZSCO occupational title correctly — *RegisteredNurseMedicalPractice* in this case.

### **6.3 Conclusion**

This chapter has focused on the steps taken to solve research issue two — occupation-specific data querying. The relevant instances for each EDO class and subclass were populated. The chapter provided a detailed description of each subclass's members, indicating which instances were derived from other resources that were recycled, or where instances were obtained through analysing the one month's trial period job advertisements from the Seek.com.au job board. It is intended to include more sources of employer demand data (e.g. from other job boards, from social media website, etc.) to incorporate into EDO as future work of this research.

The next chapter will focus on the steps taken to address research issue three — expensive and time-intensive approaches. The solution developed for this process consists of a semi-automated employer demand intelligence framework.

# Chapter 7 - Semi-Automatic Employer Demand Intelligence Tool

## 7.1 Introduction

The previous chapter provided the detailed solution for research issue two- occupation specific data and its querying. This chapter details the steps that were taken to address research issue three, expensive and time intensive approaches, by implementing the proposed solution of a semi-automatic Employer Demand Intelligence Tool (EDIT).

The focus first falls on the main processes of the Employer Demand Intelligence Framework (EDIF), and then turns to a detailed look at the steps that EDIT performs in order to produce employer demand intelligence semi-automatically. As part of the results and discussion section, a detailed example is provided to show the steps graphically, as well as the SPARQL query outcomes, pros and cons of using the GATE application, some of the metrics of EDIT and the limitations of EDIT as it currently stands.

## 7.2 Overview of the Employer Demand Intelligence Tool

As proposed in Chapter Four, this solution aims to overcome the issue whereby current employer demand related intelligence approaches are extremely resource intensive, causing huge time delays in gathering relevant and current data related to the domain. It has become imperative that a solution be developed to address this issue. As such, the solution of a semi-automatic tool called EDIT has been developed as a prototype in this research with the following requirements:

- Purpose

Develop a prototype software tool that is able to semi-automatically populate and enhance EDO based on the content of online-published job advertisements.

- Functionality

- Extract job advertisement content from html text by filtering unwanted web content.
- Recognise the existence of concepts found in the online job advertisements that match with those concepts stored in EDO.
- Provide the ability to link the recognised text to the corresponding concept in EDO.
- Develop the ability to populate job advertisement instances for EDO and assert relevant object and datatype property values.

- Input

Job advertisements collected in their original HTML format for Western Australian locations, and for the Nursing and Midwifery occupations only.

- Output

Populated and evolved/enhanced EDO and dataset

This section focuses on the detail of EDIT and its prototype implementation. It is very important to have a clear picture of the job market at any time — not only for governments for timely decision making but also for scenarios like individuals doing career planning. Furthermore, due to the volatile and dynamic nature of the job market it is necessary that the information about employer demand is up-to-date. This information can be collected from job advertisements and published on the web. However, the job advertisements usually contain unstructured information in textual form, with frequent usage of specific jargon related to the nature and field of the advertised vacancy. Additionally, the information collection involves hundreds of advertisements, which is difficult to process manually, if possible at all. The advertisements, albeit highly unstructured, are already in digital form and the development of an automated

process capable of extracting meaningful structured data and statistics for these text advertisements (which is the objective achieved by EDIT) can be extremely useful in this scenario.

EDIT is a combination of state-of-the-art software technologies that have been developed around the scientific framework of ontology. These technologies and the scientific framework have been combined into a one stop system, where a user enters a query in EDIT through a SPARQL endpoint on the internet, to retrieve desired results.

It has become somewhat of a cliché to state that a large amount of knowledge about any domain is stored as unstructured electronic text online. The cliché is nevertheless a true reflection of the reality in industry (government, companies, universities etc.) where there is an ever growing body of information on the web (Massey & Wong, 2011). EDIT has been designed to gather data from unstructured text on the web (social media, company, government and other websites) by utilising ontology-assisted information extraction techniques. Stage one of the EDIF prototype development, has been achieved by mining one of these data drenched resources on the web, namely a job board called SEEK<sup>37</sup>. As pointed out in Chapter Four, due to resource constraints, the prototype has only been developed for one stream of concepts in order to prove that the conceptual framework is viable and usable by its end users.

The main processes of EDIT for its development in stage one (this PhD research), is presented in Figure 82. The process starts where EDIT collects and stores the information from the job board SEEK for occupations under the Registered Nurse and Midwifery classification into EDIT's corpus. Unnecessary web page content such as links to other pages and graphics is removed and only the advertisement content is stored in the corpus.

---

<sup>37</sup> <http://www.seek.com.au>

The next process entails the extraction of information contained in the corpus, where the useful information which matches with the concepts in EDO, is extracted from the advertisement text. As depicted in Figure 82, in addition to the corpus creation process, the process also uses a gazetteer, which is the output (called the EDO domain knowledge) of another process. This process extracts the concepts (classes and instances) in EDO and stores them in the gazetteer. The term gazetteer, which will be further explained in the next section, is a standard term in the area of text processing to store entity names or nouns.

Finally, the information extracted from each advertisement is populated into the ontology or used to add extra components to EDO — no concepts are deleted or removed from EDO. The data relating to the generic concepts of EDO are populated into EDO as instances and property values. This extended ontology forms the knowledge base which end users can query via the SPARQL endpoint, in order to obtain answers for their employer demand intelligence interrogations.

This section provides an overview of EDIT, and how the challenges of extracting useful and relevant information from unstructured text of advertisements has been overcome by EDIT. The next section details each step of that EDIT performs in order to process the employer demand intelligence that is gathered from relevant sources on the web.

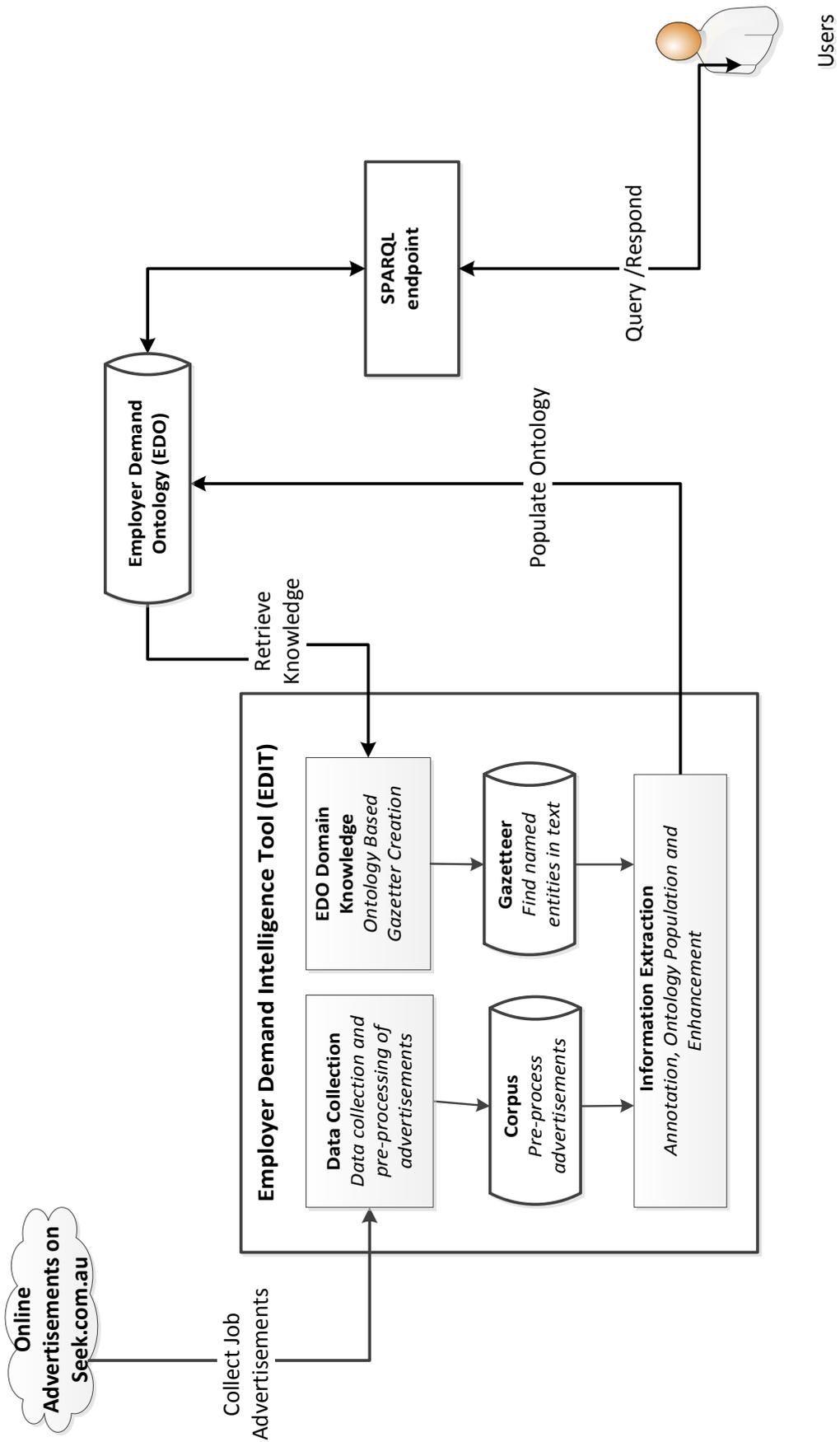


Figure 82 Main processes of the Employer Demand Intelligence Tool.

## **7.3 Detailed Processes of the Employer Demand Intelligence**

### **Tool**

This section discusses the processes that are performed within EDIT to attain employer demand intelligence. Figure 83 shows the sequence of these processes and the flow of information between them. Each process and its purpose are explained below. This description follows the actual sequence of these processes in EDIT. The details of the respective corresponding software component that was utilised in EDIT to perform each process, is also specified after the purpose of each process.

#### **7.3.1 Creation of Corpus**

This process entails the collection of data from the web and preparing it for processing by the subsequent stages. This entails two tasks:

1. Archiving job advertisements

Job advertisements that are published on the online job portal SEEK<sup>38</sup> were downloaded and stored on a daily basis for the duration of one month. The process was repeated daily and the advertisements were organised into corpora of advertisements published online in a calendar month.

Using the standard terminology in text processing, the term ‘document’ is used to refer to each job advertisement. The documents are saved in .html archive files with all the webpage graphics and links. This raw data is kept for future referencing purposes.

---

<sup>38</sup> <http://www.seek.com.au>

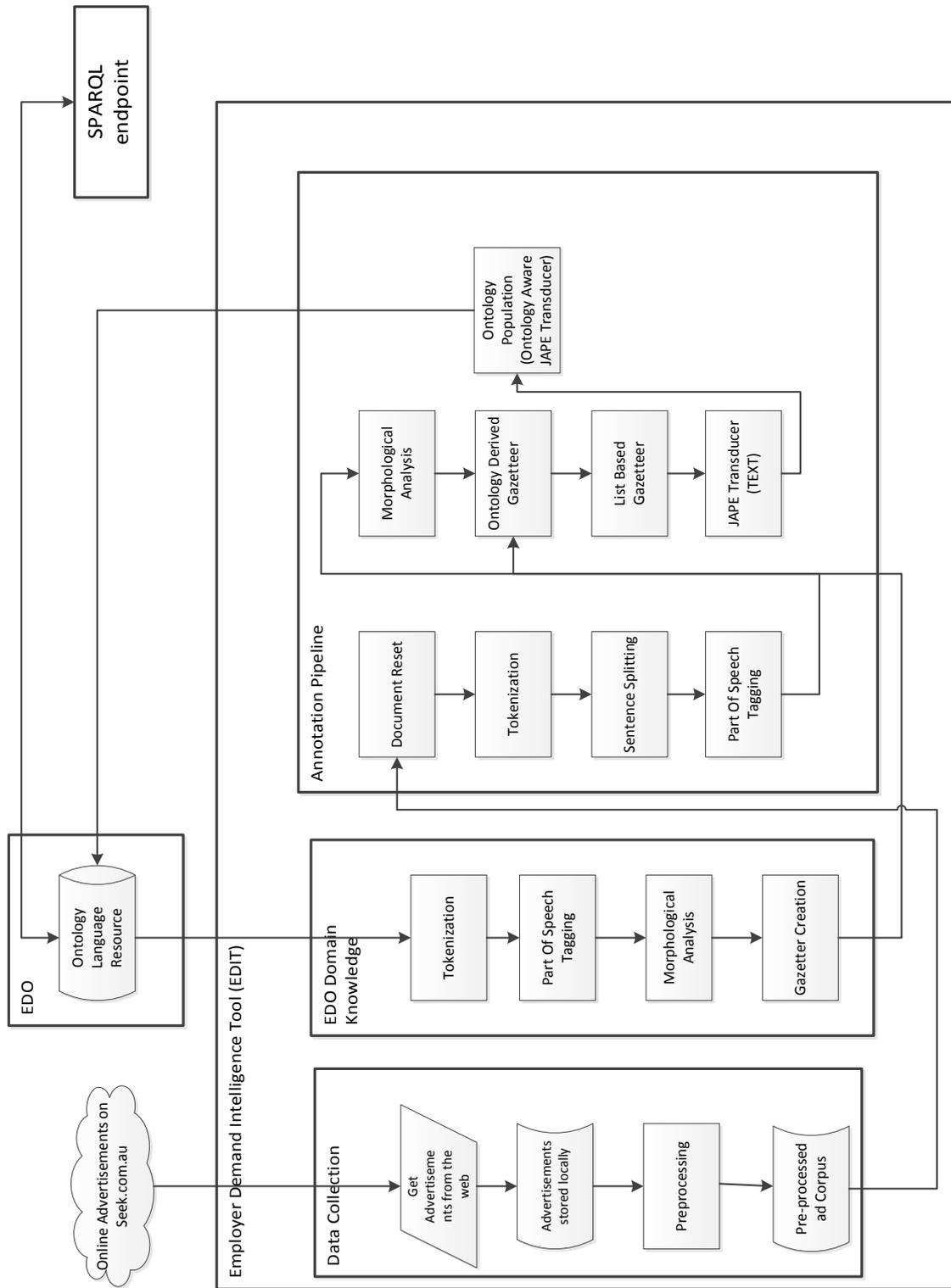


Figure 83 Employer Demand Intelligence Tool flowchart.

## 2. Pre-processing

The documents collected in the first step contained images and links to other pages which were not directly related to the advertisement. A program was written in Java using the Document Object Model (DOM) parsing to remove these images and links while retaining the job advertisements' text. The output from this stage is shown in Figure 84 as an example.

**21 Mar 2013**

- Location:  
Bunbury & South West
- Salary:  
\$20 - \$24.99 per hour
- Work type:  
Part Time
- Classification:  
Healthcare & Medical Nursing - Aged Care

Aged Carers

- \$23+ Per Hour plus Allowances
- Facilities in Bunbury, Eaton and Harvey
- Regular shifts available
- Start ASAP!!

**The Company:**

Flexi Staff Pty Ltd is a leading Western Australian owned and managed company specialising in all aspects of the recruitment and labour hire.

We are a major supplier in providing staff to all Aged Care facilities in Bunbury and the South West region and are currently seeking experienced Aged Care workers for regular, flexible and casual work.

We have shifts available in the Morning, Afternoons and Nightshifts and these are available Monday - Friday and Weekends.

**Essential Criteria:**

- Cert III Aged Care (Essential)
- Current Police Clearance or willing to obtain
- Current WA C Class Drivers Licence
- Reliable Transport
- Available to start ASAP

**Figure 84** A pre-processed job advertisement to remove unnecessary HTML elements as links and images.

### 7.3.2 Loading the Corpus and EDO into EDIT

EDIF has a component coined the Employer Demand Intelligence Tool (EDIT), which has been developed for this research by using a special purpose, open source, research-oriented text processing engine called GATE (General Architecture for Text Engineering) (Cunningham et al., 2013). GATE provides a state-of-the-art development environment which is called the GATE Developer, to develop natural language processing applications. The basic components are needed to develop a working text processing application, which are provided by GATE. A piece of software developed in GATE is called a GATE application, which consists of several processing resources which follow each other in sequence. A processing resource (PR) is a component program that performs one process on the provided input document, called a language resource (LR). The input of a processing resource is either a language resource or the output from an earlier processing resource. A sequence of processing resources is called a 'pipeline' or 'processing pipeline'. The pipeline needs to be designed for each application according to the nature of the input and the specific objective of an application.

A processing pipeline operates on one or more 'language resources' to produce the desired output. An individual document, corpus of documents or ontology which is processed in a text processing system is called a 'language resource'.

EDIT uses two language resources. The first language resource is the corpus — containing all the online job advertisements collected in the previous step — and the second is EDO. Figure 85 shows the EDO loaded into EDIT as a language resource. EDO plays a dual role in EDIT: firstly it is used to programmatically generate a gazetteer (described in the next step), and at the end of the process it is once again used by EDIT to store the information that is automatically extracted from the documents (job advertisements) by EDIT's pipeline.



Figure 85 The EDO loaded as a language resource into the EDIT (GATE application).

### 7.3.3 Creating an EDIT Pipeline

As discussed, the EDIT text processing application has been designed and implemented in GATE. It uses the knowledge contained in EDO to extract information from the job advertisements in the corpus and populates this extracted information as new instances and properties in EDO. This information can be accessed using standard SPARQL queries.

EDIT consists of the following processing resources shown in Figure 86. The sequence in which these processing resources work and the flow of information between them was also illustrated in Figure 83.

#### 1. Document Reset

A document reset processing resource is used in the beginning of an application so that a document is reset before EDIT runs on documents which have already been processed. A document reset processing resource resets the document to its original state by removing all the annotation sets except the one containing the document format analysis (as original mark-ups) (Cunningham, et al., 2013). EDIT uses the ANNIE Document Reset Processing Resource, which is a part of the GATE component called ‘A Nearly New Information Extraction Engine’ (ANNIE).

Selected Processing resources		
!	Name	Type
	 Document Reset	Document Reset PR
	 ANNIE English Tokeniser	ANNIE English Tokeniser
	 ANNIE Sentence Splitter	ANNIE Sentence Splitter
	 ANNIE POS Tagger	ANNIE POS Tagger
	 GATE Morphological analyser	GATE Morphological analyser
	 Flexible Gazetteer	Flexible Gazetteer
	 ANNIE Gazetteer_seek	ANNIE Gazetteer
	 JAPE Transducer_text	JAPE Transducer
	 JAPE Transducer_ontology	JAPE Transducer

**Figure 86 Processing resources in the EDIT pipeline (GATE application).**

## 2. Tokenizer

Meaningful elements within text such as words, phrases or symbols are called tokens. The task of identifying the tokens and breaking up a stream of text into tokens is called tokenization (Wilcock, 2009). A tokenizer can split text into tokens by searching for white space (spaces, tabs and newlines) in the text. However, the presence of punctuations, numbers and symbols etc. in the text needs to be accounted for and complex string processing with regular expressions is needed for meaningful tokenization.

GATE provides a tokenizer processing resource called the ANNIE English Tokenizer Processing Resource, which was modified and used in EDIT. The ANNIE tokenizer comprises a normal tokenizer and a JAPE transducer (refer to point seven below). The JAPE transducer utilises regular expressions and grammar rules to further process basic tokens. The transducer also amends the generic output of the tokenizer to fit the conditions of the English part-of-speech tagger. As an example, the transducer joins together into one token structures like “ ’30s”, “ ’Cause”, “ ’em”, “ ’N”, “ ’S”, “ ’s”, “ ’T”, “ ’d”, “ ’ll”, “ ’m”, “ ’re”, “ ’til”, “ ve” “. The JAPE transducer also converts negative structures like “don’t” from three tokens (“don”, “ ’ ” and “t”) into two tokens (“do” and “n’t””) (Cunningham et al., 2013).

## 3. Sentence Boundary Detection

Sentence boundary detection is performed during this phase. The detection of sentence boundaries entail finding the point in the text where the sentence starts and the point where it ends. The simplest approach for written language is to split the text at full stops (“.”). However, this is too simple, as there are other end-of-sentence markers besides fullstops, such as question marks (“?”) and exclamation marks (“!”). Moreover, other uses of full stops that do not mark the end of a sentence, such as in abbreviations (“Mr.”, “etc.”), also exist. Another scenario is where a full stop has more than one function, for example when “etc.” is the last word of a sentence — a single full stop can simultaneously mark both the end of the abbreviation and the end of

the sentence. A practical reason for exercising sentence boundary detection is that syntactic parsing is normally done on individual sentences rather than complete texts. The time that automatic syntactic parsing takes as the text length grows, seems to increase considerably in a computational context. Therefore, it is important to split the text into shorter units at the onset (Wilcock, 2009).

The ANNIE Sentence Splitter PR that EDIT employs, is a flow-through of finite-state transducers which segment the text into sentences. A gazetteer list of abbreviations helps differentiate sentence-marking full stops from other kinds. Sentences are annotated with the type ‘Sentence’, and sentence breaks (such as a full stop) are assigned ‘Split’ annotations (Cunningham et al., 2013).

#### 1. Part of Speech (POS) Tagging

The POS Tagger assigns each token with a part-of-speech label. The main parts of speech for the English language are below (Wilcock, 2009):

- Noun (N);
- Verb (V);
- Adjective (A or Adj);
- Adverb (Adv);
- Preposition (P);
- Article or Determiner (Det);
- Conjunction (Conj); and
- Interjection (Int)

Linguistically, these basic categories have been further subdivided (Wilcock, 2009):

- Nouns have been subdivided into proper nouns (John, Mary, Helsinki), common nouns (man, woman, city), and pronouns (he, she, it);
- Verbs have been subdivided into intransitive verbs (walk, with no object), transitive verbs (see, with one object), and ditransitive verbs (give, with two objects); and
- Adjectives have been subdivided into absolute (big), comparative (bigger), and superlative (biggest).

The ANNIE POS Tagger Processing Resource is utilised in EDIT by using a default lexicon and a rule-set which are based on the outcome of a large Wall Street Journal trained corpus (Cunningham et al., 2013). The processes are explained below:

#### 1. Morphological Analysis

In morphological analysis, the sentence structure is analysed for identification, analysis and description of the structure of a given language's linguistic units, such as root words and affixes. Words can be related to other words by rules (grammar). For example, the words 'cup' and 'cups' are closely related — differentiated only by the plurality morpheme "-s". The Gate Morphological Analyser Processing Resource is used in EDIT to achieve this.

#### 2. Named Entity Recognition with Gazetteers

Gazetteer is a list that is used to find named entities in text. For example a list of names of places in Australia can be used to find the mention of a particular geographical location, such as Perth or New South Wales in the text. EDIT uses two gazetteers, the first is a programmatically created flexible gazetteer from the knowledge contained in EDO by using the OntoRoot component of GATE, and the second is a manually created word-list based gazetteer.

- i. The *Flexible Gazetteer Processing Resource* is used in EDIT to employ a list of entities generated from an ontology instead of a using a list of names (like the ANNIE gazetteer).

The Flexible Gazetteer utilises the list of names produced at runtime by another EDIT component called the OntoRoot Gazetteer (Figure 87). The OntoRoot Gazetteer is a processing resource which takes an ontology as the input and constructs a list of names for the Flexible Gazetteer’s use. In EDIT, EDO is provided as the input to the OntoRoot Gazetteer and links with the Flexible Gazetteer Processing Resource. The Flexible Gazetteer Processing Resource then looks in the documents (job advertisements) for the words and phrases provided by the OntoRoot Gazetteer. The output of this processing resource contains the words or phrases which appear in the ontology as instances of classes and the location of the matched text within each document.

- ii. In addition to the Flexible Gazetteer, the ANNIE Gazetteer was also modified (coined the SEEK Gazetteer) for use in EDIT. This is a truncated version of the original ANNIE Gazetteer which is used to recognise other named entities that are not mentioned in EDO.

Loaded Processing resources	
Name	Type
 Onto Root Gazetteer	Onto Root Gazetteer
 onto	ANNIE POS Tagger
 onto	GATE Morphological analyser
 onto	ANNIE English Tokeniser

**Figure 87 Processing resources for OntoRoot Gazetteer in the GATE application.**

### 3. Regular Expression-based Processing with Transducers

In text processing, transducers are finite state machines that process annotations by using regular expressions. GATE provides a JAPE (Java Annotation Patterns Engine) to implement these transducers. The JAPE Transducer processing resource operates over annotations based on regular expressions. It is used to perform certain actions on the output of previous processing resources. Each JAPE grammar has two parts, which are separated by a ‘-->’ symbol (Figure 88). The first part is a regular expression and the second part is a Java program which performs

certain actions. The first part (the regular expression) processes the input (tokens) and when a token or sequence of tokens match with the regular expression, the Java code in the second part is executed to perform the desired processing.

EDIT employs two JAPE transducers for JAPE grammar rules to extract information from the documents, based on the SEEK Gazetteer Processing Resource, while the other is an ontology aware JAPE transducer.

The grammar shown in Figure 88 is part of the JAPE transducer's processing of the output as part of the non-ontology aware gazetteer. This process searches for the concept of 'salary' in each job advertisement.

```
Phase: Salary
Input: Token Money
options: control = first

Rule: salaryDual
({Token.string == "Salary"} {Token.string=="."} {Money}){Token.string == "-"} {Money})

:lookup
-->
:lookup.Benefit = {kind="Benefit", rule="salaryDual"}

Rule: salarySingle
({Token.string == "Salary"} {Token.string=="."} {Money})

:lookup
-->
:lookup.Benefit = {kind="Benefit", rule="salarySingle"}
```

**Figure 88** JAPE grammar extracting salary amounts from job advertisement text.

Another grammar from the same transducer is show in Figure 89, where the code looks for the dates that are contained in the job advertisement. This grammar populates the ontology with

the advertisement dates of the job advertisement, and is a part of the ontology-aware transducer (unlike the non-ontology aware transducer described for salary above).

```
// This grammar checks whether or not the document has an instance in the ontology
// and then puts the advertisement date as a data type property value.

Imports: {
import java.net.URL;
import org.apache.commons.io.FilenameUtils;
}
Phase: InsertDate
Input: Date
Rule: insertDate
({Date}):mention
-->
:mention{

String baseURI="http://www.semanticweb.org/ontologies/2011/7/EmployerDemandOntology.owl#";
URL docURL=doc.getSourceUrl();
String baseName=FilenameUtils.getBaseName(docURL.getFile())+"."+FilenameUtils.getExtension(docURL.getFile());

// initializaiton
gate.AnnotationSet lookup = (gate.AnnotationSet) bindings.get("mention");
gate.Annotation mentionAnn = (gate.Annotation) lookup.iterator().next();
FeatureMap lookupFeatures = mentionAnn.getFeatures();
gate.FeatureMap features = Factory.newFeatureMap();
features.putAll(lookupFeatures);

//find the text covered by the annotation
String theMentionText = gate.Utls.stringFor(doc, mentionAnn);

// verifying the existence of class "Advertisement" in the Ontology
OClass aClass = ontology.getOClass(ontology.createOURI(baseURI+"Vacancy"));

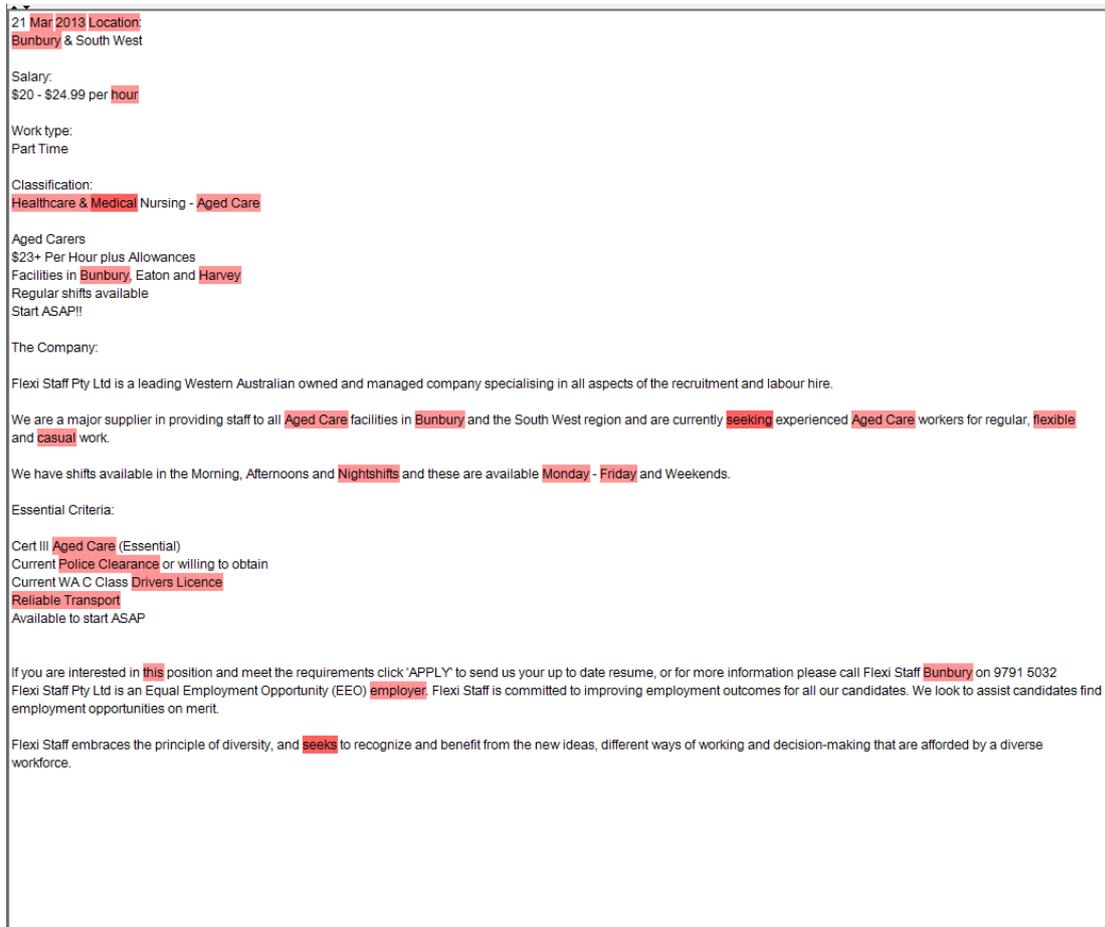
//create URI for current document
OURI jobURI = ontology.createOURI(baseURI + baseName);

//search the instance in ontology
OInstance jobInstance=ontology.getOInstance(jobURI);
//System.out.println("Instance Found:"+ jobInstance);
```

**Figure 89** Ontology aware JAPE grammar for advertisement date.

The ontology aware JAPE transducer is used to populate EDO with the information extracted by the previous processing resources. The transducer then verifies the information extracted by the Flexible Gazetteer Processing Resource and creates the necessary instances, object property values and datatype property values in EDO. In Figure 90, the advertisement text which matches the concepts contained in EDO, is highlighted. For example, the concepts *Bunbury* and *Harvey*

were identified in the job advertisement shown in the screenshot, as instances of the greater concept *LocationType*, and the requirements of *PoliceClearance* and *ReliableTransport* were identified as they are instances of the greater concept or class called *PositionRequirements*. This matching is done by the Flexible Gazetteer PR.



**Figure 90** Text appearing in job advertisements matched with the concept in the EDO by the EDIT through utilization of the Flexible Gazetteer Processing Resource.

In Figure 91, the JAPE grammar is shown, which creates an instance for each advertisement. This grammar checks whether an instance already exists in EDO for a job advertisement, and in case the job advertisement is new and has not been processed before by EDIT, it creates a new instance for it in the ontology. In Figure 92, a screenshot of EDO is shown, in which the instances created several job advertisements that are visible. More specifically, it also shows the object

properties (circled in yellow) and datatype properties (circled in green) that were created for the advertisement '9.htm'.

```

// checks whether or not the document has been processed before
// and creates an instance of class "Vacancy" for the for the current document if it does not exist
// in the ontology.

Imports: {
import java.net.URL;
import org.apache.commons.io.FilenameUtils;
}
Phase: CreateInstance
Input: Lookup
Rule: CreateInstance
({Lookup.type == instance}):mention
//({Lookup}):mention
-->
:mention{

//System.out.println("Instance?:"+doc.getFeatures());

if (doc.getFeatures().containsKey("InstanceCreated"))
{
//System.out.println("Instance exists");
}
else
{
System.out.println("Instance does not exist. Creating ...");
doc.getFeatures().put("InstanceCreated", "yes");
String baseURI="http://www.semanticweb.org/ontologies/2011/7/EmployerDemandOntology.owl#";

URL docURL=doc.getSourceUri();
String baseName=FilenameUtils.getBaseName(docURL.getFile())+"."+FilenameUtils.getExtension(docURL.getFile());
//System.out.println("===== \n Document:"+ baseName +"\n");

// verifying the existence of class "Advertisement" in the Ontology
OClass aClass = ontology.getOClass(ontology.createOURI(baseURI+"Vacancy"));
//System.out.println("ClassFound:"+ aClass +"\n");

//create uri for current document
OURI jobURI = ontology.createOURI(baseURI + baseName);

//search the instance in ontology
OInstance jobInstance=ontology.getOInstance(jobURI);
//System.out.println("Instance Found:"+ jobInstance +"\n");

//if instance does not exist create one
if (jobInstance == null)
//if (!ontology.containsOInstance(jobInstance))
{
//System.out.println("Creating Instance:" +jobURI+ "\n");
jobInstance = ontology.addOInstance(jobURI, aClass);
//System.out.println("Instance Created:" + ontology.getOInstance(jobURI));
}
}

```

**Figure 91** JAPE grammar used to create instances for job advertisements.

The screenshot displays the GATE application interface. On the left, a list of classes and instances is shown, including '5.htm' through '99.htm'. The main area is divided into two sections: 'Properties' and 'Property Values'.

**Properties:**

- hasSalary
- hasAdvertisement
- label
- hasBenefit
- hasOwner
- hasLocationType
- isEmployer
- isResponsibilityOf
- isOccupationOf
- isEmployerOf
- hasLocation
- comment
- isUniquelidOf
- isAdvisedVia
- isOwnedBy
- seeAlso
- hasOccupation
- isConditionOf
- isBenefitOf
- isDefinedBy
- isIndicationOf
- hasOrganization
- hasAttributeRequirements
- isRequirementOf
- isLocationOf
- hasEmployeeLanguage
- isSalaryLevelOf
- hasRequirement
- hasEmployeeQualificationLicenceOfRegistration
- hasAdvertisementStartDate
- versionInfo
- isPeriodOf
- hasResponsibility
- hasWebsite
- hasUniquelid

**Property Values:**

- adDate: 21 Mar 2013
- hasSalary: \$20 - \$24.99
- hasOtherRequirements: Police Clearance
- hasOtherRequirements: OwnTransport
- hasEmployeeQualificationLicenceOfRegistration: DriversLicence
- hasLocation: Blunbury
- hasLocation: Harvey
- hasAttributeRequirements: Flexible

**new datatype properties:**

- hasSalary
- hasOtherRequirements
- hasOtherRequirements
- hasEmployeeQualificationLicenceOfRegistration
- hasLocation
- hasLocation
- hasAttributeRequirements

**New object properties:**

- hasSalary
- hasOtherRequirements
- hasOtherRequirements
- hasEmployeeQualificationLicenceOfRegistration
- hasLocation
- hasLocation
- hasAttributeRequirements

Figure 92 Instances and property values added into EDO by EDIT (view is from the GATE application into EDO).

## **7.4 Results and Discussion**

To summarise, when the framework recognises the occurrence of concepts in the job advertisements, matching those contained in EDO, it creates an instance for each advertisement in the EDD. After this, the additional concepts recognised are linked to the instance of the advertisement, using object properties and datatype properties. The system has the ability to create new datatype or object properties in the ontology for those instances for which there is no matching or appropriate property. As such, an enhanced ontology evolves alongside the ontology population process.

A case study is provided to illustrate the processes that a single job advertisement undergoes during the different stages of the EDIF.

### **7.4.1 A Case Study**

#### **A. Data Collection**

An example of a typical job advertisement archived from seek.com as collected in the first step is provided in Figure 93. In addition to advertisement related content, web pages also have links to other pages on the website which are removed in the pre-processing stage to get a simplified advertisement (Figure 94).

## Aged Carers

Flexi Staff Pty Ltd - [More jobs by this advertiser](#)

### Aged Carers

- \$23+ Per Hour plus Allowances
- Facilities in Bunbury, Eaton and Harvey
- Regular shifts available
- Start ASAP!!

**The Company:**

Flexi Staff Pty Ltd is a leading Western Australian owned and managed company specialising in all aspects of the recruitment and labour hire.

We are a major supplier in providing staff to all Aged Care facilities in Bunbury and the South West region and are currently seeking experienced Aged Care workers for regular, flexible and casual work.

We have shifts available in the Morning, Afternoons and Nightshifts and these are available Monday - Friday and Weekends.

**Essential Criteria:**

- Cert III Aged Care (Essential)
- Current Police Clearance or willing to obtain
- Current WA C Class Drivers Licence
- Reliable Transport
- Available to start ASAP

If you are interested in this position and meet the requirements click 'APPLY' to send us your up to date resume, or for more information please call Flexi Staff Bunbury on 9791 5032

Flexi Staff Pty Ltd is an Equal Employment Opportunity (EEO) employer. Flexi Staff is committed to improving employment outcomes for all our candidates. We look to assist candidates find employment opportunities on merit.

Flexi Staff embraces the principle of diversity, and seeks to recognize and benefit from the new ideas, different ways of working and decision-making that are afforded by a diverse workforce.

---

**Head Office**      **Bunbury**      **Malaga**      **Rockingham**  
T: 08 9479 4781    T: 08 9791 5032    T: 08 9300 9933    T: 08 9592 7500

[www.flexistaff.com.au](http://www.flexistaff.com.au) Flexi Staff Pty Ltd ABN 83 009 440 288 ACN 009 440 288

**21 Mar 2013**

Location: Bunbury & South West

Salary: \$20 - \$24.99 per hour

Work type: Part Time

Classification: Healthcare & Medical ▶ Nursing - Aged Care

[Apply for this job](#)

Add to shortlist

Email this job

Add notes

Print job and notes

Share this job

**SEEK Learning courses**

TAFE, university or online courses

Choose occupation

[GO](#)

[Apply for this job](#)

**Figure 93** An original online job advertisement from [www.seek.com.au](http://www.seek.com.au).

**21 Mar 2013**

- Location:  
Bunbury & South West
- Salary:  
\$20 - \$24.99 per hour
- Work type:  
Part Time
- Classification:  
Healthcare & Medical Nursing - Aged Care

Aged Carers

- \$23+ Per Hour plus Allowances
- Facilities in Bunbury, Eaton and Harvey
- Regular shifts available
- Start ASAP!!

**The Company:**

Flexi Staff Pty Ltd is a leading Western Australian owned and managed company specialising in all aspects of the recruitment and labour hire.

We are a major supplier in providing staff to all Aged Care facilities in Bunbury and the South West region and are currently seeking experienced Aged Care workers for regular, flexible and casual work.

We have shifts available in the Morning, Afternoons and Nightshifts and these are available Monday - Friday and Weekends.

**Essential Criteria:**

- Cert III Aged Care (Essential)
- Current Police Clearance or willing to obtain
- Current WA C Class Drivers Licence
- Reliable Transport
- Available to start ASAP

**Figure 94** The 'cleaned' job advertisement.

## **B. Loading the Advertisement as a Corpus**

All the advertisements are loaded as a corpus language resource into the GATE system (Figure 95). Each component of the EDIT pipeline processes the advertisements in the sequence shown previously in Figure 83.

## **C. EDIT Pipeline**

In this step the EDIT pipeline performs the following seven processes on the advertisement.

1. Document Reset: This process clears the results of any previous execution of the pipeline on a document and refreshes the system for the next processes.

Index	Document name
0	 0.htm_00C89
1	 1.htm_00C8A
2	 10.htm_00C8B
3	 100.htm_00C8C
4	 101.htm_00C8D
5	 102.htm_00C8E
6	 103.htm_00C8F
7	 104.htm_00C90
8	 105.htm_00C91
9	 106.htm_00C92
10	 107.htm_00C93
11	 11.htm_00C94
12	 12.htm_00C95
13	 13.htm_00C96
14	 14.htm_00C97
15	 15.htm_00C98
16	 16.htm_00C99

**Figure 95** A collection of job advertisements loaded into EDIT's corpus.

2. Tokenization: As explained previously, this process breaks up the text into tokens (Figure 96).

21 Mar 2013 Location:  
Bunbury & South West

Salary:  
\$20 - \$24.99 per hour

Work type:  
Part Time

Classification:  
Healthcare & Medical Nursing - Aged Care

Aged Carers  
\$23+ Per Hour plus Allowances  
Facilities in Bunbury, Eaton and Harvey  
Regular shifts available  
Start ASAP!!

The Company:  
Flexi Staff Pty Ltd is a leading Western Australian owned and managed company specialising in all aspects of the recruitment and labour hire.  
We are a major supplier in providing staff to all Aged Care facilities in Bunbury and the South West region and are currently seeking experienced Aged Care workers for regular, flexible and casual work.  
We have shifts available in the Morning, Afternoons and Nightshifts and these are available Monday - Friday and Weekends.

Essential Criteria:  
Cert III Aged Care (Essential)  
Current Police Clearance or willing to obtain  
Current WA C Class Drivers Licence  
Reliable Transport  
Available to start ASAP

If you are interested in this position and meet the requirements click 'APPLY' to send us your up to date resume, or for more information please call Flexi Staff Bunbury on 9791 5032  
Flexi Staff Pty Ltd is an Equal Employment Opportunity (EEO) employer. Flexi Staff is committed to improving employment outcomes for all our candidates. We look to assist candidates find employment opportunities on merit.

Flexi Staff embraces the principle of diversity, and seeks to recognize and benefit from the new ideas, different ways of working and decision-making that are afforded by a diverse workforce.

**Figure 96** Putting the job advertisement through the tokenization process.

3. Sentence splitting: In this stage, the sentence boundaries are determined (Figure 97).

21 Mar 2013 Location:  
Bunbury & South West

Salary:  
\$20 - \$24.99 per hour

Work type:  
Part Time

Classification:  
Healthcare & Medical Nursing - Aged Care

Aged Carers  
\$23+ Per Hour plus Allowances  
Facilities in Bunbury, Eaton and Harvey  
Regular shifts available  
Start ASAP!!

The Company:  
Flexi Staff Pty Ltd is a leading Western Australian owned and managed company specialising in all aspects of the recruitment and labour hire.  
We are a major supplier in providing staff to all Aged Care facilities in Bunbury and the South West region and are currently seeking experienced Aged Care workers for regular, flexible and casual work.  
We have shifts available in the Morning, Afternoons and Nightshifts and these are available Monday - Friday and Weekends.

Essential Criteria:  
Cert III Aged Care (Essential)  
Current Police Clearance or willing to obtain  
Current WA C Class Drivers Licence  
Reliable Transport  
Available to start ASAP

If you are interested in this position and meet the requirements click 'APPLY' to send us your up to date resume, or for more information please call Flexi Staff Bunbury on 9791 5032  
Flexi Staff Pty Ltd is an Equal Employment Opportunity (EEO) employer. Flexi Staff is committed to improving employment outcomes for all our candidates. We look to assist candidates find employment opportunities on merit.

Flexi Staff embraces the principle of diversity, and seeks to recognize and benefit from the new ideas, different ways of working and decision-making that are afforded by a diverse workforce.

**Figure 97** The job advertisement is being sentence split.

4. POS tagging and morphological Analysis Figure 98: The results of these two steps are show in the figure below. POS tagging determines the usage of a word (noun, verb etc.) and the morphological analyser find the root word (e.g. manage for managed).

Flexi Staff Pty Ltd is a leading Western Australian owned and managed company specialising in all aspects of the recruitment and labour hire.

We are a major supplier in providing staff to all Aged Care facilities in Bunbury and the South West region and are currently seeking experienced Aged Care workers for regular, flexible and casual work

We have shifts available in the Morning, Afternoons and Nights shifts and these are available Monday - Friday and Weekends.

Essential Criteria:

Features
amix=s, category=v6z, kind=word, length=2, orin=lowercase, root=be, string=s}
category=DT, kind=word, length=1, orth=lowercase, root=a, string=a}
affix=ing, category=VBG, kind=word, length=7, orth=lowercase, root=lead, string=leading}
category=NNP, kind=word, length=7, orth=upperinitial, root=western, string=Western}
category=JJ, kind=word, length=10, orth=upperinitial, root=australian, string=Australian}
affix=ed, category=VBD, kind=word, length=5, orth=lowercase, root=own, string=owned}
category=CC, kind=word, length=3, orth=lowercase, root=and, string=and}
affix=ed, category=VBN, kind=word, length=7, orth=lowercase, root=manage, string=managed}
category=NN, kind=word, length=7, orth=lowercase, root=company, string=company}
affix=ing, category=VBG, kind=word, length=12, orth=lowercase, root=specialise, string=specialising}
category=IN, kind=word, length=2, orth=lowercase, root=in, string=in}
category=DT, kind=word, length=3, orth=lowercase, root=all, string=all}
affix=s, category=NNS, kind=word, length=7, orth=lowercase, root=aspect, string=aspects}
category=IN, kind=word, length=2, orth=lowercase, root=of, string=of}
category=DT, kind=word, length=3, orth=lowercase, root=the, string=the}
category=NN, kind=word, length=11, orth=lowercase, root=recruitment, string=recruitment}
category=CC, kind=word, length=3, orth=lowercase, root=and, string=and}
category=NN, kind=word, length=6, orth=lowercase, root=labour, string=labour}
category=NN, kind=word, length=4, orth=lowercase, root=hire, string=hire}
category=, kind=punctuation, length=1, root=, , string=, }
category=FRP, kind=word, length=2, orth=upperinitial, root=we, string=We}
affix=, category=VBP, kind=word, length=3, orth=lowercase, root=be, string=are}
category=DT, kind=word, length=1, orth=lowercase, root=a, string=a}
category=JJ, kind=word, length=5, orth=lowercase, root=major, string=major}
category=NN, kind=word, length=8, orth=lowercase, root=supplier, string=supplier}
category=IN, kind=word, length=2, orth=lowercase, root=in, string=in}
affix=ing, category=VBG, kind=word, length=9, orth=lowercase, root=provide, string=providing}
category=NN, kind=word, length=5, orth=lowercase, root=staff, string=staff}
category=TO, kind=word, length=2, orth=lowercase, root=to, string=to}
category=DT, kind=word, length=3, orth=lowercase, root=all, string=all}
category=NNP, kind=word, length=4, orth=upperinitial, root=aged, string=Aged}

Figure 98 POS and morphological checking of the job advertisement.

5. Recognition of concepts appearing in the text by the flexible gazetteer Figure 99.

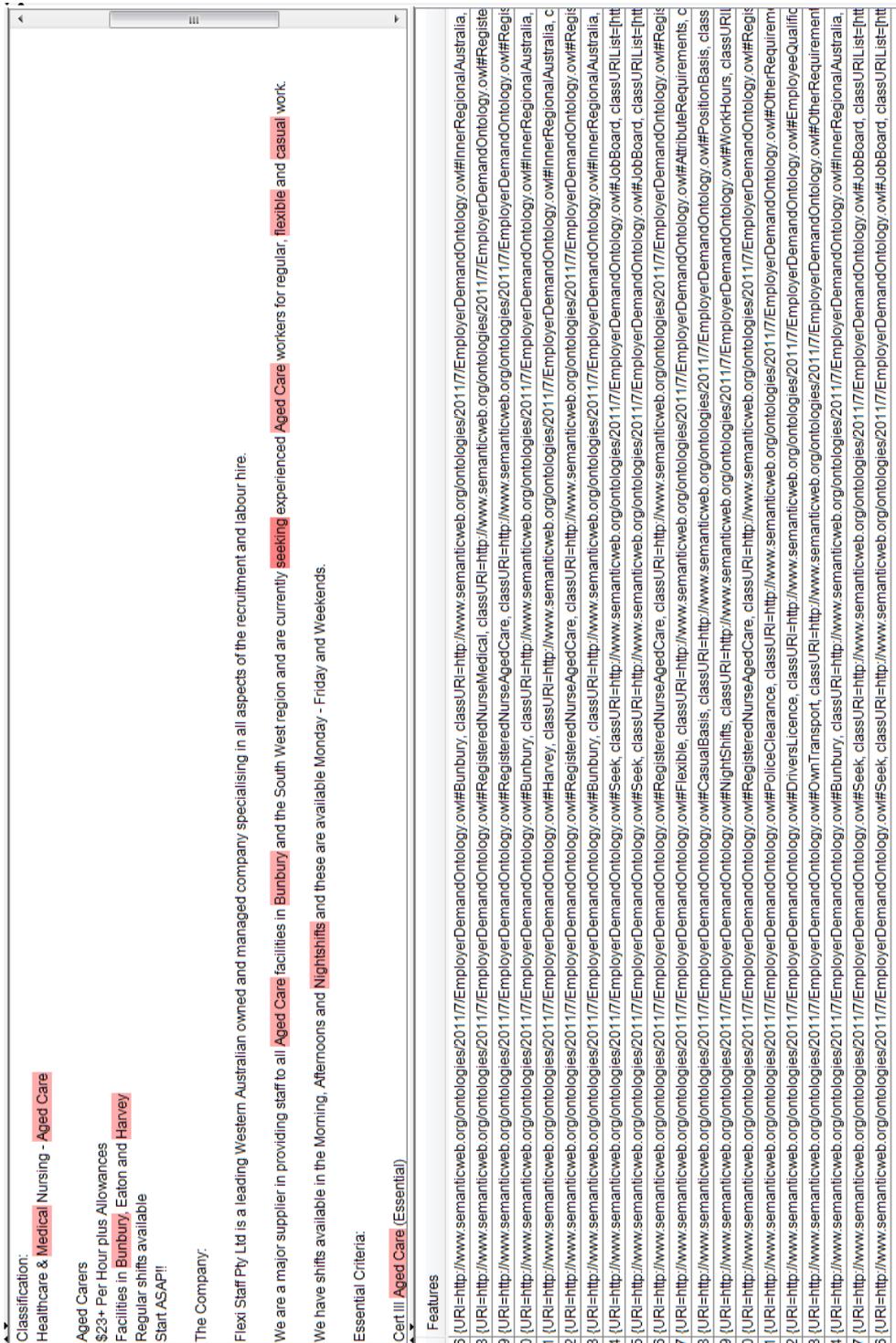


Figure 99 The flexible gazetteer recognises the concepts in the text.

6. The word-list based gazetteer along with the “JAPE Transducer Text” is used to recognise salary amounts and dates appearing in the text as information about these is not available in EDO.

7. The information extracted by the previous steps is populated into EDO by creating a knowledge instance for the advertisement with properties shown in (Figure 100).

Property Values		
adDate	21 Mar 2013	X
hasOtherRequirements	OwnTransport	X
hasOtherRequirements	PoliceClearance	X
hasLocation	Harvey	X
hasLocation	Bunbury	X
hasEmployeeQualificationLicenceOrRegistration	DriversLicence	X
hasAttributeRequirements	Flexible	X

**Figure 100** Population of extracted information through the process into EDO.

### 7.4.2 SPARQL Query Outcomes

It is necessary to detail the testing that was performed on EDIF. SPARQL queries have been run on EDIT to check that the tool produces the expected outcome and can, as such, be approved for user uptake in industry.

Once the information about a set of advertisements is inserted in the ontology by EDIT, the EDO ontology can be queried using SPARQL to get a range of quantitative data from the ontology. SPARQL is a query language that can query ontologies to produce the desired summarised results, similar to conventional SQLs which are used to query traditional databases (McCarthy, 2005; “W3C SPARQL Query Language for RDF”, 2013).

The EDD uncovers the employer demand intelligence in RDF format, which allows queries to be executed on it. If the dataset was not available in RDF format, software queries would not be possible, and users would have to spend a lot of time looking at the original sources to try and find their answers.

The following text and figures show how a variation of data has been filtered and retrieved from EDIT through the submission of SPARQL queries to the endpoint<sup>39</sup>. To make the text more readable, the following prefixes are introduced:

owl: < <http://www.w3.org/2002/07/owl#>>

rdf: < <http://www.w3.org/1999/02/22-rdf-syntax-ns#>>

rdfs: < <http://www.w3.org/2002/07/owl#>>

edo: < <http://www.semanticweb.org/ontologies/2011/7/EmployerDemandOntology.owl#>>

Figure 101 indicates the query and subsequent results to retrieve all the jobs and their locations that were advertised on 18 March 2013.

<b><u>Query</u></b>		
PREFIX edo: < <a href="http://www.semanticweb.org/ontologies/2011/7/EmployerDemandOntology.owl#">http://www.semanticweb.org/ontologies/2011/7/EmployerDemandOntology.owl#</a> >		
SELECT *		
WHERE { ?job edo:hasLocation ?Location; edo:adDate ?adDate		
Filter(?adDate= "18 Mar 2013" @en) }		
ORDER BY ?Vacancy		
<b><u>Result</u></b>		
1.htm	Perth	"18 Mar 2013"@en
10.htm	Perth	"18 Mar 2013"@en
17.htm	Perth	"18 Mar 2013"@en
56.htm	Yanchep	"18 Mar 2013"@en
56.htm	Perth	"18 Mar 2013"@en
60.htm	Perth	"18 Mar 2013"@en
60.htm	Geraldton	"18 Mar 2013"@en
81.htm	Perth	"18 Mar 2013"@en
94.htm	Perth	"18 Mar 2013"@en
94.htm	Bunbury	"18 Mar 2013"@en
97.htm	Perth	"18 Mar 2013"@en

Figure 101 A query to retrieve all the job advertisements and their locations that were advertised on 18 March 2013.

<sup>39</sup> <http://www.employerdemandintelligence.org>

Figure 102 shows the query that retrieves all the instances of job advertisements that were advertised in each location in Western Australia, and the number of advertisements for each of these geographical locations.

<b><u>Query</u></b>	
PREFIX edo: <http://www.semanticweb.org/ontologies/2011/7/EmployerDemandOntology.owl#>	
SELECT (SAMPLE(?location) AS ?location) (COUNT(?location) as ?number)	
WHERE {	
?job edo:hasLocation ?location	
}	
GROUP BY ?location	
ORDER BY ?location	
<b><u>Result</u></b>	
Albany	"8"^^<http://www.w3.org/2001/XMLSchema#integer>
Bunbury	"27"^^<http://www.w3.org/2001/XMLSchema#integer>
Busselton	"6"^^<http://www.w3.org/2001/XMLSchema#integer>
Esperance	"3"^^<http://www.w3.org/2001/XMLSchema#integer>
Geraldton	"4"^^<http://www.w3.org/2001/XMLSchema#integer>
Harvey	"2"^^<http://www.w3.org/2001/XMLSchema#integer>
Mandurah	"5"^^<http://www.w3.org/2001/XMLSchema#integer>
MargaretRiver	"1"^^<http://www.w3.org/2001/XMLSchema#integer>
Northam	"4"^^<http://www.w3.org/2001/XMLSchema#integer>
Perth	"74"^^<http://www.w3.org/2001/XMLSchema#integer>
Yanchep	"1"^^<http://www.w3.org/2001/XMLSchema#integer>

**Figure 102** A query to retrieve all the job advertisements that were advertised in the various locations in Western Australia.

Figure 103 provides the query to show the number of job advertisements that had *OtherRequirements* listed, and what each of those requirements were.

<b>Query</b>	
PREFIX edo: <http://www.semanticweb.org/ontologies/2011/7/EmployerDemandOntology.owl#>	
SELECT (SAMPLE(?req) AS ?req) (COUNT(?req) as ?number)	
WHERE {	
?job edo:hasOtherRequirements ?req	
}	
GROUP BY ?req	
<b>Result</b>	
OwnTransport	"3"^^<http://www.w3.org/2001/XMLSchema#integer>
OwnVehicle	"3"^^<http://www.w3.org/2001/XMLSchema#integer>
PoliceClearance	"56"^^<http://www.w3.org/2001/XMLSchema#integer>
WorkingRights	"6"^^<http://www.w3.org/2001/XMLSchema#integer>
WorkingWithChildrenCheck	"1"^^<http://www.w3.org/2001/XMLSchema#integer>

**Figure 103** The SPARQL query to establish the number of job advertisements that were advertised with *OtherRequirements*.

### 7.4.3 The Pros and Cons of Using GATE

GATE is a state of the art tool in its class, as it provides several Application Programming Interfaces (APIs) for all the required functionality for text processing, as well as ontology processing for this project — which no other tool provides as a single development environment. Furthermore, GATE is an open source software and has been used for many research oriented projects. Several other popular ontology population tools such as KIM<sup>40</sup>, are driven by GATE.

As well as the pros of using GATE, there were also some cons. The development of EDIT suffered considerable delays, as the GATE developer is not commercial software, but open source software, and is constantly undergoing enhancements and upgrades; therefore the official documentation is not always up-to-date. Although GATE is based on Java, the programming

<sup>40</sup> <http://www.ontotext.com/kim>

expertise required to use its components and APIs are unique to GATE, requiring knowledge of text processing techniques to be able to work with it.

#### 7.4.4 Metrics and Development Time

Table 11 provides the approximate development time of EDIT. In total, 119 days were spent on its development; however, there were several days where technical difficulty was experienced and where there was a learning curve involved in understanding how certain software aspects function. If the difficulties are excluded, the development time was brought down to only 19 days.

Activity	Number of days' development	Number of days' technical difficulty experienced	Number of days' learning curve needed
Problem Analysis	5	15	10
Basic Processing Pipeline Design	2	30	12
Gazetteer Preparation	2	1	2
Jape Grammar	5	10	20
<b>Total development time</b>	<b>19</b>	<b>56</b>	<b>44</b>

**Table 11** The approximate time taken to develop EDIT.

Table 12 provides a range of metrics that are relevant to EDIT, such as the number of job advertisements that were processed during the one month's trial period for the EDIT, and the number of instances that were populated that have the object properties *hasSalary*, *hasOtherRequirements*, and *hasLocation*.

Metric	Quantity
Number of job advertisements processed	105
Number of new data type properties created	2
Number of data type property instances asserted	131
Number of job advertisements with a date recorded	122
Number of new object properties created	6
Number of object properties asserted	495
Number of <i>hasRequirement</i> instances populated	2
Number of <i>hasSalary</i> instances populated	9
Number of <i>hasOtherRequirements</i> instances populated	69
Number of <i>hasLocation</i> instances populated	135
Number of <i>hasEmployeeQualificationLicenceOrRegistration</i> instances populated	35
Number of <i>hasAttributeRequirements</i> instances populated	244
Number of <i>hasEmployeeLanguage</i> instances populated	10

**Table 12** A variety of metrics relevant to the EDIT.

#### 7.4.5 Limitations of the EDIT

EDIT is only a prototype developed to assess the effectiveness of EDO in representing its targeted domain. Several aspects which need improvement in future work are discussed below:

- 1 The Seek Gazetteer Processing Resource has a minimal functionality which enables EDIT to find the salary of a job and the date of publication of an advertisement. However, its functionality can be enhanced through extension of the gazetteer by adding more names into

its lists and adding more JAPE grammar to the JAPE Transducer Processing Resource to process these lookups.

2. The approach followed in EDIT is to use a locally stored ontology, and it is consequently only able to extract information that has a representation in the ontology. The functionality of EDIT can be enhanced in future versions: instead of using such an isolated ontology it can be linked with the semantic web to search for a wider range of concepts.

3. EDIT is based on a fully automated process that looks for the concepts in the job advertisements that match those contained within the EDO. Therefore, sometimes it wrongly associates words which are used in the job advertisements' text in a different sense to the meaning implied by the concepts contained in EDO. These mistakes cannot be completely removed, but it can be reduced to some extent by enhancing the usage of the Part of Speech Tagger's output and incorporating more JAPE grammar.

## **7.5 Conclusion**

This chapter has focused on the development and usage of the Employer Demand Intelligence Tool (EDIT), a GATE application that has been amended to suit the needs of EDIF. EDIF is almost entirely automatic. The only time that human intervention is required now that EDIT has been developed is at run-time, and so it is referred to as a semi-automatic system.

The detailed steps of EDIF discussed in this chapter are: step one: the creation of the corpus; step two: loading the corpus and the EDO into the EDIT; and step three: creating an EDIT pipeline to process the corpus. An example of the processing of one job through EDIF has been provided to show the detailed steps of the system.

The chapter concludes with a results and discussion section, where three detailed SPARQL queries and their results are provided, along with a summary of the difficulties that were

experienced in using GATE and some metrics. This is followed by a brief outline of the limitations of EDIT, and suggested options for future work to enhance the system. Evaluation and testing of EDIF are discussed in the next chapter.

# Chapter 8 - Employer Demand Intelligence Framework Evaluation

## 8.1 Introduction

The previous chapters in this thesis focused on the developmental stages of the Employer Demand Ontology (EDO): the methodology, modelling, design and implementation of the ontology, whilst providing solutions for research issues one, two and three. In this chapter, the focus turns to ensuring EDO is validated and verified.

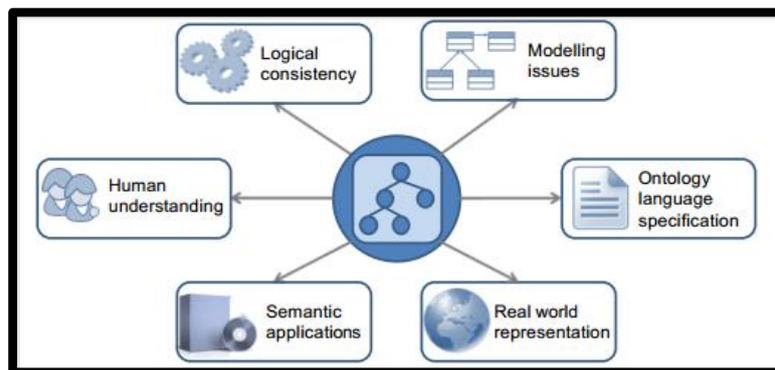
Validating and verifying an ontology is part of the final stage of the ontology lifecycle, and provides a report on whether the ontology has been developed to sufficient engineering standards (i.e., whether it is consistent and satisfiable), and whether the end-user would be able to use the ontology for the purposes that it was developed.

The chapter first looks at the fundamentals of evaluating an ontology, and then provides the specific verification and validation method used to evaluate EDO. The method consists of four detailed use case scenarios, with questions and queries that put EDO to the test. The validation of actual industry use cases is planned as future research.

## 8.2 Verification and Validation Fundamentals

The final step of the ontology development lifecycle is the validation and verification of the ontology to evaluate the ontology's quality and outcome. The increasing uptake and evolution of ontologies has provided many suggestions on how to evaluate an ontology (Strohmaier et al., 2013). Brank et al. (2005) highlighted the following four techniques: i) techniques that compare the ontology to some kind of 'golden standard'; ii) techniques that use an ontology in an application and evaluate the results as such; iii) techniques that compare it with a data source; and iv) evaluation of techniques by humans to establish how well the ontology meets a set of

predefined standards, requirements and criteria. As there is no single perfect evaluation method, the decision of which technique to use should be based on the purpose of the evaluation, the application's ontology, and the aspect of the ontology that needs to be evaluated. Figure 104 provides six dimensions to be evaluated as per Poveda-Villalon et al. (2012).



**Figure 104** Six dimensions that can be identified in ontology quality (Poveda-Villalon, et al., 2012).

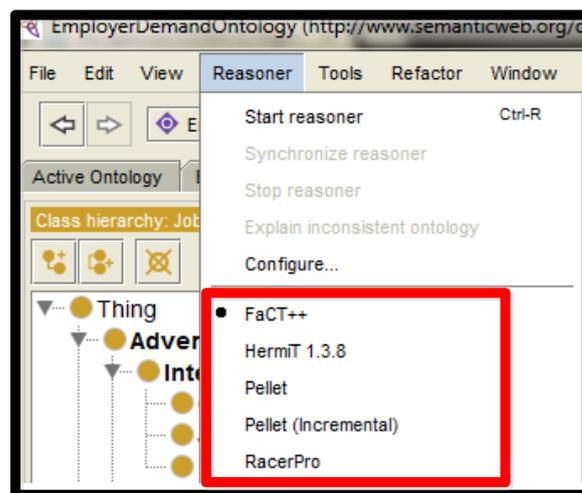
The development of EDO was based on the NeOn methodology in which some aspects of ontology evaluation had already occurred — but for an earlier phase of the ontology lifecycle than the phase that will be described in this chapter. EDO has been developed to the extent where it is now ready for verification and validation. Verification answers the question ‘Have we built the system right?’, whereas validation answers the question ‘Have we built the right system?’ (Easterbrook, 2010). Thus, verification checks whether the system is well-engineered and validation is the process that evaluates whether the customers’ actual needs have been met.

As such, the evaluation technique that is appropriate at this stage of the ontology lifecycle, is a combination of points ii) and iv) above — techniques that use the ontology in an application to evaluate the results, and techniques where evaluation is done by humans to try and assess if the ontology meets the set of predefined requirements, which have been defined at the start of the ontology lifecycle.

To evaluate EDO's final stage of its lifecycle, the following two methods have been used: i) the application of reasoners to verify the EDO for its engineering robustness; and ii) applying a set of use cases to validate the ontology's original purpose and usefulness. The following two sections will discuss each of these evaluation tasks in detail.

### **8.3 The Employer Demand Ontology Logical Consistency Verification**

Ensuring an ontology is consistent is an important part of an ontology's development and testing. If an ontology is inconsistent, no reliable conclusion can be deduced. The EDO has been reasoned to check its logical consistency. The FaCT++, HermiT, Pellet, Pellet (Incremental), RacerPro and TrOWL reasoners were all used for consistency verification (Figure 105).



**Figure 105** Some of the reasoners used in Protégé.

The time it took to reason EDO with FaCT++ was 226ms; with HermiT 1.3.8 it took 1382ms; with Pellet it took 714ms; with Pellet (Incremental) it took 1349ms; and with TrOWL it took 478ms (Figure 106). The reasoners checked the class, object property and data property hierarchies, the class assertions, the object property assertions and whether there were the same individuals contained within the ontology.

```

Protege.exe
cwd: C:\Program Files\Protege_4.1
cmd: "C:\Program Files\Java\jre6\bin\java.exe" -Xms262144000 -Xmx524288000 -Dlog
4j.configuration=file:log4j.xml -DentityExpansionLimit=100000000 -Dorg.protege.p
lugin.dir=plugins -Dfile.encoding=UTF-8 -classpath "C:\Program Files\Protege_4.1
\bin\felix.jar;C:\Program Files\Protege_4.1\bin\ProtegeLauncher.jar;C:\Program F
iles\Protege_4.1\lax.jar;" com.zerog.lax.LAX "C:/Program Files/Protege_4.1/Prote
ge.lax" "C:/Users/Chamonix/AppData/Local/Temp/lax0558.tmp"
Starting Protege 4 OWL Editor (Version 4.1.0, Build = 239)
Platform:
  Java: JUM 1.6.0_43-b01 Memory: 506M
  Language: en, Country: AU
  Framework: Apache Software Foundation (1.5)
  OS: windows7 (6.1)
  Processor: x86
Installed plugin Pellet Reasoner Plug-in
Installed plugin OWLPropViz Plug-in
Installed plugin Dlquery
Installed plugin Matrix Views
Installed plugin OWLViz
Installed plugin The Protege 4 OWL Editor
Installed plugin OntoGraf Plug-in
Installed plugin Hermit Reasoner
Installed plugin The OWL API
Installed plugin TrOWL Reasoner
Installed plugin Factplusplus Plug-in
Using OWL API version 3.2.3.1824
Rebuilding entity indices...
... rebuilt in 50 ms
Loading Anonymous-1 from file:/C:/Users/Chamonix/Documents/PhD/Methodology%20de
velopment/Ontology/EmployerDemandOntology/EmployerDemandOntology_v2.owl
Setting active ontology to <http://www.semanticweb.org/ontologies/2011/7/Employe
rDemandOntology.owl>
Rebuilding entity indices...
... rebuilt in 5 ms
... active ontology changed
Setting active ontology to <http://www.semanticweb.org/ontologies/2011/7/Employe
rDemandOntology.owl>
Rebuilding entity indices...
... rebuilt in 2 ms
... active ontology changed
FaCT++ Kernel: Reasoner for the SROIQ(D) Description Logic, 32-bit
Copyright (C) Dmitry Tsarkov, 2002-2013. Version 1.6.2 (19 February 2013)
Initializing the reasoner by performing the following steps:
  class hierarchy
  object property hierarchy
  data property hierarchy
  class assertions
  object property assertions
  same individuals
FaCT++ classified in 226ms
Initializing the reasoner by performing the following steps:
  class hierarchy
  object property hierarchy
  data property hierarchy
  class assertions
  object property assertions
  same individuals
Hermit 1.3.8 classified in 1382ms
Initializing the reasoner by performing the following steps:
  class hierarchy
  object property hierarchy
  data property hierarchy
  class assertions
  object property assertions
  same individuals
Pellet classified in 714ms
Initializing the reasoner by performing the following steps:
  class hierarchy
  object property hierarchy
  data property hierarchy
  class assertions
  object property assertions
  same individuals
Extracting 197 elements
Extracting: 100% complete in 00:00
Extracting finished in 00:00
Pellet (Incremental) classified in 1349ms
Initializing the reasoner by performing the following steps:
  class hierarchy
  object property hierarchy
  data property hierarchy
  class assertions
  object property assertions
  same individuals
TrOWL classified in 478ms

```

Figure 106 Some of the Protégé reasoners' logs while classifying the EDO.

Consistency verification through a reasoner for EDO included consistency checking, concept satisfiability, classification, and realisation. These facilities are all standard inference services conventionally provided by a reasoner. EDO does not contain any contradictory facts — logical consistency of the ontology was also checked through the reasoners mentioned above.

EDO also has the concept ‘satisfiability’, i.e. a class in the EDO is able to have instances. Defining an instance of a class causes the whole ontology to be inconsistent if a class is unsatisfiable. EDO also has a complete class hierarchy (classification service) that can be used to answer queries. The most specific class that an instance belongs to can be found in EDO (realisation service). From this verification experiment, EDO was found to be consistent and ready for user consumption.

## **8.4 The Employer Demand Ontology Validation**

As discussed, ontologies can be used to advance existing eRecruitment and eHumanResources applications, and they can also enable new uses for eRecruitment and eHumanResources. A cross section of four interesting fictional use cases has been developed to validate EDO, based on its wide stakeholder reach. These use cases have been formulated from real-world usage examples, and consist of a set of typical queries or functionalities that the end-user may like to use the agent for. Validating EDO in industry by allowing real users to access the system has been recorded as a future project in this research.

In order to query EDD for the use cases below, the DL Query function in Protégé 4.3 has been used. The resultant classes and instances provided in the use case examples can be browsed in Webprotégé available from the EDIF website<sup>41</sup>. It is also possible to query EDD via a SPARQL endpoint from this website. However, for each of the first questions of the four use cases, a random sample of five job advertisements (or if the query returned less than five results — all

---

<sup>41</sup> <http://www.employerdemandintelligence.org>

results) in their original states were provided in screenshots to support the results shown in each query. Due to space limitations in this thesis, it was not possible to supply all query results' job advertisements; it was deemed sufficient to use random sampling to support the results.

#### **8.4.1 Use case 1: Government Official Developing Policy and Funding Protocols**

##### **1. Scenario**

Thomas works at the Department of Health in Perth as the State Manager: Workforce Development. His area is responsible for looking after the whole of Western Australia's health staffing supply needs. This includes issues like ensuring that there are enough doctors in each town to serve the community; that colleges and vocational institutions have sufficient funding; to ensure students are trained in the occupational areas of need so there is sufficient supply of staff across the state; and to assign grants to groups that are able and keen to manage projects that can support the work that the Department of Health is doing towards their statewide workforce management responsibilities. One of the tasks that Thomas looks after on a quarterly basis, is to draw up an environmental scan of the health workforce in Western Australia, to inform his area's policy development practices. Amongst other tasks, he phones companies to enquire about recent skill shortages experienced in their companies over the last three months and look at the job boards online to get a feel for the number of job advertisements and types of occupations that employers are advertising at that point in time. This informs his analysis and subsequent report writing. However, Thomas is finding it increasingly difficult to make an informed judgement about the vacancies — the job boards do not provide detailed analyses about the number and types of jobs for each regional area, nor does he have the time to trawl each advertisement to get more information, and keep track of the number of advertisements being posted online every day.

## 2. Questions

To assist Thomas in his quest, he would like to access EDIT to obtain the following type of information from the system:

1. How many jobs were advertised in total this month?
2. Which occupation generally has the most requirements?

## 3. Results from EDO

As discussed, in order to retrieve a natural language question from the system, it needs to be formulated into a format that the system can understand. Each question's transition from natural language to DL query language is discussed below.

1. How many jobs were advertised in total this month for Regional and Remote Areas in Western Australia compared to those advertised for Perth's Metropolitan area?

EDO only consists of one month's job advertisements at this stage and as such it is not necessary to indicate a time period for the query. Entering the query as depicted in Figure 107, with the box 'individuals' ticked on the right of the screen, retrieves all the job advertisement instances that were advertised, and hence recorded into EDO for that month. However two queries need to be run — one for each of the two areas specified.

The first result shows that 80 instances exist in EDD that meet the requirements of the query for Perth's Metropolitan area (Figure 107). A random sample of five of the query result job advertisements has been provided in Figure 108 to Figure 112 to support the results of the query, showing their locations as one of the metropolitan areas in Western Australia.

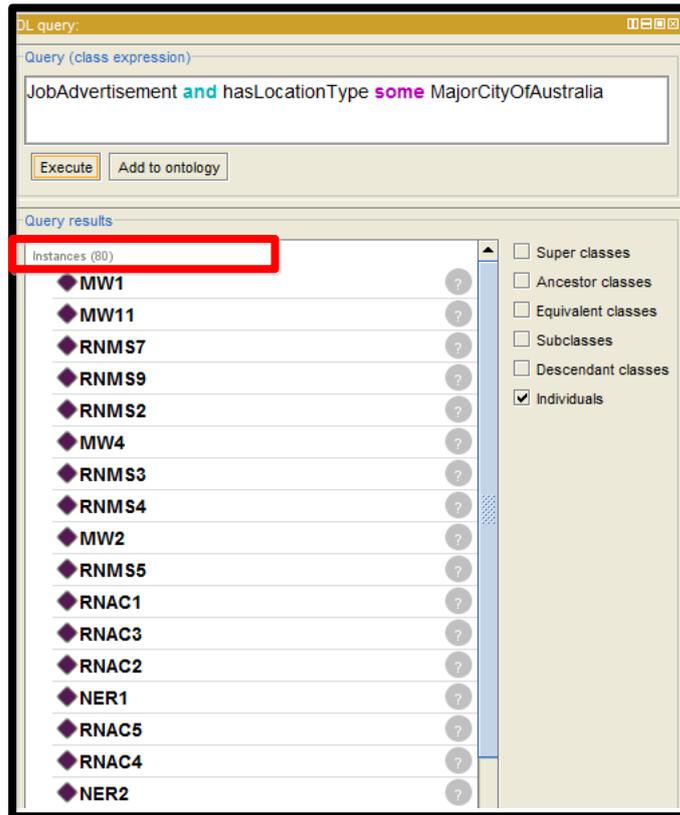


Figure 107 DL Query and result showing the number of *JobAdvertisements* advertised for the month in Western Australia’s metropolitan area.

**IPA**

## Calling Midwives to Perth

- Permanent Employment
- State of the Art Maternity Unit
- Sponsorship visas offered (if required)

IPA Healthcare is currently looking for Midwives to join a well respected private hospital in Perth. Continue your Midwifery career in this state-of-the-art maternity unit. Permanent full time positions are available in antenatal, postnatal, labour and delivery units.

This hospital is offering excellent working conditions, salary sacrificing, fantastic support, impressive salaries and staff development. For these positions you must be eligible for registration with the Nurses Board of Australia and have previous midwifery experience.

International Midwives are encouraged to apply with permanent and temporary sponsorship visas offered to successful candidates.

IPA's dedicated health recruitment team can offer advice and guidance through the registration and visa process as well as providing ongoing support throughout your relocation to Australia. Apply today we are waiting to hear from you!

Please apply online today or for more information please contact IPA:

Contact: Ginny Donley, IPA Healthcare Recruitment Consultant  
 Email: [gdonley@ipa.com.au](mailto:gdonley@ipa.com.au)  
 Phone: +61 8 9463 1921

Ginny Donley  
 Email: Please click the 'Apply Now' button below.

**Healthcare**

Figure 108 Job advertisement MW1.



## Registered Nurses - Midwives

Your Agency of choice – Because it's all about you!

- You choose when
- You choose where
- You choose how often
- You choose an agency that understands family friendly

The West Australian Nursing Agency is part of the Healthcare Australia Group - the largest provider of agency nurses within Australia.

We are looking for experienced registered nurses to join our talented team for midwifery shifts throughout the **Perth** metropolitan area.

Join WANA today to receive:

- competitive rates
- free uniform
- ongoing support and clinical supervision
- free online education
- optional income distribution facility
- regular social activities
- a friendly team who are always available to help you!

Phone our recruitment team on 1300 660 321 or email [recruitment@wana.com.au](mailto:recruitment@wana.com.au)

Visit us online at [www.wana.com.au](http://www.wana.com.au)

Successful applicants will be asked to provide:

- Working With Children Check
- Criminal Record Check
- Completion of Mandatory Reporting of Child Sexual Abuse Training
- Current Registration with the WA Nurses Board

Email: Please click the 'Apply Now' button below.

Figure 109 Job advertisement MW2.



## Perth Registered Nurses - Perth, Western Australia



Are you an experienced **general nurse** looking for challenging and regular work?

**Nursing Australia** is a leading supplier of staffing solutions to facilities across. We currently have opportunities for regular and ongoing work in both public and private acute facilities within **Perth**. As part of our preferred provider agreements with the Health Department of Western Australia, Ramsay Health Care and many other private hospitals, Aged Care and Community clients, we are well placed to keep you in regular work.

**Nursing members benefit from:**

- Agency rates of pay - paid weekly
- Professional indemnity insurance
- On line rostering
- Referral Bonanza program!
- Ability to work across all states of Australia

Essential requirements include:

- Previous relevant experience in your clinical area |
- Current Australian nursing registration.
- Current police clearance
- Standard annual competencies (or we can assist you to attain the same)

For more information please contact **Nursing Australia** on 13 10 95 or visit [www.nursingaustralia.com](http://www.nursingaustralia.com) Alternatively follow the links below to apply online  
 Email: Please click the 'Apply Now' button below.

13 10 95

[www.nursingaustralia.com](http://www.nursingaustralia.com)



Melbourne Sydney Brisbane Adelaide Perth Darwin Newcastle Gold Coast

Figure 110 Job advertisement RNMS5.

**Registered Nurses, Enrolled Nurses & Aged Care Staff REQUIRED**

**Registered Nurses, Enrolled Nurses & Aged Care Staff REQUIRED FOR IMMEDIATE START!**

Want to earn GREAT RATES?  
 Want to work FLEXIBLE HOURS/SHIFTS?  
 Do you have experience in AGED CARE?

If YES, read on....

Auscare Staffing Agency are currently seeking RNs, ENs and experienced aged care AINs to work at various locations throughout Perth's metropolitan area.

Do you have **Certificate III or IV in Aged Care** with residential aged care experience (6 months or more) and a desire to be part of a dynamic and motivated team?

If the answer is still YES, then, we have regular and casual shifts available in reputable aged care facilities - AM, PM and ND – waiting for you NOW!

**Take your pick, do a great job and earn great rates!**

**We can offer you:**

- A variety of shifts ~ work where and when you want,
- Weekly pay cycles ~ to make your life easier,
- Fantastic pay rates ~ because we love our staff,
- A great working environment ~ a real team,
- Excellent opportunities/rewards for hardworking staff!!!

**Essential criteria:**

- Solid nursing experience or a minimum 6 months aged care experience with a Certificate III in Aged Care;
- Current Manual Handling Certificate or be prepared to enrol in our FREE training course;
- The ability to provide a high standard of care;
- Current National Police Clearance [or be willing to obtain one];
- Must have transport & current driver's license
- Two professional references

If this sounds like you then send us your CV TODAY and don't miss out on a fantastic opportunity!

Please email your resume to [admin@auscarestaffing.com.au](mailto:admin@auscarestaffing.com.au)

If you have any queries please contact us on 08 6364 3917

Figure 111 Job advertisement RNAC5.

 **IPA**

**Calling ALL CN Medical / Surgical Nurses**

- Permanent Employment
- State of the Private Hospital
- Sponsorship visas offered (if required)

IPA Healthcare is currently looking for Clinical Medical & Surgical Registered Nurse to join a well respected private hospital in Perth. Continue your clinical career in this state-of-the-art hospital. Permanent full time positions are available in medical & surgical wards.

This hospital is offering excellent working conditions, salary sacrificing, fantastic support, impressive salaries and staff development. For these positions you must be eligible for registration with the Nurses Board of Australia and have previous Clinical Nurse level experience.

International nurses are encouraged to apply with permanent and temporary sponsorship visas offered to successful candidates.

IPA's dedicated health recruitment team can offer advice and guidance through the registration and visa process as well as providing ongoing support throughout your relocation to Australia. Apply today we are waiting to hear from you!

Please apply online today or for more information please contact IPA;

Contact: Ginny Donley, IPA Healthcare Recruitment Consultant  
 Email: [gdonley@ipa.com.au](mailto:gdonley@ipa.com.au)  
 Phone: +61 8 9463 1921

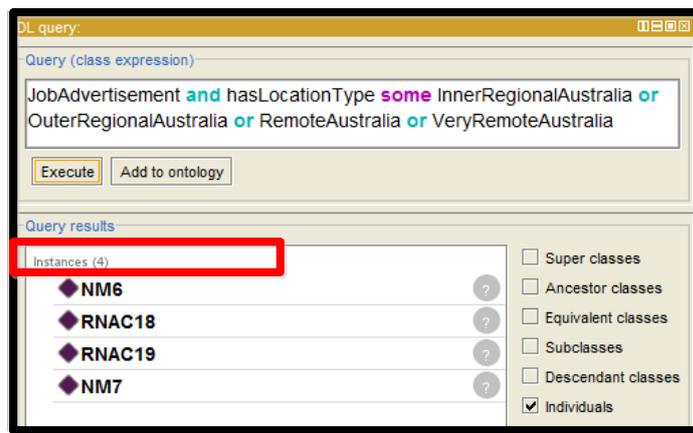
Ginny Donley  
 Email: Please click the 'Apply Now' button below.

**Healthcare**

Figure 112 Job advertisement RNMS2.

The second result, highlighted in Figure 113, shows that four instances exist in the EDD that meet the requirements of the query for Western Australia’s Regional and Remote areas.

It is clear that many more positions were being advertised for Western Australia’s metropolitan area than for any regional areas in Australia. This contradicts reports read in the literature, where it is always indicated that a greater number of vacancies exist in regional areas than in metropolitan areas, and this is an issue Thomas should be investigating.



**Figure 113** DL Query and result showing the number of JobAdvertisements advertised for the month in Western Australia’s regional areas.

2. Which nursing occupation group had the most requirements from employers?

This natural language question requires four DL queries to be run, in order to achieve the desired answer: one for each type of nursing occupation class (*MidwifeAndNursingProfessionals*) in EDD: *Midwives*, *NurseEducatorsAndResearchers*, *NurseManagers* and *RegisteredNurses*. The four DL queries are required to retrieve answers where the *JobAdvertisement* has a *PositionRequirement* of any type for each of the specified *MidwifeAndNursingProfessionals* subclasses (Figure 114 to Figure 117).

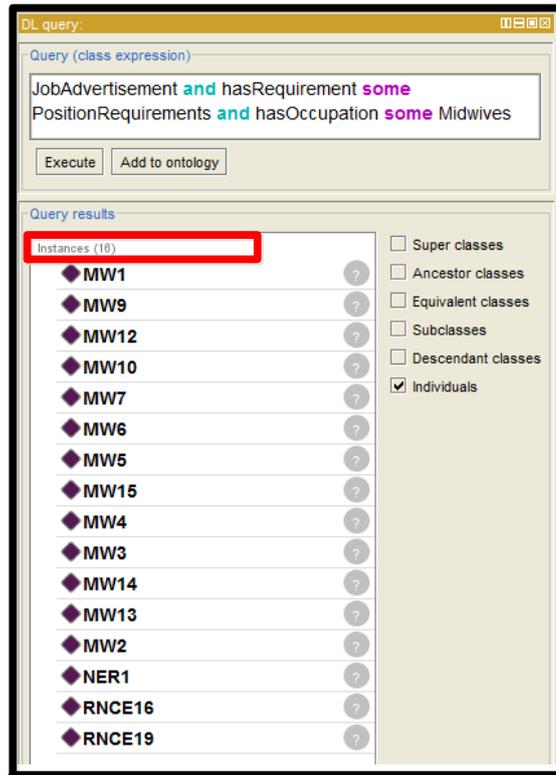


Figure 114 DL Query and result showing the number of *PositionRequirements* advertised for the occupation class *Midwives*.

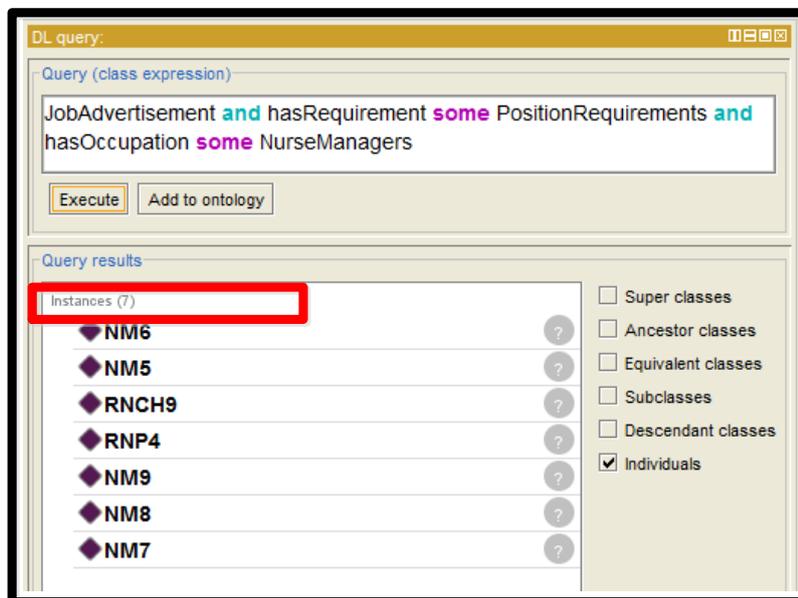


Figure 115 DL Query and result showing the number of *PositionRequirements* advertised for the occupation class *NurseManagers*.

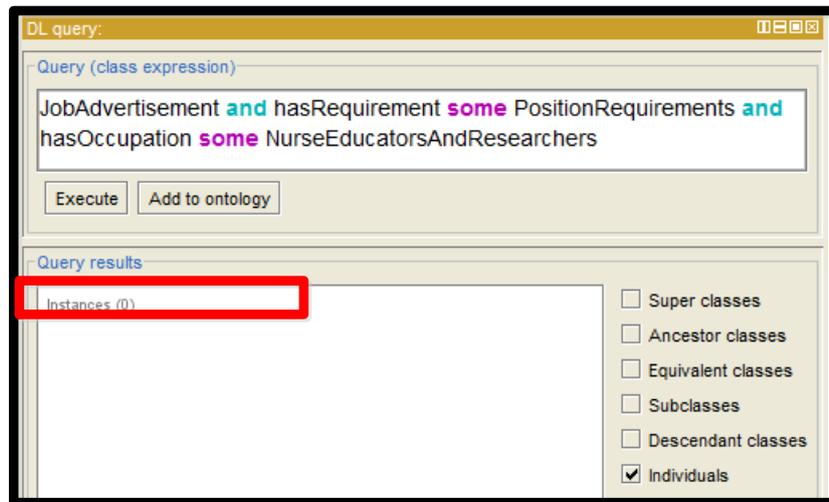


Figure 116 DL Query and result showing the number of *PositionRequirements* advertised for the occupation class *NurseEducatorsAndResearchers*.

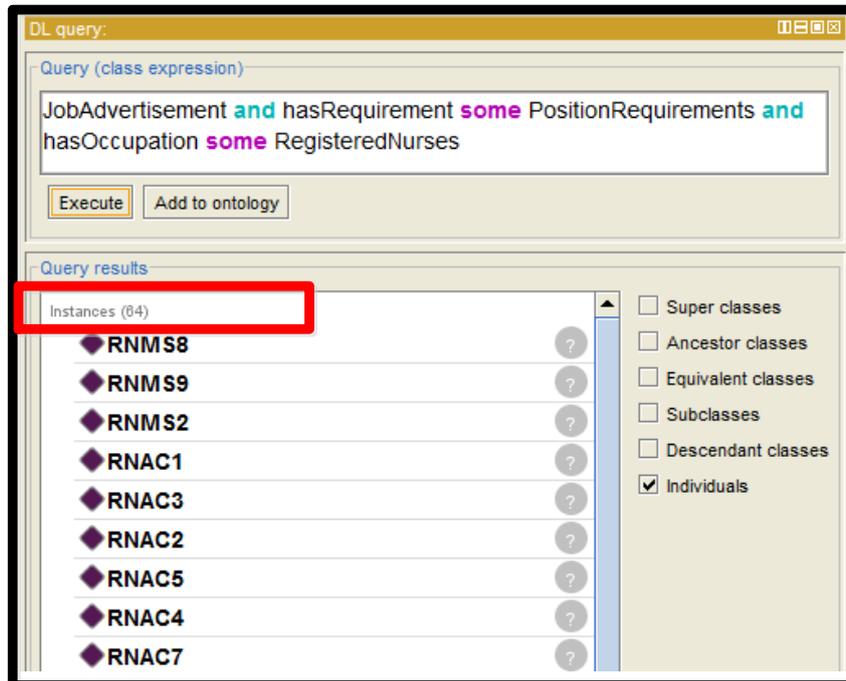


Figure 117 DL Query and result showing the number of *PositionRequirements* advertised for the occupation class *RegisteredNurses*.

The results for the four queries are as follows:

The class *Midwives* had 16 instances where *PositionRequirements* were advertised by employers.

The class *NurseManagers* had seven instances where *PositionRequirements* were advertised by employers.

The class *NurseEducatorsAndResearchers* had zero instances where *PositionRequirements* were advertised by employers.

The class *RegisteredNurses* had 64 instances where *PositionRequirements* were advertised by employers.

From the results it is evident that the class *RegisteredNurses* had the most requirements specified by employers. This is probably due to the fact that there were far more advertisements for *RegisteredNurses* recorded (67) in total, than there were for any of the other nursing occupational groups during the trial month.

## **8.4.2 Use case 2: Prospective Student Choosing Course of Study**

### **A. Scenario**

Sophi has finished her final year of school, and is about to submit her application to further her studies at one of the local universities in Perth. As her grandmother was an inspiration to Sophi all her life, Sophi has decided to follow in her footsteps and also become a Midwife who could look after those in need. Sophi has determined that two universities in Perth offer nursing as a course to study, however these universities do not provide the same areas of specialisation in nursing, and Sophi is unsure which nursing specialisation she needs to apply for (and hence which university), to pursue her goal. Sophi has had a fond relationship with her grandmother and all her grandmother's friends in the old age home where her grandmother spent her last years, but is more keen on caring for pregnant mothers and their babies than becoming a Registered Nurse in Aged Care. This has made Sophi wonder which attribute requirements

employers are looking for when recruiting Midwives, and also Nurse Managers, so she can apply for the nursing degree that best suits her personality and interests, but also will land her a job once she finishes with university.

## **B. Questions**

To assist Sophi in her quest, she would like to access EDIT to obtain the following type of information from the system:

1. What are the attribute requirements that employers are looking for when employing Midwives?
2. What are the attribute requirements that employers are looking for when employing Nurse Managers?

## **C. Results from EDO**

Each question's transition from natural language to DL query language is discussed below.

1. What are the attribute requirements that employers are looking for when employing Midwives?

The question requires two steps to be performed with the DL Query. Firstly the DL Query has to look for all *JobAdvertisement* instances advertised for *Midwives* where any type of *AttributeRequirements* was specified (Figure 118). In Figure 119 to Figure 121, the original job advertisements are displayed for the three query results returned, indicating their respective attribute requirements highlighted in yellow.

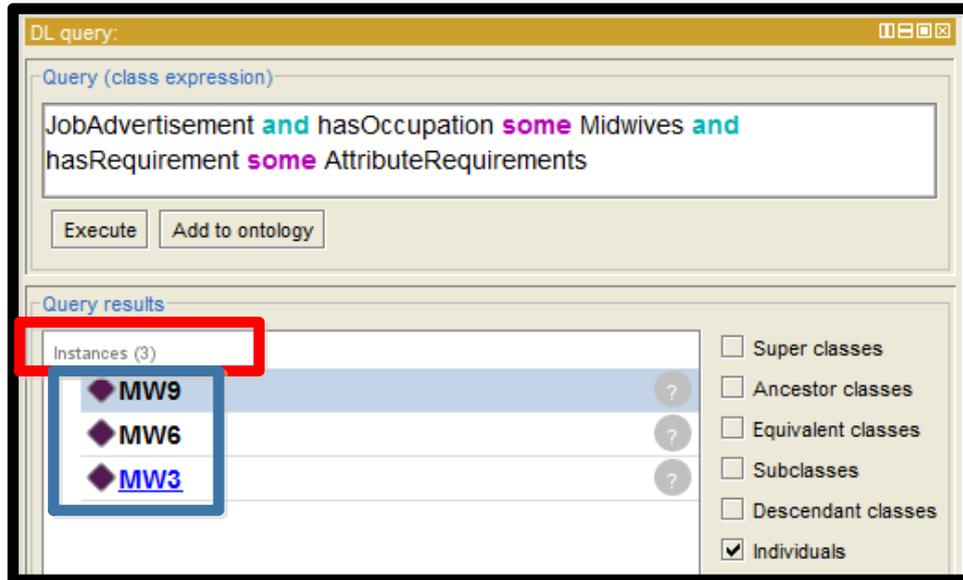


Figure 118 DL Query and result showing the *JobAdvertisement* instances that were advertised for the occupation class *Midwives* where a type of *AttributeRequirement* was specified.



Figure 119 Job advertisement *MW9* has attribute requirement *CustomerServiceCommitment*.



**Midwives for exciting positions based in Western Australia**

- Registered or eligible for Australian registration,
- Supply a National Police Clearance and
- Demonstrate knowledge of midwifery practice

**ScarletBLACK** is one of the leading providers of all grades and specialties of medical staff to both the public and private sectors.

We are inviting expression of interests from Midwives for exciting positions based in Western Australia. Our client is seeking qualified and experienced Midwives to work in their hospitals which are all modern and well equipped with supportive and dynamic working environments.

***This is an exciting opportunity to broaden your horizons!***

The Obstetrics and Gynaecology Unit offers a chance to develop new skills in midwifery practice through a well supported orientation program and an ongoing working/learning environment.

Applicants must have:

**ESSENTIAL**

- Ability to work within and contribute to the Mission/Values of the hospital.
- Registered General Nurse Division 1 endorsed as a midwife.
- Highly developed interpersonal and communication skills.
- **Demonstrated ability and commitment to maintaining confidentiality.**
- Promotes excellence in patient care.
- Demonstrated knowledge of current midwifery practices and trends, including trends in infant nutrition.
- An ability to relate in a courteous and **caring** professional manner to others, ensuring patient rights are respected.

**DESIRABLE**

- Nursing Certificate in Unit speciality. Preferably working towards a tertiary qualification in advanced clinical nursing, administration or education.
- An understanding of, and commitment to, quality improvement.

For further details please telephone Helena Kukutai on (09) 913 4635 or click on the "apply now" link and send your CV in **Word Format** quoting Reference number HK4517

***If this vacancy is not quite what you are looking for please call us to discuss your individual requirements.***

Email: Please click the 'Apply Now' button below.

**With exclusive partnerships in Australia, USA, UK, Canada, Dubai and the Middle East we are now able to offer career opportunities to healthcare professionals wanting to work overseas.**

Figure 120 Job advertisement MW6 has attribute requirements *Caring* and *ConfidentialityCommitment*.

**HAYS** Recruiting experts  
in Healthcare

**MIDWIVES**  
**MIDWIVES**

- An excellent salary and benefits package
- Great employer
- Excellent Location

Our clients, based in the land of the sun sea and sand "Australia" have vacancies arising for flexible, **enthusiastic** Midwives committed to providing a high quality standard of care. These permanent posts offer a variety of experience, both in the hospital and community settings and will involve rotation through all clinical areas, day and night duty.

They are an innovative and progressive obstetric-led unit, working towards midwifery led-care. The unit provides care for all normal and high-risk pregnancies via an integrated service, which actively encourages and supports women's choice; individualised women centred care being a high priority.

**Minimum Requirements:**

Midwifery Degree/Diploma

12 months postgrad experience

An excellent salary and benefits package are on offer to all experienced staff.

If working in Australia is something you've always dreamed of, and you are not afraid to relocate, then contact us today!

**Leona Fawl, Hays Healthcare**

**Leona.fawl@hays.com.au**

**08 92544590**

Reference Number: 771-1069876

**To apply please click the 'Apply' button below.**

**Figure 121** Job advertisement *MW3* has attribute requirement *Enthusiastic*.

The second step, requires the user to click on each of the hyperlinks provided in the query result (highlighted in blue in Figure 118), which takes the user to the screen of each job advertisement where the specific attribute requirements can be investigated (Figure 122). Job advertisement *MW6*, shows that the attribute requirements *Caring* and *ConfidentialityCommitment* were listed in the job advertisement as requirements that prospective applicants of that vacancy should have. *MW3* required prospective applicants to have *Enthusiasm*, and *MW3* required prospective applicants to have *CustomerServiceCommitment*.

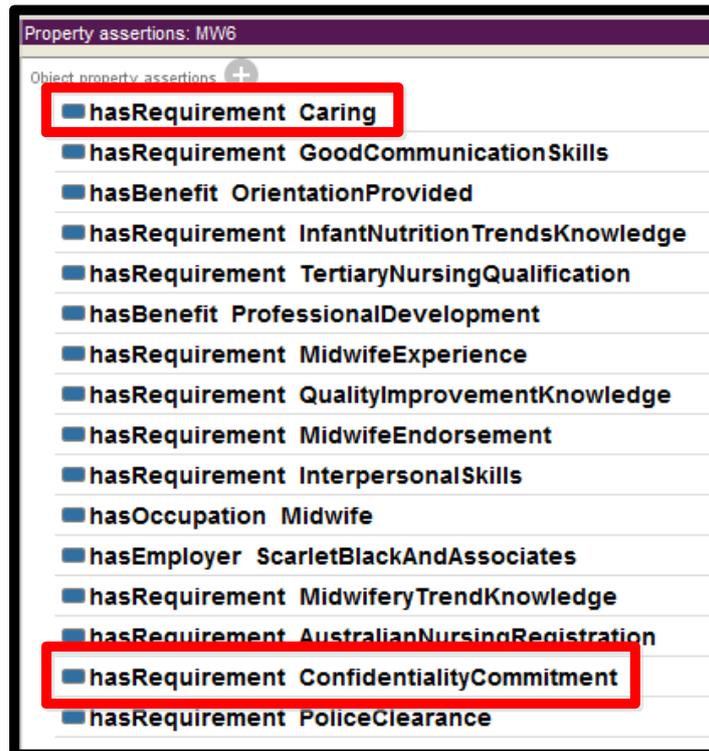


Figure 122 Detailed view of *MW6* — one of the *JobAdvertisement* instances retrieved through the query specified in the previous figure.

2. What are the attribute requirements that employers are looking for when employing Nurse Managers?

As with the previous question, this question requires two steps to be performed with the DL Query. Firstly the DL Query has to look for all *JobAdvertisement* instances advertised for *NurseManagers* where any type of *AttributeRequirements* were specified. The query and results are shown in Figure 123.

The second step requires the user to click on each of the hyperlinks provided in the query result (highlighted in blue in Figure 123) which takes the user to the screen of each job advertisement, where the specific attribute requirements can be investigated.

Job advertisement *NM5* shows that the attribute requirements *DeliveringExcellence*, *FlexibleToChangingResidentNeeds* and *Enthusiasm*, were listed in the job advertisement as

requirements prospective applicants of that vacancy should have. *NM7* required prospective applicants to have *RelateToDementia*, and *NM8* required prospective applicants to have *Proactive* and *CanDoAttitude*.

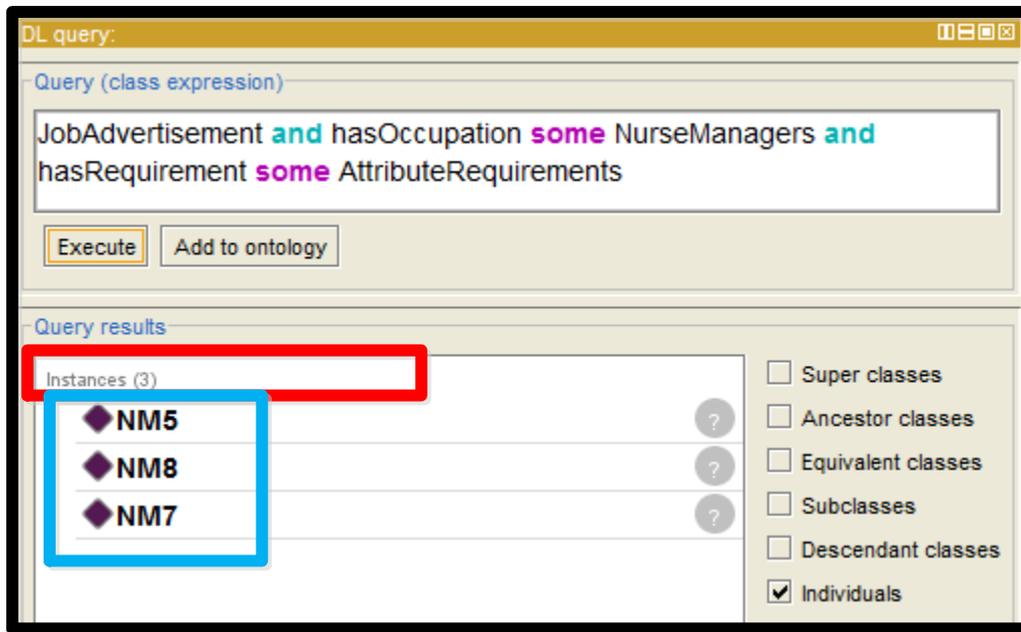


Figure 123 DL Query and result showing the *JobAdvertisement* instances that were advertised for the occupation class *NurseManagers* where a type of *AttributeRequirement* was specified.

From the above information, Sophi is able to make a comparison table (Table 13) to assist in her decision.

Midwives	Nurse Managers
Caring	Delivering Excellence
Confidentiality Commitment	Flexible to Changing Resident Needs
Enthusiasm	Enthusiasm
Customer Service Commitment	Relate to Dementia
	Proactive
	Can Do Attitude

Table 13 Attribute requirements for *Midwives* versus *NurseManagers*.

### **8.4.3 Use case 3: Human Resource Manager at a Company**

#### **A. Scenario**

Anita is a Recruitment Advisor at the New Perth Hospital, which is currently being expanded to include two brand new areas for the hospital: maternity ward and community health day visit facilities. The two new wings of the building will be completed and ready for operation in approximately six months' time, at which stage staff will be employed to cover for these new areas of operation. Anita is trying to establish what the market is offering to midwives and community health nurses in terms of employee benefits, so that she can ensure the New Perth Hospital will be competitive in the staff packages that they would like to advertise in their job advertisements, thus guaranteeing they get a sufficient number of applications for the many positions she needs to fill. At this point in time, Anita does not have the time to analyse each job advertisement one-by-one for the last month, to obtain this type of information from the job advertisements' free text sections.

#### **B. Questions**

To assist Anita in her quest, she would like to access the EDIT to obtain the following type of information from the system:

1. What type of employee benefits have been offered to Midwives and Nurse Professionals in the Perth metropolitan area?
2. How many employers are prepared to offer Visa Sponsorships to their prospective candidates?

#### **C. Results from EDO**

Each question's transition from natural language to DL query language is discussed below.

1. What type of employee benefits have been offered to Midwives and Nursing Professionals in the Perth metropolitan area?

To find a few examples of employee benefits that have been advertised in Perth, it is necessary for the DL Query to retrieve all job advertisements where the location has been included as a MajorCityOfAustralia. Figure 124 shows the query and results as a total of 54 job advertisements which included employee benefits.

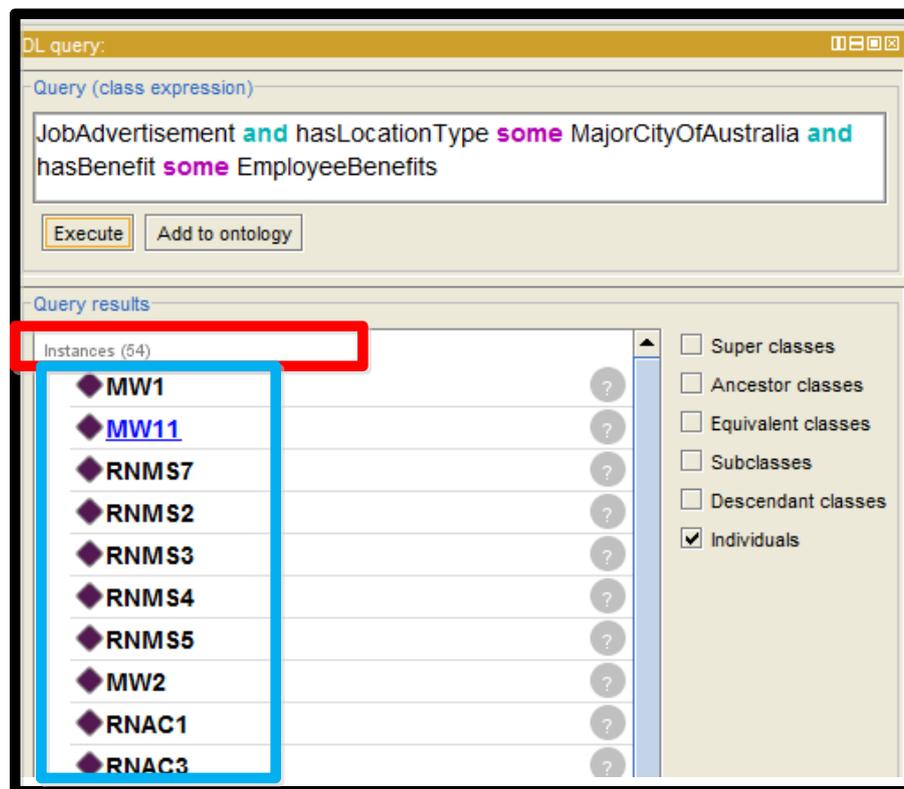
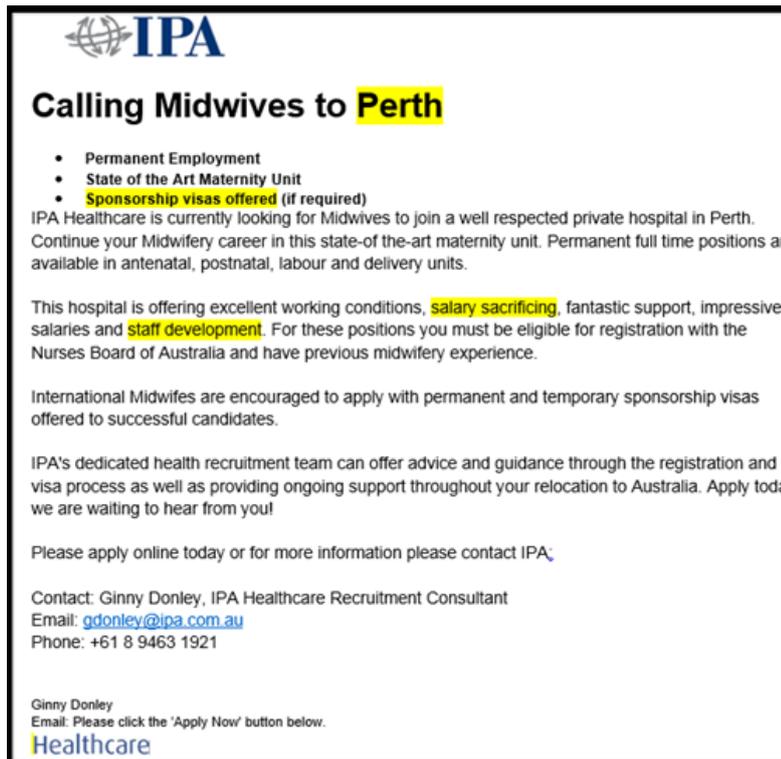


Figure 124 DL Query and result showing the *JobAdvertisement* instances that were advertised for all nursing occupation classes where a type of *AttributeRequirement* was specified.

Following the links to the first five job advertisements listed in the results window (highlighted in blue in Figure 124) provided the following examples of employee benefits on offer from employers (displayed in Figure 125 to Figure 129).



**IPA**

## Calling Midwives to Perth

- Permanent Employment
- State of the Art Maternity Unit
- Sponsorship visas offered (if required)

IPA Healthcare is currently looking for Midwives to join a well respected private hospital in Perth. Continue your Midwifery career in this state-of-the-art maternity unit. Permanent full time positions are available in antenatal, postnatal, labour and delivery units.

This hospital is offering excellent working conditions, salary sacrificing, fantastic support, impressive salaries and staff development. For these positions you must be eligible for registration with the Nurses Board of Australia and have previous midwifery experience.

International Midwives are encouraged to apply with permanent and temporary sponsorship visas offered to successful candidates.

IPA's dedicated health recruitment team can offer advice and guidance through the registration and visa process as well as providing ongoing support throughout your relocation to Australia. Apply today we are waiting to hear from you!

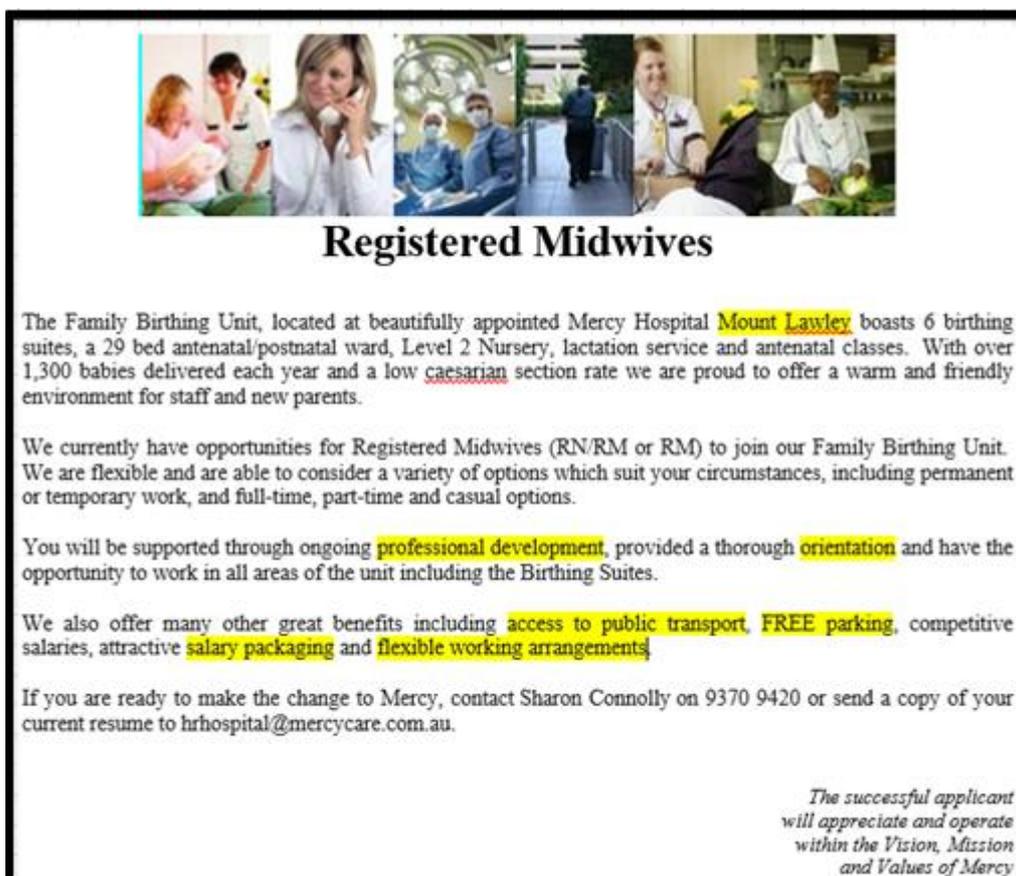
Please apply online today or for more information please contact IPA.

Contact: Ginny Donley, IPA Healthcare Recruitment Consultant  
 Email: [gdonley@ipa.com.au](mailto:gdonley@ipa.com.au)  
 Phone: +61 8 9463 1921

Ginny Donley  
 Email: Please click the 'Apply Now' button below.

**Healthcare**

Figure 125 Job advertisement MW1 with location Perth, and benefits on offer SalaryPackaging, SponsorshipVisaAvailable and ProfessionalDevelopment.



## Registered Midwives

The Family Birthing Unit, located at beautifully appointed Mercy Hospital Mount Lawley boasts 6 birthing suites, a 29 bed antenatal/postnatal ward, Level 2 Nursery, lactation service and antenatal classes. With over 1,300 babies delivered each year and a low caesarian section rate we are proud to offer a warm and friendly environment for staff and new parents.

We currently have opportunities for Registered Midwives (RN/RM or RM) to join our Family Birthing Unit. We are flexible and are able to consider a variety of options which suit your circumstances, including permanent or temporary work, and full-time, part-time and casual options.

You will be supported through ongoing professional development, provided a thorough orientation and have the opportunity to work in all areas of the unit including the Birthing Suites.

We also offer many other great benefits including access to public transport, FREE parking, competitive salaries, attractive salary packaging and flexible working arrangements.

If you are ready to make the change to Mercy, contact Sharon Connolly on 9370 9420 or send a copy of your current resume to [hrhospital@mercy.com.au](mailto:hrhospital@mercy.com.au).

*The successful applicant will appreciate and operate within the Vision, Mission and Values of Mercy*

Figure 126 Job advertisement MW11 in Mount Lawley (a suburb of Perth), and benefits offered OrientationProvided, ProfessionalDevelopment, FreeCarParking, SalaryPackaging, CloseToPublicTransport and FlexibleWorkPractices.

**Registered and Enrolled Nurses - It's time to reflect!**

- Award Winning Company - Best Employer 2010
- Shifts Available for You Now!

CPE Group is a recruitment agency who genuinely believes in providing recruitment and staffing solutions with care. Recently **Quality Assured** and awarded **2010 Best Employer**, we understand the importance of providing our team members with a great employment experience in a highly values based business. That's right - you are important to us.

With a reputation for quality Nurses, our business is growing significantly and so we need talented Registered Nurses and Midwives to join our team of caring professionals. Shifts are available in Public and Private Hospitals, as well as Aged Care in all **Perth** areas.

As a CPE Group employee you will feel a valued member of our team and can access:

- Shifts **Close to Home**
- Qualification Allowances\*
- Competitive Pay Rates
- **Weekly Pay Cycle**
- **Friendly WA Call Centre Team**
- **Free Uniform**

To be eligible to join our team you must be an RN 1.2 and above, and possess:

- Mandatory Checks
  - Manual Handling
  - Working with Children Check
  - National Police Clearance
  - Department of Health Criminal Record Screen
  - Fire Safety and Evacuation
  - Managing Aggressive Behaviour
  - Mandatory Reporting of Child Sex Abuse
  - CPR
- Valid National Registration

You could be working for CPE Group within a week of application subject to presentation of qualifications and checks together with attendance at our interviews and induction - make sure your diary is free!

So if you are a talented Nurse who genuinely believes in quality nursing and care, you'll fit right in to our award winning team so apply now - *we really look forward having you work with us.*

Queries to Emily Peters on 1300 665 082

Figure 127 Job advertisement *RNMS7* in *Perth*, offering benefits *WACallCentreTeam*, *UniformProvided*, *WeeklyPayCycles* and *ShiftsCloseToHome*.

 **IPA**

**Calling ALL CN Medical / Surgical Nurses**

- Permanent Employment
- State of the Private Hospital
- **Sponsorship visas offered** (if required)

IPA Healthcare is currently looking for Clinical Medical & Surgical Registered Nurse to join a well respected private hospital in **Perth**. Continue your clinical career in this state-of-the-art hospital. Permanent full time positions are available in medical & surgical wards.

This hospital is offering excellent working conditions, **salary sacrificing**, fantastic support, impressive salaries and **staff development**. For these positions you must be eligible for registration with the Nurses Board of Australia and have previous Clinical Nurse level experience.

International nurses are encouraged to apply with permanent and temporary sponsorship visas offered to successful candidates.

IPA's dedicated health recruitment team can offer advice and guidance through the registration and visa process as well as providing ongoing support throughout your relocation to Australia. Apply today we are waiting to hear from you!

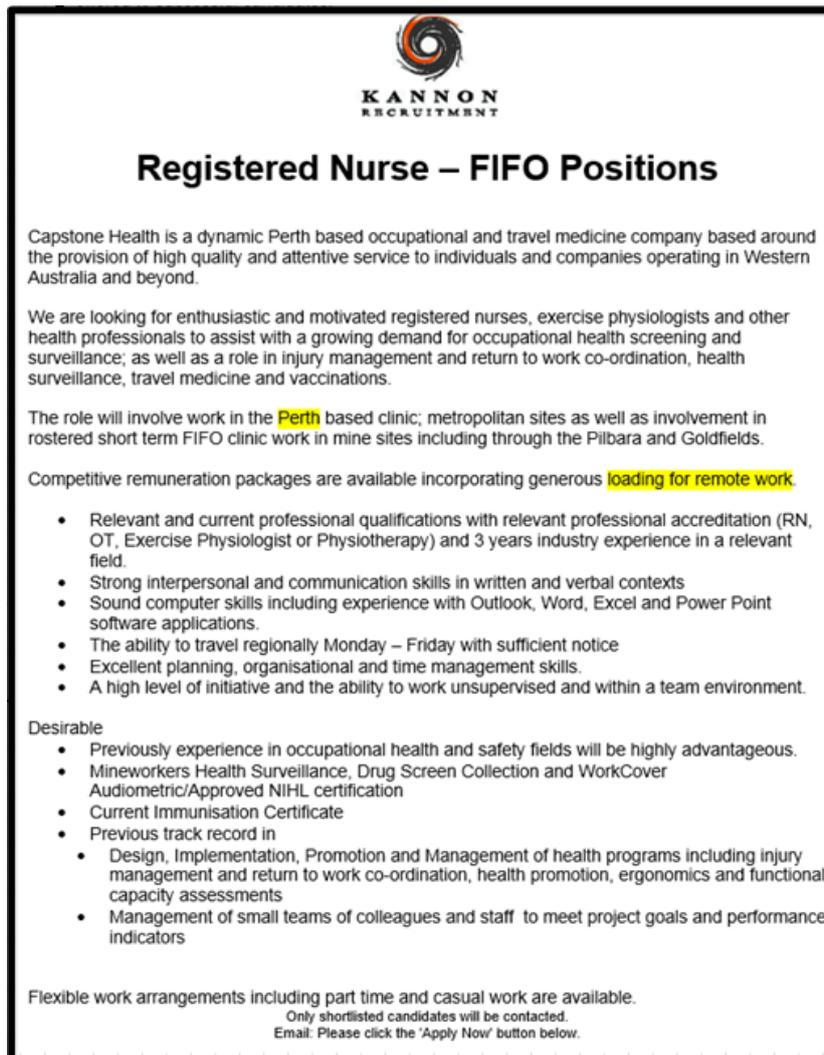
Please apply online today or for more information please contact IPA;

Contact: Ginny Donley, IPA Healthcare Recruitment Consultant  
 Email: [gdonley@ipa.com.au](mailto:gdonley@ipa.com.au)  
 Phone: +61 8 9463 1921

Ginny Donley  
 Email: Please click the 'Apply Now' button below.

**Healthcare**

Figure 128 Job advertisement *RNMS2* in *Perth*, offering benefits *SalaryPackaging*, *ProfessionalDevelopment* and *SponsorshipVisaAvailable*.



  
**KANNON**  
RECRUITMENT

## Registered Nurse – FIFO Positions

Capstone Health is a dynamic Perth based occupational and travel medicine company based around the provision of high quality and attentive service to individuals and companies operating in Western Australia and beyond.

We are looking for enthusiastic and motivated registered nurses, exercise physiologists and other health professionals to assist with a growing demand for occupational health screening and surveillance; as well as a role in injury management and return to work co-ordination, health surveillance, travel medicine and vaccinations.

The role will involve work in the **Perth** based clinic; metropolitan sites as well as involvement in rostered short term FIFO clinic work in mine sites including through the Pilbara and Goldfields.

Competitive remuneration packages are available incorporating generous **loading for remote work**.

- Relevant and current professional qualifications with relevant professional accreditation (RN, OT, Exercise Physiologist or Physiotherapy) and 3 years industry experience in a relevant field.
- Strong interpersonal and communication skills in written and verbal contexts
- Sound computer skills including experience with Outlook, Word, Excel and Power Point software applications.
- The ability to travel regionally Monday – Friday with sufficient notice
- Excellent planning, organisational and time management skills.
- A high level of initiative and the ability to work unsupervised and within a team environment.

**Desirable**

- Previously experience in occupational health and safety fields will be highly advantageous.
- Mineworkers Health Surveillance, Drug Screen Collection and WorkCover Audiometric/Approved NIHL certification
- Current Immunisation Certificate
- Previous track record in
  - Design, Implementation, Promotion and Management of health programs including injury management and return to work co-ordination, health promotion, ergonomics and functional capacity assessments
  - Management of small teams of colleagues and staff to meet project goals and performance indicators

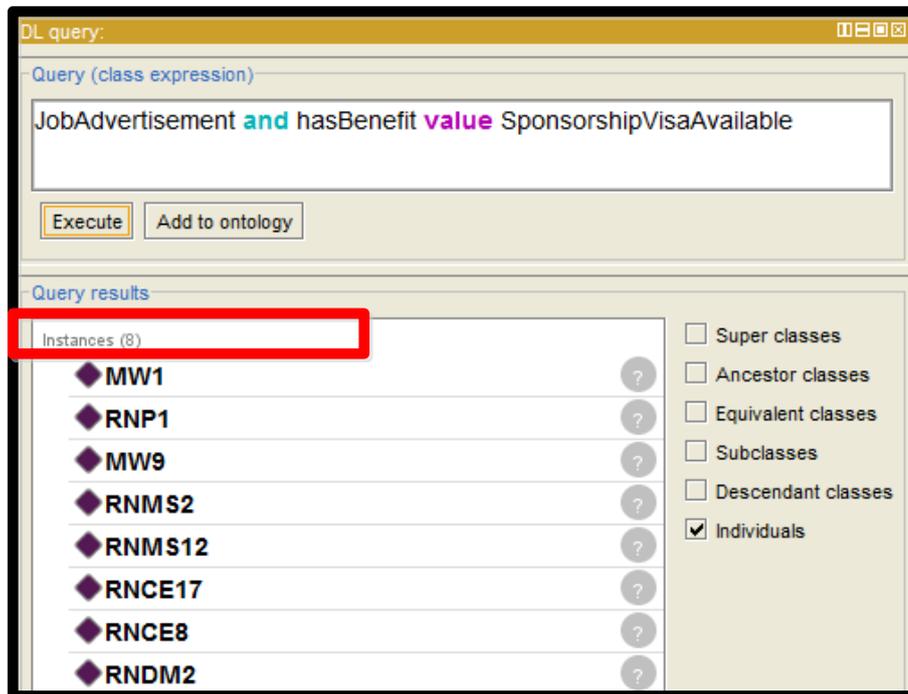
Flexible work arrangements including part time and casual work are available.  
Only shortlisted candidates will be contacted.  
Email: Please click the 'Apply Now' button below.

**Figure 129** Job advertisement *RNMS3* in *Perth*, with benefit offered *RemoteWorkAllowance*.

It is clear to Anita from a quick analysis of the first five job advertisements that many employers offer *SponsorshipVisaAvailable*, *ProfessionalDevelopment* and *SalaryPackaging* and as such will consider offering these benefits to their future employees, as the minimum to be able to attract top calibre applicants.

2. How many employers are prepared to offer Visa Sponsorships to their prospective candidates?

This question requires the DL Query to search for all instances where sponsorship visas have been offered as part of the position package (Figure 130).



**Figure 130 DL Query and result showing the *JobAdvertisement* instances that specified the employee benefit *SponsorshipVisaAvailable*.**

The query result indicates to Anita that a total of eight job advertisements included the benefit of visa sponsorships for nurses who would want to come and work in Australia from other countries.

#### 8.4.4 Use case 4: Curriculum and Career Developer at a Tertiary Institution

##### A. Scenario

Carl is a senior lecturer in Surgical Practices 410 in the Faculty of Health at the University of Perth. As a senior lecturer, one of Carl's responsibilities is to ensure that his unit's teaching content is relevant to what employers are currently looking for, in terms of graduate skill capabilities. He also needs to ensure that students are accurately informed of all the qualifications that they are expected to have once they have completed their nursing degrees at the University of Perth. Before students are able to graduate from the University of Perth's Nursing degree, they are sent for in-house practical training as part of the university's agreement with the New Perth Hospital. During their practical training period, students are able to attend

additional courses at the New Perth Hospital. This is part of the university's commitment to help students land a job at their employer of choice as soon as they have completed their practical work. The students are keen to know which employers of the top three hospitals in Perth were the largest advertiser of vacancies, so that they can direct their job seeking efforts towards those institutions.

## **B. Questions**

To assist Carl in his quest, he would like to access EDIF to obtain the following type of information from the system:

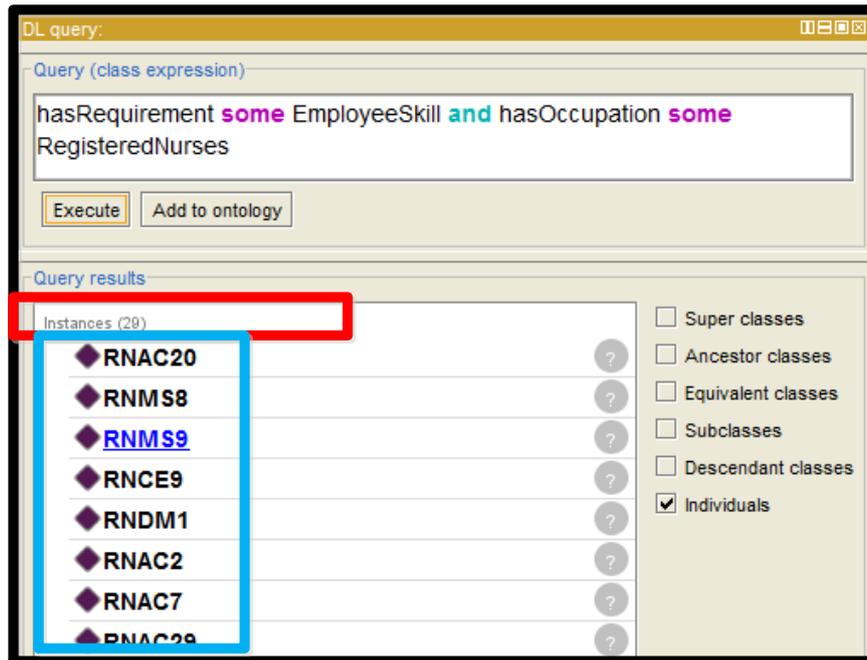
1. What types of skill requirements are employers indicating they need when employing Registered Nurses?
2. Which of the top three hospitals in Perth was the major employer during January 2011?

## **C. Results from EDO**

Each question's transition from natural language to DL query language is discussed below.

1. What types of skill requirements are employers indicating they need when employing Registered Nurses?

To find a few examples of employee skill requirements that have been advertised for Registered Nurses, it is necessary for the DL Query to retrieve all job advertisement instances in which employee skills were listed as a requirement for applicants. Figure 131 shows the query and results as a total of 29 job advertisements, which included some type of employee skills as a requirement.



**Figure 131 DL Query and result showing some of the *JobAdvertisement* instances that specified employee skills as a requirement.**

As with some of the previous examples, this query involves two steps. As such it is also necessary to follow the links to the identified job advertisements to see which employee skill requirements were listed in each of the relevant job advertisements.

Instead of providing the Protégé screenshots of these advertisements, a random five original advertisements of the above listed results have been provided below in Figure 132 to Figure 136, as examples of employee skill requirements from employers.

**Registered Nurse/ Aged Care**

Meath Care Inc. is an Aged Care provider of Residential Accommodation and Independent Living Units.

Our Kingsley Facility is a 125 bed Aged Care Facility and provides accommodation for, low, high and dementia secure services for the frail aged. The building provides an excellent home like environment for residents and a supportive work culture for staff.

We require a Registered Nurse who has a passion for caring for elderly and **can lead a small team** of dedicated care workers.

The hours are 07:00am-14:30pm Monday to Friday.

Criteria Required:

- Current Registration as a General Nurse with the Nurse Board of Western Australia.
- Current police clearance is an essential requirement of the above position.
- Previous experience of working in Aged Care with knowledge and application of Accreditation Standards in Aged Care.
- Knowledge of ACFI processes and relevant documentation.
- Proven clinical knowledge and expertise.
- Effective communication skills** essential.

Applicants should apply in writing to Carol McCourt, Facility Manager Meath Care Kingsley 18 Hocking Road, Kingsley WA 6026 or email to Carol.McCourt@meath.org.au

Figure 132 Job advertisement *RNAC20* requires *GoodCommunicationSkills* and *TeamLeadingSkills*.



**Registered Nurse**

- Full salary packaging benefits
- Working with an enthusiastic team
- Casual Position

Perth Primary Care Network (PPCN) is a not-for-profit organisation that focuses on delivering health outcomes to local communities by supporting General Practice and the broader primary health care sector.

PPCN operates a Street Doctor Mobile Medical Service that takes primary healthcare to the marginalized communities of Perth in their own environment. This service is now seeking a Registered Nurse to join their multidisciplinary team.

This position involves triaging patients and providing medical services on behalf of the General Practitioner, such as, providing first aid assistance, assessment, therapeutic care and treatment, diagnostic services, clinical data management and pathology collection. This is a casual position with various shifts available throughout the week.

We are looking for an individual with **great people skills** and who has a non judgmental approach. The successful applicant must be a Registered Nurse with previous experience within a similar role, be reliable, have the **ability to work well within a team environment**, be **computer literate**, and have a current manual driver's license, Police Clearance Check and a Working with Children Check. Experience in General Practice nursing, Mental Health nursing and drug and alcohol experience are all desirable.

Benefits of this position include working with an enthusiastic team and full salary packaging benefits, with full details available at the interview stage.

For any further queries please contact the Human Resources Coordinator, on 9376 9200 or send your resume and covering letter to [hr@ppcn.org.au](mailto:hr@ppcn.org.au) or PO Box 354 Guildford WA 6935.  
Email: Please click the 'Apply Now' button below.

Figure 133 Job advertisement *RNMS8* requires *TeamPlayerSkills*, *ComputerSkills* and *InterpersonalSkills*.

**Practice Nurse (RN) - Part Time**  
Gosnells Family Practice is looking for a experienced motivated part time practice nurse to work 3 days per week.

Demonstrating excellent **communications skills** and willing to work as a **team** and **independently**.

Gosnells Family Practice is fully computerized practice with 1 full time Doctor and 7 part time doctors.

Skills required for this role:

- \*Experience in childhood immunisation
- \*Ability to **assist GP's with minor procedure**.
- \*Experience in **wound management & dressings**.
- \*Experience in performing ECG's.
- \* **Maintain recalls and test results**.
- \*Experience with care plans.

Remuneration in accordance with experience and qualifications.

If you would like to join Gosnells Family Practice as a valued team member please email your application with cover letter and CV to:  
marta.bosnjak@medicalcentral.com.au.

Marta Bosnjak  
Practice Manager  
Gosnells Family Practice  
3/2194 Albany Hwy  
Gosnells WA 6164  
08 9398 8488  
marta.bosnjak@medicalcentral.com.au

**Figure 134** Job advertisement *RNMS9* with skills *TeamPlayerSkills*, *RecallAndTestResultMaintenanceSkills*, *WoundCareSkills*, *WorkUnsupervisedSkills*, *GoodCommunicationSkills* and *ProcedureAssistanceSkills*.



**scarletBLACK**  
and associates  
SPECIALISED HEALTH RECRUITMENT

## Emergency Department Nurses for Australia

- Fantastic opportunities
- New experiences
- Great remuneration packages

ScarletBLACK is one of the leading providers of all grades and specialties of medical staff to both the public and private sectors.

We are currently recruiting for Registered Nurses with experience in Emergency Department for our client based in Perth – Western Australia

ED Nurse's if you're looking for exciting new job opportunities in a Fantastic location, spending your off work time relaxing on the beach or lighting up the Barbie, then this could be your dream come true!

All of our hospitals encourage staff to maintain a healthy work lifestyle balance & offer many staff social and support programs as well as encouraging career development.

### Benefits

- \* Travel support
- \* Friendly staff
- \* Employee Assistance Programs
- \* Supported education programs
- \* Training and Development programs
- \* First-rate career opportunities

Applicants must be able to demonstrate **effective communication skills** and contemporary nursing practice. Recent emergency department experience is desirable.

Do you have the desire to work in a country that offers fantastic opportunities and rewards?

For further details please telephone Helena Kukutai, on (09) 913 4635 or click on the "apply now" link and send your CV in Word Format with Reference number HK4597  
*If this vacancy is not quite what you are looking for please call us to discuss your individual requirements.*

Email: Please click the 'Apply Now' button below.

With exclusive partnerships in Australia, USA, UK, Canada, Dubai and the Middle East we are now able to offer career opportunities to healthcare professionals wanting to work overseas.

Figure 135 Job advertisement RNCE9 with skill requirement *GoodCommunicationSkills*.

## Community Mental Health Nurse

As a member of the community mental health crisis team you will be delivering a high standard of mental health and crisis care in the community.

Being part of the multidisciplinary team you will be providing...

- Crisis assessments during a client's acute episode.
- Provide specialist assessment, triage and treatment in the emergency departments covering all age groups from child adolescents through to aged care.
- Facilitation of transfer from Emergency care to the Mental Health inpatient unit.
- Undertake Medication administration.
- Be involved in daily hand over of patient care plans.

Essential: 2 years post graduate experience in community mental Health, current Australia registration, **ability to work within a team** and also **autonomously**.

**Start date:** Open start date.  
**Position Type:** This is a permanent position contract also available for the right applicant.  
**How to apply:** Submit your C.V now to find out how you can secure a great opportunity.

APPLY TO WORK WITH SUGARMAN TODAY.  
PLEASE SEND YOUR C.V TO  
leigh@sugarmangroup.com.au  
or call us on  
02 9549 5700

Figure 136 Job advertisement *RNDM1* requires *WorkUnsupervisedSkills* and *TeamPlayerSkills*.

It is clear to Carl from a quick analysis of the first five job advertisements that employers seem to regularly require applicants to have *GoodCommunicationSkills*, *TeamPlayerSkills* and *WorkUnsupervisedSkills*. However there are also many other skills required, and a broader analysis of all skill requirements will involve Carl to take a closer look at the other remaining retrieved 24 job advertisements as well.

2. Which of the top three hospitals in Perth was the major employer during January 2011?

This question requires three DL Queries to be run — one for each top hospital in Perth (Figure 137 to Figure 139).

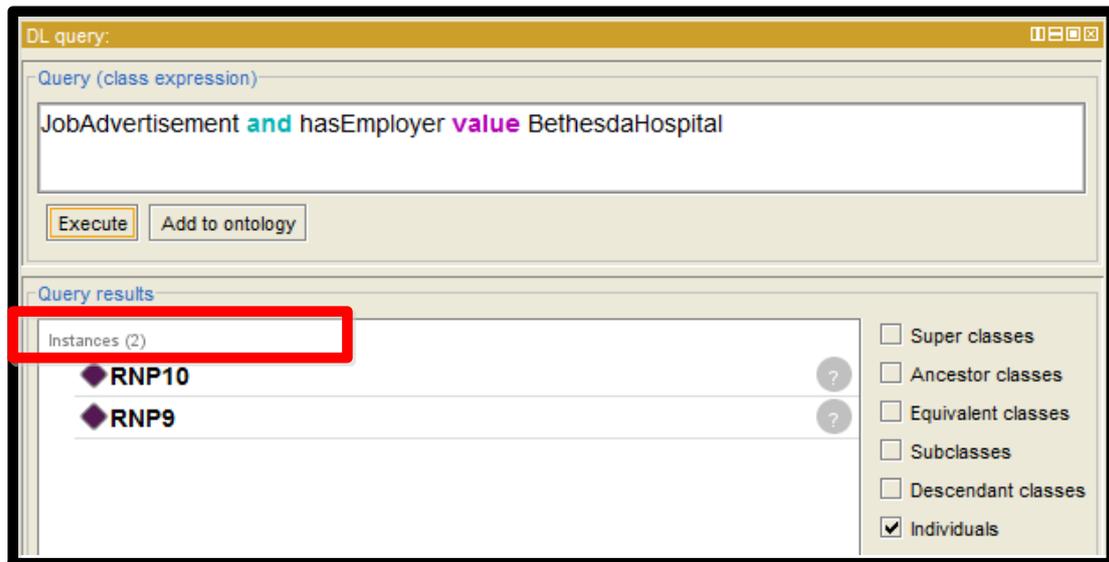


Figure 137 DL Query and result showing the number of job advertisements the employer *BethesdaHospital* advertised during January 2011.

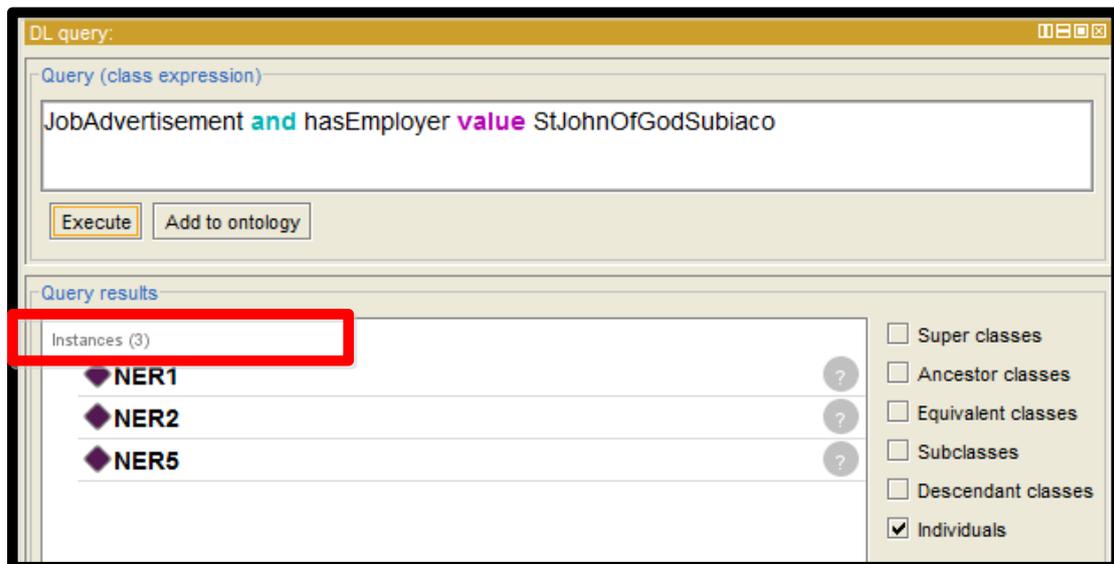
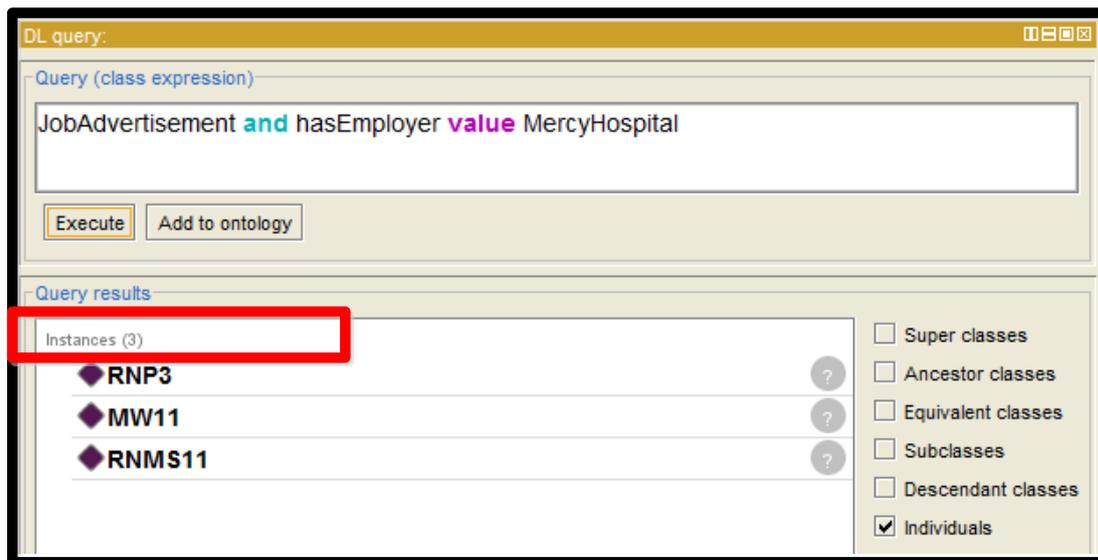


Figure 138 DL Query and result showing the number of job advertisements the employer *StJohnOfGodSubiaco* advertised during January 2011.



**Figure 139** DL Query and result showing the number of job advertisements the employer *MercyHospital* advertised during January 2011.

From the query results it would seem that both *MercyHospital* and *StJohnOfGodSubiaco* advertised three positions during January 2011. *BethesdaHospital* only advertised for two positions during this time.

#### 8.4.5 An Integrated View of Use Case Needs

Based on the range of use cases identified for EDO, a consolidated list of use case needs questions was created around centralised ideas. Some of these integrated questions are provided below, and follow a similar hierarchy on which EDO was modelled. This list is not exhaustive and merely provides a cross-section of some of the possible natural language use questions that EDO could be queried for.

##### 1. Skill shortages

1. Which occupation codes have been identified as potential skill shortages this week/month/year?
2. Which location is the most affected by skill shortages overall this week?
3. For what period of time were skill shortage occupations generally advertised for this week?

## **2. Types of requirements**

### **a. Skills**

1. Which competencies do employers expect future employees for occupation X<sup>42</sup> to have?
2. How many years' experience do employers expect future employees for occupations X to have?
3. Which qualifications do employers expect future employees for occupation X to have?
4. What kind of licence or registration do employers expect future employees for occupation X to have?
5. What kind of knowledge do employers expect future employees for occupation X to have?

### **b. Attributes**

1. What kind of attributes do employers expect future employees for occupation X to have?

## **3. Location**

1. How many positions were advertised for vacancies in regional areas this week?
2. Which regional and rural areas have had the least number of job advertisements this week?
3. What are the job advertisement statistics for regional vs. rural vs. metropolitan for this week?
4. Were there certain occupations that did not have job advertisements for specific locations this week?
5. Which area had the least advertisements for occupation X this week?

## **4. Work conditions**

1. What position bases exist for vacancies of occupation X?

---

<sup>42</sup> Occupation X refers to all occupation types in the ANZSCO (approximately 520 classifications)

2. Which types of work hours are usually on offer for a vacancy for occupation X?

## **5. Occupational clusters**

1. Which occupational type had the most advertisements this week?
2. Which occupational type had the least advertisements this week?

## **6. Interesting facts**

3. Which job was listed for the shortest period of time this week?
4. Which company posted the most job advertisements this week?
5. Which occupation has had the most advertisements this week?
6. Which state has had the most advertisements this week?
7. Which occupation generally has the most requirements?
8. Which occupation generally has the least requirements?

## **8.5 Conclusion**

In previous chapters, the first and middle phases of the ontology lifecycle were discussed. This chapter has looked at the final stage of the ontology lifecycle — verification and validation of the ontology. Even though there are several types of evaluation techniques in use today, the eventual decision as to which technique to use depends on the stage the ontology development is at, the purpose it is to be used for, and the aspect of the ontology that needs to be evaluated.

EDO evaluation used included methods to check the ontology's engineering soundness as well as its user adaptability and usefulness. The verification was done with reasoners found in the Protégé software tool, and its validation done through the implementation of four significant use cases.

Through the implementation of these evaluation techniques, EDO was found to fulfill all the requirements for the ontology specification phase to be ready for user consumption; the

ontology also met all the fictional use case criteria by successfully running a variety of queries through EDO. Validation through actual industry use case testing has been listed as a future project for this research.

The next and final chapter of the thesis recapitulates the work done throughout this research, and provides an overview of the future work planned for EDIF.

# Chapter 9 - Conclusion and Future Work

## 9.1 Introduction

This chapter provides a restatement of the research that was carried out for this PhD, with an emphasis on how the proposed solutions have addressed the research objectives that were described in Chapters One and Three. This is supported by a discussion of the research's significance. The chapter concludes with a section on future work planned for the research.

## 9.2 Recapitulation of the Thesis

In **Chapter One**, a preliminary insight into the field of Employer Demand Intelligence is provided, highlighting what the importance of employer demand intelligence is by giving the main uses of the data.

The motivation for the study has been an awareness of the ambiguity that exists around employer demand intelligence terminology and the lack of knowledge sharing amongst stakeholders in the domain. A review of current methods shows that they have too narrow a scope to cater for detailed employer demand intelligence needs, and the narrow scope and outdated information are mostly attributed to the fact that no automated system yet exists that can gather the information with ICT supported efforts. It is clear that if such a system is to be developed, it will have to be extensively tested to ensure that it will not only be used for future academic research endeavours, but also be usable by the employer demand intelligence stakeholders.

The chapter describes the objectives and scope of the research, which is to propose a conceptual framework based on an underlying knowledge representation that will address the issues identified above, and develop an Employer Demand Intelligence Framework (EDIF) that will

semi-automatically gather data from the internet, and populate this data into the proposed underlying knowledge representation basis.

The chapter concludes by providing the significance of the research and a detailed layout of how the thesis has been structured to assist the reader in the navigations through the text.

In **Chapter Two**, the current state of research in the employer demand intelligence domain is presented. Through a review of the literature, several approaches by which employer demand intelligence are identified are categorised as follows:

1. Survey based approaches
2. Manual job advertisement analysis approaches
3. Online based approaches
4. Structured approaches

In total, 21 employer demand intelligence-related studies are discussed in the literature review. After each discussion of an approach category, the issues associated with each category are deliberated. The chapter concludes by delivering a critical evaluation of the approaches in an integrated view.

In **Chapter Three**, a list of key concepts and their definitions as they relate to this research is provided, after which the problems that are currently facing the employer demand intelligence domain are discussed in detail.

The chapter then analyses the current practical problems by summarising the research issues underlying these problems in the employer demand domain as: i) no underlying knowledge representation exists; ii) current employer demand datasets do not have occupation-specific information; iii) research in the domain is being performed through manual and time-intensive approaches; and iv) a prototype system will have to be sufficiently tested for industry readiness and not only for further academic development.

The chapter presents the thesis goal and the research questions addressed:

1. How can employer demand concepts be modelled in order to represent the employer demand domain?
2. How can instances be populated that are occupation specific, linking them to a wide range of other employer demand intelligence variables?
3. How can a semi-automatic system be developed that gathers, analyses and reports on employer demand intelligence? How can an Employer Demand Intelligence Framework be evaluated for industry readiness, and to serve as the foundation for future employer demand intelligence expansion?

The chapter concludes with an overview of research methods, and a justification for the use of the science- and engineering-based method for this research, which is broadly categorised into conceptual, perceptual, and practical stages.

In **Chapter Four**, an overview is provided of the required solutions to the problems defined in Chapter Three. An enormous amount of data faces employer demand intelligence stakeholders when analysing the domain, and this data is increasing significantly every day. Ontology, proposed as the foundation for the solutions that address the research questions, is defined, and the different types of ontologies and characteristics are explained.

The backbone of the research — the conceptual framework — is comprehensively explained in this chapter, and each solution requirement is described in detail.

In **Chapter Five**, the biggest chapter of the thesis, the development of Employer Demand Ontology (EDO) is discussed in detail, including its metrics. The chapter looks at the fundamentals of ontology methodology, and describes and justifies the choice of NeOn methodology for developing EDO in this research.

The focus then turns to the modelling and design of EDO, where the EDO notation section highlights the various uses of graphical options used to describe and substantiate the differing aspects of the ontology in the text, and elaborates on the design decisions for construction of EDO.

The remainder of Chapter Five is devoted to particularising all the classes that EDO has under its two-tiered structure. This chapter addresses research issue one.

In **Chapter Six**, the manual population of the EDO, done from scratch, is detailed in depth. EDO was instantiated with records from a month's worth of job advertisements collected from the job board SEEK, as well as with existing geographical, language and occupational title hierarchies that are the accepted industry standards for those specific classes. Protégé,<sup>43</sup> an open source Java editing and knowledge acquisition software tool, was used for the development of EDO.

At the end of EDO's development and one month's trial population, EDO consisted of more than 5,000 ontological concepts. This chapter addresses research issue two.

In **Chapter Seven**, the development and workings of the Employer Demand Intelligence Framework (EDIF) as a whole, and more specifically, the Employer Demand Intelligence Tool (EDIT) as a sub-component of EDIF, are provided. The development of EDIT allowed for EDIF to become semi-automatic at run time — metrics were provided to indicate development time and metrics for the Employer Demand Data created by EDIT. The only aspect of EDIT that is yet to be fully automated is the prompting of EDIF to collect the data from the relevant sources online.

An example of a job advertisement that flows through EDIF from start to finish is provided to illustrate the workings of EDIT specifically, and EDIF more generally. EDIT was validated through

---

<sup>43</sup> <http://protege.stanford.edu/>

a range of SPARQL queries to check that the semi-automised framework functions well and is ready for industry uptake.

The chapter concludes with a discussion of the limitations of EDIF as it currently stands, and addresses research issue three.

**Chapter Eight** presents an evaluation of the various components of EDIF. After an overview of the fundamentals of verification and validation, EDIF was evaluated for the following aspects:

1. EDO was verified to check that it was logically consistent;
2. EDO was also validated through the practical application of four contemporary use cases to simulate how industry would use EDIF.

**Chapter Nine** is the current and final chapter of the thesis.

### **9.3 Significance of the Research**

The development of the Employer Demand Ontology (EDO) and the Employer Demand Intelligence Framework (EDIF) as a whole, has significance for both the employer demand and ontology domains. As is evident from the literature review done in Chapter Two, no appropriate framework, ontology or tool yet exists to meet the needs of detailed employer demand intelligence requirements; in the light of this deficiency, the importance of this PhD research is elaborated below under two sections: scientific (or theoretical) significance, and social (or practical) significance.

#### **9.3.1 Scientific Significance**

- At the time of writing this thesis, EDO is the first ontological framework developed globally that is explicitly for the employer demand intelligence domain. EDO provides a formal representational framework for employer demand concepts and relationships that can be

used to link, integrate and represent the muddled and disconnected knowledge of employer demand in a systematic and meaningful way.

- EDO has the ability to deliver an overview of different employer demand concepts and experimental findings so that these concepts and findings can be categorised accordingly, and compared with each other. As a result, the unified structure of EDO allows for potential previous undiscovered relationships, among the different concepts, to emerge and serve as motivation for employer demand researchers and stakeholders and allow further examination into these areas.
- The development of an explicit, common and formal vocabulary framework for the employer demand domain, enables interoperability between employer demand research teams from a variety of industry stakeholders, thus allowing for more competent retrieval and analysis of information, saving scarce research time and money.
- EDO serves as the foundation of the design for the ontology-driven system that this research provides — the Employer Demand Intelligence Framework (EDIF) for information retrieval and analysis purposes — which has the ability to allow future data mining. EDO can also be applied to the Semantic Web search engines, through which information can be obtained, controlled and examined in an intelligent way. The worldwide attempt to develop and replace the current form of the web with the Semantic Web fits neatly with the development and potential alignment of EDO.

### **9.3.2 Social Significance**

- EDIF can be used as the basis for the development of other applications or systems, to assist in the forecasting of economic indicators such as labour market forecasting.
- EDIF will be able to provide a much wider scope of employer demand intelligence than current tools or approaches are able to offer, and can be easily extended to include further employer demand concepts in EDO as they are deemed appropriate to the domain.

- The semi-automated structure of EDIF will provide current, up-to-the-minute intelligence about the employer demand domain. EDIF can be prompted to gather the latest data from the market at the push of a button, at any time.
- EDIF utilises information that is freely available to all via the web. The output that is delivered via EDIF is not dependent on expensive contracts, or the purchasing of data from employers or other resources.
- EDIF can be easily extended to include gathering and analysing data from an array of sources, and is not solely dependent on information from only one platform on the web.
- The semi-automated nature of EDIF frees up valuable and scarce resources that can now be put to good use in other areas of concern in employer demand domain research.
- Because EDIF is a framework that is hosted on the web, the research can be performed globally from any location, merely by accessing the internet, and as such is not location or personnel dependent.

## **9.4 Future Work**

The Employer Demand Intelligence Framework (EDIF) is an aspiring effort, and a lot of its capacity has been discussed in this thesis; however, due to resource restrictions on this research, there are some limitations and potential enhancements that need to be elucidated and marked as future work. The ontology will need periodic updating to ensure it matches current industry values. It is a common fact that ontologies are mostly continuously evolving. The tools being developed by other experts are making it easier to update ontologies- some tools allowing for semi-automatic updating. The EDIT developed in this research, is one of the tools that has proven its ability to semi-automatically update the EDIF. As the field of ontology and Web 2.0 grows in research, so will the tools too and ontologies will be able to be updated quicker and easier in due time with very little manual input. This will be pursued under the two main components of this research: future work relating to EDO, and future work relating to EDIF.

### 9.4.1 Employer Demand Intelligence Framework Future Work

- EDIF can be evaluated by real users in industry to provide their feedback on aspects like the user interface, what other concepts they would like to be included in EDO, and whether there are any other features they would like to use EDIF for.
- An algorithm could be developed that calculates each job advertisement's level of experience, or expertise required in total for a vacancy. A number of factors would have to be considered for this, such as the weight that would be assigned to each type of skill requirement indicated in the job advertisement, and the number of years' experience and level of management required of the ideal candidate to the vacancy. This could be useful for EDIF stakeholders to learn what are the overall average level of experience and other skill set requirements that employers are looking for in each occupation type. This could also be compared on a geographical level to see if some states require more highly skilled employees than others, and so on.
- Other sources of data can be incorporated into EDIF, such as other job boards, employers' own websites and social media (e.g. LinkedIn, Facebook, and MySpace).
- The SPARQL endpoint on the EDIF<sup>44</sup> website can be made more intuitive, with additional pages attached to the main page, to provide interesting insights into the employer demand intelligence that has been gathered by EDIF over time.
- Enhanced automation of EDIF: to further automate EDIF from its current semi-automatic ability, where the data collection prompt is still dependent on human intervention, would be to provide code that prompts it to collect the data from the sources on a continuous basis each day.
- Automation of EDIF, albeit an advantage in most circumstances, is disadvantageous when it comes to:

---

<sup>44</sup> <http://www.employerdemandintelligence.org>

- The identification of homographs — words that are spelt the same, but have different meanings — are incorrectly recorded in EDD. Even though the differences between a noun and a verb can be detected, these are only apparent when they are used in sentences, when the analysis of sentence structure determines this (the positioning of ‘subject’ → ‘Verb’ → ‘object’ in a sentence). There are many instances where text in job advertisements is not necessarily provided in sentence format, only as bullet points or headers, and thus EDIT cannot distinguish whether it is a noun or a verb. An example is the word ‘contract’, which could mean both ‘an agreement’, or ‘to get or acquire or incur’ something.
- Dealing with spelling errors in the text of the job advertisements to ensure the correct data is captured, and under the correct class. At this stage, human intervention is still needed to deal with this aspect.

#### **9.4.2 Employer Demand Ontology Future Work**

- The use of geospatial vocabulary as part of the location type class in EDO could also be considered.
- Precision and recall verification.
- EDO can be enhanced by adding more instance names (e.g. location or organisation names) to be used for text matching in the documents extracted from the web, allowing more data to be recorded for each advertisement on the web.
- The EDO occupation relationships have been based on the Australian Bureau of Statistics classification model for occupations, which is the model that is used across Australia and New Zealand for most data collections efforts. Future work could include adding dimensions to the EDO occupational classes, by including other factors such as the economic area that the occupations are related to, and evaluating organisation systems like

the Simple Knowledge Organisation System (SKOS) (SKOS Simple Knowledge Organization System, 2012), for possible inclusion in the EDO.

- Presently, the mapping that was done between the SEEK and ANZSCO occupational classifications is contained in a spreadsheet, and has been elucidated through this thesis. This mapping can be formalised or recorded in EDO for reference purposes. One option to do this could be to add annotations to each EDO occupational type, indicating which SEEK category each instance corresponds to.
- To allow EDO to classify a salary as either a Low, Average or High salary, certain parameters would need to be established for each occupation's general level of income by possibly using a fuzzy theory. This will allow the vacancy's salary to be assigned to the correct salary level category in the EDO. Data on the different levels of salary for each location, and for each level of expertise or experience that an employee holds are very difficult to obtain. This is due to the same challenges that this EDIF has been developed to overcome — the manual intensive effort and sheer scale of industries that would need to be covered to obtain this information. Two options to resolve these difficulties could be to use an existing employee compensation profile, such as the one developed by Payscale<sup>45</sup> for salary scales in the public domain; alternatively, EDO could develop its own salary scale over time, based on the salary levels that are being recorded for each occupation by EDIF.
- EDO needs to be further developed to include other occupational types and geographical locations to be able to provide intelligence across all industries and areas.
- EDO can be published by following the best practices from the W3C.<sup>46</sup>

---

<sup>45</sup> <http://www.payscale.com/about.asp>

<sup>46</sup> <http://www.w3.org/TR/swbp-vocab-pub/>

- EDD can be published by following best practices, similar as for EDO, and also include making an entry on the Datahub.<sup>47</sup>

### **9.4.3 Employer Demand Intelligence Tool Future Work**

- Running the data collection on a continuous basis in order to obtain information for more than one month at a time. This way, the period that a position has been advertised for on the job portal can be tracked by recording the first and the last day the job advertisement is first recorded in EDD. As such, it could provide a dashboard to EDIF stakeholders such as the government.
- A wider range of concepts can be included in the ontology that EDIF uses as its referencing source. Instead of using a locally stored ontology, EDO can be linked to a wider community of ontologies via the Semantic Web. For example, EDO can be linked to other ontologies related to the employer demand intelligence domain, such as those that may be developed by universities about their various course offerings; these could be linked to EDO to provide additional information about the types of courses that universities are offering that match those qualifications and skill sets that employers are looking for.
- For the SPARQL query examples, the dates used to query the EDD have been modelled as Literals in this research. The data types can be improved in future work.

## **9.5 Conclusion**

This chapter has provided a brief summary of the work completed for this PhD research, from a review of the literature of current employer demand intelligence-related approaches, to how the problems that need to be addressed have been challenged and responded to. This research project has shown that, while developing a comprehensive Employer Demand Intelligence Framework (EDIF) is an ambitious endeavour, it nevertheless is possible when the framework is developed incrementally. That said, there are still rather challenging areas in this research that

---

<sup>47</sup> <http://datahub.io/>

need further development, and as such, future work tasks have been outlined in the final section of this chapter to allow complete understanding of the framework's ability as it stands at the time of writing this thesis.

---

# References

- Access Economics. (2009). Employment forecasts. Retrieved from <http://www.accesseconomics.com.au/publicationsreports/publications.php>
- Advani, A., Tu, S., & Musen, M. (1998). Domain Modeling with Integrated Ontologies: Principles for Reconciliation and Reuse. *Technical Report SMI-97-0681*. Stanford Medical Informatics.
- Almeida, J., & Simões, A. (2006). *Recycling Thesauri into a Multilingual Ontology*. Paper presented at the Fifth International Conference on Language Resources and Evaluation, LREC 2006. Retrieved from <http://alfarrabio.di.uminho.pt/~albie/publications/lrec06.pdf>
- Arcodia, C., & Barker, T. (2003). A Review of Web-Based Job Advertisements for Australian Event Management Positions. *Journal of Human Resources in Hospitality & Tourism*, 1(4), 1 - 18.
- Australia. Department of Employment and Workplace Relations. (2005). Workforce Tomorrow: Adapting to a more diverse Australian Labour Market. Retrieved from [http://careers.sbc.vic.edu.au/reports/files/Workforce\\_Tomorrow.pdf](http://careers.sbc.vic.edu.au/reports/files/Workforce_Tomorrow.pdf)
- Australia. Department of Education, Employment and Workplace Relations. (2009). Skilled Occupation List (SOL) and Employer Nomination Scheme Occupation List (ENSOL). Retrieved from <http://www.immi.gov.au/allforms/pdf/1121i.pdf>
- Australia. Department of Education Employment and Workplace Relations. (2010). New List of Skilled Occupations Intended to Replace the Current Skilled Occupation List. Retrieved from <http://www.immi.gov.au/skilled/general-skilled-migration/pdf/new-list-of-occupations.pdf>
- Australia. Department of Education Employment and Workplace Relations. (2010). Vacancy report *Vacancy report*. Retrieved from [http://foi.deewr.gov.au/system/files/doc/other/child\\_care\\_vacancies\\_quarterly\\_snapshot\\_december\\_2010.pdf](http://foi.deewr.gov.au/system/files/doc/other/child_care_vacancies_quarterly_snapshot_december_2010.pdf)
- Australian Bureau of Statistics. (2001). ADGC Remoteness Classification. Retrieved from [http://www.abs.gov.au/Websitedbs/D3110122.NSF/4a255eef008309e44a255eef00061e57/f9c96fb635cce780ca256d420005dc02/\\$FILE/appendix\\_A.pdf](http://www.abs.gov.au/Websitedbs/D3110122.NSF/4a255eef008309e44a255eef00061e57/f9c96fb635cce780ca256d420005dc02/$FILE/appendix_A.pdf)
- Australian Bureau of Statistics. (2011). Remoteness Structure. Retrieved from <http://www.abs.gov.au/websitedbs/D3310114.nsf/home/remoteness+structure>
- Australian Bureau of Statistics. (2012a). About the Census Retrieved from <http://www.abs.gov.au/websitedbs/censushome.nsf/home/about?opendocument&navpos=100>
- Australian Bureau of Statistics. (2012b). What is the Census? Retrieved from <http://www.abs.gov.au/websitedbs/censushome.nsf/home/what?opendocument&navpos=110>

- 
- Australian Chamber of Commerce and Industry. ACCI Review: Leading Australian Business. (2009) (Vol. 156). Melbourne
- Baker, N. (2013). Definition of a Job Advertisement. Retrieved from [http://www.ehow.com/facts\\_6781868\\_definition-job-advertisement.html](http://www.ehow.com/facts_6781868_definition-job-advertisement.html)
- Ballinger, C. (2012). Taking a more objective view to employability. *Strategic HR Review*, 11(1).
- Benjamins, V. R., Fensel, D., & Gomez-Perez, A. (1998). *Knowledge management through ontologies*. Paper presented at the CEUR Workshop. Retrieved from [http://oa.upm.es/6239/1/Knowledge\\_Management\\_through\\_Ontologies.pdf](http://oa.upm.es/6239/1/Knowledge_Management_through_Ontologies.pdf)
- Bergman, M. (2010). A New Methodology for Building Lightweight, Domain Ontologies Retrieved from <http://www.mkbergman.com/908/a-new-methodology-for-building-lightweight-domain-ontologies/>
- Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The Semantic Web. *Scientific American*, 34-43. Retrieved from <http://www.scientificamerican.com/article.cfm?id=the-semantic-web>
- Biesalski, E., & Abecker, A. (2005). Human Resource Management with Ontologies. In K.-D. Althoff, A. Dengel, R. Bergmann, M. Nick & T. Roth-Berghofer (Eds.), *Professional Knowledge Management* (Vol. 3782, pp. 499-507). Berlin :Springer.
- Bizer, C., Heese, R., Mochol, M., Oldakowski, R., Tolksdorf, R. & Eckstein, R.. (2005). The Impact of Semantic Web Technologies on Job Recruitment Processes. In O. Ferstl, E. Sinz, S. Eckert & T. Isselhorst (Eds.), *Wirtschaftsinformatik 2005* (pp. 1367-1381): Physica-Verlag HD.
- Brank, J., Grobelnik, M., & Dunja, M. (2005). *A survey of ontology evaluation techniques*. Paper presented at the Conference on Data Mining and Data Warehouses. Retrieved from <http://eprints.pascal-network.org/archive/00001198/01/BrankEvaluationSiKDD2005.pdf>
- Bray, G. (1988). *Sub-State Data*. Paper presented at the Australian Skills Shortages and Labour Market Data, Sydney.
- Calderon, A., & Stockdale, L. (2008). *Skills shortages and future skills requirements for the labour force*. Paper presented at the 11th Australian Vocational Education and Training Research Association Conference, Adelaide, Australia.
- Celino, I., Cerizza, D., Cesarini, M., Valle, E. D., Paoli, F., Estublier, J., . . . Zhao, G. (2010). SEEMP: A Networked Marketplace for Employment Services. Retrieved from [http://dx.doi.org/10.1007/978-3-642-03507-4\\_7](http://dx.doi.org/10.1007/978-3-642-03507-4_7)
- Chamber of Minerals and Energy WA. (2009). *Developing a Growth Outlook for WA's Minerals and Energy Industry*. Perth.
- Chandrasekaran, B., & Johnson, T. (Eds.). (1993). *Generic Tasks and Task Structures: History, Critique and New Directions*.
- Chang, E., Dillon, T., & Hussain, F. (2006). *Trust and Reputation for Service Oriented Environments: Technologies for Building Business Intelligence and Consumer Confidence*. Hoboken, NJ: John Wiley & Sons.

- 
- Chen, M. M. (2013). *Lipoprotein Ontology*. Unpublished Ph.D thesis, Curtin University, Perth, Western Australia.
- Corazzon, C. (2013). *Theory and History of Ontology*. Retrieved from <http://www.ontology.co/>
- Crosby, P. B. (1979). *Quality if Free: The Art of Making Quality Certain*. New York: New American Library.
- Cunningham, H., Maynard, D., Bontcheva, K., Tablan, V., Aswani, N., Roberts, I., . . . Peters, W. (2013). Developing Language Processing Components with GATE Version 7 (a User Guide). Retrieved from <http://gate.ac.uk/sale/tao/split.html>
- Curtain, R. (1988). *Addressing Skill Shortages: The Limitations of Macro-Level Labour Market Forecasting and an Alternative Approach*. Paper presented at the Australian Skills Shortages and Labour Market Data, Sydney.
- de Hoog, R. (1998). Methodologies for building knowledge based systems: Achievements and prospects. *The Handbook of Applied Expert Systems, 1*, 11.
- Department of Employment and Workplace Relations. (2008-09). Skill Shortage Methodology. Retrieved from <http://www.workplace.gov.au/NR/ronlyres/09321BDC-74EA-444A-B20D-9B3025DBC42B/0/WorkplaceMethodologyPaper.pdf>
- Department of Employment and Workplace Relations. (2007). Perth Labour Force Region Skills in Demand Survey. Perth. Retrieved from <http://foi.deewr.gov.au/documents/perth-labour-force-region-skills-demand-survey-2007>
- Department of Employment Education Training and Youth Affairs. (1998). *Skills in Australia : trends and shortages*. Canberra.
- Department of Training and Workforce Development. (2009). WA workforce development plan: A skilled workforce for the future (pp. 7). Perth. Retrieved from <http://www.dtwd.wa.gov.au/dtwd/detcms/apprenticeships-and-training/training-and-workforce-development/wa-workforce-development-plan-a-skilled-workforce-for-the-future.en?page=5&tab=Main#toc5>
- Department of Training and Workforce Development. (2010). *Workforce development plan*. Perth: Retrieved from <http://www.dtwd.wa.gov.au/dtwd/detcms/navigation/workforce-development-plan/>
- Dorn, J., Naz, T., & Pichlmair, M. (2007). *Ontology development for human resource management*. Paper presented at the 4th International Conference on Knowledge Management. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.205.194&rep=rep1&type=pdf>
- Duerst, M. (2005). Internationalized Resource Identifiers (IRIs). Retrieved from <http://www.ietf.org/rfc/rfc3987.txt>
- Easterbrook, S. (2010). The difference between Verification and Validation. Retrieved from <http://www.easterbrook.ca/steve/2010/11/the-difference-between-verification-and-validation/>

- 
- European Centre for the Development of Vocational Training. (2010). Identifying skills needs. Retrieved from <http://www.cedefop.europa.eu/EN/identifying-skills-needs/index.aspx>
- Fernandez-Lopez, M. (1999). *Overview Of Methodologies For Building Ontologies*. Paper presented at the Proceedings of the IJCAI-99 Workshop on Ontologies and Problem-Solving Methods: Lessons Learned and Future Trends, CEUR, Aachen, Germany. Retrieved from [http://oa.upm.es/5480/1/Overview\\_Of\\_Methodologies.pdf](http://oa.upm.es/5480/1/Overview_Of_Methodologies.pdf)
- Fernandez, M., Gomez-Perez, A., & Juristo, N. (1997). *METHONTOLOGY: From Ontological Art Towards Ontological Engineering*. Paper presented at the AAAI-97 Spring Symposium on Ontological Engineering, Stanford University. Retrieved from [http://oa.upm.es/5484/1/METHONTOLOGY\\_.pdf](http://oa.upm.es/5484/1/METHONTOLOGY_.pdf)
- Forbes, D. E., Wongthongtham, P., Singh, J., & Thompson, S. (2013 (in press)). Ontology Supported Assistive Communications in Healthcare. *Communications of the Association for Information Systems*.
- Fuller, D. E., & Oxley, S. P. (1987). *An Industry Studies Framework- Methods of Assessing Industry Performance and Skill Requirements*. Paper presented at the Canberra CAE.
- Galliers, R. D. (1992). *Research issues in information systems as Organisational Processes*. Paper presented at the Third Australian Conference on Information Systems, Wollongong, Australia.
- Gangemi, A., Catenacci, C., Ciaramita, M., & Lehmann, J. (2006). *Modelling ontology evaluation and validation*. Paper presented at the 3rd European Semantic Web Conference (ESWC2006), LNCS, Budva. Retrieved from [http://www.loa.istc.cnr.it/Papers/ESWC\\_GangemiEtAl\\_Final2.pdf](http://www.loa.istc.cnr.it/Papers/ESWC_GangemiEtAl_Final2.pdf)
- García-Sánchez, F., Martínez-Béjar, R., Contreras, L., Fernández-Breis, J. T., & Castellanos-Nieves, D. (2006). An ontology-based intelligent system for recruitment. *Expert Systems with Applications*, 31(2), 248-263. doi: 10.1016/j.eswa.2005.09.023
- Goble, C. (2003). *The Grid Needs You*. Paper presented at the International Conference on Cooperative Object Oriented Information Systems.
- Gomez-Perez, A. (2001). Ontology Evaluation. *International Journal of Intelligent Systems*, 16(3), 391-409.
- Gomez-Perez, A., Fernandez-Lopez, M., & Corcho, O. (Eds.). (2003). *Ontological Engineering*.
- Gomez-Perez, A., Juristo, N., & Pazos, J. (1995). Evaluation and assessment of knowledge sharing technology. In N. J. Mars (Ed.), *Towards very large knowledge bases: Knowledge building and knowledge sharing* (pp. 289-296). [place of pub]: IOS Press.
- Gomez-Perez, A., Motta, E., & Suarez-Figueroa, M. C. Introduction to the NeOn Methodology Retrieved 21 May, 2013, from <http://www.neon-project.org/web-content/media/book-chapters/Chapter-01.pdf>
- Gómez-Pérez, A., Ramírez, J., & Villazón-Terrazas, B. (2007). An Ontology for Modelling Human Resources Management Based on Standards. Retrieved from [http://dx.doi.org/10.1007/978-3-540-74819-9\\_66](http://dx.doi.org/10.1007/978-3-540-74819-9_66)

- 
- Gordon, M., & Gordon, B. L. J. (1980). *The Shortage of skilled people: proceedings of the conference*. Newcastle, N.S.W: Institute of Industrial Economics, University of Newcastle.
- Green, F., Machin, S., & Wilkinson, D. (1998). The Meaning and Determinants of Skills Shortages. *Oxford Bulletin of Economics and Statistics*, 60(2), 165-187. doi: 10.1111/1468-0084.00093
- Gruber, T. (n.d.). What is an Ontology? Retrieved from <http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>
- Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2), 199-220. doi: <http://dx.doi.org/10.1006/knac.1993.1008>
- Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing? *International Journal of Human-Computer Studies*, 43(5-6), 907-928. doi: <http://dx.doi.org/10.1006/ijhc.1995.1081>
- Gruber, T. R. (1995). Towards Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal of Human-Computer Studies*, 43(5/6), 907-928.
- Gruninger, M. (1995). *Methodology for the design and evaluations of ontologies*. Paper presented at the IJCAI95 Workshop on Basic Ontological Issues in Knowledge Sharing, Montreal.
- Gruninger, M., & Fox, M. S. (1994). *The Design and Evaluation of Ontologies for Enterprise Engineering*. Paper presented at the Workshop on Implemented Ontologies, European Conference on Artificial Intelligence, Amsterdam. Retrieved from <http://stl.mie.utoronto.ca/publications/gruninger-onto-ecai94.pdf>
- Gruninger, M., & Lee, J. (2002). Ontology Applications and Design. *Communications of the ACM*, 45(2), 39-41.
- Gualtieri, A., & Ruffolo, M. (2005). *An Ontology-Based Framework for Representing Organizational Knowledge*. Paper presented at the I-KNOW'05, Graz, Austria.
- Guarino, N., Oberle, D., & Staab, S. (2009). What is an ontology? In ... (Eds.), *Handbook on Ontologies* (2nd ed., pp. 1-17). London: Springer.
- Hadzic, M., Wongthongtham, P., Dillon, T., & Chang, E. (2009). Ontology-Based Multi-Agent Systems. *Studies in Computational Intelligence*, 219.
- Handel, M. J. (2003). Skills Mismatch in the Labor Market. *Annual Review of Sociology*, 29(1), 135-165. doi:10.1146/annurev.soc.29.010202.100030
- Hayton, G. (1988). *Translation of Labour Market Needs into Training Responses*. Paper presented at the Australian Skills Shortages and Labour Market Data, Sydney.
- Hayton, G., & Cheyne, M. (1988). *An Evaluation of Post-Trade Courses in Fluid Power Relevant to the National Metal Industries Award*. Adelaide: TAFE National Centre for Research and Development.
- Hevner, A., March, S., Park, J., & Ram, S. (2004). Design science in information systems research. *Management Information Systems Quarterly*, 28(1), 75-105.

- 
- Horridge, M., Drummond, N., Jupp, S., Moulton, G., & Stevens, R. (2009). *A Practical Guide To Building OWL Ontologies Using Protégé 4 and CO-ODE Tools Edition 1. 2* Manchester: The University Of Manchester.
- Institute of Electrical and Electronics Engineers (IEEE). (1987). Standard Taxonomy for Software Engineering Standards. *IEEE Std 1002-1987, 0\_1*. doi: 10.1109/ieeestd.1987.122157
- Institute of Electrical and Electronics Engineers (IEEE). (1990). Standard Glossary of Software Engineering Terminology. *IEEE Std 610.12-1990, 1-84*. doi: 10.1109/ieeestd.1990.101064
- Institute of Electrical and Electronics Engineers (IEEE). (1990). Standard Glossary of Data Management Terminology. *IEEE Std 610.5-1990, 0\_1*. doi: 10.1109/ieeestd.1990.94601
- Institute of Electrical and Electronics Engineers (IEEE). (2006). Standard for Developing a Software Project Life Cycle Process. *IEEE Std 1074-2006 (Revision of IEEE Std 1074-1997), 0\_1-104*. doi: 10.1109/ieeestd.2006.219190
- ISO/ IEC 9126. (2013). Retrieved from [http://en.wikipedia.org/wiki/ISO/IEC\\_9126](http://en.wikipedia.org/wiki/ISO/IEC_9126)
- Jarrar, M., & Meersman, R. (2002). *Scalability and Knowledge Reusability in Ontology Modeling*. Paper presented at the International Conference on Infrastructure for e-Business, e-Education, e-Science, and e-Medicine (SSGRR2002). Retrieved from [http://starpc15.vub.ac.be/website/files/ssgrr02s\\_0.pdf](http://starpc15.vub.ac.be/website/files/ssgrr02s_0.pdf)
- Jennings, V. E. (1980). *Towards an Adaptable and Fully-Employed Workforce*. Paper presented at the The Shortage of Skilled People, Newcastle.
- Jobs Online. (2009). Department of Labour: New Zealand. Retrieved from <http://www.dol.govt.nz/publications/jol/index.asp>
- JobsDB Dimension. (2011). JobsDB Dimension: Introduction. Retrieved from <http://www.jobsdb.com/Dimension/>
- Juran, J. M., Gryna, F. M., & Bingham, R. S. (Eds.). (1974). *Quality Control Handbook* (3rd ed.). New York: McGraw-Hill.
- Kashyap, V., & Sheth, A. P. (Eds.). (1996). *Semantic Heterogeneity in Global Information Systems: The Role of Metadata, Context and Ontologies*. Cooperative Information Systems: Current Trends and Directions.
- Kiplagat, J. K., & Rotich, P. K. (2007). Science And Engineering Research: Assessment Methods. In M. Iskander (Ed.), *Innovations in E-learning, Instruction Technology, Assessment, and Engineering Education* (pp. 163-167). Netherlands: Springer.
- Kohn, A., Bry, F., & Manta, A. (2010). Semantic Search on Unstructured Data: Explicit Knowledge through Data Recycling (pp. 17-35): IGI Global.
- Lau, T., & Sure, Y. (2002). Introducing Ontology-based Skills Management at a large Insurance Company. *IN PROCEEDINGS OF THE MODELLIERUNG 2002*, 123-134.
- Library of Congress. Browse or Search the ISO Language Codes. (2010). Retrieved from <http://www.loc.gov/standards/iso639-2/php/langcodes-search.php>

- 
- Lv, H., & Zhu, B. (2006, Nov.). *Skill ontology-based semantic model and its matching algorithm*. Paper presented at the 7th International Conference on Computer-Aided Industrial Design and Conceptual Design. CAIDCD. Retrieved from [http://ieeexplore.ieee.org/xpl/login.jsp?tp=&number=4127158&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs\\_all.jsp%3Farnumber%3D4127158](http://ieeexplore.ieee.org/xpl/login.jsp?tp=&number=4127158&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs_all.jsp%3Farnumber%3D4127158)
- MacDonald, K., & Klinger, P. (2009, September 17). State told to build up trainees for boom, *The West Australian*.
- Maedche, A. (2003). *Ontology Learning for the Semantic Web*. *Kluwer International Series in Engineering and Computer Science*, 665. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.73.4559&rep=rep1&type=pdf>
- Marchante, A. J., Ortega, B., & Pagán, R. (2006). Determinants of skills shortages and hard-to-fill vacancies in the hospitality sector. *Tourism Management*, 27(5), 791-802. doi: DOI: 10.1016/j.tourman.2005.05.013
- Massey, L., & Wong, W. (2011). A Cognitive Based Approach to Identify Topics in Text Using the Web as a Knowledge Source. In W. Wong, W. Liu & M. Bennamoun (Eds.), *Ontology Learning and Knowledge Discovery Using the Web: Challenges and Recent Advances* (pp. 61-78). Perth, Western Australia: Information Science Reference.
- McCarthy, P. (2005). Search RDF data with SPARQL. Retrieved from <http://www.ibm.com/developerworks/xml/library/j-sparql/>
- Mochol, M., & Simperl, E. (2006). *Practical Guidelines for Building Semantic eRecruitment Applications*. Paper presented at the International Conference on Knowledge Management IKNOW 06, Special Track: Advanced Semantic Technologies AST'06, Graz, Austria. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.91.8052&rep=rep1&type=pdf>
- Mules, T. (2007). *A job vacancy series for tourism and hospitality: An exploratory study*. Paper presented at the CAUTHE 2007 Conference, Manly, Australia. Retrieved from <http://search.informit.com.au/documentSummary;dn=533742443305159;res=IELBUS>
- Neches, R., Fikes, R. E., Gruber, T. R., Senator, T., & Swarout, W. R. (1991). Enabling technology for knowledge sharing. *AI Magazine*, 12(3), 36-56.
- Negru, S., & Lohmann, S. (2013). VOWL: Visual Notation for OWL Ontologies, from <http://vowl.visualdataweb.org/v1/>
- Noy, N., & McGuinness, D. (2001). *Ontology Development 101: A Guide to Creating Your First Ontology* *Stanford Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880*. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.136.5085>
- Nunamaker, J. F. J., Chen, M., & Purdin, T. D. M. (1991). Systems Development in Information Systems Research. *Journal of Management Information Systems*, 7(3), 89-106.
- O'Callaghan, A. (2008). *Latest Trends in Recruitment & Selection*. Johannesburg: FASSET.

- 
- Obitko, M. (2007). Reasoning. *Ontologies and Semantic Web* Retrieved 8 September, 2013, from <http://www.obitko.com/tutorials/ontologies-semantic-web/reasoning.html>
- Obrst, L., Ceusters, W., Mani, I., Ray, S., & Smith, B. (2007). The evaluation of ontologies. In C. J. O. Baker & K. Cheung (Eds.), *Revolutionizing KnowledgeDiscovery in the Life Sciences* (pp. 139-158). Berlin: Springer.
- Oliver, J. M., & Turton, J. R. (1982). Is there a shortage of skilled labour? *British Journal of Industrial Relations*, 20(2), 195-200. doi: 10.1111/j.1467-8543.1982.tb00097.x
- OWL Working Group. (2012). Web Ontology Language (OWL). Retrieved from <http://www.w3.org/2001/sw/wiki/OWL>
- Poveda-Villalon, M., Suarez-Figueroa, M. C., & Gomez-Perez, A. (2012). *Validating Ontologies with OOPS!* Paper presented at the 18th International Conference on Knowledge Engineering and Knowledge Management. Retrieved from <http://oa.upm.es/14384/1/mpovedaEtAlCameraReady.pdf>
- Probyn, A. (2009, September 2). PM plans a blue-collar boom, *The West Australian*.
- Rawling, S. (1988). *Using Labour Market Data in Careers Advice and Councelling*. Paper presented at the Australian Skills Shortages and Labour Market Data, Sydney.
- Reed, S. L., & Lenat, D. B. (2002). *Mapping Ontologies into Cyc*. Paper presented at the AAAI 2002 Conference Workshop on Ontologies For The Semantic Web, Edmonton, Canada. Retrieved from <http://aaai.org/Papers/Workshops/2002/WS-02-11/WS02-11-010.pdf>
- Richardson, S. (2009). What is a Skill Shortage? *Australian Bulletin of Labour*, 35(1), 326.
- Robinson, R., & Beesley, L. (2010). Linkages between creativity and intention to quit: An occupational study of chefs. *Tourism Management*, 31(6), 765-776. doi: DOI: 10.1016/j.tourman.2009.08.003
- Robinson, R. N. S., Arcodia, C. V., Tian, C., & Charlton, P. (2010). Cookery-related employment vacancy advertisements and skills shortages. *International Journal of Contemporary Hospitality Management*, 22(4), 483-499.
- Senker, P. (1992). Skill Shortages and Britain's International Competitiveness. In D. Bosworth, P. Dutton & J. Lewis (Eds.), *Skill Shortages: Causes and Consequences* (p. 206). London: Avebury.
- Shah, C., & Burke, G. (2005). Skills Shortages: Concepts, Measurement and Policy Responses. *Australian Bulletin of Labour*, 31(1), 44-71.
- Sidhu, A. S. (2008). *Structured Vocabularies for Proteins: Protein Ontology Project*. La Trobe University, Melbourne.
- Simperl, E. (2009). Reusing ontologies on the Semantic Web: A feasibility study. *Data & Knowledge Engineering*, 68(10), 905-925. doi: 10.1016/j.datak.2009.02.002
- Skill Shortages. (2007) Retrieved from <http://www.workplace.gov.au/workplace/Publications/LabourMarketAnalysis/SkillShortages/>

- 
- Skills Australia. (2010). Australian Workforce Futures: A National Workforce Development Strategy. Retrieved from [http://www.skillsaustralia.gov.au/PDFs\\_RTFs/WWF\\_strategy.pdf](http://www.skillsaustralia.gov.au/PDFs_RTFs/WWF_strategy.pdf)
- SKOS Simple Knowledge Organization System (2012). Retrieved from <http://www.w3.org/2004/02/skos/>
- Sloan, J. (2010). *The workforce planning imperative* (1st ed.). Adelaide: Julie Sloan Management.
- Smalberger, C., Wongthongtham, P., Villazon-Terrazas, B., & Forbes, D. E. (2013 (under review)). Ontology-based Employer Demand Management. *Management Information Systems Quarterly*.
- SMN Ontology. (2008). Retrieved from <http://www.anandavala.info/ontology/SMNOntology/smn-ontology.html>
- Sowa, J. F. (2009). Building, Sharing, and Merging Ontologies. Retrieved from <http://www.jfsowa.com/ontology/ontoshar.htm>
- SPARQL. (2013) Retrieved from <http://en.wikipedia.org/wiki/SPARQL>
- Stevens, R., Goble, C. A., & Bechhofer, S. (2000). Ontology based knowledge representation for bioinformatics. *Briefings in Bioinformatics*, 1(4), 398-414.
- Storey, K. (2001). Fly-in/Fly-out and Fly-over: mining and regional development in Western Australia. *Australian Geographer*, 32(2), 133 - 148.
- Strohmaier, M., Walk, S., Pöschko, J., Lamprecht, D., Tudorache, T., Nyulas, C., . . . Noy, N. F. (2013). How ontologies are made: Studying the hidden social dynamics behind collaborative ontology engineering projects. *Web Semantics: Science, Services and Agents on the World Wide Web*. doi: <http://dx.doi.org/10.1016/j.websem.2013.04.001>
- Studer, R., Benjamins, V. R., & Fensel, D. (1998). Knowledge Engineering: Principles and Methods. *Data & Knowledge Engineering*, 25, 161-197.
- Suarez-Figueroa, M. C. (2010). NeOn Methodology for building ontology networks: Specification, scheduling and reuse. PhD thesis. Retrieved from [http://oa.upm.es/3879/2/MARIA\\_DEL\\_CARMEN\\_SUAREZ\\_DE\\_FIGUEROA\\_BAONZA.pdf](http://oa.upm.es/3879/2/MARIA_DEL_CARMEN_SUAREZ_DE_FIGUEROA_BAONZA.pdf)
- Suarez-Figueroa, M. C., Gomez-Perez, A., & Fernandez-Lopez, M. (2012). The NeOn Methodology for Ontology Engineering. In M. C. Suárez Figueroa, A. Gomez-Perez, E. Motta & A. Gangemi (Eds.), *Ontology Engineering in a Networked World*.
- Suarez-Figueroa, M. C., Gomez-Perez, A., & Villason-Terrazas, B. (2009). *How to Write and Use the Ontology Requirements Specification Document*. Paper presented at the Proceedings of the Confederated International Conferences, CoopIS, DOA, IS, and ODBASE, On the Move to Meaningful Internet Systems: Part II, Vilamoura, Portugal. Retrieved from <http://delicias.dia.fi.upm.es/wiki/images/b/b0/ORequirementsSpecification.pdf>
- Sure, Y., Staab, S., & Studer, R. (2009). Ontology Engineering Methodology. In S. Staab & R. Studer (Eds.), *Handbook on Ontologies*. Berlin: Springer-Verlag.

- 
- Sure, Y., Tempich, C., & Vrandecic, D. (2006). Ontology Engineering Methodologies. *Semantic Web Technologies: Trends and Research in Ontology-based Systems*, 171-187.
- Sweet, R. (1988). *TAFE and Labour Market Data Requirements*. Paper presented at the Australian Skills Shortages and Labour Market Data, Sydney.
- Swoogle Semantic Web Search. (2007). Retrieved from <http://swoogle.umbc.edu/>
- Terziev, I., Kiryakov, A., & Manov, D. (2005). Base upper-level ontology (BULO) guidance. In S. S. E. K. Technologies (Ed.). Retrieved from [http://proton.semanticweb.org/D1\\_8\\_1.pdf](http://proton.semanticweb.org/D1_8_1.pdf)
- Trichet, F., & Leclère, M. (2003). A Framework for Building Competency-Based Systems Dedicated to Human Resource Management. In N. Zhong, Z. Raś, S. Tsumoto & E. Suzuki (Eds.), *Foundations of Intelligent Systems* (Vol. 2871, pp. 633-639). Berlin: Springer.
- Turnbull, R. G. (1980). *The Government's Response: Australia*. Paper presented at the The Shortage of Skilled People, Newcastle.
- Uschold, M., & Gruninger, M. (1996). Ontologies: principles, methods and applications. *The Knowledge Engineering Review*, 11(02), 93-136. doi: 10.1017/S0269888900007797
- Vacancy. (2013). Cambridge Dictionaries Online: Cambridge University Press. Retrieved from <http://dictionary.cambridge.org/dictionary/business-english/vacancy>
- Van Ours, J., & Ridder, G. (1992). Vacancies and the Recruitment of New Employees. *Journal of Labor Economics*, 10(2), 138-155.
- Vaughan, G. (1988). *Occupations or Skills?* Paper presented at the Australian Skills Shortages and Labour Market Data, Sydney.
- Villazón-Terrazas, B., Ramírez, J., Suárez-Figueroa, M. C., & Gómez-Pérez, A. (2011). A network of ontology networks for building e-employment advanced systems. *Expert Systems with Applications*, 38(11), 13612-13624. doi: 10.1016/j.eswa.2011.04.125
- Vrandecic, D. (2009). Ontology Evaluation. In S. Staab & R. Studer (Eds.), *Handbook on Ontologies* (pp. 293-313). Berlin: Springer-Verlag.
- W3C SPARQL Query Language for RDF. (2013). Retrieved from <http://www.w3.org/TR/rdf-sparql-query/#initDefinitions>
- Warwick Institute for Employment Research. (2010a). Labour Market Assessment & Forecasting. Retrieved from <http://www2.warwick.ac.uk/fac/soc/ier/research/forecast>
- Warwick Institute for Employment Research. (2010b). Local Economy Forecasting Model (LEFM). Retrieved from <http://www2.warwick.ac.uk/fac/soc/ier/publications/software/lefm/>
- Western Australia. Department of Training and Workforce Development. (2009). WA workforce development plan: A skilled workforce for the future. Perth. Retrieved from <http://www.trainingwa.wa.gov.au/dtwd/detcms/apprenticeships-and-training/training-and-workforce-development/wa-workforce-development-plan-a-skilled-workforce-for-the-future.en?page=5&tab=Main#toc5>

- 
- Western Australia. Department of Training and Workforce Development. (2010). *Workforce development plan*. Perth: Retrieved from <http://www.dtwd.wa.gov.au/dtwd/detcms/navigation/workforce-development-plan/>
- Wilcock, G. (2009). *Introduction to Linguistic Annotation and Text Analytics*. [place of pub.]: Morgan & Claypool.
- Williams, B. (1988). *Higher Education Needs (with Special Reference to Engineering)*. Paper presented at the Australian Skills Shortages and Labour Market Data, Sydney.
- Wilson, R. (1988). *Occupational Vacancy Data*. Paper presented at the Australian Skills Shortages and Labour Market Data, Sydney.
- Wongthongtham, P. (2006). *A Methodology for Multi-site Distributed Software Development*. Perth: Curtin University.
- Zemmouchi-Ghomari, L. (2012). HERO Higher Education Reference Ontology. Retrieved from <http://herontology.esi.dz/content/downloads>

Every reasonable effort has been made to acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been incorrectly acknowledged or omitted.

---

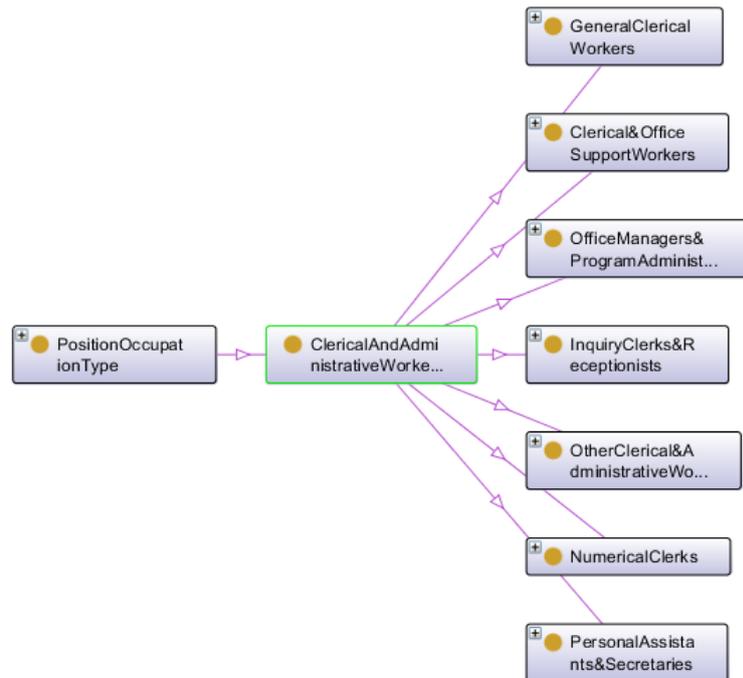
# Appendices

**Appendix 1 - OntoGraf presentations of the remaining seven  
*PositionOccupationType* subclasses of the EDO .....314**

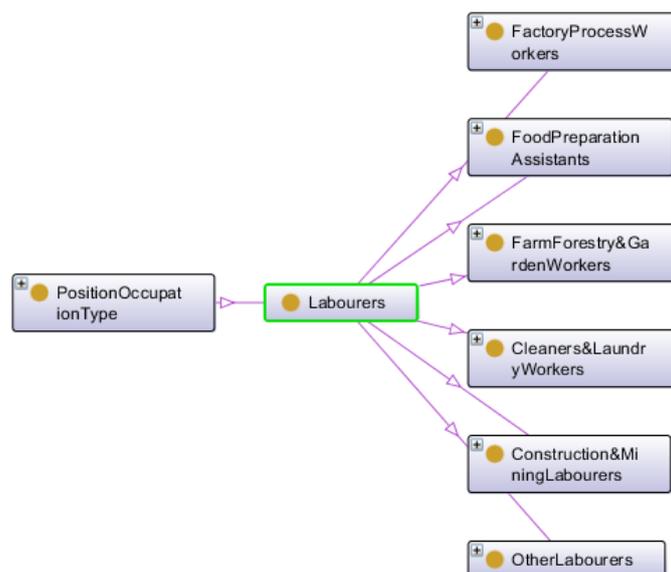
---

## Appendix 1 - OntoGraf presentations of the remaining seven *PositionOccupationType* subclasses of the EDO

### 1. Clerical and Administrative Workers

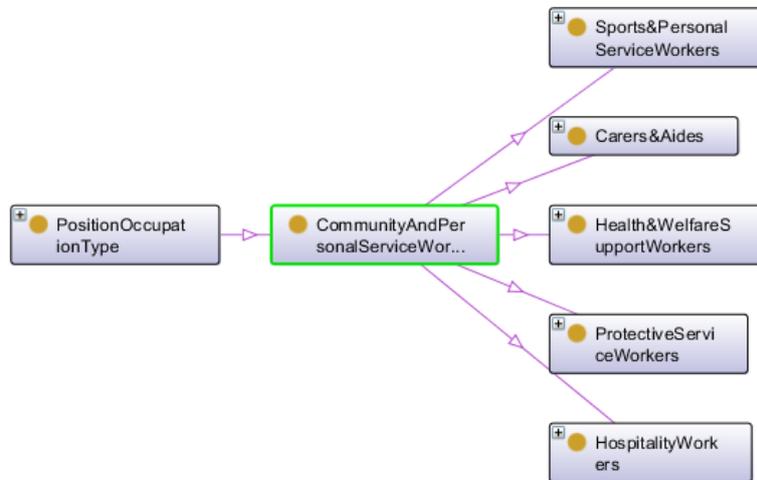


### 2. Labourers

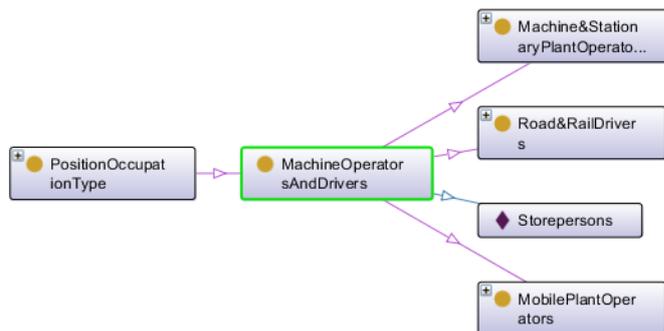


---

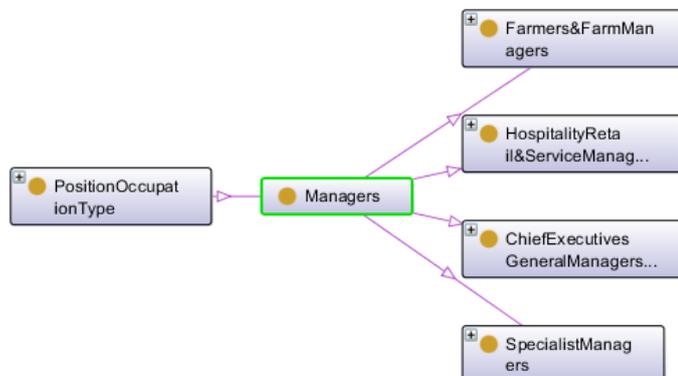
### 3. Community and Personal Service Workers



### 4. Machine Operators and Drivers

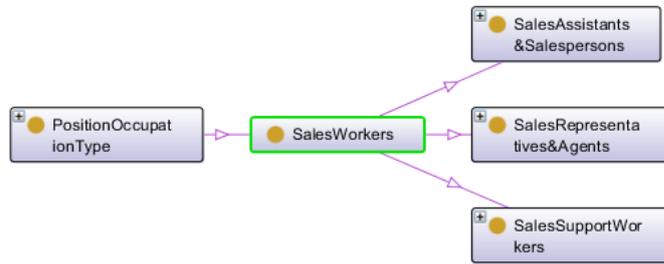


### 5. Managers



---

## 6. Sales Workers



## 7. Technicians and Trade Workers

